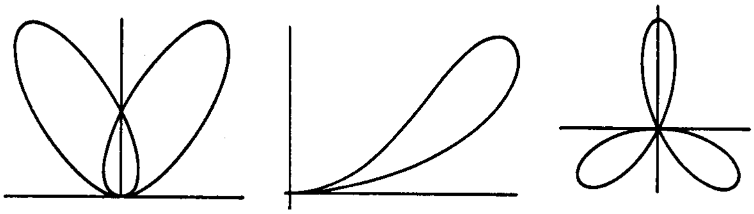
Algebraic Geometry

J.S. Milne



Version 6.02
March 19, 2017

These notes are an introduction to the theory of algebraic varieties emphasizing the similarities to the theory of manifolds. In contrast to most such accounts they study abstract algebraic varieties, and not just subvarieties of affine and projective space. This approach leads more naturally into scheme theory.


```
BibTeX information
@misc{milneAG,
  author={Milne, James S.},
  title={Algebraic Geometry (v6.02)},
  year={2017},
  note={Available at www.jmilne.org/math/},
  pages={221}
}
```

v2.01 (August 24, 1996). First version on the web.

v3.01 (June 13, 1998).

v4.00 (October 30, 2003). Fixed errors; many minor revisions; added exercises; added two sections/chapters; 206 pages.

v5.00 (February 20, 2005). Heavily revised; most numbering changed; 227 pages.

v5.10 (March 19, 2008). Minor fixes; TEXstyle changed, so page numbers changed; 241 pages.

v5.20 (September 14, 2009). Minor corrections; revised Chapters 1, 11, 16; 245 pages.

v5.22 (January 13, 2012). Minor fixes; 260 pages.

v6.00 (August 24, 2014). Major revision; 223 pages.

v6.01 (August 23, 2015). Minor fixes; 226 pages.

v6.02 (March 19, 2017). Minor fixes; 221 pages.

Available at www.jmilne.org/math/
Please send comments and corrections to me at the address on my web page.

The curves are a tacnode, a ramphoid cusp, and an ordinary triple point.

Contents

Contents **3**

Introduction **7**

1 Preliminaries from commutative algebra **11**

a. Rings and ideals, 11 ; b. Rings of fractions, 15 ; c. Unique factorization, 21 ; d. Integral dependence, 24; e. Tensor Products, 30 ; f. Transcendence bases, 33; Exercises, 33.

2 Algebraic Sets **35**

a. Definition of an algebraic set, 35 ; b. The Hilbert basis theorem, 36; c. The Zariski topology, 37; d. The Hilbert Nullstellensatz, 38; e. The correspondence between algebraic sets and radical ideals, 39 ; f. Finding the radical of an ideal, 43; g. Properties of the Zariski topology, 43; h. Decomposition of an algebraic set into irreducible algebraic sets, 44 ; i. Regular functions; the coordinate ring of an algebraic set, 47; j. Regular maps, 48; k. Hypersurfaces; finite and quasi-finite maps, 48; l. Noether normalization theorem, 50 ; m. Dimension, 52 ; Exercises, 56.

3 Affine Algebraic Varieties **57**

a. Sheaves, 57 ; b. Ringed spaces, 58; c. The ringed space structure on an algebraic set, 59 ; d. Morphisms of ringed spaces, 62 ; e. Affine algebraic varieties, 63; f. The category of affine algebraic varieties, 64; g. Explicit description of morphisms of affine varieties, 65 ; h. Subvarieties, 68; i. Properties of the regular map $\mathrm{Spm}(\alpha)$, 69; j. Affine space without coordinates, 70; k. Birational equivalence, 71; l. Noether Normalization Theorem, 72; m. Dimension, 73 ; Exercises, 77.

4 Local Study **79**

a. Tangent spaces to plane curves, 79 ; b. Tangent cones to plane curves, 81 ; c. The local ring at a point on a curve, 83; d. Tangent spaces to algebraic subsets of $\mathbb{A}^m$, 84 ; e. The differential of a regular map, 86; f. Tangent spaces to affine algebraic varieties, 87 ; g. Tangent cones, 91; h. Nonsingular points; the singular locus, 92 ; i. Nonsingularity and regularity, 94; j. Examples of tangent spaces, 95; Exercises, 96.

5 Algebraic Varieties **97**

a. Algebraic prevarieties, 97; b. Regular maps, 98; c. Algebraic varieties, 99; d. Maps from varieties to affine varieties, 101; e. Subvarieties, 101 ; f. Prevarieties obtained by patching, 102; g. Products of varieties, 103 ; h. The separation axiom revisited, 108; i. Fibred products, 110 ; j. Dimension, 111; k. Dominant maps, 113; l. Rational maps; birational equivalence, 113; m. Local study, 114; n. Étale maps, 115 ; o. Étale neighbourhoods, 118 ; p. Smooth maps, 120 ; q. Algebraic varieties as a functors, 121 ; r. Rational and unirational varieties, 124 ; Exercises, 125.

6 Projective Varieties **127**

a. Algebraic subsets of $\mathbb{P}^n$, 127; b. The Zariski topology on $\mathbb{P}^n$, 131; c. Closed subsets of $\mathbb{A}^n$ and $\mathbb{P}^n$, 132 ; d. The hyperplane at infinity, 133; e. $\mathbb{P}^n$ is an algebraic variety, 133; f. The homogeneous coordinate ring of a projective variety, 135; g. Regular functions on a projective variety, 136; h. Maps from projective varieties, 137; i. Some classical maps of projective varieties, 138 ; j. Maps to projective space, 143; k. Projective space without coordinates, 143; l. The functor defined by projective space, 144; m. Grassmann varieties, 144 ; n. Bezout's theorem, 148; o. Hilbert polynomials (sketch), 149; p. Dimensions, 150; q. Products, 152 ; Exercises, 153.

7 Complete Varieties **155**

a. Definition and basic properties, 155 ; b. Proper maps, 157; c. Projective varieties are complete, 158 ; d. Elimination theory, 159 ; e. The rigidity theorem; abelian varieties, 163; f. Chow's Lemma, 165 ; g. Analytic spaces; Chow's theorem, 167; h. Nagata's Embedding Theorem, 168 ; Exercises, 169.

8 Normal Varieties; (Quasi-)finite maps; Zariski's Main Theorem **171**

a. Normal varieties, 171 ; b. Regular functions on normal varieties, 174 ; c. Finite and quasi-finite maps, 176 ; d. The fibres of finite maps, 182; e. Zariski's main theorem, 184 ; f. Stein factorization, 189; g. Blow-ups, 190 ; h. Resolution of singularities, 190 ; Exercises, 191.

9 Regular Maps and Their Fibres **193**

a. The constructibility theorem, 193; b. The fibres of morphisms, 196; c. Flat maps and their fibres, 199 ; d. Lines on surfaces, 206; e. Bertini's theorem, 211; f. Birational classification, 211; Exercises, 212.

Solutions to the exercises **213**

Index **219**

Notations

We use the standard (Bourbaki) notations: $\mathbb{N} = \{0, 1, 2, \ldots\}$, $\mathbb{Z}$ = ring of integers, $\mathbb{R}$ = field of real numbers, $\mathbb{C}$ = field of complex numbers, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ = field of p elements, p a prime number. Given an equivalence relation, $[*]$ denotes the equivalence class containing $*$. A family of elements of a set A indexed by a second set I, denoted $(a_i)_{i \in I}$, is a function $i \mapsto a_i : I \to A$. We sometimes write $|S|$ for the number of elements in a finite set S.

Throughout, k is an algebraically closed field. Unadorned tensor products are over k. For a k-algebra R and k-module M, we often write M_R for $R \otimes M$. The dual $\mathrm{Hom}_{k\text{-linear}}(E, k)$ of a finite-dimensional k-vector space E is denoted by $E^\vee$.

All rings will be commutative with 1, and homomorphisms of rings are required to map 1 to 1.

We use Gothic (fraktur) letters for ideals:

$\mathfrak{a}$	$\mathfrak{b}$	$\mathfrak{c}$	$\mathfrak{m}$	$\mathfrak{n}$	$\mathfrak{p}$	$\mathfrak{q}$	$\mathfrak{A}$	$\mathfrak{B}$	$\mathfrak{C}$	$\mathfrak{M}$	$\mathfrak{N}$	$\mathfrak{P}$	$\mathfrak{Q}$
a	b	c	m	n	p	q	A	B	C	M	N	P	Q

Finally

$X \stackrel{\text{def}}{=} Y$	X is defined to be Y, or equals Y by definition;
$X \subset Y$	X is a subset of Y (not necessarily proper, i.e., X may equal Y);
$X \approx Y$	X and Y are isomorphic;
$X \simeq Y$	X and Y are canonically isomorphic (or there is a given or unique isomorphism).

A reference "Section 3m" is to Section m in Chapter 3; a reference "(3.45)" is to this item in chapter 3; a reference "(67)" is to (displayed) equation 67 (usually given with a page reference unless it is nearby).

Prerequisites

The reader is assumed to be familiar with the basic objects of algebra, namely, rings, modules, fields, and so on.

References

Atiyah and MacDonald 1969: Introduction to Commutative Algebra, Addison-Wesley.
CA: Milne, J.S., Commutative Algebra, v4.02, 2017.
FT: Milne, J.S., Fields and Galois Theory, v4.52, 2017.
Hartshorne 1977: Algebraic Geometry, Springer.
Shafarevich 1994: Basic Algebraic Geometry, Springer.

A reference monnnn (resp. sxnnnn) is to question nnnn on mathoverflow.net (resp. math.stackexchange.com).

We sometimes refer to the computer algebra programs

CoCoA (**Co**mputations in **Co**mmutative **A**lgebra) http://cocoa.dima.unige.it/.

Macaulay 2 (Grayson and Stillman) http://www.math.uiuc.edu/Macaulay2/.

Acknowledgements

I thank the following for providing corrections and comments on earlier versions of these notes: Jorge Nicolás Caro Montoya, Sandeep Chellapilla, Rankeya Datta, Umesh V. Dubey, Shalom Feigelstock, Tony Feng, B.J. Franklin, Sergei Gelfand, Daniel Gerig, Darij Grinberg, Lucio Guerberoff, Isac Hedén, Guido Helmers, Florian Herzig, Christian Hirsch, Cheuk-Man Hwang, Jasper Loy Jiabao, Dan Karliner, Lars Kindler, John Miller, Joaquin Rodrigues, Sean Rostami, David Rufino, Hossein Sabzrou, Jyoti Prakash Saha, Tom Savage, Nguyen Quoc Thang, Bhupendra Nath Tiwari, Israel Vainsencher, Soli Vishkautsan, Dennis Bouke Westra, Felipe Zaldivar, Luochen Zhao, and others.

QUESTION: If we try to explain to a layman what algebraic geometry is, it seems to me that the title of the old book of Enriques is still adequate: Geometrical Theory of Equations ....
GROTHENDIECK: Yes! but your "layman" should know what a system of algebraic equations is. This would cost years of study to Plato.
QUESTION: It should be nice to have a little faith that after two thousand years every good high school graduate can understand what an affine scheme is ...

From the notes of a lecture series that Grothendieck gave at SUNY at Buffalo in the summer of 1973 (in 167 pages, Grothendieck manages to cover very little).

Introduction

There is almost nothing left to discover in geometry.
Descartes, March 26, 1619

Just as the starting point of linear algebra is the study of the solutions of systems of linear equations,

$$\sum_{j=1}^{n} a_{ij} X_j = b_i, \quad i = 1, \ldots, m, \tag{1}$$

the starting point for algebraic geometry is the study of the solutions of systems of polynomial equations,

$$f_i(X_1, \ldots, X_n) = 0, \quad i = 1, \ldots, m, \quad f_i \in k[X_1, \ldots, X_n].$$

One immediate difference between linear equations and polynomial equations is that theorems for linear equations don't depend on which field k you are working over,[1] but those for polynomial equations depend on whether or not k is algebraically closed and (to a lesser extent) whether k has characteristic zero.

A better description of algebraic geometry is that it is the study of polynomial functions and the spaces on which they are defined (algebraic varieties), just as topology is the study of continuous functions and the spaces on which they are defined (topological spaces), differential topology the study of infinitely differentiable functions and the spaces on which they are defined (differentiable manifolds), and so on:

algebraic geometry	regular (polynomial) functions	algebraic varieties
topology	continuous functions	topological spaces
differential topology	differentiable functions	differentiable manifolds
complex analysis	analytic (power series) functions	complex manifolds.

The approach adopted in this course makes plain the similarities between these different areas of mathematics. Of course, the polynomial functions form a much less rich class than the others, but by restricting our study to polynomials we are able to do calculus over any field: we simply define

$$\frac{d}{dX} \sum a_i X^i = \sum i a_i X^{i-1}.$$

Moreover, calculations with polynomials are easier than with more general functions.

[1] For example, suppose that the system (1) has coefficients $a_{ij} \in k$ and that K is a field containing k. Then (1) has a solution in k^n if and only if it has a solution in K^n, and the dimension of the space of solutions is the same for both fields. (Exercise!)

Consider a nonzero differentiable function $f(x,y,z)$. In calculus, we learn that the equation

$$f(x,y,z) = C \tag{2}$$

defines a surface S in $\mathbb{R}^3$, and that the tangent plane to S at a point $P = (a,b,c)$ has equation[2]

$$\left(\frac{\partial f}{\partial x}\right)_P (x-a) + \left(\frac{\partial f}{\partial y}\right)_P (y-b) + \left(\frac{\partial f}{\partial z}\right)_P (z-c) = 0. \tag{3}$$

The inverse function theorem says that a differentiable map $\alpha\colon S \to S'$ of surfaces is a local isomorphism at a point $P \in S$ if it maps the tangent plane at P isomorphically onto the tangent plane at $P' = \alpha(P)$.

Consider a nonzero polynomial $f(x,y,z)$ with coefficients in a field k. In these notes, we shall learn that the equation (2) defines a surface in k^3, and we shall use the equation (3) to define the tangent space at a point P on the surface. However, and this is one of the essential differences between algebraic geometry and the other fields, the inverse function theorem doesn't hold in algebraic geometry. One other essential difference is that $1/X$ is not the derivative of any rational function of X, and nor is X^{np-1} in characteristic $p \neq 0$ — these functions can not be integrated in the field of rational functions $k(X)$.

These notes form a basic course on algebraic geometry. Throughout, we require the ground field to be algebraically closed in order to be able to concentrate on the geometry. Additional chapters, treating more advanced topics, can be found on my website.

The approach to algebraic geometry taken in these notes

In differential geometry it is important to define differentiable manifolds abstractly, i.e., not as submanifolds of some Euclidean space. For example, it is difficult even to make sense of a statement such as "the Gauss curvature of a surface is intrinsic to the surface but the principal curvatures are not" without the abstract notion of a surface.

Until the mid 1940s, algebraic geometry was concerned only with algebraic subvarieties of affine or projective space over algebraically closed fields. Then, in order to give substance to his proof of the congruence Riemann hypothesis for curves and abelian varieties, Weil was forced to develop a theory of algebraic geometry for "abstract" algebraic varieties over arbitrary fields,[3] but his "foundations" are unsatisfactory in two major respects:

- ◇ Lacking a sheaf theory, his method of patching together affine varieties to form abstract varieties is clumsy.[4]
- ◇ His definition of a variety over a base field k is not intrinsic; specifically, he fixes some large "universal" algebraically closed field Ω and defines an algebraic variety over k to be an algebraic variety over Ω together with a k-structure.

In the ensuing years, several attempts were made to resolve these difficulties. In 1955, Serre resolved the first by borrowing ideas from complex analysis and defining an algebraic variety over an algebraically closed field to be a topological space with a sheaf of functions that is locally affine.[5] Then, in the late 1950s Grothendieck resolved all such difficulties by developing the theory of schemes.

[2] Think of S as a level surface for the function f, and note that the equation is that of a plane through (a,b,c) perpendicular to the gradient vector $(\nabla f)_P$ of f at P.

[3] Weil, André. Foundations of algebraic geometry. American Mathematical Society, Providence, R.I. 1946.

[4] Nor did Weil use the Zariski topology in 1946.

[5] Serre, Jean-Pierre. Faisceaux algébriques cohérents. Ann. of Math. (2) 61, (1955). 197–278, commonly referred to as FAC.

In these notes, we follow Grothendieck except that, by working only over a base field, we are able to simplify his language by considering only the closed points in the underlying topological spaces. In this way, we hope to provide a bridge between the intuition given by differential geometry and the abstractions of scheme theory.

CHAPTER **1**

Preliminaries from commutative algebra

Algebraic geometry and commutative algebra are closely intertwined. For the most part, we develop the necessary commutative algebra in the context in which it is used. However, in this chapter, we review some basic definitions and results from commutative algebra.

a Rings and ideals

Basic definitions

Let A be a ring. A ***subring*** of A is a subset that contains 1_A and is closed under addition, multiplication, and the formation of negatives. An A-***algebra*** is a ring B together with a homomorphism $i_B: A \to B$. A ***homomorphism of*** A-***algebras*** $B \to C$ is a homomorphism of rings $\varphi: B \to C$ such that $\varphi(i_B(a)) = i_C(a)$ for all $a \in A$.

Elements $x_1, \ldots, x_n$ of an A-algebra B are said to ***generate*** it if every element of B can be expressed as a polynomial in the x_i with coefficients in $i_B(A)$, i.e., if the homomorphism of A-algebras $A[X_1, \ldots, X_n] \to B$ acting as i_A on A and sending X_i to x_i is surjective.

When $A \subset B$ and $x_1, \ldots, x_n \in B$, we let $A[x_1, \ldots, x_n]$ denote the A-subalgebra of B generated by the x_i.

A ring homomorphism $A \to B$ is said to be of ***finite-type***, and B is a ***finitely generated*** A-algebra if B is generated by a finite set of elements as an A-algebra.

A ring homomorphism $A \to B$ is ***finite***, and B is a ***finite***[1] A-algebra, if B is finitely generated as an A-module.

Let k be a field, and let A be a k-algebra. When $1_A \neq 0$ in A, the map $k \to A$ is injective, and we can identify k with its image, i.e., we can regard k as a subring of A. When $1_A = 0$ in a ring A, then A is the zero ring, i.e., $A = \{0\}$.

A ring is an ***integral domain*** if it is not the zero ring and if $ab = 0$ implies that $a = 0$ or $b = 0$; in other words, if $ab = ac$ and $a \neq 0$, then $b = c$.

For a ring A, $A^\times$ is the group of elements of A with inverses (the units in the ring).

Ideals

Let A be a ring. An ***ideal*** $\mathfrak{a}$ in A is a subset such that

[1] The term "module-finite" is also used.

(a) $\mathfrak{a}$ is a subgroup of A regarded as a group under addition;

(b) $a \in \mathfrak{a}$, $r \in A \Rightarrow ra \in \mathfrak{a}$.

The ***ideal generated by a subset*** S of A is the intersection of all ideals $\mathfrak{a}$ containing S — it is easy to see that this is in fact an ideal, and that it consists of all finite sums of the form $\sum r_i s_i$ with $r_i \in A$, $s_i \in S$. The ideal generated by the empty set is the zero ideal $\{0\}$. When $S = \{s_1, s_2, \ldots\}$, we write $(s_1, s_2, \ldots)$ for the ideal it generates.

Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals in A. The set $\{a+b \mid a \in \mathfrak{a}, b \in \mathfrak{b}\}$ is an ideal, denoted by $\mathfrak{a}+\mathfrak{b}$. The ideal generated by $\{ab \mid a \in \mathfrak{a}, b \in \mathfrak{b}\}$ is denoted by $\mathfrak{ab}$. Clearly $\mathfrak{ab}$ consists of all finite sums $\sum a_i b_i$ with $a_i \in \mathfrak{a}$ and $b_i \in \mathfrak{b}$, and if $\mathfrak{a} = (a_1, \ldots, a_m)$ and $\mathfrak{b} = (b_1, \ldots, b_n)$, then $\mathfrak{ab} = (a_1 b_1, \ldots, a_i b_j, \ldots, a_m b_n)$. Note that

$$\mathfrak{ab} \subset \mathfrak{a} \cap \mathfrak{b}. \tag{4}$$

The kernel of a homomorphism $A \to B$ is an ideal in A. Conversely, for any ideal $\mathfrak{a}$ in A, the set of cosets of $\mathfrak{a}$ in A forms a ring $A/\mathfrak{a}$, and $a \mapsto a + \mathfrak{a}$ is a homomorphism $\varphi\colon A \to A/\mathfrak{a}$ whose kernel is $\mathfrak{a}$. The map $\mathfrak{b} \mapsto \varphi^{-1}(\mathfrak{b})$ is a one-to-one correspondence between the ideals of $A/\mathfrak{a}$ and the ideals of A containing $\mathfrak{a}$.

An ideal $\mathfrak{p}$ is ***prime*** if $\mathfrak{p} \neq A$ and $ab \in \mathfrak{p} \Rightarrow a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. Thus $\mathfrak{p}$ is prime if and only if $A/\mathfrak{p}$ is nonzero and has the property that

$$ab = 0 \implies a = 0 \text{ or } b = 0,$$

i.e., $A/\mathfrak{p}$ is an integral domain. Note that if $\mathfrak{p}$ is prime and $a_1 \cdots a_n \in \mathfrak{p}$, then at least one of the $a_i \in \mathfrak{p}$.

An ideal $\mathfrak{m}$ in A is ***maximal*** if it is maximal among the proper ideals of A. Thus $\mathfrak{m}$ is maximal if and only if $A/\mathfrak{m}$ is nonzero and has no proper nonzero ideals, and so is a field. Note that

$$\mathfrak{m} \text{ maximal} \implies \mathfrak{m} \text{ prime}.$$

The ideals of $A \times B$ are all of the form $\mathfrak{a} \times \mathfrak{b}$ with $\mathfrak{a}$ and $\mathfrak{b}$ ideals in A and B. To see this, note that if $\mathfrak{c}$ is an ideal in $A \times B$ and $(a,b) \in \mathfrak{c}$, then $(a,0) = (1,0)(a,b) \in \mathfrak{c}$ and $(0,b) = (0,1)(a,b) \in \mathfrak{c}$. Therefore, $\mathfrak{c} = \mathfrak{a} \times \mathfrak{b}$ with

$$\mathfrak{a} = \{a \mid (a,0) \in \mathfrak{c}\}, \quad \mathfrak{b} = \{b \mid (0,b) \in \mathfrak{c}\}.$$

Ideals $\mathfrak{a}$ and $\mathfrak{b}$ in A are ***coprime*** (or ***relatively prime***) if $\mathfrak{a} + \mathfrak{b} = A$. Assume that $\mathfrak{a}$ and $\mathfrak{b}$ are coprime, and let $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$ be such that $a + b = 1$. For $x, y \in A$, let $z = ay + bx$; then

$$\begin{aligned} z &\equiv bx \equiv x \bmod \mathfrak{a} \\ z &\equiv ay \equiv y \bmod \mathfrak{b}, \end{aligned}$$

and so the canonical map

$$A \to A/\mathfrak{a} \times A/\mathfrak{b} \tag{5}$$

is surjective. Clearly its kernel is $\mathfrak{a} \cap \mathfrak{b}$, which contains $\mathfrak{ab}$. Let $c \in \mathfrak{a} \cap \mathfrak{b}$; then

$$c = c1 = ca + cb \in \mathfrak{ab}.$$

Hence, (5) is surjective with kernel $\mathfrak{ab}$. This statement extends to finite collections of ideals.

THEOREM 1.1 (CHINESE REMAINDER THEOREM). *Let $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ be ideals in a ring A. If $\mathfrak{a}_i$ is coprime to $\mathfrak{a}_j$ whenever $i \neq j$, then the map*

$$A \to A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n \tag{6}$$

is surjective, with kernel $\prod \mathfrak{a}_i = \bigcap \mathfrak{a}_i$.

PROOF. We have proved the statement for $n = 2$, and we use induction to extend it to $n > 2$. For $i \geq 2$, there exist elements $a_i \in \mathfrak{a}_1$ and $b_i \in \mathfrak{a}_i$ such that

$$a_i + b_i = 1.$$

The product $\prod_{i \geq 2}(a_i + b_i)$ lies in $\mathfrak{a}_1 + \mathfrak{a}_2 \cdots \mathfrak{a}_n$ and equals 1, and so

$$\mathfrak{a}_1 + \mathfrak{a}_2 \cdots \mathfrak{a}_n = A.$$

Therefore,

$$\begin{aligned} A/\mathfrak{a}_1 \cdots \mathfrak{a}_n &= A/\mathfrak{a}_1 \cdot (\mathfrak{a}_2 \cdots \mathfrak{a}_n) \\ &\simeq A/\mathfrak{a}_1 \times A/\mathfrak{a}_2 \cdots \mathfrak{a}_n && \text{by the } n = 2 \text{ case} \\ &\simeq A/\mathfrak{a}_1 \times A/\mathfrak{a}_2 \times \cdots \times A/\mathfrak{a}_n && \text{by induction.} \end{aligned}$$

□

We let $\operatorname{spec}(A)$ denote the set of prime ideals in a ring A and $\operatorname{spm}(A)$ the set of maximal ideals.

Noetherian rings

PROPOSITION 1.2. *The following three conditions on a ring A are equivalent:*

(a) *every ideal in A is finitely generated;*

(b) *every ascending chain of ideals $\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots$ eventually becomes constant, i.e., $\mathfrak{a}_m = \mathfrak{a}_{m+1} = \cdots$ for some m;*

(c) *every nonempty set of ideals in A has a maximal element.*

PROOF. (a) $\implies$ (b): Let $\mathfrak{a}_1 \subset \mathfrak{a}_2 \subset \cdots$ be an ascending chain of ideals. Then $\bigcup \mathfrak{a}_i$ is an ideal, and hence has a finite set $\{a_1, \ldots, a_n\}$ of generators. For some m, all the a_i belong to $\mathfrak{a}_m$, and then

$$\mathfrak{a}_m = \mathfrak{a}_{m+1} = \cdots = \bigcup \mathfrak{a}_i.$$

(b) $\implies$ (c): Let Σ be a nonempty set of ideals in A. If Σ has no maximal element, then the axiom of dependent choice[2] implies that there exists an infinite strictly ascending chain of ideals in Σ, contradicting (b).

(c) $\implies$ (a): Let $\mathfrak{a}$ be an ideal, and let Σ be the set of finitely generated ideals contained in $\mathfrak{a}$. Then Σ is nonempty because it contains the zero ideal, and so it contains a maximal element $\mathfrak{c} = (a_1, \ldots, a_r)$. If $\mathfrak{c} \neq \mathfrak{a}$, then there exists an $a \in \mathfrak{a} \smallsetminus \mathfrak{c}$, and $(a_1, \ldots, a_r, a)$ will be a finitely generated ideal in $\mathfrak{a}$ properly containing $\mathfrak{c}$. This contradicts the definition of $\mathfrak{c}$, and so $\mathfrak{c} = \mathfrak{a}$. □

[2] This says the following: let R be a binary relation on a nonempty set X, and suppose that, for each a in X, there exists a b such that aRb; then there exists a sequence $(a_n)_{n \in \mathbb{N}}$ of elements of X such that $a_n R a_{n+1}$ for all n. This axiom is strictly weaker than the axiom of choice (q.v. Wikipedia).

A ring A is ***noetherian*** if it satisfies the equivalent conditions of the proposition. On applying (c) to the set of all proper ideals containing a fixed proper ideal, we see that every proper ideal in a noetherian ring is contained in a maximal ideal. This is, in fact, true for every ring, but the proof for non-noetherian rings requires Zorn's lemma (CA 2.2).

A ring A is said to be ***local*** if it has exactly one maximal ideal $\mathfrak{m}$. Because every nonunit is contained in a maximal ideal, for a local ring $A^{\times} = A \smallsetminus \mathfrak{m}$.

PROPOSITION 1.3 (NAKAYAMA'S LEMMA). *Let A be a local ring with maximal ideal $\mathfrak{m}$, and let M be a finitely generated A-module.*

(a) *If $M = \mathfrak{m}M$, then $M = 0$.*

(b) *If N is a submodule of M such that $M = N + \mathfrak{m}M$, then $M = N$.*

PROOF. (a) Suppose that $M \neq 0$. Choose a minimal set of generators $\{e_1, \ldots, e_n\}$, $n \geq 1$, for M, and write

$$e_1 = a_1 e_1 + \cdots + a_n e_n, \quad a_i \in \mathfrak{m}.$$

Then

$$(1 - a_1)e_1 = a_2 e_2 + \cdots + a_n e_n$$

and, as $(1 - a_1)$ is a unit, $e_2, \ldots, e_n$ generate M, contradicting the minimality of the set.

(b) The hypothesis implies that $M/N = \mathfrak{m}(M/N)$, and so $M/N = 0$. □

Now let A be a local noetherian ring with maximal ideal $\mathfrak{m}$. Then $\mathfrak{m}$ is an A-module, and the action of A on $\mathfrak{m}/\mathfrak{m}^2$ factors through $k \overset{\text{def}}{=} A/\mathfrak{m}$.

COROLLARY 1.4. *Elements $a_1, \ldots, a_n$ of $\mathfrak{m}$ generate $\mathfrak{m}$ as an ideal if and only if their residues modulo $\mathfrak{m}^2$ generate $\mathfrak{m}/\mathfrak{m}^2$ as a vector space over k. In particular, the minimum number of generators for the maximal ideal is equal to the dimension of the vector space $\mathfrak{m}/\mathfrak{m}^2$.*

PROOF. If $a_1, \ldots, a_n$ generate $\mathfrak{m}$, it is obvious that their residues generate $\mathfrak{m}/\mathfrak{m}^2$. Conversely, suppose that their residues generate $\mathfrak{m}/\mathfrak{m}^2$, so that $\mathfrak{m} = (a_1, \ldots, a_n) + \mathfrak{m}^2$. Because A is noetherian, $\mathfrak{m}$ is finitely generated, and Nakayama's lemma shows that $\mathfrak{m} = (a_1, \ldots, a_n)$. □

DEFINITION 1.5. Let A be a noetherian ring.

(a) The ***height*** $\text{ht}(\mathfrak{p})$ of a prime ideal $\mathfrak{p}$ in A is the greatest length d of a chain of distinct prime ideals

$$\mathfrak{p} = \mathfrak{p}_d \supset \mathfrak{p}_{d-1} \supset \cdots \supset \mathfrak{p}_0. \tag{7}$$

(b) The ***Krull dimension*** of A is $\sup\{\text{ht}(\mathfrak{p}) \mid \mathfrak{p} \text{ a prime ideal in } A\}$.

Thus, the Krull dimension of a noetherian ring A is the supremum of the lengths of chains of prime ideals in A (the length of a chain is the number of gaps, so the length of (7) is d). For example, a field has Krull dimension 0, and conversely an integral domain of Krull dimension 0 is a field. The height of every nonzero prime ideal in a principal ideal domain is 1, and so such a ring has Krull dimension 1 (provided it is not a field).

The height of every prime ideal in a noetherian ring is finite, but the Krull dimension of the ring may be infinite because it may contain a sequence of prime ideals $\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{p}_3, \ldots$ such that $\text{ht}(\mathfrak{p}_i)$ tends to infinity (CA, p.13).

DEFINITION 1.6. A local noetherian ring of Krull dimension d is said to be ***regular*** if its maximal ideal can be generated by d elements.

It follows from Corollary 1.4 that a local noetherian ring is regular if and only if its Krull dimension is equal to the dimension of the vector space $\mathfrak{m}/\mathfrak{m}^2$.

LEMMA 1.7. *In a noetherian ring, every set of generators for an ideal contains a finite generating subset.*

PROOF. Let $\mathfrak{a}$ be an ideal in a noetherian ring A, and let S be a set of generators for $\mathfrak{a}$. An ideal maximal among those generated by a finite subset of S must contain every element of S (otherwise it wouldn't be maximal), and so equals $\mathfrak{a}$. □

In the proof of the next theorem, we use that a polynomial ring over a noetherian ring is noetherian (see Lemma 2.8).

THEOREM 1.8 (KRULL INTERSECTION THEOREM). *Let A be a noetherian local ring with maximal ideal $\mathfrak{m}$; then $\bigcap_{n\geq 1}\mathfrak{m}^n = \{0\}$.*

PROOF. Let $a_1,\ldots,a_r$ generate $\mathfrak{m}$. Then $\mathfrak{m}^n$ consists of all finite sums

$$\sum_{i_1+\cdots+i_r=n} c_{i_1\cdots i_r} a_1^{i_1}\cdots a_r^{i_r}, \quad c_{i_1\cdots i_r}\in A.$$

In other words, $\mathfrak{m}^n$ consists of the elements of A of the form $g(a_1,\ldots,a_r)$ for some homogeneous polynomial $g(X_1,\ldots,X_r)\in A[X_1,\ldots,X_r]$ of degree n. Let S_m denote the set of homogeneous polynomials f of degree m such that $f(a_1,\ldots,a_r)\in\bigcap_{n\geq 1}\mathfrak{m}^n$, and let $\mathfrak{a}$ be the ideal in $A[X_1,\ldots,X_r]$ generated by the set $\bigcup_m S_m$. According to the lemma, there exists a finite set $\{f_1,\ldots,f_s\}$ of elements of $\bigcup_m S_m$ that generates $\mathfrak{a}$. Let $d_i = \deg f_i$, and let $d = \max d_i$. Let $b\in\bigcap_{n\geq 1}\mathfrak{m}^n$; then $b\in\mathfrak{m}^{d+1}$, and so $b = f(a_1,\ldots,a_r)$ for some homogeneous polynomial f of degree $d+1$. By definition, $f\in S_{d+1}\subset\mathfrak{a}$, and so

$$f = g_1 f_1 + \cdots + g_s f_s$$

for some $g_i\in A[X_1,\ldots,X_r]$. As f and the f_i are homogeneous, we can omit from each g_i all terms not of degree $\deg f - \deg f_i$, since these terms cancel out. Thus, we may choose the g_i to be homogeneous of degree $\deg f - \deg f_i = d+1-d_i > 0$. Then $g_i(a_1,\ldots,a_r)\in\mathfrak{m}$, and so

$$b = f(a_1,\ldots,a_r) = \sum\nolimits_i g_i(a_1,\ldots,a_r)\cdot f_i(a_1,\ldots,a_r)\in\mathfrak{m}\cdot\bigcap\nolimits_{n\geq 1}\mathfrak{m}^n.$$

Thus, $\bigcap\mathfrak{m}^n = \mathfrak{m}\cdot\bigcap\mathfrak{m}^n$, and Nakayama's lemma implies that $\bigcap\mathfrak{m}^n = 0$. □

ASIDE 1.9. Let A be the ring of germs of analytic functions at $0\in\mathbb{R}$ (see p.58 for the notion of a germ of a function). Then A is a noetherian local ring with maximal ideal $\mathfrak{m} = (x)$, and $\mathfrak{m}^n$ consists of the functions f that vanish to order n at $x = 0$. The theorem says (correctly) that only the zero function vanishes to all orders at 0. By contrast, the function e^{-1/x^2} shows that the Krull intersection theorem fails for the ring of germs of infinitely differentiable functions at 0 (this ring is not noetherian).

b Rings of fractions

A ***multiplicative subset*** of a ring A is a subset S with the property:

$$1\in S, \quad a,b\in S \implies ab\in S.$$

Define an equivalence relation on $A \times S$ by

$$(a,s) \sim (b,t) \iff u(at-bs) = 0 \text{ for some } u \in S.$$

Write $\frac{a}{s}$ or a/s for the equivalence class containing (a,s), and define addition and multiplication of equivalence classes in the way suggested by the notation:

$$\tfrac{a}{s} + \tfrac{b}{t} = \tfrac{at+bs}{st}, \quad \tfrac{a}{s}\tfrac{b}{t} = \tfrac{ab}{st}.$$

It is easy to check that these do not depend on the choices of representatives for the equivalence classes, and that we obtain in this way a ring

$$S^{-1}A = \left\{\frac{a}{s} \mid a \in A, s \in S\right\}$$

and a ring homomorphism $a \mapsto \frac{a}{1}\colon A \to S^{-1}A$, whose kernel is

$$\{a \in A \mid sa = 0 \text{ for some } s \in S\}.$$

For example, if A is an integral domain an $0 \notin S$, then $a \mapsto \frac{a}{1}$ is injective, but if $0 \in S$, then $S^{-1}A$ is the zero ring.

Write i for the homomorphism $a \mapsto \frac{a}{1}\colon A \to S^{-1}A$.

PROPOSITION 1.10. *The pair $(S^{-1}A, i)$ has the following universal property: every element $s \in S$ maps to a unit in $S^{-1}A$, and any other homomorphism $A \to B$ with this property factors uniquely through i:*

$$\begin{tikzcd} A \arrow[r, "i"] \arrow[dr] & S^{-1}A \arrow[d, dotted, "\exists!"] \\ & B \end{tikzcd}$$

PROOF. If β exists,

$$s\tfrac{a}{s} = a \implies \beta(s)\beta(\tfrac{a}{s}) = \beta(a) \implies \beta(\tfrac{a}{s}) = \alpha(a)\alpha(s)^{-1},$$

and so β is unique. Define

$$\beta(\tfrac{a}{s}) = \alpha(a)\alpha(s)^{-1}.$$

Then

$$\tfrac{a}{c} = \tfrac{b}{d} \implies s(ad-bc) = 0 \text{ some } s \in S \implies \alpha(a)\alpha(d) - \alpha(b)\alpha(c) = 0$$

because $\alpha(s)$ is a unit in B, and so β is well-defined. It is obviously a homomorphism. □

As usual, this universal property determines the pair $(S^{-1}A, i)$ uniquely up to a unique isomorphism.

When A is an integral domain and $S = A \smallsetminus \{0\}$, $F = S^{-1}A$ is the field of fractions of A, which we denote $F(A)$. In this case, for any other multiplicative subset T of A not containing 0, the ring $T^{-1}A$ can be identified with the subring $\{\frac{a}{t} \in F \mid a \in A, t \in S\}$ of F.

We shall be especially interested in the following examples.

EXAMPLE 1.11. Let $h \in A$. Then $S_h = \{1, h, h^2, \ldots\}$ is a multiplicative subset of A, and we let $A_h = S_h^{-1}A$. Thus every element of A_h can be written in the form a/h^m, $a \in A$, and

$$\frac{a}{h^m} = \frac{b}{h^n} \iff h^N(ah^n - bh^m) = 0, \quad \text{some } N.$$

If h is nilpotent, then $A_h = 0$, and if A is an integral domain with field of fractions F and $h \neq 0$, then A_h is the subring of F of elements of the form a/h^m, $a \in A$, $m \in \mathbb{N}$.

EXAMPLE 1.12. Let $\mathfrak{p}$ be a prime ideal in A. Then $S_\mathfrak{p} = A \smallsetminus \mathfrak{p}$ is a multiplicative subset of A, and we let $A_\mathfrak{p} = S_\mathfrak{p}^{-1}A$. Thus each element of $A_\mathfrak{p}$ can be written in the form $\frac{a}{c}$, $c \notin \mathfrak{p}$, and

$$\frac{a}{c} = \frac{b}{d} \iff s(ad - bc) = 0, \text{ some } s \notin \mathfrak{p}.$$

The subset $\mathfrak{m} = \{\frac{a}{s} \mid a \in \mathfrak{p}, s \notin \mathfrak{p}\}$ is a maximal ideal in $A_\mathfrak{p}$, and it is the only maximal ideal, i.e., $A_\mathfrak{p}$ is a local ring.[3] When A is an integral domain with field of fractions F, $A_\mathfrak{p}$ is the subring of F consisting of elements expressible in the form $\frac{a}{s}$, $a \in A$, $s \notin \mathfrak{p}$.

LEMMA 1.13. *For every ring A and $h \in A$, the map $\sum a_i X^i \mapsto \sum \frac{a_i}{h^i}$ defines an isomorphism*

$$A[X]/(1 - hX) \xrightarrow{\simeq} A_h.$$

PROOF. If $h = 0$, both rings are zero, and so we may assume $h \neq 0$. In the ring $A[x] = A[X]/(1 - hX)$, $1 = hx$, and so h is a unit. Let $\alpha\colon A \to B$ be a homomorphism of rings such that $\alpha(h)$ is a unit in B. The homomorphism $\sum a_i X^i \mapsto \sum \alpha(a_i)\alpha(h)^{-i}\colon A[X] \to B$ factors through $A[x]$ because $1 - hX \mapsto 1 - \alpha(h)\alpha(h)^{-1} = 0$, and, because $\alpha(h)$ is a unit in B, this is the unique extension of α to $A[x]$. Therefore $A[x]$ has the same universal property as A_h, and so the two are (uniquely) isomorphic by an isomorphism that fixes elements of A and makes h^{-1} correspond to x. $\square$

Let S be a multiplicative subset of a ring A, and let $S^{-1}A$ be the corresponding ring of fractions. Any ideal $\mathfrak{a}$ in A, generates an ideal $S^{-1}\mathfrak{a}$ in $S^{-1}A$. If $\mathfrak{a}$ contains an element of S, then $S^{-1}\mathfrak{a}$ contains a unit, and so is the whole ring. Thus some of the ideal structure of A is lost in the passage to $S^{-1}A$, but, as the next proposition shows, much is retained.

PROPOSITION 1.14. *Let S be a multiplicative subset of the ring A. The map*

$$\mathfrak{p} \mapsto S^{-1}\mathfrak{p} = (S^{-1}A)\mathfrak{p}$$

is a bijection from the set of prime ideals of A disjoint from S to the set of prime ideals of $S^{-1}A$ with inverse $\mathfrak{q} \mapsto$(inverse image of $\mathfrak{q}$ in A).

PROOF. For an ideal $\mathfrak{b}$ of $S^{-1}A$, let $\mathfrak{b}^c$ be the inverse image of $\mathfrak{b}$ in A, and for an ideal $\mathfrak{a}$ of A, let $\mathfrak{a}^e = (S^{-1}A)\mathfrak{a}$ be the ideal in $S^{-1}A$ generated by the image of $\mathfrak{a}$.

For an ideal $\mathfrak{b}$ of $S^{-1}A$, certainly, $\mathfrak{b} \supset \mathfrak{b}^{ce}$. Conversely, if $\frac{a}{s} \in \mathfrak{b}$, $a \in A$, $s \in S$, then $\frac{a}{1} \in \mathfrak{b}$, and so $a \in \mathfrak{b}^c$. Thus $\frac{a}{s} \in \mathfrak{b}^{ce}$, and so $\mathfrak{b} = \mathfrak{b}^{ce}$.

For an ideal $\mathfrak{a}$ of A, certainly $\mathfrak{a} \subset \mathfrak{a}^{ec}$. Conversely, if $a \in \mathfrak{a}^{ec}$, then $\frac{a}{1} \in \mathfrak{a}^e$, and so $\frac{a}{1} = \frac{a'}{s}$ for some $a' \in \mathfrak{a}$, $s \in S$. Thus, $t(as - a') = 0$ for some $t \in S$, and so $ast \in \mathfrak{a}$. If $\mathfrak{a}$ is a prime ideal disjoint from S, this implies that $a \in \mathfrak{a}$: for such an ideal, $\mathfrak{a} = \mathfrak{a}^{ec}$.

[3] First check $\mathfrak{m}$ is an ideal. Next, if $\mathfrak{m} = A_\mathfrak{p}$, then $1 \in \mathfrak{m}$; but if $1 = \frac{a}{s}$ for some $a \in \mathfrak{p}$ and $s \notin \mathfrak{p}$, then $u(s - a) = 0$ some $u \notin \mathfrak{p}$, and so $ua = us \notin \mathfrak{p}$, which contradicts $a \in \mathfrak{p}$. Finally, $\mathfrak{m}$ is maximal because every element of $A_\mathfrak{p}$ not in $\mathfrak{m}$ is a unit.

If $\mathfrak{b}$ is prime, then certainly $\mathfrak{b}^c$ is prime. For any ideal $\mathfrak{a}$ of A, $S^{-1}A/\mathfrak{a}^e \simeq \bar{S}^{-1}(A/\mathfrak{a})$ where $\bar{S}$ is the image of S in $A/\mathfrak{a}$. If $\mathfrak{a}$ is a prime ideal disjoint from S, then $\bar{S}^{-1}(A/\mathfrak{a})$ is a subring of the field of fractions of $A/\mathfrak{a}$, and is therefore an integral domain. Thus, $\mathfrak{a}^e$ is prime.

We have shown that $\mathfrak{p} \mapsto \mathfrak{p}^e$ and $\mathfrak{q} \mapsto \mathfrak{q}^c$ are inverse bijections between the prime ideals of A disjoint from S and the prime ideals of $S^{-1}A$. □

LEMMA 1.15. *Let $\mathfrak{m}$ be a maximal ideal of a ring A, and let $\mathfrak{n} = \mathfrak{m}A_{\mathfrak{m}}$. For all n, the map*

$$a + \mathfrak{m}^n \mapsto \tfrac{a}{1} + \mathfrak{n}^n \colon A/\mathfrak{m}^n \to A_{\mathfrak{m}}/\mathfrak{n}^n \tag{8}$$

is an isomorphism. Moreover, it induces isomorphisms

$$\mathfrak{m}^r/\mathfrak{m}^n \to \mathfrak{n}^r/\mathfrak{n}^n$$

for all $r < n$.

PROOF. The second statement follows from the first, because of the exact commutative diagram ($r < n$):

$$\begin{tikzcd}
0 \arrow[r] & \mathfrak{m}^r/\mathfrak{m}^n \arrow[r] \arrow[d] & A/\mathfrak{m}^n \arrow[r] \arrow[d, "\simeq"] & A/\mathfrak{m}^r \arrow[r] \arrow[d, "\simeq"] & 0 \\
0 \arrow[r] & \mathfrak{n}^r/\mathfrak{n}^n \arrow[r] & A_{\mathfrak{m}}/\mathfrak{n}^n \arrow[r] & A_{\mathfrak{m}}/\mathfrak{n}^r \arrow[r] & 0.
\end{tikzcd}$$

Let $S = A \smallsetminus \mathfrak{m}$. Then $A_{\mathfrak{m}} = S^{-1}A$ and $\mathfrak{n}^n = \mathfrak{m}^n A_{\mathfrak{m}} = \{\frac{b}{s} \in A_{\mathfrak{m}} \mid b \in \mathfrak{m}^n, s \in S\}$. In order to show that the map (8) is injective, it suffices to show that

$$\tfrac{a}{1} = \tfrac{b}{s} \text{ with } a \in A, b \in \mathfrak{m}^n, s \in S \implies a \in \mathfrak{m}^n.$$

But if $\frac{a}{1} = \frac{b}{s}$, then $tas = tb \in \mathfrak{m}^n$ for some $t \in S$, and so $tas = 0$ in $A/\mathfrak{m}^n$. The only maximal ideal in A containing $\mathfrak{m}^m$ is $\mathfrak{m}$ (because $\mathfrak{m}' \supset \mathfrak{m}^m \implies \mathfrak{m}' \supset \mathfrak{m}$), and so the only maximal ideal in $A/\mathfrak{m}^n$ is $\mathfrak{m}/\mathfrak{m}^n$. As st is not in $\mathfrak{m}/\mathfrak{m}^n$, it must be a unit in $A/\mathfrak{m}^n$, and as $sta = 0$ in $A/\mathfrak{m}^n$, a must be 0 in $A/\mathfrak{m}^n$, i.e., $a \in \mathfrak{m}^n$.

We now prove that the map (8) is surjective. Let $\frac{a}{s} \in A_{\mathfrak{m}}$, $a \in A$, $s \in S$. Because the only maximal ideal of A containing $\mathfrak{m}^n$ is $\mathfrak{m}$, no maximal ideal contains both s and $\mathfrak{m}^n$. It follows that $(s) + \mathfrak{m}^n = A$. Therefore, there exist $b \in A$ and $q \in \mathfrak{m}^n$ such that $sb + q = 1$ in A. It follows that s is invertible in $A_{\mathfrak{m}}/\mathfrak{n}^n$, and so $\frac{a}{s}$ is the *unique* element of this ring such that $s\frac{a}{s} = a$. As $s(ba) + qa = a$, the image of ba in $A_{\mathfrak{m}}/\mathfrak{n}^n$ also has this property and therefore equals $\frac{a}{s}$ in $A_{\mathfrak{m}}/\mathfrak{n}^n$. □

PROPOSITION 1.16. *In every noetherian ring, only 0 lies in all powers of all maximal ideals.*

PROOF. Let a be an element of a noetherian ring A. If $a \neq 0$, then $\{b \mid ba = 0\}$ is a proper ideal, and so is contained in some maximal ideal $\mathfrak{m}$. Then $\frac{a}{1}$ is nonzero in $A_{\mathfrak{m}}$, and so $\frac{a}{1} \notin (\mathfrak{m}A_{\mathfrak{m}})^n$ for some n (by the Krull intersection theorem 1.8), which implies that $a \notin \mathfrak{m}^n$. □

NOTES. For more on rings of fractions, see CA §5.

Modules of fractions

Let S be a multiplicative subset of the ring A, and let M be an A-module. Define an equivalence relation on $M \times S$ by

$$(m,s) \sim (n,t) \iff u(tm - sn) = 0 \text{ for some } u \in S.$$

Write $\frac{m}{s}$ for the equivalence class containing (m,s), and define addition and scalar multiplication by the rules:

$$\frac{m}{s} + \frac{n}{t} = \frac{mt+ns}{st}, \quad \frac{a}{s}\frac{m}{t} = \frac{am}{st}, \quad m,n \in M, \quad s,t \in S, \quad a \in A.$$

It is easily checked that these do not depend on the choices of representatives for the equivalence classes, and that we obtain in this way an $S^{-1}A$-module

$$S^{-1}M = \{\tfrac{m}{s} \mid m \in M, s \in S\}$$

and a homomorphism $m \mapsto \frac{m}{1} \colon M \xrightarrow{i_S} S^{-1}M$ of A-modules whose kernel is

$$\{a \in M \mid sa = 0 \text{ for some } s \in S\}.$$

1.17. The elements of S act invertibly on $S^{-1}M$, and every homomorphism from M to an A-module N with this property factors uniquely through i_S.

$$\begin{tikzcd} M \arrow[r, "i_S"] \arrow[dr] & S^{-1}M \arrow[d, dotted, "\exists!"] \\ & N. \end{tikzcd}$$

PROPOSITION 1.18. *The functor $M \rightsquigarrow S^{-1}M$ is exact. In other words, if the sequence of A-modules*

$$M' \xrightarrow{\alpha} M \xrightarrow{\beta} M''$$

is exact, then so also is the sequence of $S^{-1}A$-modules

$$S^{-1}M' \xrightarrow{S^{-1}\alpha} S^{-1}M \xrightarrow{S^{-1}\beta} S^{-1}M''.$$

PROOF. Because $\beta \circ \alpha = 0$, we have $0 = S^{-1}(\beta \circ \alpha) = S^{-1}\beta \circ S^{-1}\alpha$. Therefore $\mathrm{Im}(S^{-1}\alpha) \subset \mathrm{Ker}(S^{-1}\beta)$. For the reverse inclusion, let $\frac{m}{s} \in \mathrm{Ker}(S^{-1}\beta)$ where $m \in M$ and $s \in S$. Then $\frac{\beta(m)}{s} = 0$ and so, for some $t \in S$, we have $t\beta(m) = 0$. Then $\beta(tm) = 0$, and so $tm = \alpha(m')$ for some $m' \in M'$. Now

$$\frac{m}{s} = \frac{tm}{ts} = \frac{\alpha(m')}{ts} \in \mathrm{Im}(S^{-1}\alpha). \qquad \square$$

PROPOSITION 1.19. *Let A be a ring, and let M be an A-module. The canonical map*

$$M \to \prod \{M_{\mathfrak{m}} \mid \mathfrak{m} \text{ a maximal ideal in } A\}$$

is injective.

PROOF. Let $m \in M$ map to zero in all $M_{\mathfrak{m}}$. The annihilator $\mathfrak{a} = \{a \in A \mid am = 0\}$ of m is an ideal in A. Because m maps to zero $M_{\mathfrak{m}}$, there exists an $s \in A \smallsetminus \mathfrak{m}$ such that $sm = 0$. Therefore $\mathfrak{a}$ is not contained in $\mathfrak{m}$. Since this is true for all maximal ideals $\mathfrak{m}$, $\mathfrak{a} = A$, and so it contains 1. Now $m = 1m = 0$. $\square$

COROLLARY 1.20. *An A-module $M=0$ if $M_{\mathfrak{m}}=0$ for all maximal ideals $\mathfrak{m}$ in A.*

PROOF. Immediate consequence of the lemma. □

PROPOSITION 1.21. *Let A be a ring. A sequence of A-modules*

$$M' \xrightarrow{\alpha} M \xrightarrow{\beta} M'' \qquad (*)$$

is exact if and only if

$$M'_{\mathfrak{m}} \xrightarrow{\alpha_{\mathfrak{m}}} M_{\mathfrak{m}} \xrightarrow{\alpha_{\mathfrak{m}}} M''_{\mathfrak{m}} \qquad (**)$$

is exact for all maximal ideals $\mathfrak{m}$.

PROOF. The necessity is a special case of (1.18). For the sufficiency, let $N=\mathrm{Ker}(\beta)/\mathrm{Im}(\alpha)$. Because the functor $M \rightsquigarrow M_{\mathfrak{m}}$ is exact,

$$N_{\mathfrak{m}}=\mathrm{Ker}(\beta_{\mathfrak{m}})/\mathrm{Im}(\alpha_{\mathfrak{m}}).$$

If (**) is exact for all $\mathfrak{m}$, then $N_{\mathfrak{m}}=0$ for all $\mathfrak{m}$, and so $N=0$ (by 1.20). But this means that (*) is exact. □

COROLLARY 1.22. *A homomorphism $M\to N$ of A-modules is injective (resp. surjective) if and only if $M_{\mathfrak{m}}\to N_{\mathfrak{m}}$ is injective (resp. surjective) for all maximal ideals $\mathfrak{m}$.*

PROOF. Apply the proposition to $0\to M\to N$ (resp. $M\to N\to 0$). □

Direct limits

A ***directed set*** is a pair $(I,\le)$ consisting of a set I and a partial ordering $\le$ on I such that for all $i,j\in I$, there exists a $k\in I$ with $i,j\le k$.

Let $(I,\le)$ be a directed set, and let A be a ring. A ***direct system*** of A-modules indexed by $(I,\le)$ is a family $(M_i)_{i\in I}$ of A-modules together with a family $(\alpha^i_j\colon M_i\to M_j)_{i\le j}$ of A-linear maps such that $\alpha^i_i=\mathrm{id}_{M_i}$ and $\alpha^j_k\circ\alpha^i_j=\alpha^i_k$ all $i\le j\le k$.[4] An A-module M together with A-linear maps $\alpha^i\colon M_i\to M$ such that $\alpha^i=\alpha^j\circ\alpha^i_j$ for all $i\le j$ is the ***direct limit*** of the system (M_i,α^j_i) if

(a) $M=\bigcup_{i\in I}\alpha^i(M_i)$, and

(b) $m_i\in M_i$ maps to zero in M if and only if it maps to zero in M_j for some $j\ge i$.

Direct limits of A-algebras are defined similarly.

PROPOSITION 1.23. *For every multiplicative subset S of A, $S^{-1}A\simeq\varinjlim A_h$, where h runs over the elements of S (partially ordered by division).*

PROOF. When $h|h'$, say, $h'=hg$, there is a canonical homomorphism $\frac{a}{h}\mapsto\frac{ag}{h'}\colon A_h\to A_{h'}$, and so the rings A_h form a direct system indexed by the set S. When $h\in S$, the homomorphism $A\to S^{-1}A$ extends uniquely to a homomorphism $\frac{a}{h}\mapsto\frac{a}{h}\colon A_h\to S^{-1}A$ (1.10), and these homomorphisms are compatible with the maps in the direct system. Now it is easy to see that $S^{-1}A$ satisfies the conditions to be the direct limit of the A_h. □

[4]Regard I as a category with $\mathrm{Hom}(a,b)$ empty unless $a\le b$, in which case it contains a single element. Then a direct system is a functor from I to the category of A-modules.

c Unique factorization

Let A be an integral domain. An element a of A is ***irreducible*** if it is not zero, not a unit, and admits only trivial factorizations, i.e.,

$$a = bc \implies b \text{ or } c \text{ is a unit.}$$

An element a is said to be ***prime*** if (a) is a prime ideal, i.e.,

$$a|bc \implies a|b \text{ or } a|c.$$

An integral domain A is called a ***unique factorization domain*** (or a ***factorial domain***) if every nonzero nonunit in A can be written as a finite product of irreducible elements in exactly one way up to units and the order of the factors. Principal ideal domains, for example, $\mathbb{Z}$ and $k[X]$, are unique factorization domains,

PROPOSITION 1.24. *Let A be an integral domain, and let a be an element of A that is neither zero nor a unit. If a is prime, then a is irreducible, and the converse holds when A is a unique factorization domain.*

PROOF. Assume that a is prime. If $a = bc$, then a divides bc and so a divides b or c. Suppose the first, and write $b = aq$. Now $a = bc = aqc$, which implies that $qc = 1$ because A is an integral domain, and so c is a unit. Therefore a is irreducible.

For the converse, assume that a is irreducible and that A is a unique factorization domain. If $a|bc$, then

$$bc = aq, \text{ some } q \in A.$$

On writing each of b, c, and q as a product of irreducible elements, and using the uniqueness of factorizations, we see that a differs from one of the irreducible factors of b or c by a unit. Therefore a divides b or c. □

COROLLARY 1.25. *Let A be an integral domain. If A is a unique factorization domain, then every prime ideal of height 1 is principal.*

PROOF. Let $\mathfrak{p}$ be a prime ideal of height 1. Then $\mathfrak{p}$ contains a nonzero element, and hence an irreducible element a. We have $\mathfrak{p} \supset (a) \supset (0)$. As (a) is prime and $\mathfrak{p}$ has height 1, we must have $\mathfrak{p} = (a)$. □

PROPOSITION 1.26. *Let A be an integral domain in which every nonzero nonunit element is a finite product of irreducible elements. If every irreducible element of A is prime, then A is a unique factorization domain.*

PROOF. Suppose that

$$a_1 \cdots a_m = b_1 \cdots b_n \tag{9}$$

with the a_i and b_i irreducible elements in A. As a_1 is prime, it divides one of the b_i, which we may suppose to be b_1. As b_1 is irreducible, $b_1 = ua_1$ for some unit u. On cancelling a_1 from both sides of (9), we obtain the equality

$$a_2 \cdots a_m = (ub_2)b_3 \cdots b_n.$$

Continuing in this fashion, we find that the two factorizations are the same up to units and the order of the factors. □

PROPOSITION 1.27 (GAUSS'S LEMMA). *Let A be a unique factorization domain with field of fractions F. If $f(X) \in A[X]$ factors into the product of two nonconstant polynomials in $F[X]$, then it factors into the product of two nonconstant polynomials in $A[X]$.*

PROOF. Let $f = gh$ in $F[X]$. For suitable $c, d \in A$, the polynomials $g_1 = cg$ and $h_1 = dh$ have coefficients in A, and so we have a factorization

$$cdf = g_1 h_1 \text{ in } A[X].$$

If an irreducible element p of A divides cd, then, looking modulo (p), we see that

$$0 = \overline{g_1} \cdot \overline{h_1} \text{ in } (A/(p))[X].$$

According to Proposition 1.24, (p) is prime, and so $(A/(p))[X]$ is an integral domain. Therefore, p divides all the coefficients of at least one of the polynomials g_1, h_1, say g_1, so that $g_1 = pg_2$ for some $g_2 \in A[X]$. Thus, we have a factorization

$$(cd/p)f = g_2 h_1 \text{ in } A[X].$$

Continuing in this fashion, we can remove all the irreducible factors of cd, and so obtain a factorization of f in $A[X]$. □

Let A be a unique factorization domain. A nonzero polynomial

$$f = a_0 + a_1 X + \cdots + a_m X^m$$

in $A[X]$ is said to be ***primitive*** if the coefficients a_i have no common factor (other than units). Every polynomial f in $F[X]$ can be written $f = c(f) \cdot f_1$ with $c(f) \in F$ and f_1 primitive. The element $c(f)$, which is well-defined up to multiplication by a unit, is called the ***content*** of f. Note that $f \in A[X]$ if and only if $c(f) \in A$.

LEMMA 1.28. *The product of two primitive polynomials is primitive.*

PROOF. Let

$$f = a_0 + a_1 X + \cdots + a_m X^m$$
$$g = b_0 + b_1 X + \cdots + b_n X^n,$$

be primitive polynomials, and let p be an irreducible element of A. Let a_{i_0}, $i_0 \le m$, be the first coefficient of f not divisible by p, and let b_{j_0}, $j_0 \le n$, the first coefficient of g not divisible by p. Then all the terms in the sum $\sum_{i+j=i_0+j_0} a_i b_j$ are divisible by p, except $a_{i_0} b_{j_0}$, which is not divisible by p. Therefore, p doesn't divide the $(i_0 + j_0)$th-coefficient of fg. We have shown that no irreducible element of A divides all the coefficients of fg, which must therefore be primitive. □

PROPOSITION 1.29. *Let A be a unique factorization domain with field of fractions F. For polynomials $f, g \in F[X]$,*

$$c(fg) = c(f) \cdot c(g);$$

hence every factor in $A[X]$ of a primitive polynomial is primitive.

PROOF. Let $f = c(f) \cdot f_1$ and $g = c(g) \cdot g_1$ with f_1 and g_1 primitive. Then

$$fg = c(f) \cdot c(g) \cdot f_1 g_1$$

with $f_1 g_1$ primitive, and so $c(fg) = c(f)c(g)$. □

COROLLARY 1.30. *The irreducible elements in $A[X]$ are the irreducible elements a of A and the nonconstant primitive polynomials f such that f is irreducible in $F[X]$.*

PROOF. Obvious from (1.27) and (1.29). □

PROPOSITION 1.31. *If A is a unique factorization domain, then so also is $A[X]$.*

PROOF. We shall check that A satisfies the conditions of (1.26).

Let $f \in A[X]$, and write $f = c(f) f_1$. Then $c(f)$ is a product of irreducible elements in A, and f_1 is a product of irreducible primitive polynomials. This shows that f is a product of irreducible elements in $A[X]$.

Let a be an irreducible element of A. If a divides fg, then it divides $c(fg) = c(f)c(g)$. As a is prime (1.24), it divides $c(f)$ or $c(g)$, and hence also f or g.

Let f be an irreducible primitive polynomial in $A[X]$. Then f is irreducible in $F[X]$, and so if f divides the product gh of $g, h \in A[X]$, then it divides g or h in $F[X]$. Suppose the first, and write $fq = g$ with $q \in F[X]$. Then $c(q) = c(f)c(q) = c(fq) = c(g) \in A$, and so $q \in A[X]$. Therefore f divides g in $A[X]$.

We have shown that every element of $A[X]$ is a product of irreducible elements and that every irreducible element of $A[X]$ is prime, and so $A[X]$ is a unique factorization domain (1.26). □

Polynomial rings

Let k be a field. The elements of the polynomial ring $k[X_1, \ldots, X_n]$ are finite sums

$$\sum c_{a_1 \cdots a_n} X_1^{a_1} \cdots X_n^{a_n}, \quad c_{a_1 \cdots a_n} \in k, \quad a_j \in \mathbb{N},$$

with the obvious notions of equality, addition, and multiplication. In particular, the monomials form a basis for $k[X_1, \ldots, X_n]$ as a k-vector space.

The ***degree***, $\deg(f)$, of a nonzero polynomial f is the largest total degree of a monomial occurring in f with nonzero coefficient. Since $\deg(fg) = \deg(f) + \deg(g)$, $k[X_1, \ldots, X_n]$ is an integral domain and $k[X_1, \ldots, X_n]^\times = k^\times$. An element f of $k[X_1, \ldots, X_n]$ is irreducible if it is nonconstant and $f = gh \implies g$ or h is constant.

THEOREM 1.32. *The ring $k[X_1, \ldots, X_n]$ is a unique factorization domain.*

PROOF. Note that

$$k[X_1, \ldots, X_{n-1}][X_n] = k[X_1, \ldots, X_n].$$

This simply says that every polynomial f in n symbols $X_1, \ldots, X_n$ can be expressed uniquely as a polynomial in X_n with coefficients in $k[X_1, \ldots, X_{n-1}]$,

$$f(X_1, \ldots, X_n) = a_0(X_1, \ldots, X_{n-1}) X_n^r + \cdots + a_r(X_1, \ldots, X_{n-1}).$$

Since, as we noted, $k[X]$ is a unique factorization domain, the theorem follows by induction from Proposition 1.31. □

COROLLARY 1.33. *A nonzero proper principal ideal (f) in $k[X_1, \ldots, X_n]$ is prime if and only if f is irreducible.*

PROOF. Special case of (1.24). □

d Integral dependence

Let A be a subring of a ring B. An element α of B is said to be[5] ***integral*** over A if it is a root of a monic[6] polynomial with coefficients in A, i.e., if it satisfies an equation

$$\alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0, \quad a_i \in A.$$

If every element of B is integral over A, then B is said to be ***integral*** over A.

In the next proof, we shall need to apply a variant of Cramer's rule: if $x_1, \ldots, x_m$ is a solution to the system of linear equations

$$\sum_{j=1}^{m} c_{ij} x_j = 0, \quad i = 1, \ldots, m,$$

with coefficients in a ring A, then

$$\det(C) \cdot x_j = 0, \quad j = 1, \ldots, m, \tag{10}$$

where C is the matrix of coefficients. To prove this, expand out the left hand side of

$$\det \begin{pmatrix} c_{11} & \ldots & c_{1\,j-1} & \sum_i c_{1i} x_i & c_{1\,j+1} & \ldots & c_{1m} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_{m1} & \ldots & c_{m\,j-1} & \sum_i c_{mi} x_i & c_{m\,j+1} & \ldots & c_{mm} \end{pmatrix} = 0$$

using standard properties of determinants.

An A-module M is ***faithful*** if $aM = 0$, $a \in A$, implies that $a = 0$.

PROPOSITION 1.34. *Let A be a subring of a ring B. An element α of B is integral over A if and only if there exists a faithful $A[\alpha]$-submodule M of B that is finitely generated as an A-module.*

PROOF. $\Rightarrow$: Suppose that

$$\alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0, \quad a_i \in A.$$

Then the A-submodule M of B generated by 1, α, ..., α^{n-1} has the property that $\alpha M \subset M$, and it is faithful because it contains 1.

$\Leftarrow$: Let M be a faithful $A[\alpha]$-submodule of B admitting a finite set $\{e_1, \ldots, e_n\}$ of generators as an A-module. Then, for each i,

$$\alpha e_i = \sum a_{ij} e_j, \text{ some } a_{ij} \in A.$$

We can rewrite this system of equations as

$$\begin{aligned} (\alpha - a_{11})e_1 - a_{12}e_2 - a_{13}e_3 - \cdots &= 0 \\ -a_{21}e_1 + (\alpha - a_{22})e_2 - a_{23}e_3 - \cdots &= 0 \\ \cdots &= 0. \end{aligned}$$

[5]More generally, if $f\colon A \to B$ is an A-algebra, an element α of B is ***integral*** over A if it satisfies an equation

$$\alpha^n + f(a_1)\alpha^{n-1} + \cdots + f(a_n) = 0, \quad a_i \in A.$$

Thus, α is integral over A if and only if it is integral over the subring $f(A)$ of B.

[6]A polynomial is ***monic*** if its leading coefficient is 1, i.e., $f(X) = X^n +$ terms of degree less than n.

Let C be the matrix of coefficients on the left-hand side. Then Cramer's formula tells us that $\det(C)\cdot e_i = 0$ for all i. As M is faithful and the e_i generate M, this implies that $\det(C) = 0$. On expanding out the determinant, we obtain an equation

$$\alpha^n + c_1\alpha^{n-1} + c_2\alpha^{n-2} + \cdots + c_n = 0, \quad c_i \in A. \qquad \square$$

PROPOSITION 1.35. *An A-algebra B is finite if it is generated as an A-algebra by a finite set of elements each of which is integral over A.*

PROOF. Suppose that $B = A[\alpha_1, \ldots, \alpha_m]$ and that

$$\alpha_i^{n_i} + a_{i1}\alpha_i^{n_i-1} + \cdots + a_{in_i} = 0, \quad a_{ij} \in A, \quad i = 1, \ldots, m.$$

Any monomial in the α_i divisible by some $\alpha_i^{n_i}$ is equal (in B) to a linear combination of monomials of lower degree. Therefore, B is generated as an A-module by the finite set of monomials $\alpha_1^{r_1}\cdots\alpha_m^{r_m}$, $1 \le r_i < n_i$. $\square$

COROLLARY 1.36. *An A-algebra B is finite if and only if it is finitely generated and integral over A.*

PROOF. $\Leftarrow$: Immediate consequence of (1.35).

$\Rightarrow$: We may replace A with its image in B. Then B is a faithful $A[\alpha]$-module for all $\alpha \in B$ (because $1_B \in B$), and so (1.34) shows that every element of B is integral over A. As B is finitely generated as an A-module, it is certainly finitely generated as an A-algebra. $\square$

PROPOSITION 1.37. *Consider rings $A \subset B \subset C$. If B is integral over A and C is integral over B, then C is integral over A.*

PROOF. Let $\gamma \in C$. Then

$$\gamma^n + b_1\gamma^{n-1} + \cdots + b_n = 0$$

for some $b_i \in B$. Now $A[b_1, \ldots, b_n]$ is finite over A (see 1.35), and $A[b_1, \ldots, b_n][\gamma]$ is finite over $A[b_1, \ldots, b_n]$, and so it is finite over A. Therefore γ is integral over A by (1.34). $\square$

THEOREM 1.38. *Let A be a subring of a ring B. The elements of B integral over A form an A-subalgebra of B.*

PROOF. Let α and β be two elements of B integral over A. Then $A[\alpha, \beta]$ is finitely generated as an A-module (1.35). It is stable under multiplication by $\alpha \pm \beta$ and $\alpha\beta$ and it is faithful as an $A[\alpha \pm \beta]$-module and as an $A[\alpha\beta]$-module (because it contains 1_A). Therefore (1.34) shows that $\alpha \pm \beta$ and $\alpha\beta$ are integral over A. $\square$

DEFINITION 1.39. Let A be a subring of the ring B. The ***integral closure*** of A in B is the subring of B consisting of the elements integral over A.

PROPOSITION 1.40. *Let A be an integral domain with field of fractions F, and let α be an element of some field containing F. If α is algebraic over F, then there exists a $d \in A$ such that $d\alpha$ is integral over A.*

PROOF. By assumption, α satisfies an equation

$$\alpha^m + a_1\alpha^{m-1} + \cdots + a_m = 0, \quad a_i \in F.$$

Let d be a common denominator for the a_i, so that $da_i \in A$ for all i, and multiply through the equation by d^m:

$$(d\alpha)^m + a_1 d (d\alpha)^{m-1} + \cdots + a_m d^m = 0.$$

As $a_1 d, \ldots, a_m d^m \in A$, this shows that $d\alpha$ is integral over A. □

COROLLARY 1.41. *Let A be an integral domain and let E be an algebraic extension of the field of fractions of A. Then E is the field of fractions of the integral closure of A in E.*

PROOF. In fact, the proposition shows that every element of E is a quotient β/d with β integral over A and $d \in A$. □

DEFINITION 1.42. An integral domain A is is said to be ***integrally closed*** if it is equal to its integral closure in its field of fractions F, i.e., if

$$\alpha \in F, \quad \alpha \text{ integral over } A \implies \alpha \in A.$$

A ***normal domain*** is an integrally closed integral domain.

PROPOSITION 1.43. *Unique factorization domains are integrally closed.*

PROOF. Let A be a unique factorization domain, and let a/b be an element of its field of fractions. If $a/b \notin A$, then b divisible by some prime element p not dividing a. If a/b is integral over A, then it satisfies an equation

$$(a/b)^n + a_1(a/b)^{n-1} + \cdots + a_n = 0, \quad a_i \in A.$$

On multiplying through by b^n, we obtain the equation

$$a^n + a_1 a^{n-1} b + \cdots + a_n b^n = 0.$$

The element p then divides every term on the left except a^n, and hence divides a^n. Since it doesn't divide a, this is a contradiction. □

Let $F \subset E$ be fields, and let $\alpha \in E$ be algebraic over F. The ***minimum polynomial*** of α over F is the monic polynomial of smallest degree in $F[X]$ having α as a root. If f is the minimum polynomial of α, then the homomorphism $X \mapsto \alpha\colon F[X] \to F[\alpha]$ defines an isomorphism $F[X]/(f) \to F[\alpha]$, i.e., $F[x] \simeq F[\alpha]$, $x \leftrightarrow \alpha$.

PROPOSITION 1.44. *Let A be a normal domain, and let E be a finite extension of the field of fractions F of A. An element of E is integral over A if and only if its minimum polynomial over F has coefficients in A.*

PROOF. Let $\alpha \in E$ be integral over A, so that

$$\alpha^m + a_1\alpha^{m-1} + \cdots + a_m = 0, \quad \text{some } a_i \in A, \quad m > 0.$$

Let $f(X)$ be the minimum polynomial of α over F, and let α' be a conjugate of α, i.e., a root of f in some splitting field of f. Then f is also the minimum polynomial of α' over F, and so (see above), there is an F-isomorphism

$$\sigma\colon F[\alpha] \to F[\alpha'], \quad \sigma(\alpha) = \alpha'.$$

On applying σ to the above equation we obtain the equation

$$\alpha'^m + a_1\alpha'^{m-1} + \cdots + a_m = 0,$$

which shows that α' is integral over A. As the coefficients of f are polynomials in the conjugates of α, it follows from (1.38) that the coefficients of $f(X)$ are integral over A. They lie in F, and A is integrally closed, and so they lie in A. This proves the "only if" part of the statement, and the "if" part is obvious. □

COROLLARY 1.45. *Let $A \subset F \subset E$ be as in the proposition, and let α be an element of E integral over A. Then $\mathrm{Nm}_{E/F}(\alpha) \in A$, and α divides $\mathrm{Nm}_{E/F}(\alpha)$ in $A[\alpha]$.*

PROOF. Let

$$f(X) = X^m + a_1X^{m-1} + \cdots + a_m$$

be the minimum polynomial of α over F. Then $\mathrm{Nm}(\alpha) = (-1)^{mn}a_m^n$ where $n = [E\colon F[\alpha]]$ (see FT, 5.45), and so $\mathrm{Nm}(\alpha) \in A$. Because $f(\alpha) = 0$,

$$\begin{aligned} 0 &= a_m^{n-1}(\alpha^m + a_1\alpha^{m-1} + \cdots + a_m) \\ &= \alpha(a_m^{n-1}\alpha^{m-1} + \cdots + a_m^{n-1}a_{m-1}) + (-1)^{mn}\,\mathrm{Nm}(\alpha), \end{aligned}$$

and so α divides $\mathrm{Nm}_{E/F}(\alpha)$ in $A[\alpha]$. □

COROLLARY 1.46. *Let A be a normal domain with field of fractions F, and let $f(X)$ be a monic polynomial in $A[X]$. Then every monic factor of $f(X)$ in $F[X]$ has coefficients in A.*

PROOF. It suffices to prove this for an irreducible monic factor g of f in $F[X]$. Let α be a root of g in some extension field of F. Then g is the minimum polynomial of α. As α is a root of f, it is integral over A, and so g has coefficients in A. □

PROPOSITION 1.47. *Let $A \subset B$ be rings, and let A' be the integral closure of A in B. For any multiplicative subset S of A, $S^{-1}A'$ is the integral closure of $S^{-1}A$ in $S^{-1}B$.*

PROOF. Let $\frac{b}{s} \in S^{-1}A'$ with $b \in A'$ and $s \in S$. Then

$$b^n + a_1b^{n-1} + \cdots + a_n = 0$$

for some $a_i \in A$, and so

$$\left(\frac{b}{s}\right)^n + \frac{a_1}{s}\left(\frac{b}{s}\right)^{n-1} + \cdots + \frac{a_n}{s^n} = 0.$$

Therefore b/s is integral over $S^{-1}A$. This shows that $S^{-1}A'$ is contained in the integral closure of $S^{-1}B$.

For the converse, let b/s ($b \in B$, $s \in S$) be integral over $S^{-1}A$. Then

$$\left(\frac{b}{s}\right)^n + \frac{a_1}{s_1}\left(\frac{b}{s}\right)^{n-1} + \cdots + \frac{a_n}{s_n} = 0.$$

for some $a_i \in A$ and $s_i \in S$. On multiplying this equation by $s^n s_1 \cdots s_n$, we find that $s_1 \cdots s_n b \in A'$, and therefore that $\frac{b}{s} = \frac{s_1 \cdots s_n b}{s s_1 \cdots s_n} \in S^{-1}A'$. □

COROLLARY 1.48. *Let $A \subset B$ be rings, and let S be a multiplicative subset of A. If A is integrally closed in B, then $S^{-1}A$ is integrally closed in $S^{-1}B$.*

PROOF. Special case of the proposition in which $A' = A$. □

PROPOSITION 1.49. *The following conditions on an integral domain A are equivalent:*

(a) *A is integrally closed;*

(b) *$A_{\mathfrak{p}}$ is integrally closed for all prime ideals $\mathfrak{p}$;*

(c) *$A_{\mathfrak{m}}$ is integrally closed for all maximal ideals $\mathfrak{m}$.*

PROOF. The implication (a)⇒(b) follows from (1.48), and (b)⇒(c) is obvious. It remains to prove (c)⇒(a). If c is integral over A, then it is integral over each $A_{\mathfrak{m}}$, and hence lies in each $A_{\mathfrak{m}}$. It follows that the ideal consisting of the $a \in A$ such that $ac \in A$ is not contained in any maximal ideal $\mathfrak{m}$, and therefore equals A. Hence $1 \cdot c \in A$. □

Let E/F be a finite extension of fields. Then

$$(\alpha, \beta) \mapsto \mathrm{Tr}_{E/F}(\alpha\beta) \colon E \times E \to F \tag{11}$$

is a symmetric bilinear form on E regarded as a vector space over F.

LEMMA 1.50. *If E/F is separable, then the trace pairing (11) is nondegenerate.*

PROOF. Let $\beta_1, \ldots, \beta_m$ be a basis for E as an F-vector space. We have to show that the discriminant $\det(\mathrm{Tr}(\beta_i\beta_j))$ of the trace pairing is nonzero. Let $\sigma_1, \ldots, \sigma_m$ be the distinct F-homomorphisms of E into some large Galois extension Ω of F. Recall (FT 5.45) that

$$\mathrm{Tr}_{L/K}(\beta) = \sigma_1\beta + \cdots + \sigma_m\beta \tag{12}$$

By direct calculation, we have

$$\begin{aligned}
\det(\mathrm{Tr}(\beta_i\beta_j)) &= \det(\textstyle\sum_k \sigma_k(\beta_i\beta_j)) && \text{(by 12)}\\
&= \det(\textstyle\sum_k \sigma_k(\beta_i)\cdot\sigma_k(\beta_j))\\
&= \det(\sigma_k(\beta_i))\cdot\det(\sigma_k(\beta_j))\\
&= \det(\sigma_k(\beta_i))^2.
\end{aligned}$$

Suppose that $\det(\sigma_i\beta_j) = 0$. Then there exist $c_1, \ldots, c_m \in \Omega$ such that

$$\sum_i c_i\sigma_i(\beta_j) = 0 \text{ all } j.$$

By linearity, it follows that $\sum_i c_i\sigma_i(\beta) = 0$ for all $\beta \in E$, but this contradicts Dedekind's theorem on the independence of characters (FT 5.14). □

PROPOSITION 1.51. *Let A be a normal domain with field of fractions F, and let B be the integral closure of A in a separable extension E of F of degree m. There exist free A-submodules M and M' of E such that*

$$M \subset B \subset M'. \tag{13}$$

If A is noetherian, then B is a finite A-algebra.

PROOF. Let $\{\beta_1,...,\beta_m\}$ be a basis for E over F. According to (1.40), there exists a $d \in A$ such that $d \cdot \beta_i \in B$ for all i. Clearly $\{d \cdot \beta_1,\ldots,d \cdot \beta_m\}$ is still a basis for E as a vector space over F, and so we may assume to begin with that each $\beta_i \in B$. Because the trace pairing is nondegenerate, there is a dual basis $\{\beta_1',...,\beta_m'\}$ of E over F with the property that $\mathrm{Tr}(\beta_i \cdot \beta_j') = \delta_{ij}$ for all i,j. We shall show that

$$A\beta_1 + A\beta_2 + \cdots + A\beta_m \subset B \subset A\beta_1' + A\beta_2' + \cdots + A\beta_m'.$$

Only the second inclusion requires proof. Let $\beta \in B$. Then β can be written uniquely as a linear combination $\beta = \sum b_j \beta_j'$ of the β_j' with coefficients $b_j \in F$, and we have to show that each $b_j \in A$. As β_i and β are in B, so also is $\beta \cdot \beta_i$, and so $\mathrm{Tr}(\beta \cdot \beta_i) \in A$ (1.44). But

$$\mathrm{Tr}(\beta \cdot \beta_i) = \mathrm{Tr}(\sum_j b_j \beta_j' \cdot \beta_i) = \sum_j b_j \,\mathrm{Tr}(\beta_j' \cdot \beta_i) = \sum_j b_j \cdot \delta_{ij} = b_i.$$

Hence $b_i \in A$.

If A is Noetherian, then M' is a Noetherian A-module, and so B is finitely generated as an A-module. □

LEMMA 1.52. *Let A be a subring of a field K. If K is integral over A, then A is also a field.*

PROOF. Let a be a nonzero element of A. Then $a^{-1} \in K$, and it is integral over A:

$$(a^{-1})^n + a_1(a^{-1})^{n-1} + \cdots + a_n = 0, \quad a_i \in A.$$

On multiplying through by a^{n-1}, we find that

$$a^{-1} + a_1 + \cdots + a_n a^{n-1} = 0,$$

from which it follows that $a^{-1} \in A$. □

THEOREM 1.53 (GOING-UP THEOREM). *Let $A \subset B$ be rings with B integral over A.*

(a) *For every prime ideal $\mathfrak{p}$ of A, there is a prime ideal $\mathfrak{q}$ of B such that $\mathfrak{q} \cap A = \mathfrak{p}$.*

(b) *Let $\mathfrak{p} = \mathfrak{q} \cap A$; then $\mathfrak{p}$ is maximal if and only if $\mathfrak{q}$ is maximal.*

PROOF. (a) If S is a multiplicative subset of a ring A, then the prime ideals of $S^{-1}A$ are in one-to-one correspondence with the prime ideals of A not meeting S (see 1.14). It therefore suffices to prove (a) after A and B have been replaced by $S^{-1}A$ and $S^{-1}B$, where $S = A - \mathfrak{p}$. Thus we may assume that A is local, and that $\mathfrak{p}$ is its unique maximal ideal. In this case, for all proper ideals $\mathfrak{b}$ of B, $\mathfrak{b} \cap A \subset \mathfrak{p}$ (otherwise $\mathfrak{b} \supset A \ni 1$). To complete the proof of (a), I shall show that for all maximal ideals $\mathfrak{n}$ of B, $\mathfrak{n} \cap A = \mathfrak{p}$.

Consider $B/\mathfrak{n} \supset A/(\mathfrak{n} \cap A)$. Here $B/\mathfrak{n}$ is a field, which is integral over its subring $A/(\mathfrak{n} \cap A)$, and $\mathfrak{n} \cap A$ will be equal to $\mathfrak{p}$ if and only if $A/(\mathfrak{n} \cap A)$ is a field. This follows from Lemma 1.52.

(b) The ring $B/\mathfrak{q}$ contains $A/\mathfrak{p}$, and it is integral over $A/\mathfrak{p}$. If $\mathfrak{q}$ is maximal, then Lemma 1.52 shows that $\mathfrak{p}$ is also. For the converse, note that any integral domain integral over a field is a field because it is a union of integral domains finite over the field, which are automatically fields (left multiplication by an element is injective, and hence surjective, being a linear map of a finite-dimensional vector space). □

COROLLARY 1.54. *Let $A \subset B$ be rings with B integral over A. Let $\mathfrak{p} \subset \mathfrak{p}'$ be prime ideals of A, and let $\mathfrak{q}$ be a prime ideal of B such that $\mathfrak{q} \cap A = \mathfrak{p}$. Then there exists a prime ideal $\mathfrak{q}'$ of B containing $\mathfrak{q}$ and such that $\mathfrak{q}' \cap A = \mathfrak{p}'$:*

$$\begin{tikzcd} B \arrow[d, dash] & \mathfrak{q} \arrow[d, dash] & \subset & \mathfrak{q}' \arrow[d, dash] \\ A & \mathfrak{p} & \subset & \mathfrak{p}'. \end{tikzcd}$$

PROOF. We have $A/\mathfrak{p} \subset B/\mathfrak{q}$, and $B/\mathfrak{q}$ is integral over $A/\mathfrak{p}$. According to the (1.53), there exists a prime ideal $\mathfrak{q}''$ in $B/\mathfrak{q}$ such that $\mathfrak{q}'' \cap (A/\mathfrak{p}) = \mathfrak{p}'/\mathfrak{p}$. The inverse image $\mathfrak{q}'$ of $\mathfrak{q}''$ in B has the required properties. □

ASIDE 1.55. Let A be a noetherian integral domain, and let B be the integral closure of A in a finite extension E of the field of fractions F of A. Is B always a finite A-algebra? When A is integrally closed and E is separable over F, or A is a finitely generated k-algebra, then the answer is yes (1.51, 8.3). However, in 1935, Akizuki found an example of a noetherian integral domain whose integral closure in its field of fractions is not finite (according to Matsumura 1986, finding the example cost him a year's hard struggle). F.K. Schmidt found another example at about the same time (see Olberding, Trans. AMS, 366 (2014), no. 8, 4067–4095).

e Tensor Products

Tensor products of modules

Let A be a ring, and let M, N, and P be A-modules. A map $\phi: M \times N \to P$ of A-modules is said to be A-***bilinear*** if

$$\begin{aligned} \phi(x+x', y) &= \phi(x,y) + \phi(x',y), && x, x' \in M, \quad y \in N \\ \phi(x, y+y') &= \phi(x,y) + \phi(x,y'), && x \in M, \quad y, y' \in N \\ \phi(ax, y) &= a\phi(x,y), && a \in A, \quad x \in M, \quad y \in N \\ \phi(x, ay) &= a\phi(x,y), && a \in A, \quad x \in M, \quad y \in N, \end{aligned}$$

i.e., if ϕ is A-linear in each variable.

An A-module T together with an A-bilinear map

$$\phi: M \times N \to T$$

is called the ***tensor product*** of M and N over A if it has the following universal property: every A-bilinear map

$$\phi': M \times N \to T'$$

factors uniquely through ϕ.

$$\begin{tikzcd} M \times N \arrow[r, "\phi"] \arrow[dr, "\phi'"'] & T \arrow[d, dotted, "\exists!\text{ linear}"] \\ & T'. \end{tikzcd}$$

As usual, the universal property determines the tensor product uniquely up to a unique isomorphism. We write it $M \otimes_A N$. Note that

$$\mathrm{Hom}_{A\text{-bilinear}}(M \times N, T) \simeq \mathrm{Hom}_{A\text{-linear}}(M \otimes_A N, T).$$

CONSTRUCTION

Let M and N be A-modules, and let $A^{(M\times N)}$ be the free A-module with basis $M\times N$. Thus each element $A^{(M\times N)}$ can be expressed uniquely as a finite sum

$$\sum a_i(x_i,y_i),\quad a_i\in A,\quad x_i\in M,\quad y_i\in N.$$

Let P be the submodule of $A^{(M\times N)}$ generated by the following elements

$$\begin{aligned}(x+x',y)-(x,y)-(x',y),&\quad x,x'\in M,\quad y\in N\\ (x,y+y')-(x,y)-(x,y'),&\quad x\in M,\quad y,y'\in N\\ (ax,y)-a(x,y),&\quad a\in A,\quad x\in M,\quad y\in N\\ (x,ay)-a(x,y),&\quad a\in A,\quad x\in M,\quad y\in N,\end{aligned}$$

and define

$$M\otimes_A N=A^{(M\times N)}/P.$$

Write $x\otimes y$ for the class of (x,y) in $M\otimes_A N$. Then

$$(x,y)\mapsto x\otimes y\colon M\times N\to M\otimes_A N$$

is A-bilinear — we have imposed the fewest relations necessary to ensure this. Every element of $M\otimes_A N$ can be written as a finite sum[7]

$$\sum a_i(x_i\otimes y_i),\quad a_i\in A,\quad x_i\in M,\quad y_i\in N,$$

and all relations among these symbols are generated by the following relations

$$\begin{aligned}(x+x')\otimes y&=x\otimes y+x'\otimes y\\ x\otimes(y+y')&=x\otimes y+x\otimes y'\\ a(x\otimes y)&=(ax)\otimes y=x\otimes ay.\end{aligned}$$

The pair $(M\otimes_A N,(x,y)\mapsto x\otimes y)$ has the correct universal property because any bilinear map $\phi'\colon M\times N\to T'$ defines an A-linear map $A^{(M\times N)}\to T'$, which factors through $A^{(M\times N)}/K$, and gives a commutative triangle.

Tensor products of algebras

Let A and B be k-algebras. A k-algebra C together with homomorphisms $i\colon A\to C$ and $j\colon B\to C$ is called the ***tensor product*** of A and B if it has the following universal property: for every pair of homomorphisms (of k-algebras) $\alpha\colon A\to R$ and $\beta\colon B\to R$, there is a unique homomorphism $\gamma\colon C\to R$ such that $\gamma\circ i=\alpha$ and $\gamma\circ j=\beta$:

$$\begin{tikzcd} A \arrow[r, "i"] \arrow[dr, "\alpha"'] & C \arrow[d, dotted, "\exists!\,\gamma"] & B \arrow[l, "j"'] \arrow[dl, "\beta"] \\ & R. & \end{tikzcd}$$

If it exists, the tensor product, is uniquely determined up to a unique isomorphism by this property. We write it $A\otimes_k B$. Note that

$$\mathrm{Hom}_k(A\otimes_k B,R)\simeq\mathrm{Hom}_k(A,R)\times\mathrm{Hom}_k(B,R).$$

[7]"An element of the tensor product of two vector spaces is not necessarily a tensor product of two vectors, but sometimes a sum of such. This might be considered a mathematical shenanigan but if you start with the state vectors of two quantum systems it exactly corresponds to the notorious notion of entanglement which so displeased Einstein." Georges Elencwajg on mathoverflow.net.

CONSTRUCTION

Regard A and B as k-vector spaces, and form the tensor product $A \otimes_k B$. There is a multiplication map $A \otimes_k B \times A \otimes_k B \to A \otimes_k B$ for which

$$(a \otimes b)(a' \otimes b') = aa' \otimes bb'.$$

This makes $A \otimes_k B$ into a ring, and the homomorphism

$$c \mapsto c(1 \otimes 1) = c \otimes 1 = 1 \otimes c$$

makes it into a k-algebra. The maps

$$a \mapsto a \otimes 1 : A \to C \text{ and } b \mapsto 1 \otimes b : B \to C$$

are homomorphisms, and they make $A \otimes_k B$ into the tensor product of A and B in the above sense.

EXAMPLE 1.56. The algebra B, together with the given map $k \to B$ and the identity map $B \to B$, has the universal property characterizing $k \otimes_k B$. In terms of the constructive definition of tensor products, the map $c \otimes b \mapsto cb : k \otimes_k B \to B$ is an isomorphism.

EXAMPLE 1.57. The ring $k[X_1, \ldots, X_m, X_{m+1}, \ldots, X_{m+n}]$, together with the obvious inclusions

$$k[X_1, \ldots, X_m] \hookrightarrow k[X_1, \ldots, X_{m+n}] \hookleftarrow k[X_{m+1}, \ldots, X_{m+n}]$$

is the tensor product of $k[X_1, \ldots, X_m]$ and $k[X_{m+1}, \ldots, X_{m+n}]$. To verify this we only have to check that, for every k-algebra R, the map

$$\mathrm{Hom}_{k\text{-alg}}(k[X_1, \ldots, X_{m+n}], R) \to \mathrm{Hom}_{k\text{-alg}}(k[X_1, \ldots], R) \times \mathrm{Hom}_{k\text{-alg}}(k[X_{m+1}, \ldots], R)$$

induced by the inclusions is a bijection. But this map can be identified with the bijection

$$R^{m+n} \to R^m \times R^n.$$

In terms of the constructive definition of tensor products, the map

$$f \otimes g \mapsto fg : k[X_1, \ldots, X_m] \otimes_k k[X_{m+1}, \ldots, X_{m+n}] \to k[X_1, \ldots, X_{m+n}]$$

is an isomorphism.

REMARK 1.58. (a) If (b_α) is a family of generators (resp. basis) for B as a k-vector space, then $(1 \otimes b_\alpha)$ is a family of generators (resp. basis) for $A \otimes_k B$ as an A-module.

(b) Let $k \hookrightarrow \Omega$ be fields. Then

$$\Omega \otimes_k k[X_1, \ldots, X_n] \simeq \Omega[1 \otimes X_1, \ldots, 1 \otimes X_n] \simeq \Omega[X_1, \ldots, X_n].$$

If $A = k[X_1, \ldots, X_n]/(g_1, \ldots, g_m)$, then

$$\Omega \otimes_k A \simeq \Omega[X_1, \ldots, X_n]/(g_1, \ldots, g_m).$$

(c) If A and B are algebras of k-valued functions on sets S and T respectively, then $(f \otimes g)(x, y) = f(x)g(y)$ realizes $A \otimes_k B$ as an algebra of k-valued functions on $S \times T$.

f Transcendence bases

We review the theory of transcendence bases. For the proofs, see Chapter 9 of my notes *Fields and Galois Theory.*

1.59. Elements $\alpha_1, \ldots, \alpha_n$ of a k-algebra A are said to be ***algebraically dependent*** over k there exists a nonzero polynomial $f(X_1, \ldots, X_n) \in k[X_1, \ldots, X_n]$ such that $f(\alpha_1, \ldots, \alpha_n) = 0$. Otherwise, the α_i are said to be ***algebraically independent*** over k.

Now let Ω be a field containing k.

1.60. For a subset A of Ω, we let $k(A)$ denote the smallest subfield of Ω containing k and A. For example, if $A = \{x_1, \ldots, x_m\}$, then $k(A)$ consists of the quotients $\frac{f(x_1, \ldots, x_m)}{g(x_1, \ldots, x_m)}$ with $f, g \in k[X_1, \ldots, X_m]$. A subset B of Ω is ***algebraically dependent*** on A if each element of B is algebraic over $k(A)$.

1.61 (FUNDAMENTAL THEOREM). Let $A = \{\alpha_1, \ldots, \alpha_m\}$ and $B = \{\beta_1, \ldots, \beta_n\}$ be two subsets of Ω. Assume

(a) A is algebraically independent (over k);

(b) A is algebraically dependent on B (over k).

Then $m \leq n$.

The reader should note the similarity of this to the statement in linear algebra with "algebraically" replaced by "linearly".

1.62. A ***transcendence basis*** for Ω over k is an algebraically independent set A such that Ω is algebraic over $k(A)$.

1.63. Assume that there is a finite subset $A \subset \Omega$ such that Ω is algebraic over $k(A)$. Then

(a) every maximal algebraically independent subset of Ω is a transcendence basis;

(b) every subset A minimal among those such that Ω is algebraic over $k(A)$ is a transcendence basis;

(c) all transcendence bases for Ω over k have the same finite number of elements (called the ***transcendence degree***, tr. $\deg_k \Omega$, of Ω over k).

1.64. Let $k \subset L \subset \Omega$ be fields. Then

$$\text{tr.deg}_k \Omega = \text{tr.deg}_k L + \text{tr.deg}_L \Omega.$$

More precisely, if A is a transcendence basis for L/k and B is a transcendence basis for Ω/L, then $A \cup B$ is a transcendence basis for Ω/k.

Exercises

1-1. Let k be an infinite field (not necessarily algebraically closed). Show that an $f \in k[X_1, \ldots, X_n]$ that is identically zero on k^n is the zero polynomial (i.e., has all its coefficients zero).

1-2. Find a minimal set of generators for the ideal

$$(X+2Y, 3X+6Y+3Z, 2X+4Y+3Z)$$

in $k[X,Y,Z]$. What standard algorithm in linear algebra will allow you to answer this question for any ideal generated by homogeneous linear polynomials? Find a minimal set of generators for the ideal

$$(X+2Y+1, 3X+6Y+3X+2, 2X+4Y+3Z+3).$$

1-3. A ring A is said to be ***normal*** if $A_{\mathfrak{p}}$ is a normal integral domain for all prime ideals $\mathfrak{p}$ in A. Show that a noetherian ring is normal if and only if it is a finite product of normal integral domains.

1-4. Prove the statement in (1.64).

CHAPTER 2

Algebraic Sets

a Definition of an algebraic set

An ***algebraic subset*** $V(S)$ of k^n is the set of common zeros of some collection S of polynomials in $k[X_1,\ldots,X_n]$:

$$V(S)=\{(a_1,\ldots,a_n)\in k^n \mid f(a_1,\ldots,a_n)=0 \text{ all } f\in S\}.$$

We refer to $V(S)$ as the ***zero set*** of S. Note that

$$S\subset S' \implies V(S)\supset V(S');$$

— more equations means fewer solutions.

Recall that the ideal $\mathfrak{a}$ generated by a set S consists of the finite sums

$$\sum f_i g_i, \quad f_i\in k[X_1,\ldots,X_n], \quad g_i\in S.$$

Such a sum $\sum f_i g_i$ is zero at every point at which the g_i are all zero, and so $V(S)\subset V(\mathfrak{a})$, but the reverse conclusion is also true because $S\subset\mathfrak{a}$. Thus $V(S)=V(\mathfrak{a})$ — the zero set of S is the same as the zero set of the ideal generated by S. Therefore the algebraic subsets of k^n can also be described as the zero sets of ideals in $k[X_1,\ldots,X_n]$.

An empty set of polynomials imposes no conditions, and so $V(\emptyset)=k^n$. Therefore k^n is an algebraic subset. Alternatively, k^n is the zero set of the zero ideal (0). We write $\mathbb{A}^n$ for k^n regarded as an algebraic set.

Examples

2.1. If S is a set of homogeneous linear equations,

$$a_{i1}X_1+\cdots+a_{in}X_n=0, \qquad i=1,\ldots,m,$$

then $V(S)$ is a subspace of k^n. If S is a set of nonhomogeneous linear equations,

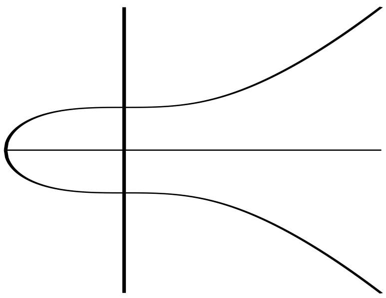
$$a_{i1}X_1+\cdots+a_{in}X_n=d_i, \qquad i=1,\ldots,m,$$

then $V(S)$ is either empty or is the translate of a subspace of k^n.

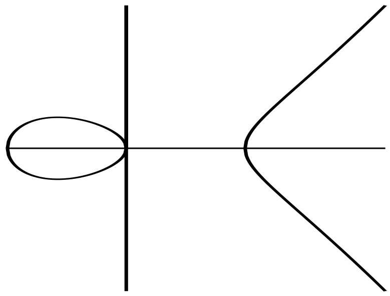
2.2. If S consists of the single equation

$$Y^2 = X^3 + aX + b, \quad 4a^3 + 27b^2 \neq 0,$$

then $V(S)$ is an ***elliptic curve***. For example:



$Y^2 = X^3 + 1$



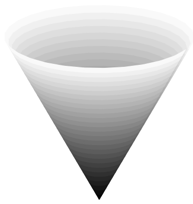
$Y^2 = X(X^2 - 1)$

We generally visualize algebraic sets as though the field k were $\mathbb{R}$, i.e., we draw the ***real locus*** of the curve. However, this can be misleading — see the examples (4.11) and (4.17) below.

2.3. If S consists of the single equation

$$Z^2 = X^2 + Y^2,$$

then $V(S)$ is a cone.



2.4. A nonzero constant polynomial has no zeros, and so the empty set is algebraic.

2.5. The proper algebraic subsets of k are the finite subsets, because a polynomial $f(X)$ in one variable X has only finitely many roots.

2.6. Some generating sets for an ideal will be more useful than others for determining what the algebraic set is. For example, the ideal

$$\mathfrak{a} = (X^2 + Y^2 + Z^2 - 1,\ X^2 + Y^2 - Y,\ X - Z)$$

can be generated by[1]

$$X - Z,\ Y^2 - 2Y + 1,\ Z^2 - 1 + Y.$$

The middle polynomial has (double) root 1, from which it follows that $V(\mathfrak{a})$ consists of the single point $(0,1,0)$.

b The Hilbert basis theorem

In our definition of an algebraic set, we didn't require the set S of polynomials to be finite, but the Hilbert basis theorem shows that, in fact, every algebraic set is the zero set of a finite set of polynomials. More precisely, the theorem states that every ideal in $k[X_1, \ldots, X_n]$ can be generated by a finite set of elements, and we have already observed that a set of generators of an ideal has the same zero set as the ideal.

[1] This is, in fact, a Gröbner basis for the ideal.

THEOREM 2.7 (HILBERT BASIS THEOREM). *The ring $k[X_1,\ldots,X_n]$ is noetherian.*

As we noted in the proof of (1.32),

$$k[X_1,\ldots,X_n] = k[X_1,\ldots,X_{n-1}][X_n].$$

Thus an induction argument shows that the theorem follows from the next lemma.

LEMMA 2.8. *If A is noetherian, then so also is $A[X]$.*

PROOF. We shall show that every ideal in $A[X]$ is finitely generated. Recall that for a polynomial

$$f(X) = a_0X^r + a_1X^{r-1} + \cdots + a_r, \quad a_i \in A, \quad a_0 \neq 0,$$

a_0 is called the leading coefficient of f.

Let $\mathfrak{a}$ be a proper ideal in $A[X]$, and let $\mathfrak{a}(i)$ denote the set of elements of A that occur as the leading coefficient of a polynomial in $\mathfrak{a}$ of degree i (we also include 0). Clearly, $\mathfrak{a}(i)$ is an ideal in A, and $\mathfrak{a}(i) \subset \mathfrak{a}(i+1)$ because, if $cX^i + \cdots \in \mathfrak{a}$, then $X(cX^i + \cdots) \in \mathfrak{a}$.

Let $\mathfrak{b}$ be an ideal of $A[X]$ contained in $\mathfrak{a}$. Then $\mathfrak{b}(i) \subset \mathfrak{a}(i)$, and if equality holds for all i, then $\mathfrak{b} = \mathfrak{a}$. To see this, let f be a polynomial in $\mathfrak{a}$. Because $\mathfrak{b}(\deg f) = \mathfrak{a}(\deg f)$, there exists a $g \in \mathfrak{b}$ such that $\deg(f-g) < \deg(f)$. In other words, $f = g + f_1$ with $g \in \mathfrak{b}$ and $\deg(f_1) < \deg(f)$. Similarly, $f_1 = g_1 + f_2$ with $g_1 \in \mathfrak{b}$ and $\deg(f_2) < \deg(f_1)$. Continuing in this fashion, we find that $f = g + g_1 + g_2 + \cdots \in \mathfrak{b}$..

As A is noetherian, the sequence

$$\mathfrak{a}(1) \subset \mathfrak{a}(2) \subset \cdots \subset \mathfrak{a}(i) \subset \cdots$$

eventually becomes constant, say $\mathfrak{a}(d) = \mathfrak{a}(d+1) = \ldots$ (and then $\mathfrak{a}(d)$ contains the leading coefficient of *every* polynomial in $\mathfrak{a}$). For each $i \leq d$, there exists a finite generating set $\{a_{i1}, a_{i2}, \ldots, a_{in_i}\}$ of $\mathfrak{a}(i)$, and for each (i,j), there exists an $f_{ij} \in \mathfrak{a}$ with leading coefficient a_{ij}. The ideal $\mathfrak{b}$ of $A[X]$ generated by the (finitely many) f_{ij} is contained in $\mathfrak{a}$ and has the property that $\mathfrak{b}(i) = \mathfrak{a}(i)$ for all i. Therefore $\mathfrak{b} = \mathfrak{a}$, and $\mathfrak{a}$ is finitely generated. □

ASIDE 2.9. One may ask how many elements are needed to generate a given ideal $\mathfrak{a}$ in $k[X_1,\ldots,X_n]$, or, what is not quite the same thing, how many equations are needed to define a given algebraic set V. For $n = 1$, the ring $k[X]$ is a principal ideal domain, which means that every ideal is generated by a single element. Also, if V is a linear subspace of k^n, then linear algebra shows that it is the zero set of $n - \dim(V)$ polynomials. All one can say in general, is that *at least* $n - \dim(V)$ polynomials are needed to define V (see 3.45), but often more are required. Determining exactly how many is an area of active research — see (3.55).

c The Zariski topology

Recall that, for ideals $\mathfrak{a}$ and $\mathfrak{b}$ in $k[X_1,\ldots,X_n]$,

$$\mathfrak{a} \subset \mathfrak{b} \implies V(\mathfrak{a}) \supset V(\mathfrak{b}).$$

PROPOSITION 2.10. *There are the following relations:*

(a) $V(0) = k^n$; $V(k[X_1,\ldots,X_n]) = \emptyset$;

(b) $V(\mathfrak{ab}) = V(\mathfrak{a} \cap \mathfrak{b}) = V(\mathfrak{a}) \cup V(\mathfrak{b})$;

(c) $V(\sum_{i\in I} \mathfrak{a}_i) = \bigcap_{i\in I} V(\mathfrak{a}_i)$ *for every family of ideals* $(\mathfrak{a}_i)_{i\in I}$.

PROOF. (a) The is obvious.

(b) Note that

$$\mathfrak{ab} \subset \mathfrak{a} \cap \mathfrak{b} \subset \mathfrak{a}, \mathfrak{b} \implies V(\mathfrak{ab}) \supset V(\mathfrak{a} \cap \mathfrak{b}) \supset V(\mathfrak{a}) \cup V(\mathfrak{b}).$$

For the reverse inclusions, observe that if $a \notin V(\mathfrak{a}) \cup V(\mathfrak{b})$, then there exist $f \in \mathfrak{a}$, $g \in \mathfrak{b}$ such that $f(a) \neq 0$, $g(a) \neq 0$; but then $(fg)(a) \neq 0$, and so $a \notin V(\mathfrak{ab})$.

(c) Recall that, by definition, $\sum \mathfrak{a}_i$ consists of all finite sums of the form $\sum f_i$, $f_i \in \mathfrak{a}_i$. Thus (c) is obvious. □

Statements (a), (b), and (c) show that the algebraic subsets of $\mathbb{A}^n$ satisfy the axioms to be the closed subsets for a topology on $\mathbb{A}^n$: both the whole space and the empty set are algebraic; a finite union of algebraic sets is algebraic; an arbitrary intersection of algebraic sets is algebraic. Thus, there is a topology on $\mathbb{A}^n$ for which the closed subsets are exactly the algebraic subsets — this is called the ***Zariski topology*** on $\mathbb{A}^n$. The induced topology on a subset V of $\mathbb{A}^n$ is called the ***Zariski topology*** on V.

The Zariski topology has many strange properties, but it is nevertheless of great importance. For the Zariski topology on k, the closed subsets are just the finite sets and the whole space, and so the topology is not Hausdorff (in fact, there are no disjoint nonempty open subsets at all). We shall see in (2.68) below that the proper closed subsets of k^2 are finite unions of points and curves. Note that the Zariski topologies on $\mathbb{C}$ and $\mathbb{C}^2$ are much coarser (have fewer open sets) than the complex topologies.

d The Hilbert Nullstellensatz

We wish to examine the relation between the algebraic subsets of $\mathbb{A}^n$ and the ideals of $k[X_1, \ldots, X_n]$ more closely, but first we must answer the question of when a collection S of polynomials has a common zero, i.e., when the system of equations

$$g(X_1, \ldots, X_n) = 0, \quad g \in S,$$

is "consistent". Obviously, equations

$$g_i(X_1, \ldots, X_n) = 0, \quad i = 1, \ldots, m$$

are inconsistent if there exist $f_i \in k[X_1, \ldots, X_n]$ such that $\sum f_i g_i = 1$, i.e., if $1 \in (g_1, \ldots, g_m)$ or, equivalently, $(g_1, \ldots, g_m) = k[X_1, \ldots, X_n]$. The next theorem provides a converse to this.

THEOREM 2.11 (HILBERT NULLSTELLENSATZ). [2] *Every proper ideal $\mathfrak{a}$ in $k[X_1, \ldots, X_n]$ has a zero in k^n.*

A point $P = (a_1, \ldots, a_n)$ in k^n defines a homomorphism "evaluate at P"

$$k[X_1, \ldots, X_n] \to k, \quad f(X_1, \ldots, X_n) \mapsto f(a_1, \ldots, a_n),$$

whose kernel contains $\mathfrak{a}$ if $P \in V(\mathfrak{a})$. Conversely, from a homomorphism $\varphi \colon k[X_1, \ldots, X_n] \to k$ of k-algebras whose kernel contains $\mathfrak{a}$, we obtain a point P in $V(\mathfrak{a})$, namely,

$$P = (\varphi(X_1), \ldots, \varphi(X_n)).$$

[2] Nullstellensatz = zero-points-theorem.

Thus, to prove the theorem, we have to show that there exists a k-algebra homomorphism $k[X_1,\ldots,X_n]/\mathfrak{a}\to k$.

Since every proper ideal is contained in a maximal ideal (see p.14), it suffices to prove this for a maximal ideal $\mathfrak{m}$. Then $K \overset{\text{def}}{=} k[X_1,\ldots,X_n]/\mathfrak{m}$ is a field, and it is finitely generated as an algebra over k (with generators $X_1+\mathfrak{m},\ldots,X_n+\mathfrak{m}$). To complete the proof, we must show $K=k$. The next lemma accomplishes this.

In the statement of the lemma, we need to allow k to be arbitrary in order to make the induction in the proof work (but the generality will be useful later — see Chapter 10). We shall also need to use that $k[X]$ has infinitely many distinct monic irreducible polynomials. When k is infinite, the polynomials $X-a$, $a\in k$, are distinct and irreducible. When k is finite, we can adapt Euclid's argument: if $p_1,\ldots,p_r$ are monic irreducible polynomials in $k[X]$, then $p_1\cdots p_r+1$ is divisible by a monic irreducible polynomial distinct from $p_1,\ldots,p_r$.

LEMMA 2.12 (ZARISKI'S LEMMA). *Let $k\subset K$ be fields, not necessarily algebraically closed. If K is finitely generated as an algebra over k, then K is algebraic over k. (Hence $K=k$ if k is algebraically closed.)*

In other words, if K is finitely generated as a ring over k, then it is finitely generated as a module.

PROOF. We shall prove this by induction on r, the minimum number of elements required to generate K as a k-algebra. The case $r=0$ being trivial, we may suppose that

$$K=k[x_1,\ldots,x_r],\quad r\geq 1.$$

If K is not algebraic over k, then at least one x_i, say x_1, is not algebraic over k. Then, $k[x_1]$ is a polynomial ring in one symbol over k, and its field of fractions $k(x_1)$ is a subfield of K. Clearly K is generated as a $k(x_1)$-algebra by $x_2,\ldots,x_r$, and so the induction hypothesis implies that $x_2,\ldots,x_r$ are algebraic over $k(x_1)$. From (1.40), we see that there exists a $c\in k[x_1]$ such that $cx_2,\ldots,cx_r$ are integral over $k[x_1]$.

Let $f\in k(x_1)$. Then $f\in K=k[x_1,\ldots,x_r]$ and so, for a sufficiently large N, $c^Nf\in k[x_1,cx_2,\ldots,cx_r]$. Therefore c^Nf is integral over $k[x_1]$ by (1.38), which implies that $c^Nf\in k[x_1]$ because $k[x_1]$ is integrally closed in $k(x_1)$ (1.43). But this contradicts the fact that that $k[x_1]$ has infinitely many distinct monic irreducible polynomials that can occur as the denominator of an f in $k(x_1)$. □

e The correspondence between algebraic sets and radical ideals

The ideal attached to a subset of k^n

For a subset W of k^n, we write $I(W)$ for the set of polynomials that are zero on W:

$$I(W)=\{f\in k[X_1,\ldots,X_n]\mid f(P)=0 \text{ all } P\in W\}.$$

Clearly, it is an ideal in $k[X_1,\ldots,X_n]$. There are the following relations:

(a) $V\subset W \implies I(V)\supset I(W)$;

(b) $I(\emptyset)=k[X_1,\ldots,X_n]$; $I(k^n)=0$;

(c) $I(\bigcup W_i) = \bigcap I(W_i)$.

Only the statement $I(k^n) = 0$ is (perhaps) not obvious. It says that every nonzero polynomial in $k[X_1,\ldots,X_n]$ is nonzero at some point of k^n. This is true with k any infinite field (see Exercise 1-1). Alternatively, it follows from the strong Hilbert Nullstellensatz (see 2.19 below).

EXAMPLE 2.13. Let P be the point $(a_1,\ldots,a_n)$, and let

$$\mathfrak{m}_P = (X_1 - a_1,\ldots,X_n - a_n).$$

Clearly $I(P) \supset \mathfrak{m}_P$, but $\mathfrak{m}_P$ is a maximal ideal, because "evaluation at $(a_1,\ldots,a_n)$" defines an isomorphism

$$k[X_1,\ldots,X_n]/(X_1 - a_1,\ldots,X_n - a_n) \to k.$$

As $I(P)$ is a proper ideal, it must equal $\mathfrak{m}_P$.

PROPOSITION 2.14. *Let W be a subset of k^n. Then $VI(W)$ is the smallest algebraic subset of k^n containing W. In particular, $VI(W) = W$ if W is an algebraic set.*

PROOF. Certainly $VI(W)$ is an algebraic set containing W. Let $V = V(\mathfrak{a})$ be another algebraic set containing W. Then $\mathfrak{a} \subset I(W)$, and so $V(\mathfrak{a}) \supset VI(W)$. □

Radicals of ideals

The ***radical*** of an ideal $\mathfrak{a}$ in a ring A is

$$\mathrm{rad}(\mathfrak{a}) \stackrel{\mathrm{def}}{=} \{f \mid f^r \in \mathfrak{a}, \text{ some } r \in \mathbb{N}\}.$$

PROPOSITION 2.15. *Let $\mathfrak{a}$ be an ideal in a ring A.*
(a) The radical of $\mathfrak{a}$ is an ideal.
(b) $\mathrm{rad}(\mathrm{rad}(\mathfrak{a})) = \mathrm{rad}(\mathfrak{a})$.

PROOF. (a) If $a \in \mathrm{rad}(\mathfrak{a})$, then clearly $fa \in \mathrm{rad}(\mathfrak{a})$ for all $f \in A$. Suppose that $a, b \in \mathrm{rad}(\mathfrak{a})$, with say $a^r \in \mathfrak{a}$ and $b^s \in \mathfrak{a}$. When we expand $(a+b)^{r+s}$ using the binomial theorem, we find that every term has a factor a^r or b^s, and so lies in $\mathfrak{a}$.
(b) If $a^r \in \mathrm{rad}(\mathfrak{a})$, then $a^{rs} = (a^r)^s \in \mathfrak{a}$ for some s. □

An ideal is said to be ***radical*** if it equals its radical. Thus $\mathfrak{a}$ is radical if and only if the ring $A/\mathfrak{a}$ is ***reduced***, i.e., without nonzero ***nilpotent*** elements. Since integral domains are reduced, prime ideals (*a fortiori*, maximal ideals) are radical. Note that $\mathrm{rad}(\mathfrak{a})$ is radical (2.15b), and hence is the smallest radical ideal containing $\mathfrak{a}$.

If $\mathfrak{a}$ and $\mathfrak{b}$ are radical, then $\mathfrak{a} \cap \mathfrak{b}$ is radical, but $\mathfrak{a} + \mathfrak{b}$ need not be: consider, for example, $\mathfrak{a} = (X^2 - Y)$ and $\mathfrak{b} = (X^2 + Y)$; they are both prime ideals in $k[X,Y]$, but $X^2 \in \mathfrak{a} + \mathfrak{b}$, $X \notin \mathfrak{a} + \mathfrak{b}$. (See 2.22 below.)

The strong Nullstellensatz

For a polynomial f and point $P \in k^n$, $f^r(P) = f(P)^r$. Therefore f^r is zero on the same set as f, and it follows that the ideal $I(W)$ is radical for every subset $W \subset k^n$. In particular, $IV(\mathfrak{a}) \supset \mathrm{rad}(\mathfrak{a})$. The next theorem states that these two ideals are equal.

THEOREM 2.16 (STRONG NULLSTELLENSATZ). *For every ideal $\mathfrak{a}$ in $k[X_1,\ldots,X_n]$,*

$$IV(\mathfrak{a}) = \mathrm{rad}(\mathfrak{a});$$

in particular, $IV(\mathfrak{a}) = \mathfrak{a}$ if $\mathfrak{a}$ is a radical ideal.

PROOF. We have already noted that $IV(\mathfrak{a}) \supset \mathrm{rad}(\mathfrak{a})$. For the reverse inclusion, we have to show that if a polynomial h vanishes on $V(\mathfrak{a})$, then $h^N \in \mathfrak{a}$ for some $N > 0$. We may assume $h \neq 0$. Let $g_1,\ldots,g_m$ generate $\mathfrak{a}$, and consider the system of $m+1$ equations in $n+1$ symbols,

$$\begin{cases} g_i(X_1,\ldots,X_n) &= 0, \quad i = 1,\ldots,m \\ 1 - Yh(X_1,\ldots,X_n) &= 0. \end{cases}$$

If $(a_1,\ldots,a_n,b)$ satisfies the first m equations, then $(a_1,\ldots,a_n) \in V(\mathfrak{a})$; consequently, $h(a_1,\ldots,a_n) = 0$, and $(a_1,\ldots,a_n,b)$ doesn't satisfy the last equation. Therefore, the equations are inconsistent, and so, according to the original Nullstellensatz, there exist $f_i \in k[X_1,\ldots,X_n,Y]$ such that

$$1 = \sum_{i=1}^{m} f_i \cdot g_i + f_{m+1} \cdot (1 - Yh)$$

(in the ring $k[X_1,\ldots,X_n,Y]$). On applying the homomorphism

$$\begin{cases} X_i \mapsto X_i \\ Y \mapsto h^{-1} \end{cases} : k[X_1,\ldots,X_n,Y] \to k(X_1,\ldots,X_n)$$

to the above equality, we obtain the identity

$$1 = \sum_{i=1}^{m} f_i(X_1,\ldots,X_n,h^{-1}) \cdot g_i(X_1,\ldots,X_n) \tag{*}$$

in $k(X_1,\ldots,X_n)$. Clearly

$$f_i(X_1,\ldots,X_n,h^{-1}) = \frac{\text{polynomial in } X_1,\ldots,X_n}{h^{N_i}}$$

for some N_i. Let N be the largest of the N_i. On multiplying (*) by h^N we obtain an equation

$$h^N = \sum_{i=1}^{m} (\text{polynomial in } X_1,\ldots,X_n) \cdot g_i(X_1,\ldots,X_n),$$

which shows that $h^N \in \mathfrak{a}$. □

COROLLARY 2.17. *The map $\mathfrak{a} \mapsto V(\mathfrak{a})$ defines a one-to-one correspondence between the set of radical ideals in $k[X_1,\ldots,X_n]$ and the set of algebraic subsets of k^n; its inverse is I.*

PROOF. We know that $IV(\mathfrak{a}) = \mathfrak{a}$ if $\mathfrak{a}$ is a radical ideal (2.16), and that $VI(W) = W$ if W is an algebraic set (2.14). Therefore, I and V are inverse bijections. □

COROLLARY 2.18. *The radical of an ideal in $k[X_1,\ldots,X_n]$ is equal to the intersection of the maximal ideals containing it.*

PROOF. Let $\mathfrak{a}$ be an ideal in $k[X_1,\ldots,X_n]$. Because maximal ideals are radical, every maximal ideal containing $\mathfrak{a}$ also contains $\mathrm{rad}(\mathfrak{a})$, and so

$$\mathrm{rad}(\mathfrak{a}) \subset \bigcap_{\mathfrak{m}\supset\mathfrak{a}} \mathfrak{m}.$$

For each $P=(a_1,\ldots,a_n)\in k^n$, the ideal $\mathfrak{m}_P=(X_1-a_1,\ldots,X_n-a_n)$ is maximal in $k[X_1,\ldots,X_n]$, and

$$f\in\mathfrak{m}_P \iff f(P)=0$$

(see 2.13). Thus

$$\mathfrak{m}_P\supset\mathfrak{a} \iff P\in V(\mathfrak{a}).$$

If $f\in\mathfrak{m}_P$ for all $P\in V(\mathfrak{a})$, then f is zero on $V(\mathfrak{a})$, and so $f\in IV(\mathfrak{a})=\mathrm{rad}(\mathfrak{a})$. We have shown that

$$\mathrm{rad}(\mathfrak{a})\supset\bigcap_{P\in V(\mathfrak{a})}\mathfrak{m}_P\supset\bigcap_{\mathfrak{m}\supset\mathfrak{a}}\mathfrak{m}.$$

□

Remarks

2.19. Because $V(0)=k^n$,

$$I(k^n)=IV(0)=\mathrm{rad}(0)=0;$$

in other words, only the zero polynomial is zero on the whole of k^n. In fact, this holds whenever k is infinite (Exercise 1-1).

2.20. The one-to-one correspondence in Corollary 2.17 is order reversing. Therefore the maximal proper radical ideals correspond to the minimal nonempty algebraic sets. But the maximal proper radical ideals are simply the maximal ideals in $k[X_1,\ldots,X_n]$, and the minimal nonempty algebraic sets are the one-point sets. As

$$I((a_1,\ldots,a_n))=(X_1-a_1,\ldots,X_n-a_n)$$

(see 2.13), this shows that the maximal ideals of $k[X_1,\ldots,X_n]$ are exactly the ideals $(X_1-a_1,\ldots,X_n-a_n)$ with $(a_1,\ldots,a_n)\in k^n$.

2.21. The algebraic set $V(\mathfrak{a})$ is empty if and only if $\mathfrak{a}=k[X_1,\ldots,X_n]$ (Nullstellensatz, 2.11).

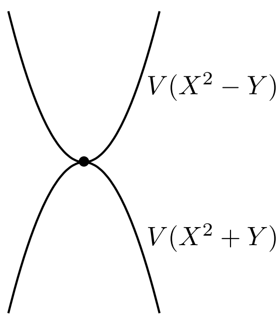
2.22. Let W and W' be algebraic sets. As $W\cap W'$ is the largest algebraic subset contained in both W and W', $I(W\cap W')$ must be the smallest radical ideal containing both $I(W)$ and $I(W')$:

$$I(W\cap W')=\mathrm{rad}(I(W)+I(W')).$$

For example, let $W=V(X^2-Y)$ and $W'=V(X^2+Y)$; then

$$I(W\cap W')=\mathrm{rad}(X^2,Y)=(X,Y)$$

(assuming characteristic $\neq 2$). Note that $W\cap W'=\{(0,0)\}$, but when realized as the intersection of $Y=X^2$ and $Y=-X^2$, it has "multiplicity 2".



2.23. Let $\mathcal{P}$ be the set of subsets of k^n and let $\mathcal{Q}$ be the set of subsets of $k[X_1,\ldots,X_n]$. Then $I:\mathcal{P}\to\mathcal{Q}$ and $V:\mathcal{Q}\to\mathcal{P}$ define a simple Galois correspondence between $\mathcal{P}$ and $\mathcal{Q}$ (see FT, 7.18). It follows that I and V define a one-to-one correspondence between $I(\mathcal{P})$ and $V(\mathcal{Q})$. But the strong Nullstellensatz shows that $I(\mathcal{P})$ consists exactly of the radical ideals, and (by definition) $V(\mathcal{Q})$ consists of the algebraic subsets. Thus we recover Corollary 2.17.

ASIDE 2.24. The algebraic subsets of $\mathbb{A}^n$ capture only part of the ideal theory of $k[X_1,\ldots,X_n]$ because two ideals with the same radical correspond to the same algebraic subset. There is a finer notion of an algebraic scheme over k for which the closed algebraic subschemes of $\mathbb{A}^n$ are in one-to-one correspondence with the ideals in $k[X_1,\ldots,X_n]$ (see Chapter 11 on my website).

f Finding the radical of an ideal

Typically, an algebraic set V will be defined by a finite set of polynomials $\{g_1,\ldots,g_s\}$, and then we shall need to find $I(V)=\mathrm{rad}(g_1,\ldots,g_s)$.

PROPOSITION 2.25. *A polynomial $h\in \mathrm{rad}(\mathfrak{a})$ if and only if $1\in(\mathfrak{a},1-Yh)$ (the ideal in $k[X_1,\ldots,X_n,Y]$ generated by the elements of $\mathfrak{a}$ and $1-Yh$).*

PROOF. We saw that $1\in(\mathfrak{a},1-Yh)$ implies $h\in\mathrm{rad}(\mathfrak{a})$ in the course of proving (2.16). Conversely, from the identities

$$1=Y^Nh^N+(1-Y^Nh^N)=Y^Nh^N+(1-Yh)\cdot(1+Yh+\cdots+Y^{N-1}h^{N-1})$$

we see that, if $h^N\in\mathfrak{a}$, then $1\in\mathfrak{a}+(1-Yh)$. □

Given a set of generators of an ideal, there is an algorithm for deciding whether or not a polynomial belongs to the ideal, and hence an algorithm for deciding whether or not a polynomial belongs to the radical of the ideal. There are even algorithms for finding a set of generators for the radical. These algorithms have been implemented in the computer algebra systems CoCoA and Macaulay 2.

g Properties of the Zariski topology

We now examine more closely the Zariski topology on $\mathbb{A}^n$ and on an algebraic subset of $\mathbb{A}^n$. Proposition 2.14 says that, for each subset W of $\mathbb{A}^n$, $VI(W)$ is the closure of W, and (2.17) says that there is a one-to-one correspondence between the closed subsets of $\mathbb{A}^n$ and the radical ideals of $k[X_1,\ldots,X_n]$. Under this correspondence, the closed subsets of an algebraic set V correspond to the radical ideals of $k[X_1,\ldots,X_n]$ containing $I(V)$.

PROPOSITION 2.26. *Let V be an algebraic subset of $\mathbb{A}^n$.*

(a) The points of V are closed for the Zariski topology.

(b) Every ascending chain of open subsets $U_1\subset U_2\subset\cdots$ of V eventually becomes constant. Equivalently, every descending chain of closed subsets of V eventually becomes constant.

(c) Every open covering of V has a finite subcovering.

PROOF. (a) We have seen that $\{(a_1,\ldots,a_n)\}$ is the algebraic set defined by the ideal $(X_1 - a_1,\ldots,X_n - a_n)$.

(b) We prove the second statement. A sequence $V_1 \supset V_2 \supset \cdots$ of closed subsets of V gives rise to a sequence of radical ideals $I(V_1) \subset I(V_2) \subset \ldots$, which eventually becomes constant because $k[X_1,\ldots,X_n]$ is noetherian.

(c) Given an open covering of V, let $\mathcal{U}$ be the collection of open subsets of V that can be expressed as a finite union of sets in the covering. If $\mathcal{U}$ does not contain V, then there exists an infinite ascending chain of sets in $\mathcal{U}$ (axiom of dependent choice), contradicting (b). □

A topological space whose points are closed is said to be a ***T_1-space***; the condition means that, for every pair of distinct points, each has an open neighbourhood not containing the other. A topological space having the property (b) is said to be ***noetherian***. The condition is equivalent to the following: every nonempty set of closed subsets of V has a minimal element. A space having property (c) is said to be ***quasicompact*** (by Bourbaki at least; others call it compact, but Bourbaki requires a compact space to be Hausdorff). The proof of (c) shows that every noetherian space is quasicompact. Since an open subset of a noetherian space is again noetherian, it will also be quasicompact.

h Decomposition of an algebraic set into irreducible algebraic sets

A topological space is said to be ***irreducible*** if it is not the union of two proper closed subsets. Equivalent conditions: every pair of nonempty open subsets has nonempty intersection; every nonempty open subset is dense. By convention, the empty space is not irreducible. Obviously, every nonempty open subset of an irreducible space is irreducible.

In a Hausdorff topological space, any two points have disjoint open neighbourhoods. Therefore, the only irreducible Hausdorff spaces are those consisting of a single point.

PROPOSITION 2.27. *An algebraic set W is irreducible if and only if $I(W)$ is prime.*

PROOF. Let W be an irreducible algebraic set, and let $fg \in I(W)$ — we have to show that either f or g is in $I(W)$. At each point of W, either f is zero or g is zero, and so $W \subset V(f) \cup V(g)$. Hence

$$W = (W \cap V(f)) \cup (W \cap V(g)).$$

As W is irreducible, one of these sets, say $W \cap V(f)$, must equal W. But then $f \in I(W)$.

Let W be an algebraic set such that $I(W)$ is prime, and let $W = V(\mathfrak{a}) \cup V(\mathfrak{b})$ with $\mathfrak{a}$ and $\mathfrak{b}$ radical ideals — we have to show that W equals $V(\mathfrak{a})$ or $V(\mathfrak{b})$. The ideal $\mathfrak{a} \cap \mathfrak{b}$ is radical, and $V(\mathfrak{a} \cap \mathfrak{b}) = V(\mathfrak{a}) \cup V(\mathfrak{b})$ (2.10); hence $I(W) = \mathfrak{a} \cap \mathfrak{b}$. If $W \neq V(\mathfrak{a})$, then there exists an $f \in \mathfrak{a} \smallsetminus I(W)$. Let $g \in \mathfrak{b}$. Then $fg \in \mathfrak{a} \cap \mathfrak{b} = I(W)$, and so $g \in I(W)$ (because $I(W)$ is prime). We conclude that $\mathfrak{b} \subset I(W)$, and so $V(\mathfrak{b}) \supset V(I(W)) = W$. □

SUMMARY 2.28. There are one-to-one correspondences (ideals in $k[X_1,\ldots,X_n]$; algebraic subsets of $\mathbb{A}^n$):

$$\begin{aligned} \text{radical ideals} &\leftrightarrow \text{algebraic subsets} \\ \text{prime ideals} &\leftrightarrow \text{irreducible algebraic subsets} \\ \text{maximal ideals} &\leftrightarrow \text{one-point sets.} \end{aligned}$$

EXAMPLE 2.29. Let $f \in k[X_1, \ldots, X_n]$. We saw (1.32) that $k[X_1, \ldots, X_n]$ is a unique factorization domain, and so (f) is a prime ideal if and only if f is irreducible (1.33). Thus

$$f \text{ is irreducible } \implies V(f) \text{ is irreducible.}$$

On the other hand, suppose f factors as

$$f = \prod f_i^{m_i}, \quad f_i \text{ distinct irreducible polynomials.}$$

Then

$$\begin{aligned} (f) &= \bigcap \left(f_i^{m_i}\right) & &\left(f_i^{m_i}\right) \text{ distinct primary ideals} \\ \operatorname{rad}(f) &= \bigcap (f_i) & &(f_i) \text{ distinct prime ideals} \\ V(f) &= \bigcup V(f_i) & &V(f_i) \text{ distinct irreducible algebraic sets.} \end{aligned}$$

LEMMA 2.30. *Let W be an irreducible topological space. If $W = W_1 \cup \ldots \cup W_r$ with each W_i closed, then W is equal to one of the W_i.*

PROOF. When $r = 2$, the statement is the definition of "irreducible". Suppose that $r > 2$. Then $W = W_1 \cup (W_2 \cup \ldots \cup W_r)$, and so $W = W_1$ or $W = (W_2 \cup \ldots \cup W_r)$; if the latter, then $W = W_2$ or $W_3 \cup \ldots \cup W_r$, etc. □

PROPOSITION 2.31. *Let V be a noetherian topological space. Then V is a finite union of irreducible closed subsets, $V = V_1 \cup \ldots \cup V_m$. If the decomposition is irredundant in the sense that there are no inclusions among the V_i, then the V_i are uniquely determined up to order.*

PROOF. Suppose that V cannot be written as a finite union of irreducible closed subsets. Then, because V is noetherian, there will be a nonempty closed subset W of V that is minimal among those that cannot be written in this way. But W itself cannot be irreducible, and so $W = W_1 \cup W_2$, with W_1 and W_2 proper closed subsets of W. Because W was minimal, each W_i is a finite union of irreducible closed subsets. Hence W is also, which is a contradiction.

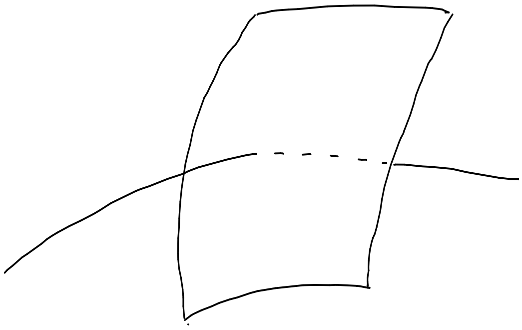
Suppose that

$$V = V_1 \cup \ldots \cup V_m = W_1 \cup \ldots \cup W_n$$

are two irredundant decompositions of V. Then $V_i = \bigcup_j (V_i \cap W_j)$, and so, because V_i is irreducible, $V_i = V_i \cap W_j$ for some j. Consequently, there is a function $f\colon \{1, \ldots, m\} \to \{1, \ldots, n\}$ such that $V_i \subset W_{f(i)}$ for each i. Similarly, there is a function $g\colon \{1, \ldots, n\} \to \{1, \ldots, m\}$ such that $W_j \subset V_{g(j)}$ for each j. Since $V_i \subset W_{f(i)} \subset V_{gf(i)}$, we must have $gf(i) = i$ and $V_i = W_{f(i)}$; similarly $fg = \mathrm{id}$. Thus f and g are bijections, and the decompositions differ only in the numbering of the sets. □

The V_i given uniquely by the proposition are called the ***irreducible components*** of V. They are exactly the maximal irreducible closed subsets of V.[3] In Example 2.29, the $V(f_i)$ are the irreducible components of $V(f)$.

[3]In fact, they are exactly the maximal irreducible subsets of V, because the closure of an irreducible subset is also irreducible.



An algebraic set with two irreducible components.

COROLLARY 2.32. *The radical of an ideal $\mathfrak{a}$ in $k[X_1,\ldots,X_n]$ is a finite intersection of prime ideals, $\mathfrak{a}=\mathfrak{p}_1\cap\ldots\cap\mathfrak{p}_n$. If there are no inclusions among the $\mathfrak{p}_i$, then the $\mathfrak{p}_i$ are uniquely determined up to order (and they are exactly the minimal prime ideals containing $\mathfrak{a}$).*

PROOF. Write $V(\mathfrak{a})$ as a union of its irreducible components, $V(\mathfrak{a})=\bigcup_{i=1}^{n}V_i$, and let $\mathfrak{p}_i=I(V_i)$. Then $\mathrm{rad}(\mathfrak{a})=\mathfrak{p}_1\cap\ldots\cap\mathfrak{p}_n$ because they are both radical ideals and

$$V(\mathrm{rad}(\mathfrak{a}))=V(\mathfrak{a})=\bigcup V(\mathfrak{p}_i)\overset{2.10\mathrm{b}}{=}V(\textstyle\bigcap_i\mathfrak{p}).$$

The uniqueness similarly follows from the proposition. □

Remarks

2.33. An irreducible topological space is connected, but a connected topological space need not be irreducible. For example, $V(X_1X_2)$ is the union of the coordinate axes in $\mathbb{A}^2$, which is connected but not irreducible. An algebraic subset V of $\mathbb{A}^n$ is disconnected if and only if there exist ideals $\mathfrak{a}$ and $\mathfrak{b}$ such that $\mathfrak{a}\cap\mathfrak{b}=I(V)$ and $\mathfrak{a}+\mathfrak{b}=k[X_1,\ldots,X_n]$, so that

$$\begin{cases} V=V(\mathfrak{a}\cap\mathfrak{b})=V(\mathfrak{a})\cup V(\mathfrak{b})\\ \emptyset=V(\mathfrak{a}+\mathfrak{b})=V(\mathfrak{a})\cap V(\mathfrak{b}).\end{cases}$$

2.34. A Hausdorff space is noetherian if and only if it is finite, in which case its irreducible components are the one-point sets.

2.35. In $k[X_1,\ldots,X_n]$, a principal ideal (f) is radical if and only if f is square-free, in which case f is a product of distinct irreducible polynomials, $f=f_1\ldots f_r$, and $(f)=(f_1)\cap\ldots\cap(f_r)$.

2.36. In a noetherian ring, every proper ideal $\mathfrak{a}$ has a decomposition into primary ideals: $\mathfrak{a}=\bigcap\mathfrak{q}_i$ (see CA §19). For radical ideals, this becomes a simpler decomposition into prime ideals, as in the corollary. For an ideal (f) with $f=\prod f_i^{m_i}$, the primary decomposition is the decomposition $(f)=\bigcap(f_i^{m_i})$ noted in Example 2.29.

i Regular functions; the coordinate ring of an algebraic set

Let V be an algebraic subset of $\mathbb{A}^n$, and let $I(V) = \mathfrak{a}$. The ***coordinate ring of*** V is

$$k[V] = k[X_1, \ldots, X_n]/\mathfrak{a}.$$

This is a finitely generated k-algebra. It is reduced (because $\mathfrak{a}$ is radical), but not necessarily an integral domain.

An $f \in k[X_1, \ldots, X_n]$ defines a function

$$P \mapsto f(P): V \to k.$$

Functions of this form are said to be ***regular.*** Two polynomials $f, g \in k[X_1, \ldots, X_n]$ define the same regular function on V if only if they define the same element of $k[V]$, and so $k[V]$ is the ring of regular functions on V. The coordinate function

$$x_i: V \to k, (a_1, \ldots, a_n) \mapsto a_i$$

is regular, and $k[V] = k[x_1, \ldots, x_n]$. In other words, the coordinate ring of an algebraic set V is the k-algebra generated by the coordinate functions on V.

For an ideal $\mathfrak{b}$ in $k[V]$, set

$$V(\mathfrak{b}) = \{P \in V \mid f(P) = 0, \text{ all } f \in \mathfrak{b}\}$$

— it is a closed subset of V. Let $W = V(\mathfrak{b})$. The quotient maps

$$k[X_1, \ldots, X_n] \to k[V] = \frac{k[X_1, \ldots, X_n]}{\mathfrak{a}} \to k[W] = \frac{k[V]}{\mathfrak{b}}$$

send a regular function on k^n to its restriction to V, and then to its restriction to W.

Write π for the quotient map $k[X_1, \ldots, X_n] \to k[V]$. Then $\mathfrak{b} \mapsto \pi^{-1}(\mathfrak{b})$ is a bijection from the set of ideals of $k[V]$ to the set of ideals of $k[X_1, \ldots, X_n]$ containing $\mathfrak{a}$, under which radical, prime, and maximal ideals correspond to radical, prime, and maximal ideals (because each of these conditions can be checked on the quotient ring, and $k[X_1, \ldots, X_n]/\pi^{-1}(\mathfrak{b}) \simeq k[V]/\mathfrak{b}$). Clearly

$$V(\pi^{-1}(\mathfrak{b})) = V(\mathfrak{b}),$$

and so $\mathfrak{b} \mapsto V(\mathfrak{b})$ is a bijection from the set of radical ideals in $k[V]$ to the set of algebraic sets contained in V.

Now (2.28) holds for ideals in $k[V]$ and algebraic subsets of V.

For $h \in k[V]$, set

$$D(h) = \{a \in V \mid h(a) \neq 0\}.$$

It is an open subset of V, because its complement is the closed set $V((h))$. It is empty if and only if h is zero (2.19).

PROPOSITION 2.37. *(a) The points of V are in one-to-one correspondence with the maximal ideals of $k[V]$.*

(b) The closed subsets of V are in one-to-one correspondence with the radical ideals of $k[V]$.

(c) The sets $D(h)$, $h \in k[V]$, are a base for the topology on V, i.e., each $D(h)$ is open, and every open set is a (finite) union of this form.

PROOF. (a) and (b) are obvious from the above discussion. For (c), we have already observed that $D(h)$ is open. Every open subset $U \subset V$ is the complement of a set of the form $V(\mathfrak{b})$, with $\mathfrak{b}$ an ideal in $k[V]$. If $f_1, \ldots, f_m$ generate $\mathfrak{b}$, then $U = \bigcup D(f_i)$. □

The $D(h)$ are called the ***basic*** (or ***principal***) ***open subsets*** of V. We sometimes write V_h for $D(h)$. Note that

$$\begin{aligned} D(h) \subset D(h') &\iff V(h) \supset V(h') \\ &\iff \mathrm{rad}((h)) \subset \mathrm{rad}((h')) \\ &\iff h^r \in (h') \text{ some } r \\ &\iff h^r = h'g, \text{ some } g. \end{aligned}$$

Some of this should look familiar: if V is a topological space, then the zero set of a family of continuous functions $f: V \to \mathbb{R}$ is closed, and the set where a continuous function is nonzero is open.

Let V be an irreducible algebraic set. Then $I(V)$ is a prime ideal, and so $k[V]$ is an integral domain. Let $k(V)$ be its field of fractions — $k(V)$ is called the ***field of rational functions*** on V.

j Regular maps

Let $W \subset k^m$ and $V \subset k^n$ be algebraic sets. Let x_i denote the ith coordinate function

$$(b_1, \ldots, b_n) \mapsto b_i : V \to k$$

on V. The ith ***component*** of a map $\varphi: W \to V$ is

$$\varphi_i = x_i \circ \varphi.$$

Thus, φ is the map

$$P \mapsto \varphi(P) = (\varphi_1(P), \ldots, \varphi_n(P)): W \to V \subset k^n.$$

DEFINITION 2.38. A continuous map $\varphi: W \to V$ of algebraic sets is ***regular*** if each of its components φ_i is a regular function.

As the coordinate functions generate $k[V]$, a continuous map φ is regular if and only if $f \circ \varphi$ is a regular function on W for every regular function f on V. Thus a regular map $\varphi: W \to V$ of algebraic sets defines a homomorphism $f \mapsto f \circ \varphi: k[V] \to k[W]$ of k-algebras, which we sometimes denote φ^*.

k Hypersurfaces; finite and quasi-finite maps

A ***hypersurface*** in $\mathbb{A}^{n+1}$ is the algebraic set H defined by a single nonzero nonconstant polynomial:

$$H: \quad f(T_1, \ldots, T_n, X) = 0.$$

We examine the regular map $H \to \mathbb{A}^n$ defined by the projection

$$(t_1, \ldots, t_n, x) \mapsto (t_1, \ldots, t_n).$$

We can write f in the form

$$f = a_0 X^m + a_1 X^{m-1} + \cdots + a_m, \quad a_i \in k[T_1, \ldots, T_m], \quad a_0 \neq 0, \quad m \in \mathbb{N}.$$

We assume that $m \neq 0$, i.e., that X occurs in f (otherwise, H is a cylinder over a hypersurface in $\mathbb{A}^n$). The fibre of the map $H \to \mathbb{A}^n$ over $(t_1, \ldots, t_n) \in k^n$ is the set of points $(t_1, \ldots, t_n, c)$ such that c is a root of the polynomial

$$a_0(t) X^m + a_1(t) X^{m-1} + \cdots + a_m(t), \quad a_i(t) \stackrel{\text{def}}{=} a_i(t_1, \ldots, t_n) \in k.$$

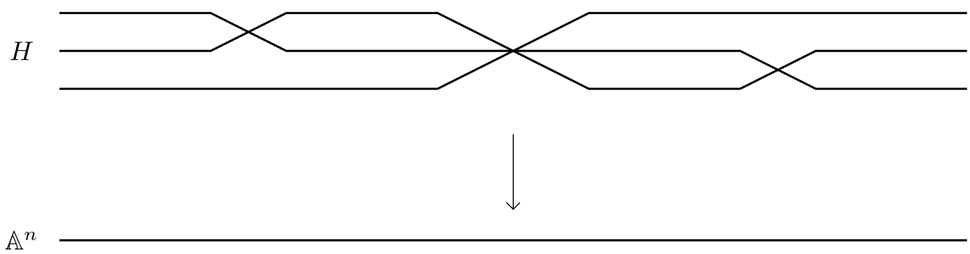
Suppose first that $a_0 \in k$, so that $a_0(t)$ is a nonzero constant independent of t. Then the fibre over t consists of the roots of the polynomial

$$a_0 X^m + a_1(t) X^{m-1} + \cdots + a_m(t), \tag{14}$$

in $k[X]$. Counting multiplicities, there are exactly m of these. More precisely, let D be the discriminant of the polynomial

$$a_0 X^m + a_1 X^{m-1} + \cdots + a_m.$$

Then $D \in k[X_1, \ldots, X_m]$, and the fibre has exactly m points over the open subset $D \neq 0$, and fewer then m points over the closed subset $D = 0$.[4] We can picture it schematically as follows ($m = 3$):



Now drop the condition that a_0 is constant. For certain t, the degree of (14) may drop, which means that some roots have "disappeared off to infinity". Consider, for example, $f(T, X) = TX - 1$; for each $t \neq 0$, there is one point $(t, 1/t)$, but there is no point with $t = 0$ (see the figure p.69). Worse, for certain t all coefficients may be zero, in which case the fibre is a line. There is a nested collection of closed subsets of $\mathbb{A}^n$ such that the number of points in the fibre (counting multiplicities) drops as you pass to a smaller subset, except that over the smallest subset the fibre may be a full line.

DEFINITION 2.39. Let $\varphi\colon W \to V$ be a regular map of algebraic subsets, and let $\varphi^*\colon k[V] \to k[W]$ be the map $f \mapsto f \circ \varphi$.

(a) The map φ is dominant if $\varphi(W)$ is dense in V.

(b) The map φ is quasi-finite if $\varphi^{-1}(P)$ is finite for all $P \in V$.

(c) The map φ is finite if $k[W]$ is a finite $k[V]$-algebra.

[4] I'm ignoring the possibility that D is identically zero. Then the open set where $D \neq 0$ is empty. This case occurs when the characteristic is $p \neq 0$, and f is a polynomial in $T_1, \ldots, T_n$, and X^p.

As $k[W]$ is finitely generated as a k-algebra, a fortiori as a $k[V]$-algebra, to say that $k[W]$ is a finite $k[V]$-algebra means that it is integral over $k[V]$ (1.36).

The map $H \to \mathbb{A}^n$ just considered is finite if and only if a_0 is constant, and quasi-finite if and only if the polynomials $a_0, \ldots, a_m$ have no common zero in k^n.

PROPOSITION 2.40. *A regular map $\varphi: W \to V$ is dominant if and only if $\varphi^*: k[V] \to k[W]$ is injective.*

PROOF. Let $f \in k[V]$. If the image of φ is dense, then

$$f \circ \varphi = 0 \implies f = 0.$$

On the other hand, if the image of φ is not dense, then its closure Z is a proper closed subset of V, and so there exists a nonzero regular function f zero on Z. Then $f \circ \varphi = 0$. □

PROPOSITION 2.41. *A dominant finite map is surjective.*

PROOF. Let $\varphi: W \to V$ be dominant and finite. Then $\varphi^*: k[V] \to k[W]$ is injective, and $k[W]$ is integral over the image of $k[V]$. According to the going-up theorem (1.53), for every maximal ideal $\mathfrak{m}$ of $k[V]$ there exists a maximal ideal $\mathfrak{n}$ of $k[W]$ such that $\mathfrak{m} = \mathfrak{n} \cap k[V]$. Because of the correspondence between points and maximal ideals, this implies that φ is surjective. □

l Noether normalization theorem

Let H be a hypersurface in $\mathbb{A}^{n+1}$. We show that, after a linear change of coordinates, the projection map $(x_1, \ldots, x_{n+1}) \mapsto (x_1, \ldots, x_n): \mathbb{A}^{n+1} \to \mathbb{A}^n$ defines a *finite* map $H \to \mathbb{A}^n$.

PROPOSITION 2.42. *Let*

$$H: \quad f(X_1, \ldots, X_{n+1}) = 0$$

be a hypersurface in $\mathbb{A}^{n+1}$. There exist $c_1, \ldots, c_n \in k$ such that the map $H \to \mathbb{A}^n$ defined by

$$(x_1, \ldots, x_{n+1}) \mapsto (x_1 - c_1 x_{n+1}, \ldots, x_n - c_n x_{n+1})$$

is finite.

PROOF. Let $c_1, \ldots, c_n \in k$. In terms of the coordinates $x_i' = x_i - c x_{n+1}$, the hyperplane H is the zero set of

$$f(X_1 + c_1 X_{n+1}, \ldots, X_n + c_n X_{n+1}, X_{n+1}) = a_0 X_{n+1}^m + a_1 X_{n+1}^{m-1} + \cdots.$$

The next lemma shows that the c_i can be chosen so that a_0 is a nonzero constant. This implies that the map $H \to \mathbb{A}^n$ defined by $(x_1, \ldots, x_{n+1}) \mapsto (x_1', \ldots, x_n')$ is finite. □

LEMMA 2.43. *Let k be an infinite field (not necessarily algebraically closed), and let $f \in k[X_1, \ldots, X_n, T]$. There exist $c_1, \ldots, c_n \in k$ such that*

$$f(X_1 + c_1 T, \ldots, X_n + c_n T, T) = a_0 T^m + a_1 T^{m-1} + \cdots + a_m$$

with $a_0 \in k^\times$ and all $a_i \in k[X_1, \ldots, X_n]$.

PROOF. Let F be the homogeneous part of highest degree of f and let $r = \deg(F)$. Then

$$F(X_1 + c_1T, \ldots, X_n + c_nT, T) = F(c_1, \ldots, c_n, 1)T^r + \text{terms of degree} < r \text{ in } T,$$

because the polynomial $F(X_1 + c_1T, \ldots, X_n + c_nT, T)$ is still homogeneous of degree r in $X_1, \ldots, X_n, T$, and so the coefficient of the monomial T^r can be obtained by setting each X_i equal to zero in F and T to 1. As $F(X_1, \ldots, X_n, T)$ is a nonzero *homogeneous* polynomial, $F(X_1, \ldots, X_n, 1)$ is a nonzero polynomial, and so we can choose the c_i so that $F(c_1, \ldots, c_n, 1) \neq 0$ (Exercise 1-1). Now

$$f(X_1 + c_1T, \ldots, X_n + c_nT, T) = F(c_1, \ldots, c_n, 1)T^r + \text{terms of degree} < r \text{ in } T,$$

with $F(c_1, \ldots, c_n, 1) \in k^\times$, as required. □

In fact, *every* algebraic set V admits a finite surjective map to $\mathbb{A}^d$ for some d.

THEOREM 2.44. *Let V be an algebraic set. For some natural number d, there exists a finite surjective map $\varphi\colon V \to \mathbb{A}^d$.*

This follows from the next statement applied to $A = k[V]$: the regular functions $x_1, \ldots, x_d$ define a map $V \to \mathbb{A}^d$, which is finite and surjective because $k[x_1, \ldots, x_d] \to A$ is finite and injective.

THEOREM 2.45 (NOETHER NORMALIZATION THEOREM). *Let A be a finitely generated k-algebra. There exist elements $x_1, \ldots, x_d \in A$ that are algebraically independent over k, and such that A is finite over $k[x_1, \ldots, x_d]$.*

It is not necessary to assume that A is reduced in (2.45), nor that k is algebraically closed, although the proof we give requires it to be infinite (for the general proof, see CA 8.1).

Let $A = k[x_1, \ldots, x_n]$. We prove the theorem by induction on n. If the x_i are algebraically independent, there is nothing to prove. Otherwise, the next lemma shows that A is finite over a subring $B = k[y_1, \ldots, y_{n-1}]$. By induction, B is finite over a subring $C = k[z_1, \ldots, z_d]$ with $z_1, \ldots, z_d$ algebraically independent, and A is finite over C.

LEMMA 2.46. *Let $A = k[x_1, \ldots, x_n]$ be a finitely generated k-algebra, and let $\{x_1, \ldots, x_d\}$ be a maximal algebraically independent subset of $\{x_1, \ldots, x_n\}$. If $n > d$, then there exist $c_1, \ldots, c_d \in k$ such that A is finite over $k[x_1 - c_1x_n, \ldots, x_d - c_dx_n, x_{d+1}, \ldots, x_{n-1}]$.*

PROOF. By assumption, the set $\{x_1, \ldots, x_d, x_n\}$ is algebraically dependent, and so there exists a nonzero $f \in k[X_1, \ldots, X_d, T]$ such that

$$f(x_1, \ldots, x_d, x_n) = 0. \tag{15}$$

Because $\{x_1, \ldots, x_d\}$ is algebraically independent, T occurs in f, and so

$$f(X_1, \ldots, X_d, T) = a_0T^m + a_1T^{m-1} + \cdots + a_m$$

with $a_i \in k[X_1, \ldots, X_d]$, $a_0 \neq 0$, and $m > 0$.

If $a_0 \in k$, then (15) shows that x_n is integral over $k[x_1, \ldots, x_d]$. Hence $x_1, \ldots, x_n$ are integral over $k[x_1, \ldots, x_{n-1}]$, and so A is finite over $k[x_1, \ldots, x_{n-1}]$.

If $a_0 \notin k$, then, for a suitable choice of $(c_1, \ldots, c_d) \in k$, the polynomial

$$g(X_1, \ldots, X_d, T) \stackrel{\text{def}}{=} f(X_1 + c_1T, \ldots, X_d + c_dT, T)$$

takes the form

$$g(X_1,\dots,X_d,T) = bT^r + b_1T + \cdots + b_r$$

with $b \in k^\times$ (see 2.43). As

$$g(x_1 - c_1x_n,\dots,x_d - c_dx_n,x_n) = 0 \tag{16}$$

this shows that x_n is integral over $k[x_1 - c_1x_n,\dots,x_d - c_dx_n]$, and so A is finite over $k[x_1 - c_1x_n,\dots,x_d - c_dx_n,x_{d+1},\dots,x_{n-1}]$ as before. □

Remarks

2.47. For an irreducible algebraic subset V of $\mathbb{A}^n$, the above argument can be modified to prove the following more precise statement:

> Let $x_1,\dots,x_n$ be the coordinate functions on V; after possibly renumbering the coordinates, we may suppose that $\{x_1,\dots,x_d\}$ is a maximal algebraically independent subset of $\{x_1,\dots,x_n\}$; then there exist $c_{ij} \in k$ such that the map
>
> $$(x_1,\dots,x_n) \mapsto (x_1 - \sum_{d+1\le j\le n} c_{1j}x_j,\dots,x_d - \sum_{d+1\le j\le n} c_{1j}x_j)\colon \mathbb{A}^n \to \mathbb{A}^d$$
>
> induces a finite surjective map $V \to \mathbb{A}^d$.

Indeed, Lemma 2.46 shows that there exist $c_1,\dots,c_n \in k$ such that $k[V]$ is finite over $k[x_1 - c_1x_n,\dots,x_d - c_dx_n,x_{d+1},\dots,x_{n-1}]$. Now $\{x_1,\dots,x_d\}$ is algebraically dependent on $\{x_1 - c_1x_n,\dots,x_d - c_dx_n\}$. If the second set were not algebraically independent, we could drop one of its elements, but this would contradict (1.61). Therefore $\{x_1 - c_1x_n,\dots,x_d - c_dx_n\}$ is a maximal algebraically independent subset of $\{x_1 - c_1x_n,\dots,x_d - c_dx_n,x_{d+1},\dots,x_{n-1}\}$ and we can repeat the argument.

m Dimension

The dimension of a topological space

Let V be a noetherian topological space whose points are closed.

DEFINITION 2.48. The ***dimension*** of V is the supremum of the lengths of the chains

$$V_0 \supset V_1 \supset \cdots \supset V_d$$

of distinct irreducible closed subsets (the length of the displayed chain is d).

2.49. Let $V_1,\dots,V_m$ be the irreducible components of V. Then (obviously)

$$\dim(V) = \max_i(\dim(V_i)).$$

2.50. Assume that V is irreducible, and let W be a proper closed subspace of V. Then every chain $W_0 \supset W_1 \supset \cdots$ in W extends to a chain $V \supset W_0 \supset \cdots$, and so $\dim(W) \le \dim(V) + 1$. If $\dim(V) < \infty$, then $\dim(W) < \dim(V)$.

Thus an irreducible topological space V has dimension 0 if and only if it is a point; it has dimension ≤ 1 if and only if every proper closed subset is a point; and, inductively, V has dimension $\le n$ if and only if every proper closed subset has dimension $\le n - 1$.

The dimension of an algebraic set

DEFINITION 2.51. The ***dimension*** of an algebraic set is its dimension as a topological space (2.48).

Because of the correspondence between the prime ideals in $k[V]$ and irreducible closed subsets of V (2.28),

$$\dim(V) = \text{Krull dimension of } k[V].$$

Note that, if $V_1,\ldots,V_m$ are the irreducible components of V, then

$$\dim V = \max_i \dim(V_i).$$

When the V_i all have the same dimension d, we say that V has ***pure dimension*** d. A one-dimensional algebraic set is called a ***curve***; a two-dimensional algebraic set is called a ***surface***; and an n-dimensional algebraic set is called an n***-fold***.

THEOREM 2.52. *Let V be an irreducible algebraic set. Then*

$$\dim(V) = \text{tr. deg}_k k(V).$$

The proof will occupy the rest of this subsection.

Let A be an arbitrary commutative ring. Let $x \in A$, and let $S_{\{x\}}$ denote the multiplicative subset of A consisting of the elements of the form

$$x^n(1-ax), \quad n \in \mathbb{N}, \quad a \in A.$$

The ***boundary*** $A_{\{x\}}$ of A at x is defined to be the ring of fractions $S_{\{x\}}^{-1}A$.

We write $\dim(A)$ for the Krull dimension of A.

PROPOSITION 2.53. *Let A be a ring and let $n \in \mathbb{N}$. Then*

$$\dim(A) \leq n \iff \text{for all } x \in A,\ \dim(A_{\{x\}}) \leq n-1.$$

PROOF. Recall (1.14) that $\operatorname{Spec}(S^{-1}A) \simeq \{\mathfrak{p} \in \operatorname{Spec}(A) \mid \mathfrak{p} \cap S = \emptyset\}$. We begin with two observations.

(a) *For every $x \in A$ and maximal ideal $\mathfrak{m} \subset A$, $\mathfrak{m} \cap S_{\{x\}} \neq \emptyset$.* Indeed, if $x \in \mathfrak{m}$, then $x \in \mathfrak{m} \cap S_{\{x\}}$; otherwise x is invertible modulo $\mathfrak{m}$, and so there exists an $a \in A$ such that $1-ax \in \mathfrak{m}$.

(b) *Let $\mathfrak{m}$ be a maximal ideal, and let $\mathfrak{p}$ be a prime ideal contained in $\mathfrak{m}$; for every $x \in \mathfrak{m} \smallsetminus \mathfrak{p}$, we have $\mathfrak{p} \cap S_{\{x\}} = \emptyset$.* Indeed, if $x^n(1-ax) \in \mathfrak{p}$, then $1-ax \in \mathfrak{p}$ (as $x \notin \mathfrak{p}$); hence $1-ax \in \mathfrak{m}$, and so $1 \in \mathfrak{m}$, which is a contradiction.

Statement (a) shows that every chain of prime ideals beginning with a maximal ideal is shortened when passing from A to $A_{\{x\}}$, while statement (b) shows that a maximal chain of length n is shortened only to $n-1$ when x is chosen appropriately. From this, the proposition is follows. □

PROPOSITION 2.54. *Let A be an integral domain with field of fractions $F(A)$, and let k be a subfield of A. Then*

$$\text{tr. deg}_k F(A) \geq \dim(A).$$

PROOF. If tr. $\deg_k F(A) = \infty$, there is nothing to prove, and so we assume that tr. $\deg_k F(A) = n \in \mathbb{N}$. We argue by induction on n. We can replace k with its algebraic closure in A without changing tr. $\deg_k F(A)$. Let $x \in A$. If $x \notin k$, then it is transcendental over k, and so

$$\text{tr. } \deg_{k(x)} F(A) = n - 1$$

by (1.64); since $k(x) \subset A_{\{x\}}$, this implies (by induction) that $\dim(A_{\{x\}}) \leq n-1$. If $x \in k$, then $0 = 1 - x^{-1}x \in S_{\{x\}}$, and so $A_{\{x\}} = 0$; again $\dim(A_{\{x\}}) \leq n-1$. Now (2.53) shows that $\dim(A) \leq n$. □

COROLLARY 2.55. *The polynomial ring $k[X_1, \ldots, X_n]$ has Krull dimension n.*

PROOF. The existence of the sequence of prime ideals

$$(X_1, \ldots, X_n) \supset (X_1, \ldots, X_{n-1}) \supset \cdots \supset (X_1) \supset (0)$$

shows that $k[X_1, \ldots, X_n]$ has Krull dimension at least n. Now (2.54) completes the proof. □

COROLLARY 2.56. *Let A be an integral domain and let k be a subfield of A. If A is finitely generated as a k-algebra, then*

$$\textit{tr. } \deg_k F(A) = \dim(A).$$

PROOF. According to the Noether normalization theorem (2.45), A is integral over a polynomial subring $k[x_1, \ldots, x_n]$. Clearly $n =$tr. $\deg_k F(A)$. The going up theorem (1.54), implies that a chain of prime ideals in $k[x_1, \ldots, x_n]$ lifts to a chain in A, and so $\dim(A) \geq \dim(k[x_1, \ldots, x_n]) = n$. Now (2.54) completes the proof. □

ASIDE 2.57. Let V be an algebraic set. The d in Theorem 2.45 is the dimension of V. Thus, V has dimension d if and only if there exists a finite surjective map $V \to \mathbb{A}^d$.

ASIDE 2.58. In linear algebra, we justify saying that a vector space V has dimension n by proving that its elements are parametrized by n-tuples. It is not true in general that the points of an algebraic set of dimension n are parametrized by n-tuples. The most we can say is that there exists a finite-to-one map to k^n (see 2.45).

ASIDE 2.59. The inequality in (2.54) may be strict, for rather trivial reasons — see mo79959.

NOTES. The above proof of (2.55) is based on that in Coquand and Lombardi, Amer. Math. Monthly 112 (2005), no. 9, 826–829.

Examples

Let V be an irreducible algebraic set and W an algebraic subset of V. If W is irreducible, then its ***codimension*** in V is

$$\operatorname{codim}_V W = \dim V - \dim W.$$

If every irreducible component of W has codimension d in V, then W is said to have ***pure codimension*** d in V.

EXAMPLE 2.60. Let $V = \mathbb{A}^n$. Then $k(V) = k(X_1, \ldots, X_n)$, which has transcendence basis $X_1, \ldots, X_n$ over k, and so $\dim(V) = n$.

EXAMPLE 2.61. If V is a linear subspace of k^n (or a translate of a linear subspace), then the dimension of V as an algebraic set is the same as its dimension in the sense of linear algebra — in fact, $k[V]$ is canonically isomorphic to $k[X_{i_1},\ldots,X_{i_d}]$ where the X_{i_j} are the "free" variables in the system of linear equations defining V.

More specifically, let $\mathfrak{c}$ be an ideal in $k[X_1,\ldots,X_n]$ generated by linear forms $\ell_1,\ldots,\ell_r$, which we may assume to be linearly independent. Let $X_{i_1},\ldots,X_{i_{n-r}}$ be such that

$$\{\ell_1,\ldots,\ell_r,X_{i_1},\ldots,X_{i_{n-r}}\}$$

is a basis for the linear forms in $X_1,\ldots,X_n$. Then

$$k[X_1,\ldots,X_n]/\mathfrak{c}\simeq k[X_{i_1},\ldots,X_{i_{n-r}}].$$

This is obvious if the forms are $X_1,\ldots,X_r$. In the general case, because $\{X_1,\ldots,X_n\}$ and $\{\ell_1,\ldots,\ell_r,X_{i_1},\ldots,X_{i_{n-r}}\}$ are both bases for the linear forms, each element of one set can be expressed as a linear combination of the elements of the other. Therefore,

$$k[X_1,\ldots,X_n]=k[\ell_1,\ldots,\ell_r,X_{i_1},\ldots,X_{i_{n-r}}],$$

and so

$$\begin{aligned} k[X_1,\ldots,X_n]/\mathfrak{c} &= k[\ell_1,\ldots,\ell_r,X_{i_1},\ldots,X_{i_{n-r}}]/\mathfrak{c}\\ &\simeq k[X_{i_1},\ldots,X_{i_{n-r}}]. \end{aligned}$$

EXAMPLE 2.62. Every nonempty algebraic set contains a point, which is a closed irreducible subset. Therefore an irreducible algebraic set has dimension 0 if and only if it consists of a single point.

EXAMPLE 2.63. If V is irreducible and W is a proper algebraic subset of V, then $\dim(V)>\dim(W)$ (2.50).

EXAMPLE 2.64. A hypersurface in $\mathbb{A}^n$ has dimension $n-1$. It suffices to prove this for an irreducible hypersurface H. Such an H is the zero set of an irreducible polynomial f (see 2.29). Let

$$k[x_1,\ldots,x_n]=k[X_1,\ldots,X_n]/(f),\quad x_i=X_i+(f),$$

and let $k(x_1,\ldots,x_n)$ be the field of fractions of $k[x_1,\ldots,x_n]$. As f is not the zero polynomial, some X_i, say, X_n, occurs in it. Then X_n occurs in every nonzero multiple of f, and so no nonzero polynomial in $X_1,\ldots,X_{n-1}$ belongs to (f). This means that $x_1,\ldots,x_{n-1}$ are algebraically independent. On the other hand, x_n is algebraic over $k(x_1,\ldots,x_{n-1})$, and so $\{x_1,\ldots,x_{n-1}\}$ is a transcendence basis for $k(x_1,\ldots,x_n)$ over k. (Alternatively, use 2.57.)

EXAMPLE 2.65. Let $F(X,Y)$ and $G(X,Y)$ be nonconstant polynomials with no common factor. Then $V(F(X,Y))$ has dimension 1 by (2.64), and so $V(F(X,Y))\cap V(G(X,Y))$ must have dimension zero; it is therefore a finite set.

PROPOSITION 2.66. *Let W be a closed set of codimension 1 in an algebraic set V; if $k[V]$ is a unique factorization domain, then $I(W)=(f)$ for some $f\in k[V]$.*

PROOF. Let $W_1,\ldots,W_s$ be the irreducible components of W; then $I(W)=\bigcap I(W_i)$, and so if we can prove $I(W_i)=(f_i)$, then $I(W)=(f_1\cdots f_r)$. Thus we may suppose that W is irreducible. Let $\mathfrak{p}=I(W)$; it is a prime ideal, and it is not zero because otherwise

$\dim(W) = \dim(V)$. Therefore it contains an irreducible polynomial f. From (1.33) we know (f) is prime. If $(f) \neq \mathfrak{p}$, then we have

$$\mathfrak{p} \supset (f) \supset (0) \qquad \text{(distinct prime ideals)}$$

and hence

$$W = V(\mathfrak{p}) \subset V(f) \subset V \qquad \text{(distinct irreducible closed subsets).}$$

But then (2.63)

$$\dim(W) < \dim(V(f)) < \dim V,$$

which contradicts the hypothesis. □

COROLLARY 2.67. *The closed sets of codimension* 1 *in* $\mathbb{A}^n$ *are exactly the hyperplanes.*

PROOF. Combine (2.64) and (2.66). □

EXAMPLE 2.68. We classify the irreducible algebraic sets V of $\mathbb{A}^2$. If V has dimension 2, then (by 2.63) it can't be a proper subset of $\mathbb{A}^2$, so it is $\mathbb{A}^2$. If V has dimension 1, then $V = V(f)$ where f is any irreducible polynomial in $I(V)$ (see 2.66 and its proof). Finally, if V has dimension zero, then it is a point. Correspondingly, the following is a complete list of the prime ideals in $k[X,Y]$:

$$(0), \quad (f) \text{ with } f \text{ irreducible}, \quad (X-a, Y-b) \text{ with } a,b \in k.$$

Exercises

2-1. Find $I(W)$, where $W = (X^2, XY^2)$. Check that it is the radical of (X^2, XY^2).

2-2. Identify k^{mn} with the set of $m \times n$ matrices, and let $r \in \mathbb{N}$. Show that the set of matrices with rank $\leq r$ is an algebraic subset of k^{mn}.

2-3. Let $V = \{(t, t^2, \ldots, t^n) \mid t \in k\}$. Show that V is an algebraic subset of k^n, and that $k[V] \approx k[X]$ (polynomial ring in one variable). (Assume k has characteristic zero.)

2-4. Let $f_1, \ldots, f_m \in \mathbb{Q}[X_1, \ldots, X_n]$. If the f_i have no common zero in $\mathbb{C}$, prove that there exist $g_1, \ldots, g_m \in \mathbb{Q}[X_1, \ldots, X_n]$ such that $f_1 g_1 + \cdots + f_m g_m = 1$. (Hint: linear algebra).

2-5. Let $k \subset K$ be algebraically closed fields, and let $\mathfrak{a}$ be an ideal in $k[X_1, \ldots, X_n]$. Show that if $f \in K[X_1, \ldots, X_n]$ vanishes on $V(\mathfrak{a})$, then it vanishes on $V_K(\mathfrak{a})$. Deduce that the zero set $V(\mathfrak{a})$ of $\mathfrak{a}$ in k^n is dense in the zero set $V_K(\mathfrak{a})$ of $\mathfrak{a}$ in K^n. [Hint: Choose a basis $(e_i)_{i \in I}$ for K as a k-vector space, and write $f = \sum e_i f_i$ (finite sum) with $f_i \in k[X_1, \ldots, X_n]$.]

2-6. Let A and B be (not necessarily commutative) $\mathbb{Q}$-algebras of finite dimension over $\mathbb{Q}$, and let $\mathbb{Q}^{\mathrm{al}}$ be the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Show that if there exists a $\mathbb{C}$-algebra homomorphism $\mathbb{C} \otimes_{\mathbb{Q}} A \to \mathbb{C} \otimes_{\mathbb{Q}} B$, then there exists a $\mathbb{Q}^{\mathrm{al}}$-algebra homomorphism $\mathbb{Q}^{\mathrm{al}} \otimes_{\mathbb{Q}} A \to \mathbb{Q}^{\mathrm{al}} \otimes_{\mathbb{Q}} B$. (Hint: The proof takes only a few lines.)

2-7. Let A be finite dimensional k-algebra, where k is an infinite field, and let M and N be A-modules. Show that if $k^{\mathrm{al}} \otimes_k M$ and $k^{\mathrm{al}} \otimes_k N$ are isomorphic $k^{\mathrm{al}} \otimes_k A$-modules, then M and N are isomorphic A-modules.

2-8. Show that the subset $\{(z, e^z) \mid z \in \mathbb{C}\}$ is not an algebraic subset of $\mathbb{C}^2$.

CHAPTER 3

Affine Algebraic Varieties

In this chapter, we define the structure of a ringed space on an algebraic set. In this way, we are led to the notion of an affine algebraic variety — roughly speaking, this is an algebraic set with no preferred embedding into $\mathbb{A}^n$. This is in preparation for Chapter 5, where we define an algebraic variety to be a ringed space that is a finite union of affine algebraic varieties satisfying a natural separation axiom.

a Sheaves

Let k be a field (not necessarily algebraically closed).

DEFINITION 3.1. Let V be a topological space, and suppose that, for every open subset U of V we have a set $\mathcal{O}_V(U)$ of functions $U \to k$. Then $U \rightsquigarrow \mathcal{O}_V(U)$ is a ***sheaf of*** k***-algebras*** if the following statements hold for every open U in V:

(a) $\mathcal{O}_V(U)$ is a k-subalgebra of the algebra of all k-valued functions on U, i.e., $\mathcal{O}_V(U)$ contains the constant functions and, if f, g lie in $\mathcal{O}_V(U)$, then so also do $f+g$ and fg;

(b) the restriction of an f in $\mathcal{O}_V(U)$ to an open subset U' of U is in $\mathcal{O}_V(U')$;

(c) a function $f: U \to k$ lies in $\mathcal{O}_V(U)$ if there exists an open covering $U = \bigcup_{i \in I} U_i$ of U such that $f|U_i \in \mathcal{O}_V(U_i)$ for all $i \in I$.

Conditions (b) and (c) say that a function $f: U \to k$ lies in $\mathcal{O}_V(U)$ if and only if every point P of U has a neighbourhood U_P such that $f|U_P$ lies in $\mathcal{O}_V(U_P)$; in other words, they say that the condition for f to lie in $\mathcal{O}_V(U)$ is *local*.

Note that, for *disjoint* open subsets U_i of V, condition (c) says that $\mathcal{O}_V(U) \simeq \prod_i \mathcal{O}_V(U_i)$.

Examples

3.2. Let V be a topological space, and for each open subset U of V let $\mathcal{O}_V(U)$ be the set of all continuous real-valued functions on U. Then $\mathcal{O}_V$ is a sheaf of $\mathbb{R}$-algebras.

3.3. Recall that a function $f: U \to \mathbb{R}$ on an open subset U of $\mathbb{R}^n$ is said to be ***smooth*** (or ***infinitely differentiable***) if its partial derivatives of all orders exist and are continuous. Let V be an open subset of $\mathbb{R}^n$, and for each open subset U of V, let $\mathcal{O}_V(U)$ be the set of all smooth functions on U. Then $\mathcal{O}_V$ is a sheaf of $\mathbb{R}$-algebras.

3.4. Recall that a function $f: U \to \mathbb{C}$ on an open subset U of $\mathbb{C}^n$, is said to be ***analytic*** (or ***holomorphic***) if it is described by a convergent power series in a neighbourhood of each point of U. Let V be an open subset of $\mathbb{C}^n$, and for each open subset U of V, let $\mathcal{O}_V(U)$ be the set of all analytic functions on U. Then $\mathcal{O}_V$ is a sheaf of $\mathbb{C}$-algebras.

3.5. Let V be a topological space, and, for each open subset U of V, let $\mathcal{O}_V(U)$ be the set of all constant functions $U \to k$. If V is reducible, then $\mathcal{O}_V$ is *not* a sheaf: let U_1 and U_2 be disjoint open subsets of V, and let f be the function on $U_1 \cup U_2$ that takes the constant value 0 on U_1 and the constant value 1 on U_2; then f is not in $\mathcal{O}_V(U_1 \cup U_2)$, and so condition (3.1c) fails. When "constant" is replaced with "locally constant", $\mathcal{O}_V$ becomes a sheaf of k-algebras (in fact, the smallest such sheaf).

3.6. Let V be a topological space, and, for each open subset U of V, let $\mathcal{O}_V(U)$ be the set of *all* functions $U \to k$. The $\mathcal{O}_V$ is a sheaf of k-algebras. By definition, all our sheaves of k-algebras are subsheaves of this one. In Chapter 11, we encounter a more general notion of a sheaf of k-algebras.

b Ringed spaces

A pair $(V, \mathcal{O}_V)$ consisting of a topological space V and a sheaf of k-algebras on V will be called a ***k-ringed space*** (or just a ***ringed space*** when the k is understood). For historical reasons, we sometimes write $\Gamma(U, \mathcal{O}_V)$ for $\mathcal{O}_V(U)$ and call its elements the ***sections*** of $\mathcal{O}_V$ over U.

Let $(V, \mathcal{O}_V)$ be a k-ringed space. For each open subset U of V, the restriction $\mathcal{O}_V|U$ of $\mathcal{O}_V$ to U,

$$U' \rightsquigarrow \Gamma(U', \mathcal{O}_V),\ U' \text{ open in } U,$$

is a sheaf of k-algebras on U.

Let $(V, \mathcal{O}_V)$ be a k-ringed space, and let $P \in V$. A ***germ*** of a function at P is an equivalence class of pairs (U, f) with U an open neighbourhood of P and $f \in \mathcal{O}_V(U)$; two pairs (U, f) and (U', f') are equivalent if the functions f and f' agree on some open neighbourhood of P in $U \cap U'$. The germs of functions at P form a k-algebra $\mathcal{O}_{V,P}$, called the ***stalk*** of $\mathcal{O}_V$ at P. In other words,

$$\mathcal{O}_{V,P} = \varinjlim \mathcal{O}_V(U) \text{ (direct limit over open neighbourhoods } U \text{ of } P).$$

In the interesting cases, $\mathcal{O}_{V,P}$ is a local ring with maximal ideal $\mathfrak{m}_P$ the set of germs that are zero at P. We often write $\mathcal{O}_P$ for $\mathcal{O}_{V,P}$.

EXAMPLE 3.7. Let $\mathcal{O}_V$ be the sheaf of holomorphic functions on $V = \mathbb{C}$, and let $c \in \mathbb{C}$. A power series $\sum_{n \geq 0} a_n (z-c)^n$, $a_n \in \mathbb{C}$, is said to be ***convergent*** if it converges on some open neighbourhood of c. The set of such power series is a $\mathbb{C}$-algebra, and I claim that it is canonically isomorphic to the stalk $\mathcal{O}_{V,c}$ of $\mathcal{O}_V$ at c.

To prove this, let f be a holomorphic function on a neighbourhood U of c. Then f has a unique power series expansion $f = \sum a_n (z-c)^n$ in some (possibly smaller) open neighbourhood of c (Cartan 1963[1], II 2.6). Moreover, another holomorphic function f' on a neighbourhood U' of c defines the same power series if and only if f and f' agree on some

[1]Cartan, Henri. Elementary theory of analytic functions of one or several complex variables. Hermann, Paris; Addison-Wesley; 1963.

neighbourhood of c contained in $U \cap U'$ (ibid. I 4.3). Thus we have a well-defined injective map from the ring of germs of holomorphic functions at c to the ring of convergent power series, which is obviously surjective.

c The ringed space structure on an algebraic set

Let V be an algebraic subset of k^n. Recall that the basic open subsets of V are those of the form

$$D(h) = \{Q \mid h(Q) \neq 0\}, \quad h \in k[V].$$

A pair $g, h \in k[V]$ with $h \neq 0$ defines a function

$$Q \mapsto \frac{g(Q)}{h(Q)} \colon D(h) \to k.$$

A function on an open subset of V is regular if it is locally of this form. More formally:

DEFINITION 3.8. Let U be an open subset of V. A function $f \colon U \to k$ is ***regular*** at $P \in U$ if there exist $g, h \in k[V]$ with $h(P) \neq 0$ such that $f(Q) = g(Q)/h(Q)$ for all Q in some neighbourhood of P. A function $f \colon U \to k$ is ***regular*** if it is regular at every $P \in U$.

Let $\mathcal{O}_V(U)$ denote the set of regular functions on an open subset U of V.

PROPOSITION 3.9. *The map $U \rightsquigarrow \mathcal{O}_V(U)$ is a sheaf of k-algebras on V.*

PROOF. We have to check the conditions of (3.1).

(a) Clearly, a constant function is regular. Suppose f and f' are regular on U, and let $P \in U$. By assumption, there exist $g, g', h, h' \in k[V]$, with $h(P) \neq 0 \neq h'(P)$ such that f and f' agree with $\frac{g}{h}$ and $\frac{g'}{h'}$ respectively on a neighbourhood U' of P. Then $f + f'$ agrees with $\frac{gh'+g'h}{hh'}$ on U', and so $f + f'$ is regular at P. Similarly, ff' agrees with $\frac{gg'}{hh'}$ on U', and so is regular at P.

(b,c) The condition for f to be regular is local. □

We next determine $\mathcal{O}_V(U)$ when U is a basic open subset of V.

LEMMA 3.10. *Let $g, h \in k[V]$ with $h \neq 0$. The function*

$$P \mapsto g(P)/h(P)^m \colon D(h) \to k$$

is zero if and only if and only if $gh = 0$ in $k[V]$.

PROOF. If g/h^m is zero on $D(h)$, then gh is zero on V because h is zero on the complement of $D(h)$. Therefore gh is zero in $k[V]$. Conversely, if $gh = 0$, then $g(P)h(P) = 0$ for all $P \in V$, and so $g(P) = 0$ for all $P \in D(h)$. □

Let $k[V]_h$ denote the ring $k[V]$ with h inverted (see 1.11). The lemma shows that the map $k[V]_h \to \mathcal{O}_V(D(h))$ sending g/h^m to the regular function $P \mapsto g(P)/h(P)^m$ is well-defined and injective.

PROPOSITION 3.11. *The above map $k[V]_h \to \mathcal{O}_V(D(h))$ is an isomorphism of k-algebras.*

PROOF. It remains to show that every regular function f on $D(h)$ arises from an element of $k[V]_h$. By definition, we know that there is an open covering $D(h) = \bigcup V_i$ and elements $g_i, h_i \in k[V]$ with h_i nowhere zero on V_i such that $f|V_i = \frac{g_i}{h_i}$. We may assume that each set V_i is basic, say, $V_i = D(a_i)$ for some $a_i \in k[V]$. By assumption $D(a_i) \subset D(h_i)$, and so $a_i^N = h_i g_i'$ for some $N \in \mathbb{N}$ and $g_i' \in k[V]$ (see p.48). On $D(a_i)$,

$$f = \frac{g_i}{h_i} = \frac{g_i g_i'}{h_i g_i'} = \frac{g_i g_i'}{a_i^N}.$$

Note that $D(a_i^N) = D(a_i)$. Therefore, after replacing g_i with $g_i g_i'$ and h_i with a_i^N, we can assume that $V_i = D(h_i)$.

We now have that $D(h) = \bigcup D(h_i)$ and that $f|D(h_i) = \frac{g_i}{h_i}$. Because $D(h)$ is quasicompact, we can assume that the covering is finite. As $\frac{g_i}{h_i} = \frac{g_j}{h_j}$ on $D(h_i) \cap D(h_j) = D(h_i h_j)$,

$$h_i h_j (g_i h_j - g_j h_i) = 0, \text{ i.e., } h_i h_j^2 g_i = h_i^2 h_j g_j \qquad (*)$$

— this follows from Lemma 3.10 if $h_i h_j \neq 0$ and is obvious otherwise. Because $D(h) = \bigcup D(h_i) = \bigcup D(h_i^2)$,

$$V((h)) = V((h_1^2, \ldots, h_m^2)),$$

and so h lies in $\operatorname{rad}(h_1^2, \ldots, h_m^2)$: there exist $a_i \in k[V]$ such that

$$h^N = \sum_{i=1}^{m} a_i h_i^2. \qquad (**)$$

for some N. I claim that f is the function on $D(h)$ defined by $\frac{\sum a_i g_i h_i}{h^N}$.

Let P be a point of $D(h)$. Then P will be in one of the $D(h_i)$, say $D(h_j)$. We have the following equalities in $k[V]$:

$$h_j^2 \sum_{i=1}^{m} a_i g_i h_i \overset{(*)}{=} \sum_{i=1}^{m} a_i g_j h_i^2 h_j \overset{(**)}{=} g_j h_j h^N.$$

But $f|D(h_j) = \frac{g_j}{h_j}$, i.e., $f h_j$ and g_j agree as functions on $D(h_j)$. Therefore we have the following equality of functions on $D(h_j)$:

$$h_j^2 \sum_{i=1}^{m} a_i g_i h_i = f h_j^2 h^N.$$

Since h_j^2 is never zero on $D(h_j)$, we can cancel it, to find that, as claimed, the function $f h^N$ on $D(h_j)$ equals that defined by $\sum a_i g_i h_i$. □

On taking $h = 1$ in the proposition, we see that the definition of a regular function on V in this section agrees with that in Section 2i.

COROLLARY 3.12. *For every $P \in V$, there is a canonical isomorphism $\mathcal{O}_P \to k[V]_{\mathfrak{m}_P}$, where $\mathfrak{m}_P$ is the maximal ideal $I(P)$.*

PROOF. In the definition of the germs of a sheaf at P, it suffices to consider pairs (f, U) with U lying in a some basis for the neighbourhoods of P, for example, the basis provided by the basic open subsets. Therefore,

$$\mathcal{O}_P = \varinjlim_{h(P)\neq 0} \Gamma(D(h), \mathcal{O}_V) \overset{(3.11)}{\simeq} \varinjlim_{h \notin \mathfrak{m}_P} k[V]_h \overset{(1.23)}{\simeq} k[V]_{\mathfrak{m}_P}.$$

□

Remarks

3.13. Let V be an algebraic set and let P be a point on V. Proposition 1.14 shows that there is a one-to-one correspondence between the prime ideals of $k[V]$ contained in $\mathfrak{m}_P$ and the prime ideals of $\mathcal{O}_P$. In geometric terms, this says that there is a one-to-one correspondence between the irreducible closed subsets of V passing through P and the prime ideals in $\mathcal{O}_P$. The irreducible components of V passsing through P correspond to the minimal prime ideals in $\mathcal{O}_P$. The ideal $\mathfrak{p}$ corresponding to an irreducible closed subset Z consists of the elements of $\mathcal{O}_P$ represented by a pair (U, f) with $f|_{Z\cap U} = 0$.

3.14. The local ring $\mathcal{O}_{V,P}$ is an integral domain if P lies on a single irreducible component of V. As $\mathcal{O}_{V,P}$ depends only on $(U, \mathcal{O}_V|U)$ for U an open neighbourhood of P, we may suppose that V itself is irreducible, in which case the statement follows from (3.12). On the other hand, if P lies on more than one irreducible component of V, then $\mathcal{O}_P$ contains more than one minimal prime ideal (3.13), and so the ideal (0) can't be prime.

3.15. Let V be an algebraic subset of k^n, and let $A = k[V]$. Propositions 2.37 and 3.11 allow us to describe $(V, \mathcal{O}_V)$ purely in terms of A:

- ◇ V is the set of maximal ideals in A.
- ◇ For each $f \in A$, let $D(f) = \{\mathfrak{m} \mid f \notin \mathfrak{m}\}$; the topology on V is that for which the sets $D(f)$ form a base.
- ◇ For $f \in A_h$ and $\mathfrak{m} \in D(h)$, let $f(\mathfrak{m})$ denote the image of f in $A_h/\mathfrak{m}A_h \simeq k$; in this way A_h becomes identified with a k-algebra of functions $D(h) \to k$, and $\mathcal{O}_V$ is the unique sheaf of k-valued functions on V such that $\Gamma(D(h), \mathcal{O}_V) = A_h$ for all $h \in A$.

3.16. When V is irreducible, all the rings attached to it can be identified with subrings of the field $k(V)$. For example,

$$\begin{aligned}\Gamma(D(h), \mathcal{O}_V) &= \left\{\tfrac{g}{h^N} \in k(V) \mid g \in k[V], \quad N \in \mathbb{N}\right\}\\ \mathcal{O}_P &= \left\{\tfrac{g}{h} \in k(V) \mid h(P) \neq 0\right\}\\ \Gamma(U, \mathcal{O}_V) &= \textstyle\bigcap_{P\in U} \mathcal{O}_P\\ &= \textstyle\bigcap \Gamma(D(h_i), \mathcal{O}_V) \text{ if } U = \bigcup D(h_i).\end{aligned}$$

Note that every element of $k(V)$ defines a function on some dense open subset of V. Following tradition, we call the elements of $k(V)$ ***rational functions*** on V.[2]

Examples

3.17. The ring of regular functions on $\mathbb{A}^n$ is $k[X_1, \ldots, X_n]$. For a nonzero polynomial $h(X_1, \ldots, X_n)$, the ring of regular functions on $D(h)$ is

$$\left\{\tfrac{g}{h^N} \in k(X_1, \ldots, X_n) \mid g \in k[X_1, \ldots, X_n], \quad N \in \mathbb{N}\right\}.$$

For a point $P = (a_1, \ldots, a_n)$, the local ring at P is

$$\mathcal{O}_P = \left\{\tfrac{g}{h} \in k(X_1, \ldots, X_n) \mid h(P) \neq 0\right\} = k[X_1, \ldots, X_n]_{(X_1-a_1, \ldots, X_n-a_n)},$$

and its maximal ideal consists of those g/h with $g(P) = 0$.

[2] The terminology is similar to that of "meromorphic function", which is also not a function on the whole space.

3.18. Let $U = \mathbb{A}^2 \smallsetminus \{(0,0)\}$. It is an open subset of $\mathbb{A}^2$, but it is not a basic open subset because its complement $\{(0,0)\}$ has dimension 0, and therefore can't be of the form $V((f))$ (see 2.64). Since $U = D(X) \cup D(Y)$, the ring of regular functions on U is

$$\Gamma(U, \mathcal{O}_{\mathbb{A}^2}) = k[X,Y]_X \cap k[X,Y]_Y$$

(intersection inside $k(X,Y)$). Thus, a regular function f on U can be expressed

$$f = \frac{g(X,Y)}{X^N} = \frac{h(X,Y)}{Y^M}.$$

We may assume that $X \nmid g$ and $Y \nmid h$. On multiplying through by $X^N Y^M$, we find that

$$g(X,Y)Y^M = h(X,Y)X^N.$$

Because X doesn't divide the left hand side, it can't divide the right hand side either, and so $N = 0$. Similarly, $M = 0$, and so $f \in k[X,Y]$. We have shown that every regular function on U extends uniquely to a regular function on $\mathbb{A}^2$:

$$\Gamma(U, \mathcal{O}_{\mathbb{A}^2}) = k[X,Y] = \Gamma(\mathbb{A}^2, \mathcal{O}_{\mathbb{A}^2}).$$

d Morphisms of ringed spaces

A ***morphism of k-ringed spaces*** $(V, \mathcal{O}_V) \to (W, \mathcal{O}_W)$ is a continuous map $\varphi\colon V \to W$ such that

$$f \in \mathcal{O}_W(U) \implies f \circ \varphi \in \mathcal{O}_V(\varphi^{-1}U)$$

for all open subsets U of W. Then, for every pair of open subsets $U \subset W$ and $U' \subset V$ with $\varphi(U') \subset U$, we get a homomorphism of k-algebras

$$f \mapsto f \circ \varphi \colon \mathcal{O}_W(U') \to \mathcal{O}_V(U),$$

and these homomorphisms are compatible with restriction to smaller open subsets. Sometimes we write $\varphi^*(f)$ for $f \circ \varphi$. A morphism of ringed spaces is an ***isomorphism*** if it is bijective and its inverse is also a morphism of ringed spaces (in particular, it is a homeomorphism).

If U is an open subset of V, then the inclusion $U \hookrightarrow V$ is a morphism of k-ringed spaces $(U, \mathcal{O}_V|U) \to (V, \mathcal{O}_V)$.

A morphism of ringed spaces maps germs of functions to germs of functions. More precisely, a morphism $\varphi\colon (V, \mathcal{O}_V) \to (W, \mathcal{O}_W)$ induces a k-algebra homomorphism

$$\mathcal{O}_{W,\varphi(P)} \to \mathcal{O}_{V,P}$$

for each $P \in V$, which sends the germ represented by (U, f) to the germ represented by $(\varphi^{-1}(U), f \circ \varphi)$. In the interesting cases, $\mathcal{O}_{V,P}$ is a local ring with maximal ideal $\mathfrak{m}_P$ consisting of the germs represented by pairs (U, f) with $f(P) = 0$. Therefore, the homomorphism $\mathcal{O}_{W,\varphi(P)} \to \mathcal{O}_{V,P}$ defined by φ maps $\mathfrak{m}_{\varphi(P)}$ into $\mathfrak{m}_P$: it is a local homomorphism of local rings.

Examples

3.19. Let V and W be topological spaces endowed with their sheaves $\mathcal{O}_V$ and $\mathcal{O}_W$ of continuous real valued functions (3.2). Every continuous map $\varphi: V \to W$ is a morphism of ringed structures $(V, \mathcal{O}_V) \to (W, O_W)$.

3.20. Let V and W be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, and let x_i be the coordinate function $(a_1, \ldots, a_n) \mapsto a_i: V \to \mathbb{R}$. Recall from advanced calculus that a map

$$\varphi: V \to W \subset \mathbb{R}^m$$

is said to be smooth if each of its component functions $\varphi_i \overset{\text{def}}{=} x_i \circ \varphi: V \to \mathbb{R}$ is smooth. If φ is smooth, then $f \circ \varphi$ is smooth for every smooth function $f: W \to \mathbb{R}$. Since a similar statement is true for functions f on open subsets of W, we see that a continuous map $\varphi: V \to W$ is smooth if and only if it is a morphism of the associated ringed spaces (3.3).

3.21. Same as (3.20), but replace $\mathbb{R}$ with $\mathbb{C}$ and "smooth" with "analytic".

e Affine algebraic varieties

We have just seen that every algebraic set $V \subset k^n$ gives rise to a k-ringed space $(V, \mathcal{O}_V)$. A k-ringed space isomorphic to one of this form is called an ***affine algebraic variety over*** k. We usually denote an affine algebraic variety $(V, \mathcal{O}_V)$ by V.

Let $(V, \mathcal{O}_V)$ and $(W, \mathcal{O}_W)$ be affine algebraic varieties. A map $\varphi: V \to W$ is ***regular*** (or a ***morphism of affine algebraic varieties***) if it is a morphism of k-ringed spaces. With these definitions, the affine algebraic varieties become a category. We usually shorten "affine algebraic variety" to "affine variety".

In particular, the regular functions define the structure of an affine variety on every algebraic set. We usually regard $\mathbb{A}^n$ as an affine variety. The affine varieties we have constructed so far have all been embedded in $\mathbb{A}^n$. We now explain how to construct affine varieties with no preferred embedding.

An ***affine*** k***-algebra*** is a reduced finitely generated k-algebra. For such an algebra A, there exist $x_i \in A$ such that $A = k[x_1, \ldots, x_n]$, and the kernel of the homomorphism

$$X_i \mapsto x_i: k[X_1, \ldots, X_n] \to A$$

is a radical ideal. Therefore (2.18) implies that the intersection of the maximal ideals in A is 0. Moreover, (2.12) implies that, for every maximal ideal $\mathfrak{m} \subset A$, the map $k \to A \to A/\mathfrak{m}$ is an isomorphism. Thus we can identify $A/\mathfrak{m}$ with k. For $f \in A$, we write $f(\mathfrak{m})$ for the image of f in $A/\mathfrak{m} = k$, i.e., $f(\mathfrak{m}) = f \pmod{\mathfrak{m}}$. This allows us to identify elements of A with functions $\operatorname{spm}(A) \to k$.

We attach a ringed space $(V, \mathcal{O}_V)$ to A by letting V be the set of maximal ideals in A. For $f \in A$, let

$$D(f) = \{\mathfrak{m} \mid f(\mathfrak{m}) \neq 0\} = \{\mathfrak{m} \mid f \notin \mathfrak{m}\}.$$

Since $D(fg) = D(f) \cap D(g)$, there is a topology on V for which the $D(f)$ form a base. A pair of elements $g, h \in A$, $h \neq 0$, defines a function

$$\mathfrak{m} \mapsto \frac{g(\mathfrak{m})}{h(\mathfrak{m})}: D(h) \to k.$$

For U an open subset of V, we define $\mathcal{O}_V(U)$ to be the set of functions $f: U \to k$ that are of this form in some neighbourhood of each point of U.

PROPOSITION 3.22. *The pair $(V, \mathcal{O}_V)$ is an affine algebraic variety with $\Gamma(D(h), \mathcal{O}_V) \simeq A_h$ for each $h \in A \smallsetminus \{0\}$.*

PROOF. Represent A as a quotient $k[X_1, \ldots, X_n]/\mathfrak{a} = k[x_1, \ldots, x_n]$. Then $(V, \mathcal{O}_V)$ is isomorphic to the k-ringed space attached to the algebraic set $V(\mathfrak{a})$ (see 3.15). □

We write $\operatorname{spm}(A)$ for the topological space V, and $\operatorname{Spm}(A)$ for the k-ringed space $(V, \mathcal{O}_V)$.

ASIDE 3.23. We have attached to an affine k-algebra A an affine variety whose underlying topological space is the set of maximal ideals in A. It may seem strange to be describing a topological space in terms of maximal ideals in a ring, but the analysts have been doing this for more than 70 years. Gel'fand and Kolmogorov in 1939[3] proved that if S and T are compact topological spaces, and the rings of real-valued continuous functions on S and T are isomorphic (just as rings), then S and T are homeomorphic. The proof begins by showing that, for such a space S, the map

$$P \mapsto \mathfrak{m}_P \stackrel{\text{def}}{=} \{f : S \to \mathbb{R} \mid f(P) = 0\}$$

is one-to-one correspondence between the points in the space and maximal ideals in the ring.

f The category of affine algebraic varieties

For each affine k-algebra A, we have an affine variety $\operatorname{Spm}(A)$, and conversely, for each affine variety $(V, \mathcal{O}_V)$, we have an affine k-algebra $k[V] = \Gamma(V, \mathcal{O}_V)$. We now make this correspondence into an equivalence of categories.

Let $\alpha: A \to B$ be a homomorphism of affine k-algebras. For every $h \in A$, $\alpha(h)$ is invertible in $B_{\alpha(h)}$, and so the homomorphism $A \to B \to B_{\alpha(h)}$ extends to a homomorphism

$$\frac{g}{h^m} \mapsto \frac{\alpha(g)}{\alpha(h)^m} : A_h \to B_{\alpha(h)}.$$

For every maximal ideal $\mathfrak{n}$ of B, $\mathfrak{m} = \alpha^{-1}(\mathfrak{n})$ is maximal in A because $A/\mathfrak{m} \to B/\mathfrak{n} = k$ is an injective map of k-algebras which implies that $A/\mathfrak{m} = k$. Thus α defines a map

$$\varphi : \operatorname{spm} B \to \operatorname{spm} A, \quad \varphi(\mathfrak{n}) = \alpha^{-1}(\mathfrak{n}) = \mathfrak{m}.$$

For $\mathfrak{m} = \alpha^{-1}(\mathfrak{n}) = \varphi(\mathfrak{n})$, we have a commutative diagram:

$$\begin{array}{ccc} A & \xrightarrow{\alpha} & B \\ \downarrow & & \downarrow \\ A/\mathfrak{m} & \xrightarrow{\simeq} & B/\mathfrak{n}. \end{array}$$

Recall that the image of an element f of A in $A/\mathfrak{m} \simeq k$ is denoted $f(\mathfrak{m})$. Therefore, the commutativity of the diagram means that, for $f \in A$,

$$f(\varphi(\mathfrak{n})) = \alpha(f)(\mathfrak{n}), \text{ i.e., } f \circ \varphi = \alpha \circ f. \tag{*}$$

Since $\varphi^{-1}D(f) = D(f \circ \varphi)$ (obviously), it follows from (*) that

$$\varphi^{-1}(D(f)) = D(\alpha(f)),$$

[3] On rings of continuous functions on topological spaces, Doklady 22, 11-15. See also Allen Shields, Banach Algebras, 1939–1989, Math. Intelligencer, Vol 11, no. 3, p15.

and so φ is continuous.

Let f be a regular function on $D(h)$, and write $f = g/h^m$, $g \in A$. Then, from (*) we see that $f \circ \varphi$ is the function on $D(\alpha(h))$ defined by $\alpha(g)/\alpha(h)^m$. In particular, it is regular, and so $f \mapsto f \circ \varphi$ maps regular functions on $D(h)$ to regular functions on $D(\alpha(h))$. It follows that $f \mapsto f \circ \varphi$ sends regular functions on any open subset of $\mathrm{spm}(A)$ to regular functions on the inverse image of the open subset. Thus α defines a morphism of ringed spaces $\mathrm{Spm}(B) \to \mathrm{Spm}(A)$.

Conversely, by definition, a morphism of $\varphi\colon (V, \mathcal{O}_V) \to (W, \mathcal{O}_W)$ of affine algebraic varieties defines a homomorphism of the associated affine k-algebras $k[W] \to k[V]$. Since these maps are inverse, we have shown:

PROPOSITION 3.24. *For all affine algebras A and B,*

$$\mathrm{Hom}_{k\text{-}alg}(A, B) \xrightarrow{\simeq} \mathrm{Mor}(\mathrm{Spm}(B), \mathrm{Spm}(A));$$

for all affine varieties V and W,

$$\mathrm{Mor}(V, W) \xrightarrow{\simeq} \mathrm{Hom}_{k\text{-}alg}(k[W], k[V]).$$

In terms of categories, Proposition 3.24 can be restated as:

PROPOSITION 3.25. *The functor $A \rightsquigarrow \mathrm{Spm}\, A$ is a (contravariant) equivalence from the category of affine k-algebras to the category of affine algebraic varieties over k; the functor $(V, \mathcal{O}_V) \rightsquigarrow \Gamma(V, \mathcal{O}_V)$ is a quasi-inverse.*

g Explicit description of morphisms of affine varieties

PROPOSITION 3.26. *Let $V \subset k^m$ and $W \subset k^n$ be algebraic subsets. The following conditions on a continuous map $\varphi\colon V \to W$ are equivalent:*

(a) *φ is a morphism of ringed spaces $(V, \mathcal{O}_V) \to (W, \mathcal{O}_W)$;*

(b) *the components $\varphi_1, \ldots, \varphi_m$ of φ are regular functions on V;*

(c) *$f \in k[W] \implies f \circ \varphi \in k[V]$.*

PROOF. (a) $\implies$ (b). By definition $\varphi_i = y_i \circ \varphi$ where y_i is the coordinate function

$$(b_1, \ldots, b_n) \mapsto b_i\colon W \to k.$$

Hence this implication follows directly from the definition of a regular map.

(b) $\implies$ (c). The map $f \mapsto f \circ \varphi$ is a k-algebra homomorphism from the ring of all functions $W \to k$ to the ring of all functions $V \to k$, and (b) says that the map sends the coordinate functions y_i on W into $k[V]$. Since the y_i generate $k[W]$ as a k-algebra, this implies that it sends $k[W]$ into $k[V]$.

(c) $\implies$ (a). The map $f \mapsto f \circ \varphi$ is a homomorphism $\alpha\colon k[W] \to k[V]$. It therefore defines a map $\mathrm{spm}(k[V]) \to \mathrm{spm}(k[W])$, and it remains to show that this coincides with φ when we identify $\mathrm{spm}(k[V])$ with V and $\mathrm{spm}(k[W])$ with W. Let $P \in V$, let $Q = \varphi(P)$, and let $\mathfrak{m}_P$ and $\mathfrak{m}_Q$ be the ideals of elements of $k[V]$ and $k[W]$ that are zero at P and Q respectively. Then, for $f \in k[W]$,

$$\alpha(f) \in \mathfrak{m}_P \iff f(\varphi(P)) = 0 \iff f(Q) = 0 \iff f \in \mathfrak{m}_Q.$$

Therefore $\alpha^{-1}(\mathfrak{m}_P) = \mathfrak{m}_Q$, which is what we needed to show. □

The equivalence of (a) and (b) means that $\varphi: V \to W$ is a regular map of algebraic sets (in the sense of Chapter 2) if and only if it is a regular map of the associated affine algebraic varieties.

Now consider equations

$$\begin{aligned} Y_1 &= f_1(X_1, \ldots, X_m) \\ &\ldots \\ Y_n &= f_n(X_1, \ldots, X_m). \end{aligned}$$

On the one hand, they define a regular map $\varphi: \mathbb{A}^m \to \mathbb{A}^n$, namely,

$$(a_1, \ldots, a_m) \mapsto (f_1(a_1, \ldots, a_m), \ldots, f_n(a_1, \ldots, a_m)).$$

On the other hand, they define a homomorphism $\alpha: k[Y_1, \ldots, Y_n] \to k[X_1, \ldots, X_n]$ of k-algebras, namely, that sending Y_i to $f_i(X_1, \ldots, X_n)$. This map coincides with $g \mapsto g \circ \varphi$, because

$$\alpha(g)(P) = g(\ldots, f_i(P), \ldots) = g(\varphi(P)).$$

Now consider closed subsets $V(\mathfrak{a}) \subset \mathbb{A}^m$ and $V(\mathfrak{b}) \subset \mathbb{A}^n$ with $\mathfrak{a}$ and $\mathfrak{b}$ radical ideals. I claim that φ maps $V(\mathfrak{a})$ into $V(\mathfrak{b})$ if and only if $\alpha(\mathfrak{b}) \subset \mathfrak{a}$. Indeed, suppose $\varphi(V(\mathfrak{a})) \subset V(\mathfrak{b})$, and let $g \in \mathfrak{b}$; for $Q \in V(\mathfrak{b})$,

$$\alpha(g)(Q) = g(\varphi(Q)) = 0,$$

and so $\alpha(f) \in IV(\mathfrak{b}) = \mathfrak{b}$. Conversely, suppose $\alpha(\mathfrak{b}) \subset \mathfrak{a}$, and let $P \in V(\mathfrak{a})$; for $f \in \mathfrak{a}$,

$$f(\varphi(P)) = \alpha(f)(P) = 0,$$

and so $\varphi(P) \in V(\mathfrak{a})$. When these conditions hold, φ is the morphism of affine varieties $V(\mathfrak{a}) \to V(\mathfrak{b})$ corresponding to the homomorphism $k[Y_1, \ldots, Y_n]/\mathfrak{b} \to k[X_1, \ldots, X_m]/\mathfrak{a}$ defined by α.

Thus, we see that the regular maps

$$V(\mathfrak{a}) \to V(\mathfrak{b})$$

are all of the form

$$P \mapsto (f_1(P), \ldots, f_n(P)), \quad f_i \in k[X_1, \ldots, X_m].$$

In particular, they all extend to regular maps $\mathbb{A}^m \to \mathbb{A}^n$.

Examples of regular maps

3.27. Let R be a k-algebra. To give a k-algebra homomorphism $k[X] \to R$ is the same as giving an element (the image of X under the homomorphism):

$$\mathrm{Hom}_{k\text{-alg}}(k[X], R) \simeq R$$

Therefore

$$\mathrm{Mor}(V, \mathbb{A}^1) \overset{3.24}{\simeq} \mathrm{Hom}_{k\text{-alg}}(k[X], k[V]) \simeq k[V].$$

In other words, the regular maps $V \to \mathbb{A}^1$ are simply the regular functions on V (as we would hope).

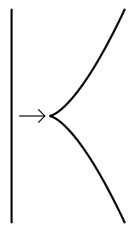
3.28. Let $\mathbb{A}^0 = \operatorname{Spm} k$. Then $\mathbb{A}^0$ consists of a single point and $\Gamma(\mathbb{A}^0, \mathcal{O}_{\mathbb{A}^0}) = k$. Every map $\mathbb{A}^0 \to V$ from $\mathbb{A}^0$ to an affine variety, and so $\operatorname{Mor}(\mathbb{A}^0, V) \simeq V$. Alternatively,

$$\operatorname{Mor}(\mathbb{A}^0, V) \simeq \operatorname{Hom}_{k\text{-alg}}(k[V], k) \simeq V$$

where the last map sends $\alpha: k[V] \to k$ to the point corresponding to the maximal ideal $\operatorname{Ker}(\alpha)$.

3.29. Consider the regular map $t \mapsto (t^2, t^3): \mathbb{A}^1 \to \mathbb{A}^2$. This is bijective onto its image,

$$V: \quad Y^2 = X^3,$$



but it is not an isomorphism onto its image because the inverse map is not regular. In view of (3.25), to prove this it suffices to show that $t \mapsto (t^2, t^3)$ does not induce an isomorphism on the rings of regular functions. We have $k[\mathbb{A}^1] = k[T]$ and $k[V] = k[X,Y]/(Y^2 - X^3) = k[x,y]$. The map on rings is

$$x \mapsto T^2, \quad y \mapsto T^3, \quad k[x,y] \to k[T],$$

which is injective, but its image is $k[T^2, T^3] \neq k[T]$. In fact, $k[x,y]$ is not integrally closed: $(y/x)^2 - x = 0$, and so (y/x) is integral over $k[x,y]$, but $y/x \notin k[x,y]$ (it maps to T under the inclusion $k(x,y) \hookrightarrow k(T)$).

3.30. Let k have characteristic $p \neq 0$, and consider the regular map

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

This is a bijection, but it is not an isomorphism because the corresponding map on rings,

$$X_i \mapsto X_i^p: k[X_1, \ldots, X_n] \to k[X_1, \ldots, X_n],$$

is not surjective.

This is the famous ***Frobenius map.*** Take k to be the algebraic closure of $\mathbb{F}_p$, and write F for the map. Recall that for each $m \geq 1$ there is a unique subfield $\mathbb{F}_{p^m}$ of k of degree m over $\mathbb{F}_p$, and that its elements are the solutions of $X^{p^m} = X$ (FT 4.21). The fixed points of F^m are precisely the points of $\mathbb{A}^n$ with coordinates in $\mathbb{F}_{p^m}$. Let $f(X_1, \ldots, X_n)$ be a polynomial with coefficients in $\mathbb{F}_{p^m}$, say,

$$f = \sum c_{i_1 \cdots i_n} X_1^{i_1} \cdots X_n^{i_n}, \quad c_{i_1 \cdots i_n} \in \mathbb{F}_{p^m}.$$

If $f(a_1, \ldots, a_n) = 0$, then

$$0 = \left(\sum c_\alpha a_1^{i_1} \cdots a_n^{i_n}\right)^{p^m} = \sum c_\alpha a_1^{p^m i_1} \cdots a_n^{p^m i_n},$$

and so $f(a_1^{p^m}, \ldots, a_n^{p^m}) = 0$. Here we have used that the binomial theorem takes the simple form $(X+Y)^{p^m} = X^{p^m} + Y^{p^m}$ in characteristic p. Thus F^m maps $V(f)$ into itself, and its fixed points are the solutions of

$$f(X_1, \ldots, X_n) = 0$$

in $\mathbb{F}_{p^m}$.

ASIDE 3.31. In one of the most beautiful pieces of mathematics of the second half of the twentieth century, Grothendieck defined a cohomology theory (étale cohomology) and proved a fixed point formula that allowed him to express the number of solutions of a system of polynomial equations with coordinates in $\mathbb{F}_{p^m}$ as an alternating sum of traces of operators on finite-dimensional vector spaces, and Deligne used this to obtain very precise estimates for the number of solutions. See my article *The Riemann hypothesis over finite fields: from Weil to the present day* and my notes *Lectures on Étale Cohomology*.

h Subvarieties

Let A be an affine k-algebra. For any ideal $\mathfrak{a}$ in A, we define

$$\begin{aligned} V(\mathfrak{a}) &= \{P \in \operatorname{spm}(A) \mid f(P) = 0 \text{ all } f \in \mathfrak{a}\} \\ &= \{\mathfrak{m} \text{ maximal ideal in } A \mid \mathfrak{a} \subset \mathfrak{m}\}. \end{aligned}$$

This is a closed subset of $\operatorname{spm}(A)$, and every closed subset is of this form.

Now let $\mathfrak{a}$ be a radical ideal in A, so that $A/\mathfrak{a}$ is again reduced. Corresponding to the homomorphism $A \to A/\mathfrak{a}$, we get a regular map

$$\operatorname{Spm}(A/\mathfrak{a}) \to \operatorname{Spm}(A)$$

The image is $V(\mathfrak{a})$, and $\operatorname{spm}(A/\mathfrak{a}) \to V(\mathfrak{a})$ is a homeomorphism. Thus every closed subset of $\operatorname{spm}(A)$ has a natural ringed structure making it into an affine algebraic variety. We call $V(\mathfrak{a})$ with this structure a ***closed subvariety*** of V.

PROPOSITION 3.32. *Let $(V, \mathcal{O}_V)$ be an affine variety and let h be a nonzero element of $k[V]$. Then*

$$(D(h), \mathcal{O}_V|D(h)) \simeq \operatorname{Spm}(A_h);$$

in particular, it is an affine variety.

PROOF. The map $A \to A_h$ defines a morphism $\operatorname{spm}(A_h) \to \operatorname{spm}(A)$. The image is $D(h)$, and it is routine (using (1.13)) to verify the rest of the statement. □

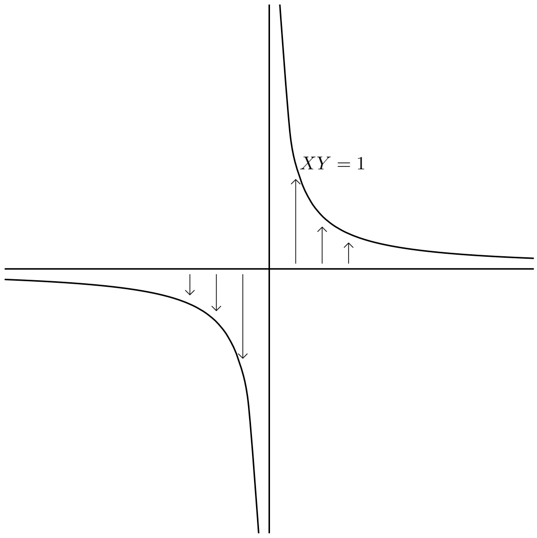
If $V = V(\mathfrak{a}) \subset \mathbb{A}^n$, then

$$(a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_n, h(a_1, \ldots, a_n)^{-1}) \colon D(h) \to \mathbb{A}^{n+1},$$

defines an isomorphism of $D(h)$ onto $V(\mathfrak{a}, 1 - hX_{n+1})$. For example, there is an isomorphism of affine varieties

$$a \mapsto (a, 1/a) \colon \mathbb{A}^1 \smallsetminus \{0\} \to V \subset \mathbb{A}^2,$$

with V equal to the subvariety $XY = 1$ of $\mathbb{A}^2$.



By an ***open affine (subset)*** U of an affine algebraic variety V, we mean an open subset U such that $(U, \mathcal{O}_V|U)$ is an affine algebraic variety. Thus, the proposition says that, for all nonzero $h \in \Gamma(V, \mathcal{O}_V)$, the open subset of V where h is nonzero is an open affine. An open affine subset of an irreducible affine algebraic variety V is irreducible with the same dimension as V (2.52).

REMARK 3.33. We have seen that all closed subsets and all *basic* open subsets of an affine variety V are again affine varieties with their natural ringed structure, but this is not true for all open subsets of V. For an open affine subset U, the natural map $U \to \operatorname{spm} \Gamma(U, \mathcal{O}_V)$ is a bijection. However, for

$$U = \mathbb{A}^2 \smallsetminus (0,0) = D(X) \cup D(Y) \subset \mathbb{A}^2,$$

we know that $\Gamma(U, \mathcal{O}_{\mathbb{A}^2}) = k[X,Y]$ (see 3.18), but $U \to \operatorname{spm} k[X,Y]$ is not a bijection, because the ideal (X,Y) is not in the image. Clearly $(U, \mathcal{O}_{\mathbb{A}^2}|U)$ is a union of affine algebraic varieties, and in Chapter 5 we shall recognize it as a (nonaffine) algebraic variety.

i Properties of the regular map Spm(α)

PROPOSITION 3.34. *Let $\alpha: A \to B$ be a homomorphism of affine k-algebras, and let*

$$\varphi: \operatorname{Spm}(B) \to \operatorname{Spm}(A)$$

be the corresponding morphism of affine varieties.

(a) *The image of φ is dense for the Zariski topology if and only if α is injective.*

(b) *The morphism φ is an isomorphism from $\operatorname{Spm}(B)$ onto a closed subvariety of $\operatorname{Spm}(A)$ if and only if α is surjective.*

PROOF. (a) Let $f \in A$. If the image of φ is dense, then

$$f \circ \varphi = 0 \implies f = 0.$$

On the other hand, if the image of φ is not dense, then the closure of its image is a proper closed subset of $\mathrm{Spm}(A)$, and so there is a nonzero function $f \in A$ that is zero on it. Then $f \circ \varphi = 0$. (See 2.40.)

(b) If α is surjective, then it defines an isomorphism $A/\mathfrak{a} \to B$ where $\mathfrak{a}$ is the kernel of α. This induces an isomorphism of $\mathrm{Spm}(B)$ with its image in $\mathrm{Spm}(A)$. The converse follows from the description of the closed subvarieties of $\mathrm{Spm}(A)$ in the last section. □

A regular map $\varphi\colon V \to W$ of affine algebraic varieties is said to be a ***dominant*** if its image is dense in W. The proposition then says that:

$$\varphi \text{ is dominant} \iff f \mapsto f \circ \varphi\colon \Gamma(W, \mathcal{O}_W) \to \Gamma(V, \mathcal{O}_V) \text{ is injective.}$$

A regular map $\varphi\colon V \to W$ of affine algebraic varieties is said to be a ***closed immersion*** if it is an isomorphism of V onto a closed subvariety of W. The proposition then says that

$$\varphi \text{ is a closed immersion} \iff f \mapsto f \circ \varphi\colon \Gamma(W, \mathcal{O}_W) \to \Gamma(V, \mathcal{O}_V) \text{ is surjective.}$$

j Affine space without coordinates

Let E be a vector space over k of dimension n. The set $\mathbb{A}(E)$ of points of E has a natural structure of an algebraic variety: the choice of a basis for E defines a bijection $\mathbb{A}(E) \to \mathbb{A}^n$, and the inherited structure of an affine algebraic variety on $\mathbb{A}(E)$ is independent of the choice of the basis (because the bijections defined by two different bases differ by an automorphism of $\mathbb{A}^n$).

We now give an intrinsic definition of the affine variety $\mathbb{A}(E)$. Let V be a finite-dimensional vector space over a field k. The ***tensor algebra*** of V is

$$T^*V = \bigoplus_{i \geq 0} V^{\otimes i} = k \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots$$

with multiplication defined by

$$(v_1 \otimes \cdots \otimes v_i) \cdot (v_1' \otimes \cdots \otimes v_j') = v_1 \otimes \cdots \otimes v_i \otimes v_1' \otimes \cdots \otimes v_j'.$$

It is a noncommutative k-algebra, and the choice of a basis $e_1, \ldots, e_n$ for V defines an isomorphism

$$e_1 \cdots e_i \mapsto e_1 \otimes \cdots \otimes e_i \colon k\{e_1, \ldots, e_n\} \to T^*(V)$$

to T^*V from the k-algebra of noncommuting polynomials in the symbols $e_1, \ldots, e_n$.

The ***symmetric algebra*** $S^*(V)$ of V is defined to be the quotient of T^*V by the two-sided ideal generated by the elements

$$v \otimes w - w \otimes v, \quad v, w \in V.$$

This algebra is generated as a k-algebra by commuting elements (namely, the elements of $V = V^{\otimes 1}$), and so is commutative. The choice of a basis $e_1, \ldots, e_n$ for V defines an isomorphism

$$e_1 \cdots e_i \mapsto e_1 \otimes \cdots \otimes e_i \colon k[e_1, \ldots, e_n] \to S^*(V)$$

to $S^*(V)$ from the commutative polynomial ring in the symbols $e_1, \ldots, e_n$. This shows that $S^*(V)$ is an affine k-algebra. The pair $(S^*(V), i)$ consisting of $S^*(V)$ and the natural

k-linear map $i: V \to S^*(V)$ has the following universal property: every k-linear map $V \to A$ from V into a k-algebra A extends uniquely to a k-algebra homomorphism $S^*(V) \to A$:

$$\begin{tikzcd} V \arrow[r, "i"] \arrow[dr, "k\text{-linear}"'] & S^*(V) \arrow[d, dotted, "\exists!\, k\text{-algebra}"] \\ & A. \end{tikzcd} \tag{17}$$

As usual, this universal property determines the pair $(S^*(V), i)$ uniquely up to a unique isomorphism.

We now define $\mathbb{A}(E)$ to be $\mathrm{Spm}(S^*(E^\vee))$ where E is the dual vector space. For an affine k-algebra A,

$$\begin{aligned} \mathrm{Mor}(\mathrm{Spm}(A), \mathbb{A}(E)) &\simeq \mathrm{Hom}_{k\text{-algebra}}(S^*(E^\vee), A) && (3.24) \\ &\simeq \mathrm{Hom}_{k\text{-linear}}(E^\vee, A) && (17) \\ &\simeq E \otimes_k A && \text{(linear algebra)}. \end{aligned}$$

In particular,

$$\mathbb{A}(E)(k) \simeq E.$$

Moreover, the choice of a basis $e_1, \ldots, e_n$ for E determines a (dual) basis $f_1, \ldots, f_n$ of $E^\vee$, and hence an isomorphism of k-algebras $k[f_1, \ldots, f_n] \to S^*(E^\vee)$. The map of algebraic varieties defined by this homomorphism is the isomorphism

$$\mathbb{A}(E) \to \mathbb{A}^n$$

whose map on the underlying sets is the isomorphism $E \to k^n$ defined by the basis of E.

k Birational equivalence

Recall that if V is irreducible, then $k[V]$ is an integral domain, and we write $k(V)$ for its field of fractions. If U is an open affine subvariety of V, then $k[V] \subset k[U] \subset k(V)$, and so $k(V)$ is also the field of fractions of $k[U]$.

DEFINITION 3.35. Two irreducible affine algebraic varieties V and W over k are ***birationally equivalent*** if $k(V) \approx k(W)$.

PROPOSITION 3.36. *Irreducible affine varieties V and W are birationally equivalent if and only if there exist open affine subvarieties U_V and U_W of V and W respectively such that $U_V \approx U_W$.*

PROOF. Let V and W be birationally equivalent irreducible affine varieties, and let $A = k[V]$ and $B = k[W]$. We use the isomorphism to identify $k(V)$ and $k(W)$. This allows us to suppose that A and B have a common field of fractions K. Let $x_1, \ldots, x_n$ generate B as k-algebra. As K is the field of fractions of A, there exists a $d \in A$ such that $dx_i \in A$ for all i; then $B \subset A_d$. The same argument shows that there exists an $e \in B$ such $A_d \subset B_e$. Now

$$B \subset A_d \subset B_e \implies B_e \subset A_{de} \subset (B_e)_e = B_e,$$

and so $A_{de} = B_e$. This shows that the open subvarieties $D(de) \subset V$ and $D(e) \subset W$ are isomorphic. We have proved the "only if" part, and the "if" part is obvious. □

THEOREM 3.37. *Every irreducible affine algebraic variety of dimension d is birationally equivalent to a hypersurface in $\mathbb{A}^{d+1}$.*

PROOF. Let V be an irreducible variety of dimension d. According to (3.38) below, there exist rational functions $x_1,\ldots,x_{d+1}$ on V such that $k(V) = k(x_1,\ldots,x_d,x_{d+1})$. Let $f \in k[X_1,\ldots,X_{d+1}]$ be an irreducible polynomial satisfied by the x_i, and let H be the hypersurface $f = 0$. Then $k(V) \approx k(H)$ and so V and H are birationally equivalent. □

Let F be a field. A polynomial $f \in F[X]$ is ***separable*** if it has distinct roots in any splitting field; equivalently, if $\gcd(f, f') = 1$. Let E be an algebraic extension of F. An element α of E is ***separable*** over F if its minimum polynomial over F is separable. The field E is ***separable*** over F if every element of E is separable over F. A finite extension E of F is separable over F if $\mathrm{Hom}_F(E,\Omega)$ has $[E\colon F]$ elements for every suitably large extension Ω of F, or if $E = F[\alpha]$ with α separable over F.

PROPOSITION 3.38. *Let Ω be a finitely generated field extension of k of transcendence degree d. If k is perfect, then there exist $x_1,\ldots,x_{d+1} \in \Omega$ such that $\Omega = k(x_1,\ldots,x_{d+1})$. After renumbering, $\{x_1,\ldots,x_d\}$ will be a transcendence basis for Ω over k and x_{d+1} will be separable over $k(x_1,\ldots,x_d)$.*

PROOF. Let $\Omega = k(x_1,\ldots,x_n)$. After renumbering, we may suppose that $x_1,\ldots,x_d$ are algebraically independent, hence a transcendence basis (1.63).

If F has characteristic zero, then $x_{d+1},\ldots,x_n$ are separable over $k(x_1,\ldots x_d)$, and so the primitive element theorem (FT, 5.1) shows that there exists a $y \in \Omega$ for which $\Omega = k(x_1,\ldots,x_d,y)$.

Thus, we may assume that k has characteristic $p \neq 0$. Because k is perfect, every polynomial in $X_1^p,\ldots,X_n^p$ with coefficients in k is a pth power in $k[X_1,\ldots,X_n]$:

$$\sum a_{i_1\cdots i_n} X_1^{i_1 p} \ldots X_n^{i_n p} = \left(\sum a_{i_1\cdots i_n}^{\frac{1}{p}} X_1^{i_1} \ldots X_n^{i_n}\right)^p. \tag{18}$$

Let n be the minimum order of a generating set for Ω over k. We shall assume that $n > d+1$ and obtain a contradiction. Let $\Omega = k(x_1,\ldots,x_n)$, as in the first paragraph of the proof. Then $f(x_1,\ldots,x_{d+1}) = 0$ for some nonzero irreducible polynomial $f(X_1,\ldots,X_{d+1})$ with coefficients in k. Not all polynomials $\partial f/\partial X_i$ are zero, for otherwise f would be a polynomial in $X_1^p,\ldots,X_{d+1}^p$, and hence a pth power. After renumbering $\{x_1,\ldots,x_{d+1}\}$, we may suppose that $\partial f/\partial X_{d+1} \neq 0$. Now x_{d+1} is separably algebraic over $k(x_1,\ldots,x_d)$ and x_{d+2} is algebraic over $k(x_1,\ldots,x_{d+1})$ (hence over $k(x_1,\ldots,x_d)$), and so the primitive element theorem (FT 5.1) shows that there exists a $y \in \Omega$ for which $k(x_1,\ldots,x_{d+2}) = k(x_1,\ldots,x_d,y)$. Now $\Omega = k(x_1,\ldots,x_d,y,x_{d+3},\ldots,x_n)$, contradicting the minimality of n.

We have shown that $\Omega = k(z_1,\ldots,z_{d+1})$ for some $z_i \in \Omega$. The argument in the last paragraph shows that, after renumbering, z_{d+1} will be separably algebraic over $k(z_1,\ldots,z_d)$, and this implies that $\{z_1,\ldots,z_d\}$ is a transcendence basis for Ω over k (1.63). □

l Noether Normalization Theorem

DEFINITION 3.39. The ***dimension*** of an affine algebraic variety is the dimension of the underlying topological space (2.48).

DEFINITION 3.40. A regular map $\varphi: W \to V$ of affine algebraic varieties is ***finite*** if the map $\varphi^*: k[V] \to k[W]$ makes $k[V]$ a finite $k[W]$-algebra.

THEOREM 3.41. *Let V be an affine algebraic variety of dimension n. Then there exists a finite map $V \to \mathbb{A}^n$.*

PROOF. Immediate consequence of (2.45). □

m Dimension

By definition, the dimension d of an affine variety V is the maximum length of a chain

$$V_0 \supset V_1 \supset \cdots$$

of distinct closed irreducible affine subvarieties. In this section, we prove that it is the length of *every maximal* chain of such subvarieties.

THEOREM 3.42. *Let V be an irreducible affine variety, and let f be a nonzero regular function on V. If f has a zero in V, then its zero set is of pure codimension 1.*

The Noether normalization theorem allows us to deduce this from the special case $V = \mathbb{A}^n$, proved in (2.64).

PROOF. [4]Let $Z_1, \ldots, Z_n$ be the irreducible components of $V(f)$. We have to show that $\dim Z_i = \dim V - 1$ for each i. There exists a point $P \in Z_i$ not contained in any other Z_j. Because the Z_j are closed, there exists an open affine neighbourhood U of P in V not meeting any Z_j with $j \neq i$. Now $V(f|U) = Z_i \cap U$, which is irreducible. Therefore, on replacing V with U, we may assume that $V(f)$ is irreducible.

As $V(f)$ is irreducible, the radical of (f) is a prime ideal $\mathfrak{p}$ in $k[V]$. According to the Noether normalization theorem (2.45), there exists an inclusion $k[\mathbb{A}^d] \hookrightarrow k[V]$ realizing $k[V]$ as a finite $k[\mathbb{A}^d]$-algebra. Let $f_0 = \mathrm{Nm}_{k(V)/k(\mathbb{A}^d)} f$. Then $f_0 \in k[\mathbb{A}^d]$ and f divides f_0 in $k[V]$ (see 1.45). Hence $f_0 \in (f) \subset \mathfrak{p}$, and so $\mathrm{rad}(f_0) \subset \mathfrak{p} \cap k[\mathbb{A}^d]$. We claim that, in fact,

$$\mathrm{rad}(f_0) = \mathfrak{p} \cap k[\mathbb{A}^d].$$

Let $g \in \mathfrak{p} \cap k[\mathbb{A}^d]$. Then $g \in \mathfrak{p} \overset{\text{def}}{=} \mathrm{rad}(f)$, and so $g^m = fh$ for some $h \in k[V]$, $m \in \mathbb{N}$. Taking norms, we find that

$$g^{me} = \mathrm{Nm}(fh) = f_0 \cdot \mathrm{Nm}(h) \in (f_0),$$

where $e = [k(V) : k(\mathbb{A}^n)]$, and so $g \in \mathrm{rad}(f_0)$, as claimed.

The inclusion $k[\mathbb{A}^d] \hookrightarrow k[V]$ therefore induces an inclusion

$$k[\mathbb{A}^d]/\mathrm{rad}(f_0) \hookrightarrow k[V]/\mathfrak{p}.$$

This makes $k[V]/\mathfrak{p}$ into a finite algebra over $k[\mathbb{A}^d]/\mathrm{rad}(f_0)$, and so the fields of fractions of these two k-algebras have the same transcendence degree:

$$\dim V(\mathfrak{p}) = \dim V(f_0).$$

Clearly $f \neq 0 \Rightarrow f_0 \neq 0$, and $f_0 \in \mathfrak{p} \Rightarrow f_0$ is nonconstant. Therefore $\dim V(f_0) = d - 1$ by (2.64). □

[4]This proof was found by John Tate.

We can restate Theorem 3.42 as follows: let V be a closed irreducible subvariety of $\mathbb{A}^n$ and let $F \in k[X_1,\ldots,X_n]$; then

$$V \cap V(F) = \begin{cases} V & \text{if } F \text{ is identically zero on } V \\ \emptyset & \text{if } F \text{ has no zeros on } V \\ \text{pure codimension 1} & \text{otherwise.} \end{cases}$$

COROLLARY 3.43. *Let V be an irreducible affine variety, and let Z be a maximal proper irreducible closed subset of V. Then* $\dim(Z) = \dim(V) - 1$.

PROOF. Because Z is a proper closed subset of V, there exists a nonzero regular function f on V vanishing on Z. Let $V(f)$ be the zero set of f in V. Then $Z \subset V(f) \subset V$, and Z must be an irreducible component of $V(f)$ for otherwise it wouldn't be maximal in V. Thus Theorem 3.42 shows that $\dim Z = \dim V - 1$. □

COROLLARY 3.44. *Let V be an irreducible affine variety. Every maximal (i.e., nonrefinable) chain*

$$V = V_0 \supset V_1 \supset \cdots \supset V_d \tag{19}$$

of distinct irreducible closed subsets of V has length $d = \dim(V)$.

PROOF. The last set V_d must be a point and each V_i must be maximal in V_{i-1}, and so, from (3.43), we find that

$$\dim V_0 = \dim V_1 + 1 = \dim V_2 + 2 = \cdots = \dim V_d + d = d. \qquad \square$$

COROLLARY 3.45. *Let V be an irreducible affine variety, and let $f_1,\ldots,f_r$ be regular functions on V. Every irreducible component Z of $V(f_1,\ldots f_r)$ has codimension at most r:*

$$\operatorname{codim}(Z) \leq r.$$

For example, if the f_i have no common zero on V, so that $V(f_1,\ldots,f_r)$ is empty, then there are no irreducible components, and the statement is vacuously true.

PROOF. We use induction on r. Because Z is a irreducible closed subset of $V(f_1,\ldots,f_{r-1})$, it is contained in some irreducible component Z' of $V(f_1,\ldots f_{r-1})$. By induction, $\operatorname{codim}(Z') \leq r-1$. Also Z is an irreducible component of $Z' \cap V(f_r)$ because

$$Z \subset Z' \cap V(f_r) \subset V(f_1,\ldots,f_r)$$

and Z is a maximal irreducible closed subset of $V(f_1,\ldots,f_r)$. If f_r vanishes identically on Z', then $Z = Z'$ and $\operatorname{codim}(Z) = \operatorname{codim}(Z') \leq r-1$; otherwise, the theorem shows that Z has codimension 1 in Z', and $\operatorname{codim}(Z) = \operatorname{codim}(Z') + 1 \leq r$. □

EXAMPLE 3.46. In the setting of (3.45), the components of $V(f_1,\ldots,f_r)$ need not all have the same dimension, and it is possible for all of them to have codimension $< r$ without any of the f_i being redundant. For example, let V be the cone

$$X_1X_4 - X_2X_3 = 0$$

in $\mathbb{A}^4$. Then $V(X_1) \cap V$ is the union of two planes:

$$V(X_1) \cap V = \{(0,0,*,*)\} \cup \{(0,*,0,*)\}.$$

Both of these have codimension 1 in V (as required by (3.42)). Similarly, $V(X_2)\cap V$ is the union of two planes,

$$V(X_2)\cap V=\{(0,0,*,*)\}\cup\{(*,0,*,0)\}.$$

However $V(X_1,X_2)\cap V$ consists of a single plane $\{(0,0,*,*)\}$: it still has codimension 1 in V, but it requires both X_1 and X_2 to define it.

PROPOSITION 3.47. *Let Z be an irreducible closed subvariety of codimension r in an affine variety V. Then there exist regular functions $f_1,\ldots,f_r$ on V such that Z is an irreducible component of $V(f_1,\ldots,f_r)$ and all irreducible components of $V(f_1,\ldots,f_r)$ have codimension r.*

PROOF. We know that there exists a chain of irreducible closed subsets

$$V\supset Z_1\supset\cdots\supset Z_r=Z$$

with codim $Z_i=i$. We shall show that there exist $f_1,\ldots,f_r\in k[V]$ such that, for all $s\le r$, Z_s is an irreducible component of $V(f_1,\ldots,f_s)$ and all irreducible components of $V(f_1,\ldots,f_s)$ have codimension s.

We prove this by induction on s. For $s=1$, take any $f_1\in I(Z_1)$, $f_1\neq 0$, and apply Theorem 3.42. Suppose $f_1,\ldots,f_{s-1}$ have been chosen, and let $Y_1,Y_2,\ldots,Y_m$, be the irreducible components of $V(f_1,\ldots,f_{s-1})$, numbered so that $Z_{s-1}=Y_1$. We seek an element f_s that is identically zero on Z_s but is not identically zero on any Y_i—for such an f_s, all irreducible components of $Y_i\cap V(f_s)$ will have codimension s, and Z_s will be an irreducible component of $Y_1\cap V(f_s)$. But no Y_i is contained in Z_s because Z_s has smaller dimension than Y_i, and so $I(Z_s)$ is not contained in any of the ideals $I(Y_i)$. Now the prime avoidance lemma (see below) tells us that there exist an $f_s\in I(Z_s)\smallsetminus\left(\bigcup_i I(Y_i)\right)$, and this is the function we want. □

LEMMA 3.48 (PRIME AVOIDANCE LEMMA). *If an ideal $\mathfrak{a}$ of a ring A is not contained in any of the prime ideals $\mathfrak{p}_1,\ldots,\mathfrak{p}_r$, then it is not contained in their union.*

PROOF. We may assume that none of the prime ideals $\mathfrak{p}_i$ is contained in a second, because then we could omit it. For a fixed i, choose an $f_i\in\mathfrak{a}\smallsetminus\mathfrak{p}_i$ and, for each $j\neq i$, choose an $f_j\in\mathfrak{p}_j\smallsetminus\mathfrak{p}_i$. Then $h_i\overset{\text{def}}{=}\prod_{j=1}^r f_j$ lies in each $\mathfrak{p}_j$ with $j\neq i$ and $\mathfrak{a}$, but not in $\mathfrak{p}_i$ (here we use that $\mathfrak{p}_i$ is prime). The element $\sum_{i=1}^r h_i$ is therefore in $\mathfrak{a}$ but not in any $\mathfrak{p}_i$. □

EXAMPLE 3.49. When V is an affine variety whose coordinate ring is a unique factorization domain, every closed subset Z of codimension 1 is of the form $V(f)$ for some $f\in k[V]$ (see 2.66). The condition that $k[V]$ be a unique factorization domain is definitely needed. Again consider the cone,

$$V\colon X_1X_4-X_2X_3=0$$

in $\mathbb{A}^4$ and let Z and Z' be the planes

$$Z=\{(*,0,*,0)\}\qquad Z'=\{(0,*,0,*)\}.$$

Then $Z\cap Z'=\{(0,0,0,0)\}$, which has codimension 2 in Z'. If $Z=V(f)$ for some regular function f on V, then $V(f|Z')=\{(0,\ldots,0)\}$, which has codimension 2, in violation of 3.42. Thus Z is not of the form $V(f)$, and so

$$k[X_1,X_2,X_3,X_4]/(X_1X_4-X_2X_3)$$

is not a unique factorization domain.

Restatement in terms of affine algebras

We restate some of these results in terms of affine algebras.

3.50. Theorem 3.42 says the following: let A be an affine k-algebra; if A is an integral domain and $f \in A$ is neither zero nor a unit, then every prime ideal $\mathfrak{p}$ minimal among those containing (f) has height 1 (principal ideal theorem).

3.51. Corollary 3.44 says the following: let A be an affine k-algebra; if A is integral domain, then every maximal chain

$$\mathfrak{p}_d \supset \mathfrak{p}_{d-1} \supset \cdots \supset \mathfrak{p}_0$$

of distinct prime ideals has length equal to the Krull dimension of A. In particular, every maximal ideal in A has height $\dim(A)$.

3.52. Let A be an affine k-algebra; if A is an integral domain and every prime ideal of height 1 in A is principal, then A is a unique factorization domain. In order to prove this, it suffices to show that every irreducible element f of A is prime (1.26). Let $\mathfrak{p}$ be minimal among the prime ideals containing (f). According to (3.50), $\mathfrak{p}$ has height 1, and so it is principal, say $\mathfrak{p} = (g)$. As $(f) \subset (g)$, $f = gq$ for some $q \in A$. Because f is irreducible, q is a unit, and so $(f) = (g) = \mathfrak{p}$ — the element f is prime.

3.53. Proposition 3.47 says the following: let A be an affine k-algebra, and let $\mathfrak{p}$ be a prime ideal of height r in A. If $\mathfrak{p}$ has height r, then there exist elements $f_1, \ldots, f_r \in A$ such that $\mathfrak{p}$ is minimal among the prime ideals containing $(f_1, \ldots, f_r)$.

ASIDE 3.54. Statements (3.50) and 3.53) are true for all noetherian rings (CA 21.3, 21.8). However, (3.51) may fail. For example, as we noted on p.14 a noetherian ring may have infinite Krull dimension. Moreover, a noetherian ring may have finite Krull dimension d without all of its maximal ideals having height d. For example, let $A = R[X]$ where $R = k[t]_{(t)}$ is a discrete valuation ring with maximal ideal (t). The Krull dimension of A is 2, and $(t, X) \supset (t) \supset (0)$ is a maximal chain of prime ideals, but the ideal $(tX - 1)$ is maximal (because $A/(tX-1) \simeq R_t$, see 1.13) and of height 1 (because it is in $k[t, X]$ and A is obtained from $k[t, X]$ by inverting the elements of $k[t] \smallsetminus (t)$).

ASIDE 3.55. Proposition 3.47 shows that a curve C in $\mathbb{A}^3$ is an irreducible component of $V(f_1, f_2)$ for some $f_1, f_2 \in k[X, Y, Z]$. In fact $C = V(f_1, f_2, f_3)$ for suitable polynomials f_1, f_2, and f_3 — this is an exercise in Shafarevich 1994 (I 6, Exercise 8; see also Hartshorne 1977, I, Exercise 2.17). Apparently, it is not known whether two polynomials always suffice to define a curve in $\mathbb{A}^3$ — see Kunz 1985, p136. The union of two skew lines in $\mathbb{P}^3$ can't be defined by two polynomials (ibid. p140), but it is unknown whether all connected curves in $\mathbb{P}^3$ can be defined by two polynomials. Macaulay (the man, not the program) showed that for every $r \geq 1$, there is a curve C in $\mathbb{A}^3$ such that $I(C)$ requires at least r generators (see the same exercise in Hartshorne for a curve whose ideal can't be generated by 2 elements).[5]

In general, a closed variety V of codimension r in $\mathbb{A}^n$ (resp. $\mathbb{P}^n$) is said to be a ***set-theoretic complete intersection*** if there exist r polynomials $f_i \in k[X_1, \ldots, X_n]$ (resp. homogeneous polynomials $f_i \in k[X_0, \ldots, X_n]$) such that

$$V = V(f_1, \ldots, f_r).$$

[5] In 1882 Kronecker proved that every algebraic subset in $\mathbb{P}^n$ can be cut out by $n+1$ polynomial equations. In 1891 Vahlen asserted that the result was best possible by exhibiting a curve in $\mathbb{P}^3$ which he claimed was not the zero locus of 3 equations. It was only 50 years later, in 1941, that Perron gave 3 equations defining Vahlen's curve, thus refuting Vahlen's claim which had been accepted for half a century. Finally, in 1973 Eisenbud and Evans proved that n equations always suffice to describe (set-theoretically) an algebraic subset of $\mathbb{P}^n$ (mo35468 Georges Elencwajg).

Such a variety is said to be an ***ideal-theoretic complete intersection*** if the f_i can be chosen so that $I(V) = (f_1, \ldots, f_r)$. Chapter V of Kunz's book is concerned with the question of when a variety is a complete intersection. Obviously there are many ideal-theoretic complete intersections, but most of the varieties one happens to be interested in turn out not to be. For example, no abelian variety of dimension > 1 is an ideal-theoretic complete intersection (being an ideal-theoretic complete intersection imposes constraints on the cohomology of the variety, which are not fulfilled in the case of abelian varieties).

Let P be a point on an irreducible variety $V \subset \mathbb{A}^n$. Then (3.47) shows that there is a neighbourhood U of P in $\mathbb{A}^n$ and functions $f_1, \ldots, f_r$ on U such that $U \cap V = V(f_1, \ldots, f_r)$ (zero set in U). Thus $U \cap V$ is a set-theoretic complete intersection in U. One says that V is a ***local complete intersection*** at $P \in V$ if there is an open affine neighbourhood U of P in $\mathbb{A}^n$ such that the ideal $I(V \cap U)$ can be generated by r regular functions on U. Note that

$$\text{ideal-theoretic complete intersection} \Rightarrow \text{local complete intersection at all } \mathfrak{p}.$$

It is not difficult to show that a variety is a local complete intersection at every nonsingular point (cf. 4.36).

Exercises

3-1. Show that a map between affine varieties can be continuous for the Zariski topology without being regular.

3-2. Let $V = \mathrm{Spm}(A)$, and let $Z = \mathrm{Spm}(A/\mathfrak{a}) \subset \mathrm{Spm}(A)$. Show that a function f on an open subset U of Z is regular if and only if, for each $P \in U$, there exists a regular function f' on an open neighbourhood U' of P in V such that f and f' agree on $U' \cap U$.

3-3. Find the image of the regular map

$$(x, y) \mapsto (x, xy) \colon \mathbb{A}^2 \to \mathbb{A}^2$$

and verify that it is neither open nor closed.

3-4. Show that the circle $X^2 + Y^2 = 1$ is isomorphic (as an affine variety) to the hyperbola $XY = 1$, but that neither is isomorphic to $\mathbb{A}^1$. (Assume $\mathrm{char}(k) \neq 2$.)

3-5. Let C be the curve $Y^2 = X^2 + X^3$, and let φ be the regular map

$$t \mapsto (t^2 - 1, t(t^2 - 1)) \colon \mathbb{A}^1 \to C.$$

Is φ an isomorphism?

CHAPTER 4

Local Study

Geometry is the art of drawing correct conclusions from incorrect figures. (La géométrie est l'art de raisonner juste sur des figures fausses.)
Descartes

In this chapter, we examine the structure of an affine algebraic variety near a point. We begin with the case of a plane curve, since the ideas in the general case are the same but the proofs are more complicated.

a Tangent spaces to plane curves

Consider the curve V in the plane defined by a nonconstant polynomial $F(X,Y)$,

$$V : F(X,Y) = 0$$

We assume that $F(X,Y)$ has no multiple factors, so that $(F(X,Y))$ is a radical ideal and $I(V) = (F(X,Y))$. We can factor F into a product of irreducible polynomials, $F(X,Y) = \prod F_i(X,Y)$, and then $V = \bigcup V(F_i)$ expresses V as a union of its irreducible components (see 2.29). Each component $V(F_i)$ has dimension 1 (by 2.64) and so V has pure dimension 1. If $F(X,Y)$ itself is irreducible, then

$$k[V] = k[X,Y]/(F(X,Y)) = k[x,y]$$

is an integral domain. Moreover, if $F \neq X - c$, then x is transcendental over k and y is algebraic over $k(x)$, and so x is a transcendence basis for $k(V)$ over k. Similarly, if $F \neq Y - c$, then y is a transcendence basis for $k(V)$ over k.

Let (a,b) be a point on V. If we were doing calculus, we would say that the tangent space at $P = (a,b)$ is defined by the equation

$$\frac{\partial F}{\partial X}(a,b)(X-a) + \frac{\partial F}{\partial Y}(a,b)(Y-b) = 0. \tag{20}$$

This is the equation of a line unless both $\frac{\partial F}{\partial X}(a,b)$ and $\frac{\partial F}{\partial Y}(a,b)$ are zero, in which case it is the equation of a plane.

We are not doing calculus, but we can define $\frac{\partial}{\partial X}$ and $\frac{\partial}{\partial Y}$ by

$$\frac{\partial}{\partial X}\left(\sum a_{ij}X^iY^j\right) = \sum i a_{ij}X^{i-1}Y^j, \quad \frac{\partial}{\partial Y}\left(\sum a_{ij}X^iY^j\right) = \sum j a_{ij}X^iY^{j-1},$$

and then the same definition applies.

DEFINITION 4.1. The ***tangent space*** $T_P V$ to V at $P = (a,b)$ is the algebraic subset defined by equation (20).

If $\frac{\partial F}{\partial X}(a,b)$ and $\frac{\partial F}{\partial Y}(a,b)$ are not both zero, then $T_P(V)$ is a line through (a,b), and we say that P is a ***nonsingular*** or ***smooth*** point of V. Otherwise, $T_P(V)$ has dimension 2, and we say that P is ***singular*** or ***multiple.*** The curve V is said to be ***nonsingular*** or ***smooth*** if all its points are nonsingular.

Examples

For each of the following examples, the reader is invited to sketch the curve. Assume that $\mathrm{char}(k) \neq 2,3$.

4.2. $X^m + Y^m = 1$. The tangent space at (a,b) is given by the equation

$$ma^{m-1}(X-a) + mb^{m-1}(Y-b) = 0.$$

All points on the curve are nonsingular unless the characteristic of k divides m, in which case $X^m + Y^m - 1$ has multiple factors,

$$X^m + Y^m - 1 = X^{m_0 p} + Y^{m_0 p} - 1 = (X^{m_0} + Y^{m_0} - 1)^p.$$

4.3. $Y^2 = X^3$ (sketched in 4.12 below). The tangent space at (a,b) is given by the equation

$$-3a^2(X-a) + 2b(Y-b) = 0.$$

The only singular point is $(0,0)$.

4.4. $Y^2 = X^2(X+1)$ (sketched in 4.10 below). Here again only $(0,0)$ is singular.

4.5. $Y^2 = X^3 + aX + b$. In (2.2) we sketch two nonsingular examples of such curves, and in (4.10) and (4.11) we sketch two singular example. The singular points of the curve are the common zeros of the polynomials

$$Y^2 - X^3 - aX - b, \quad 2Y, \quad 3X^2 + a,$$

which consist of the points $(0,c)$ with c a common zero of

$$X^3 + aX + b, \quad 3X^2 + a.$$

As $3X^2 + a$ is the derivative of $X^3 + aX + b$, we see that V is singular if and only if $X^3 + aX + b$ has a multiple root.

4.6. $V = V(FG)$ where FG has no multiple factors (so F and G are coprime). Then $V = V(F) \cup V(G)$, and a point (a,b) is singular if and only if it is a singular point of $V(F)$, a singular point of $V(G)$, or a point of $V(F) \cap V(G)$. This follows immediately from the product rule:

$$\frac{\partial(FG)}{\partial X} = F \cdot \frac{\partial G}{\partial X} + \frac{\partial F}{\partial X} \cdot G, \quad \frac{\partial(FG)}{\partial Y} = F \cdot \frac{\partial G}{\partial Y} + \frac{\partial F}{\partial Y} \cdot G.$$

The singular locus

PROPOSITION 4.7. *The nonsingular points of a plane curve form a dense open subset of the curve.*

PROOF. Let $V = V(F)$ where F is a nonconstant polynomial in $k[X,Y]$ without multiple factors. It suffices to show that the nonsingular points form a dense open subset of each irreducible component of V, and so we may assume that V (hence F) is irreducible. It suffices to show that the set of singular points is a proper closed subset. Since it is the set of common zeros of the polynomials

$$F, \quad \frac{\partial F}{\partial X}, \quad \frac{\partial F}{\partial Y},$$

it is obviously closed. It will be proper unless $\partial F/\partial X$ and $\partial F/\partial Y$ are both identically zero on V, and hence both multiples of F, but, as they have lower degree than F, this is impossible unless they are both zero. Clearly $\partial F/\partial X = 0$ if and only if F is a polynomial in Y (k of characteristic zero) or is a polynomial in X^p and Y (k of characteristic p). A similar remark applies to $\partial F/\partial Y$. Thus if $\partial F/\partial X$ and $\partial F/\partial Y$ are both zero, then F is constant (characteristic zero) or a polynomial in X^p, Y^p, and hence a pth power (characteristic p). These are contrary to our assumptions. □

Thus the singular points form a proper closed subset, called the ***singular locus***.

ASIDE 4.8. In common usage, "singular" means uncommon or extraordinary as in "he spoke with singular shrewdness". Thus the proposition says that singular points (mathematical sense) are singular (usual sense).

b Tangent cones to plane curves

A polynomial $F(X,Y)$ can be written (uniquely) as a finite sum

$$F = F_0 + F_1 + F_2 + \cdots + F_m + \cdots \tag{21}$$

with each F_m a homogeneous polynomial of degree m. The term F_1 will be denoted F_ℓ and called the ***linear form*** of F, and the first nonzero term on the right of (21) (the homogeneous summand of F of least degree) will be denoted F_* and called the ***leading form*** of F.

If $P = (0,0)$ is on the curve V defined by F, then $F_0 = 0$ and (21) becomes

$$F = aX + bY + \text{higher degree terms},$$

and the equation of the tangent space is

$$aX + bY = 0.$$

DEFINITION 4.9. Let $F(X,Y)$ be a polynomial without square factors, and let V be the curve defined by F. If $(0,0) \in V$, then the ***geometric tangent cone*** to V at $(0,0)$ is the zero set of F_*. The ***tangent cone*** is the pair $(V(F_*), k[X,Y]/F_*)$. To obtain the tangent cone at any other point, translate to the origin, and then translate back.

Note that the geometric tangent cone at a point on a curve always has dimension 1. While the tangent space tells you whether a point is nonsingular or not, the tangent cone also gives you information on the natural of a singularity.

In general we can factor F_* as

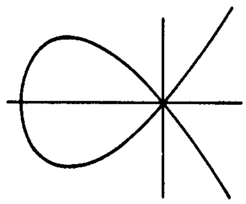
$$F_*(X,Y) = cX^{r_0}\prod\nolimits_i (Y - a_i X)^{r_i}.$$

Then $\deg F_* = \sum r_i$ is called the ***multiplicity*** of the singularity, $\mathrm{mult}_P(V)$. A multiple point is ***ordinary*** if its tangents are nonmultiple, i.e., $r_i = 1$ all i. An ordinary double point is called a ***node***. There are many names for special types of singularities — see any book, especially an old book, on algebraic curves.

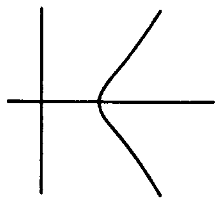
Examples

The following examples are adapted from Walker, Robert J., Algebraic Curves. Princeton Mathematical Series, vol. 13. Princeton University Press, Princeton, N. J., 1950 (reprinted by Dover 1962).

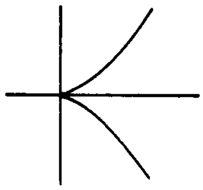
4.10. $F(X,Y) = X^3 + X^2 - Y^2$. The tangent cone at $(0,0)$ is defined by $Y^2 - X^2$. It is the pair of lines $Y = \pm X$, and the singularity is a node.



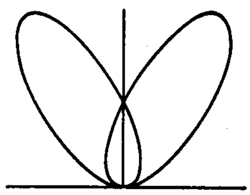
4.11. $F(X,Y) = X^3 - X^2 - Y^2$. The origin is an isolated point of the real locus. It is again a node, but the tangent cone is defined by $Y^2 + X^2$, which is the pair of lines $Y = \pm iX$. In this case, the real locus of the tangent cone is just the point (0,0).



4.12. $F(X,Y) = X^3 - Y^2$. Here the origin is a cusp. The tangent cone is defined by Y^2, which is the X-axis (doubled).

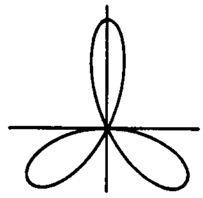


4.13. $F(X,Y) = 2X^4 - 3X^2Y + Y^2 - 2Y^3 + Y^4$. The origin is again a double point, but this time it is a tacnode. The tangent cone is again defined by Y^2.

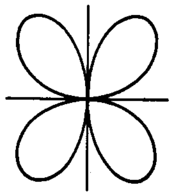


4.14. $F(X,Y) = X^4 + X^2Y^2 - 2X^2Y - XY^2 - Y^2$. The origin is again a double point, but this time it is a ramphoid cusp. The tangent cone is again defined by Y^2.

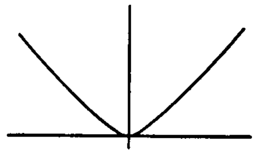
4.15. $F(X,Y) = (X^2+Y^2)^2+3X^2Y-Y^3$. The origin is an ordinary triple point. The tangent cone is defined by $3X^2Y-Y^3$, which is the triple of lines $Y=0$, $Y=\pm\sqrt{3}X$.



4.16. $F(X,Y) = (X^2+Y^2)^3-4X^2Y^2$. The origin has multiplicity 4. The tangent cone is defined by $4X^2Y^2$, which is the union of the X and Y axes, each doubled.



4.17. $F(X,Y) = X^6-X^2Y^3-Y^5$. The tangent cone is defined by $X^2Y^3+Y^5$, which consists of a triple line $Y^3=0$ and a pair of lines $Y=\pm iX$.



ASIDE 4.18. Note that the real locus of the algebraic curve in (4.17) is smooth even though the curve itself is singular. Another example of such a curve is $Y^3+2X^2Y-X^4=0$. This is singular at $(0,0)$, but its real locus is the image of $\mathbb{R}$ under the analytic map $t\mapsto(t^3+2t, t(t^3+2))$, which is injective, proper, and immersive, and hence an embedding into $\mathbb{R}^2$ with closed image. See Milnor, J., Singular points of complex hypersurfaces. PUP, 1968, or mo98366 (Elencwajg).

c The local ring at a point on a curve

PROPOSITION 4.19. *Let P be a point on a plane curve V, and let $\mathfrak{m}$ be the corresponding maximal ideal in $k[V]$. If P is nonsingular, then $\dim_k(\mathfrak{m}/\mathfrak{m}^2)=1$, and otherwise $\dim_k(\mathfrak{m}/\mathfrak{m}^2)=2$.*

PROOF. Assume first that $P=(0,0)$. Then $\mathfrak{m}=(x,y)$ in $k[V]=k[X,Y]/(F(X,Y))=k[x,y]$. Note that $\mathfrak{m}^2=(x^2,xy,y^2)$, and

$$\mathfrak{m}/\mathfrak{m}^2 = (X,Y)/(\mathfrak{m}^2+F(X,Y)) = (X,Y)/(X^2,XY,Y^2,F(X,Y)).$$

In this quotient, every element is represented by a linear polynomial $cx+dy$, and the only relation is $F_\ell(x,y)=0$. Clearly $\dim_k(\mathfrak{m}/\mathfrak{m}^2)=1$ if $F_\ell\neq 0$, and $\dim_k(\mathfrak{m}/\mathfrak{m}^2)=2$ otherwise. Since $F_\ell=0$ is the equation of the tangent space, this proves the proposition in this case.

The same argument works for an arbitrary point (a,b) except that one uses the variables $X'=X-a$ and $Y'=Y-b$; in essence, one translates the point to the origin. □

We explain what the condition $\dim_k(\mathfrak{m}/\mathfrak{m}^2)=1$ means for the local ring $\mathcal{O}_P=k[V]_\mathfrak{m}$. Let $\mathfrak{n}$ be the maximal ideal $\mathfrak{m}k[V]_\mathfrak{m}$ of this local ring. The map $\mathfrak{m}\to\mathfrak{n}$ induces an isomorphism $\mathfrak{m}/\mathfrak{m}^2\to\mathfrak{n}/\mathfrak{n}^2$ (see 1.15), and so we have

$$P \text{ nonsingular} \iff \dim_k \mathfrak{m}/\mathfrak{m}^2=1 \iff \dim_k \mathfrak{n}/\mathfrak{n}^2=1.$$

Nakayama's lemma (1.3) shows that the last condition is equivalent to $\mathfrak{n}$ being a principal ideal. As $\mathcal{O}_P$ has Krull dimension one (2.64), for its maximal ideal to be principal means that it is a regular local ring of dimension 1 (see 1.6). Thus, for a point P on a curve,

$$P \text{ nonsingular} \iff \mathcal{O}_P \text{ regular}.$$

PROPOSITION 4.20. *Every regular local ring of dimension one is a principal ideal domain.*

PROOF. Let A be such a ring, and let $\mathfrak{m} = (\pi)$ be its maximal ideal. According to the Krull intersection theorem (1.8), $\bigcap_{r\geq 0} \mathfrak{m}^r = (0)$. Let $\mathfrak{a}$ be a proper nonzero ideal in A. As $\mathfrak{a}$ is finitely generated, there exists an $r \in \mathbb{N}$ such that $\mathfrak{a} \subset \mathfrak{m}^r$ but $\mathfrak{a} \not\subset \mathfrak{m}^{r+1}$. Therefore, there exists an $a = c\pi^r \in \mathfrak{a}$ such that $a \notin \mathfrak{m}^{r+1}$. The second condition implies that $c \notin \mathfrak{m}$, and so it is a unit. Therefore $(\pi^r) = (a) \subset \mathfrak{a} \subset (\pi^r)$, and so $\mathfrak{a} = (\pi^r) = \mathfrak{m}^r$. We have shown that all ideals in A are principal.

By assumption, there exists a prime ideal $\mathfrak{p}$ properly contained in $\mathfrak{m}$. Then $A/\mathfrak{p}$ is an integral domain. As $\pi \notin \mathfrak{p}$, it is not nilpotent in $A/\mathfrak{p}$, and hence not nilpotent in A.

Let a and b be nonzero elements of A. There exist $r, s \in \mathbb{N}$ such that $a \in \mathfrak{m}^r \smallsetminus \mathfrak{m}^{r+1}$ and $b \in \mathfrak{m}^s \smallsetminus \mathfrak{m}^{s+1}$. Then $a = u\pi^r$ and $b = v\pi^s$ with u and v units, and $ab = uv\pi^{r+s} \neq 0$. Hence A is an integral domain. □

It follows from the elementary theory of principal ideal domains that the following conditions on a principal ideal domain A are equivalent:

(a) A has exactly one nonzero prime ideal;

(b) A has exactly one prime element up to associates;

(c) A is local and is not a field.

A ring satisfying these conditions is called a ***discrete valuation ring***.

THEOREM 4.21. *A point P on a plane algebraic curve is nonsingular if and only if $\mathcal{O}_P$ is regular, in which case it is a discrete valuation ring.*

PROOF. The statement summarizes the above discussion. □

d Tangent spaces to algebraic subsets of $\mathbb{A}^m$

Before defining tangent spaces at points of an algebraic subset of $\mathbb{A}^m$ we review some terminology from linear algebra (which should be familiar from advanced calculus).

LINEAR ALGEBRA

For a vector space k^m, let X_i be the ith coordinate function $\mathbf{a} \mapsto a_i$. Thus $X_1, \ldots, X_m$ is the dual basis to the standard basis for k^m. A linear form $\sum a_i X_i$ can be regarded as an element of the dual vector space $(k^m)^\vee = \mathrm{Hom}(k^m, k)$.

Let $A = (a_{ij})$ be an $n \times m$ matrix. It defines a linear map $\alpha: k^m \to k^n$, by

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \mapsto A \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j} a_j \\ \vdots \\ \sum_{j=1}^m a_{nj} a_j \end{pmatrix}.$$

Write $X_1, \ldots, X_m$ for the coordinate functions on k^m and $Y_1, \ldots, Y_n$ for the coordinate functions on k^n. Then

$$Y_i \circ \alpha = \sum_{j=1}^m a_{ij} X_j.$$

This says that the ith coordinate of $\alpha(\mathbf{a})$ is

$$\sum_{j=1}^m a_{ij} (X_j \mathbf{a}) = \sum_{j=1}^m a_{ij} a_j.$$

TANGENT SPACES

DEFINITION 4.22. Let $V \subset k^m$ be an algebraic subset of k^m, and let $\mathfrak{a} = I(V)$. The ***tangent space*** $T_{\mathbf{a}}(V)$ to V at a point $\mathbf{a} = (a_1, \ldots, a_m)$ of V is the subspace of the vector space with origin $\mathbf{a}$ cut out by the linear equations

$$\sum_{i=1}^{m} \left.\frac{\partial F}{\partial X_i}\right|_{\mathbf{a}} (X_i - a_i) = 0, \qquad F \in \mathfrak{a}. \tag{22}$$

In other words, $T_{\mathbf{a}}(\mathbb{A}^m)$ is the vector space of dimension m with origin $\mathbf{a}$, and $T_{\mathbf{a}}(V)$ is the subspace of $T_{\mathbf{a}}(\mathbb{A}^m)$ defined by the equations (22).

Write $(dX_i)_{\mathbf{a}}$ for $(X_i - a_i)$; then the $(dX_i)_{\mathbf{a}}$ form a basis for the dual vector space $T_{\mathbf{a}}(\mathbb{A}^m)^\vee$ to $T_{\mathbf{a}}(\mathbb{A}^m)$ — in fact, they are the coordinate functions on $T_{\mathbf{a}}(\mathbb{A}^m)^\vee$. As in advanced calculus, we define the ***differential*** of a polynomial $F \in k[X_1, \ldots, X_m]$ at $\mathbf{a}$ by the equation:

$$(dF)_{\mathbf{a}} = \sum_{i=1}^{m} \left.\frac{\partial F}{\partial X_i}\right|_{\mathbf{a}} (dX_i)_{\mathbf{a}}.$$

It is again a linear form on $T_{\mathbf{a}}(\mathbb{A}^m)$. In terms of differentials, $T_{\mathbf{a}}(V)$ is the subspace of $T_{\mathbf{a}}(\mathbb{A}^m)$ defined by the equations:

$$(dF)_{\mathbf{a}} = 0, \quad F \in \mathfrak{a}. \tag{23}$$

I claim that, in (22) and (23), it suffices to take the F to lie in a generating subset for $\mathfrak{a}$. The product rule for differentiation shows that if $G = \sum_j H_j F_j$, then

$$(dG)_{\mathbf{a}} = \sum_j H_j(\mathbf{a}) \cdot (dF_j)_{\mathbf{a}} + F_j(\mathbf{a}) \cdot (dH_j)_{\mathbf{a}}.$$

If $F_1, \ldots, F_r$ generate $\mathfrak{a}$ and $\mathbf{a} \in V(\mathfrak{a})$, so that $F_j(\mathbf{a}) = 0$ for all j, then this equation becomes

$$(dG)_{\mathbf{a}} = \sum_j H_j(\mathbf{a}) \cdot (dF_j)_{\mathbf{a}}.$$

Thus $(dF_1)_{\mathbf{a}}, \ldots, (dF_r)_{\mathbf{a}}$ generate the k-vector space $\{(dF)_{\mathbf{a}} \mid F \in \mathfrak{a}\}$.

DEFINITION 4.23. A point $\mathbf{a}$ on an algebraic set V is ***nonsingular*** (or ***smooth***) if it lies on a single irreducible component W of V and the dimension of the tangent space at $\mathbf{a}$ is equal to the dimension of W; otherwise it is ***singular*** (or ***multiple***).

Thus, a point $\mathbf{a}$ on an irreducible algebraic set V is nonsingular if and only if $\dim T_{\mathbf{a}}(V) = \dim V$. As in the case of plane curves, a point on V is nonsingular if and only if it lies on a single irreducible component of V, and is nonsingular on it.

Let $\mathfrak{a} = (F_1, \ldots, F_r)$, and let

$$J = \mathrm{Jac}(F_1, \ldots, F_r) = \left(\frac{\partial F_i}{\partial X_j}\right) = \begin{pmatrix} \frac{\partial F_1}{\partial X_1}, & \ldots, & \frac{\partial F_1}{\partial X_m} \\ \vdots & & \vdots \\ \frac{\partial F_r}{\partial X_1}, & \ldots, & \frac{\partial F_r}{\partial X_m} \end{pmatrix}.$$

Then the equations defining $T_{\mathbf{a}}(V)$ as a subspace of $T_{\mathbf{a}}(\mathbb{A}^m)$ have matrix $J(\mathbf{a})$. Therefore, linear algebra shows that

$$\dim_k T_{\mathbf{a}}(V) = m - \operatorname{rank} J(\mathbf{a}),$$

and so $\mathbf{a}$ is nonsingular if and only if the rank of $\mathrm{Jac}(F_1,\ldots,F_r)(\mathbf{a})$ is equal to $m-\dim(V)$. For example, if V is a hypersurface, say $I(V)=(F(X_1,\ldots,X_m))$, then

$$\mathrm{Jac}(F)(\mathbf{a})=\left(\frac{\partial F}{\partial X_1}(\mathbf{a}),\ldots,\frac{\partial F}{\partial X_m}(\mathbf{a})\right),$$

and $\mathbf{a}$ is nonsingular if and only if not all of the partial derivatives $\frac{\partial F}{\partial X_i}$ vanish at $\mathbf{a}$.

We can regard J as a matrix of regular functions on V. For each r,

$$\{\mathbf{a}\in V\mid \operatorname{rank} J(\mathbf{a})\leq r\}$$

is closed in V, because it is the set where certain determinants vanish. Therefore, there is an open subset U of V on which $\operatorname{rank} J(\mathbf{a})$ attains its maximum value, and the rank jumps on closed subsets. Later (4.37) we shall show that the maximum value of $\operatorname{rank} J(\mathbf{a})$ is $m-\dim V$, and so the nonsingular points of V form a nonempty open subset of V.

e The differential of a regular map

Consider a regular map

$$\varphi\colon\mathbb{A}^m\to\mathbb{A}^n,\quad \mathbf{a}\mapsto(P_1(a_1,\ldots,a_m),\ldots,P_n(a_1,\ldots,a_m)).$$

We think of φ as being given by the equations

$$Y_i=P_i(X_1,\ldots,X_m),\ i=1,\ldots n.$$

It corresponds to the map of rings $\varphi^*\colon k[Y_1,\ldots,Y_n]\to k[X_1,\ldots,X_m]$ sending Y_i to $P_i(X_1,\ldots,X_m)$, $i=1,\ldots n$.

Let $\mathbf{a}\in\mathbb{A}^m$, and let $\mathbf{b}=\varphi(\mathbf{a})$. Define $(d\varphi)_{\mathbf{a}}\colon T_{\mathbf{a}}(\mathbb{A}^m)\to T_{\mathbf{b}}(\mathbb{A}^n)$ to be the map such that

$$(dY_i)_{\mathbf{b}}\circ(d\varphi)_{\mathbf{a}}=\sum\left.\frac{\partial P_i}{\partial X_j}\right|_{\mathbf{a}}(dX_j)_{\mathbf{a}},$$

i.e., relative to the standard bases, $(d\varphi)_{\mathbf{a}}$ is the map with matrix

$$\mathrm{Jac}(P_1,\ldots,P_n)(\mathbf{a})=\begin{pmatrix}\frac{\partial P_1}{\partial X_1}(\mathbf{a}), & \ldots, & \frac{\partial P_1}{\partial X_m}(\mathbf{a})\\ \vdots & & \vdots\\ \frac{\partial P_n}{\partial X_1}(\mathbf{a}), & \ldots, & \frac{\partial P_n}{\partial X_m}(\mathbf{a})\end{pmatrix}.$$

For example, suppose $\mathbf{a}=(0,\ldots,0)$ and $\mathbf{b}=(0,\ldots,0)$, so that $T_{\mathbf{a}}(\mathbb{A}^m)=k^m$ and $T_{\mathbf{b}}(\mathbb{A}^n)=k^n$, and

$$P_i=\sum_{j=1}^{m}c_{ij}X_j+\text{(higher terms)},\ i=1,\ldots,n.$$

Then $Y_i\circ(d\varphi)_{\mathbf{a}}=\sum_j c_{ij}X_j$, and the map on tangent spaces is given by the matrix (c_{ij}), i.e., it is simply $\mathbf{t}\mapsto(c_{ij})\mathbf{t}$.

Let $F\in k[X_1,\ldots,X_m]$. We can regard F as a regular map $\mathbb{A}^m\to\mathbb{A}^1$, whose differential will be a linear map

$$(dF)_{\mathbf{a}}\colon T_{\mathbf{a}}(\mathbb{A}^m)\to T_{\mathbf{b}}(\mathbb{A}^1),\qquad \mathbf{b}=F(\mathbf{a}).$$

When we identify $T_{\mathbf{b}}(\mathbb{A}^1)$ with k, we obtain an identification of the differential of F (F regarded as a regular map) with the differential of F (F regarded as a regular function).

LEMMA 4.24. *Let $\varphi\colon \mathbb{A}^m \to \mathbb{A}^n$ be as at the start of this subsection. If φ maps $V = V(\mathfrak{a}) \subset k^m$ into $W = V(\mathfrak{b}) \subset k^n$, then $(d\varphi)_{\mathbf{a}}$ maps $T_{\mathbf{a}}(V)$ into $T_{\mathbf{b}}(W)$, $\mathbf{b} = \varphi(\mathbf{a})$.*

PROOF. We are given that

$$f \in \mathfrak{b} \Rightarrow f \circ \varphi \in \mathfrak{a},$$

and have to prove that

$$f \in \mathfrak{b} \Rightarrow (df)_{\mathbf{b}} \circ (d\varphi)_{\mathbf{a}} \text{ is zero on } T_{\mathbf{a}}(V).$$

The chain rule holds in our situation:

$$\frac{\partial f}{\partial X_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial Y_j}\frac{\partial Y_j}{\partial X_i}, \quad Y_j = P_j(X_1,\ldots,X_m), \quad f = f(Y_1,\ldots,Y_n).$$

If φ is the map given by the equations

$$Y_j = P_j(X_1,\ldots,X_m), \qquad j = 1,\ldots,n,$$

then the chain rule implies

$$d(f \circ \varphi)_{\mathbf{a}} = (df)_{\mathbf{b}} \circ (d\varphi)_{\mathbf{a}}, \quad \mathbf{b} = \varphi(\mathbf{a}).$$

Let $\mathbf{t} \in T_{\mathbf{a}}(V)$; then

$$(df)_{\mathbf{b}} \circ (d\varphi)_{\mathbf{a}}(\mathbf{t}) = d(f \circ \varphi)_{\mathbf{a}}(\mathbf{t}),$$

which is zero if $f \in \mathfrak{b}$ because then $f \circ \varphi \in \mathfrak{a}$. Thus $(d\varphi)_{\mathbf{a}}(\mathbf{t}) \in T_{\mathbf{b}}(W)$. □

We therefore get a map $(d\varphi)_{\mathbf{a}}\colon T_{\mathbf{a}}(V) \to T_{\mathbf{b}}(W)$. The usual rules from advanced calculus hold. For example,

$$(d\psi)_{\mathbf{b}} \circ (d\varphi)_{\mathbf{a}} = d(\psi \circ \varphi)_{\mathbf{a}}, \quad \mathbf{b} = \varphi(\mathbf{a}).$$

f Tangent spaces to affine algebraic varieties

The definition (4.22) of the tangent space at a point on an algebraic set uses the embedding of the algebraic set into $\mathbb{A}^n$. In this section, we give an intrinsic definition of the tangent space at a point of an affine algebraic variety that makes clear that it depends only on the local ring at the point.

Dual numbers

For an affine algebraic variety V and a k-algebra R (not necessarily an affine k-algebra), we define $V(R)$ to be $\mathrm{Hom}_{k\text{-alg}}(k[V], R)$. For example, if $V \subset \mathbb{A}^n$ and $\mathfrak{a} = I(V)$, then

$$V(R) = \{(a_1,\ldots,a_n) \in R^n \mid f(a_1,\ldots,a_n) = 0 \text{ for all } f \in \mathfrak{a}\}.$$

A homomorphism $R \to S$ of k-algebras defines a map $V(R) \to V(S)$ of sets.

The ***ring of dual numbers*** is $k[\varepsilon] = k[X]/(X^2)$ where $\varepsilon = X + (X^2)$. Thus $k[\varepsilon] = k \oplus k\varepsilon$ as a k-vector space, and

$$(a + b\varepsilon)(a' + b'\varepsilon) = aa' + (ab' + a'b)\varepsilon, \quad a,b,a',b' \in k.$$

Note that there is a k-algebra homomorphism $\varepsilon \mapsto 0\colon k[\varepsilon] \to k$.

DEFINITION 4.25. Let P be a point on an affine algebraic variety V over k. The tangent space to V at P is

$$T_P(V) = \{P' \in V(k[\varepsilon]) \mid P' \mapsto P \text{ under } V(k[\varepsilon]) \to V(k)\}.$$

Thus an element of $T_P(V)$ is a homomorphism of k-algebras $\alpha\colon k[V] \to k[\varepsilon]$ whose composite with $k[\varepsilon] \xrightarrow{\varepsilon\mapsto 0} k$ is the point P. To say that $k[V] \to k$ is the point P means that its kernel is $\mathfrak{m}_P$, and so $\mathfrak{m}_P = \alpha^{-1}((\varepsilon))$.

PROPOSITION 4.26. *Let V be an algebraic subset of $\mathbb{A}^n$, and let $V' = (V, \mathcal{O}_V)$ be V equipped with its canonical structure of an affine algebraic variety. Let $P \in V$. Then*

$$T_P(V) \text{ (as defined in 4.22)} \simeq T_P(V') \text{ (as defined in 4.25).}$$

PROOF. Let $I(V) = \mathfrak{a}$ and let $P = (a_1, \ldots, a_n)$. On rewriting a polynomial $F(X_1, \ldots, X_n)$ in terms of the variables $X_i - a_i$, we obtain the (trivial Taylor) formula,

$$F(X_1, \ldots, X_n) = F(a_1, \ldots, a_n) + \sum \left.\frac{\partial F}{\partial X_i}\right|_{\mathbf{a}} (X_i - a_i) + R$$

with R a finite sum of products of at least two terms $(X_i - a_i)$.

According to (4.25), $T_P(V')$ consists of the elements $\mathbf{a} + \varepsilon\mathbf{b}$ of $k[\varepsilon]^n = k^n \oplus k^n\varepsilon$ lying in $V(k[\varepsilon])$. Let $F \in \mathfrak{a}$. On setting X_i equal to $a_i + \varepsilon b_i$ in the above formula, we obtain:

$$F(a_1 + \varepsilon b_1, \ldots, a_n + \varepsilon b_n) = \varepsilon \left(\sum \left.\frac{\partial F}{\partial X_i}\right|_{\mathbf{a}} b_i \right).$$

Thus, $(a_1 + \varepsilon b_1, \ldots, a_n + \varepsilon b_n)$ lies in $V(k[\varepsilon])$ if and only if $(b_1, \ldots, b_n) \in T_{\mathbf{a}}(V)$. □

We can restate this as follows. Let V be an affine algebraic variety, and let $P \in V$. Choose an embedding $V \hookrightarrow \mathbb{A}^n$, and let P map to $(a_1, \ldots, a_n)$. Then the point

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n)\varepsilon$$

of $\mathbb{A}^n(k[\varepsilon])$ is an element of $T_P(V)$ (definition 4.25) if and only if $(b_1, \ldots, b_n)$ is an element of $T_P(V)$ (definition 4.22).

It is possible to define $\mathrm{Spm}(k[\varepsilon])$ as a k-ringed space in a slightly more general sense than we defined it: it is a one-point space e with the sheaf $\mathcal{O}$ such that $\Gamma(e, \mathcal{O}) = k[\varepsilon]$. Geometrically, we can think of $\mathrm{Spm}(k[\varepsilon])$ as being the point $\mathrm{Spm}(k)$ equipped with a tangent vector. To give a morphism of k-ringed spaces $\mathrm{Spm}(k[\varepsilon]) \to V$ amounts to giving a point P of V (the image of $\mathrm{Spm}(k)$) together with a tangent vector at P.

PROPOSITION 4.27. *Let V be an affine variety, and let $P \in V$. There is a canonical isomorphism*

$$T_P(V) \simeq \mathrm{Hom}(\mathcal{O}_P, k[\varepsilon]) \text{ (local homomorphisms of local } k\text{-algebras).}$$

PROOF. By definition, an element of $T_P(V)$ is a homomorphism $\alpha\colon k[V] \to k[\varepsilon]$ such that $\alpha^{-1}((\varepsilon)) = \mathfrak{m}_P$. Therefore α maps elements of $k[V] \smallsetminus \mathfrak{m}_P$ into $(k[\varepsilon] \smallsetminus (\varepsilon)) = k[\varepsilon]^\times$, and so α extends (uniquely) to a homomorphism $\alpha'\colon \mathcal{O}_P \to k[\varepsilon]$. By construction, α' is a local homomorphism of local k-algebras, and clearly every such homomorphism arises in this way from an element of $T_P(V)$. □

Derivations

DEFINITION 4.28. Let A be a k-algebra and M an A-module. A ***k-derivation*** is a map $D\colon A \to M$ such that

(a) $D(c) = 0$ for all $c \in k$;

(b) $D(f+g) = D(f) + D(g)$;

(c) $D(fg) = f \cdot Dg + g \cdot Df$ (Leibniz's rule).

Note that the conditions imply that D is k-linear (but not A-linear). We write $\mathrm{Der}_k(A, M)$ for the k-vector space of all k-derivations $A \to M$.

For example, let A be a local k-algebra with maximal ideal $\mathfrak{m}$, and assume that $A/\mathfrak{m} = k$. For $f \in A$, let $f(\mathfrak{m})$ denote the image of f in $A/\mathfrak{m}$. Then $f - f(\mathfrak{m}) \in \mathfrak{m}$, and the map

$$f \mapsto df \overset{\text{def}}{=} f - f(\mathfrak{m}) \mod \mathfrak{m}^2$$

is a k-derivation $A \to \mathfrak{m}/\mathfrak{m}^2$ because, mod $\mathfrak{m}^2$,

$$\begin{aligned} 0 &= (f - f(\mathfrak{m}))(g - g(\mathfrak{m})) \\ &= -fg + f(\mathfrak{m})g(\mathfrak{m}) + f \cdot (g - g(\mathfrak{m})) + g(f - f(\mathfrak{m})) \\ &= -d(fg) + f \cdot dg + g \cdot df. \end{aligned}$$

PROPOSITION 4.29. *Let $(A, \mathfrak{m})$ be as above. There are canonical isomorphisms*

$$\mathrm{Hom}_{\text{local } k\text{-algebra}}(A, k[\varepsilon]) \to \mathrm{Der}_k(A, k) \to \mathrm{Hom}_{k\text{-linear}}(\mathfrak{m}/\mathfrak{m}^2, k)$$

PROOF. The composite $k \xrightarrow{c \mapsto c} A \xrightarrow{f \mapsto f(\mathfrak{m})} k$ is the identity map, and so, when regarded as k-vector space, A decomposes into

$$A = k \oplus \mathfrak{m}, \quad f \leftrightarrow (f(\mathfrak{m}), f - f(\mathfrak{m})).$$

Let $\alpha\colon A \to k[\varepsilon]$ be a local homomorphism of k-algebras, and write $\alpha(a) = a_0 + D_\alpha(a)\varepsilon$. Because α is a homomorphism of k-algebras, $a_0 = a(\mathfrak{m})$. We have

$$\begin{aligned} \alpha(ab) &= (ab)_0 + D_\alpha(ab)\varepsilon, \text{ and} \\ \alpha(a)\alpha(b) &= (a_0 + D_\alpha(a)\varepsilon)(b_0 + D_\alpha(b)\varepsilon) = a_0 b_0 + (a_0 D_\alpha(b) + b_0 D_\alpha(a))\varepsilon. \end{aligned}$$

On comparing these expressions, we see that D_α satisfies Leibniz's rule, and therefore is a k-derivation $\mathcal{O}_P \to k$. Conversely, if $D\colon A \to k$ is a k-derivation, then

$$\alpha\colon a \mapsto a(\mathfrak{m}) + D(a)\varepsilon$$

is a local homomorphism of k-algebras $A \to k[\varepsilon]$, and all such homomorphisms arise in this way.

A derivation $D\colon A \to k$ is zero on k and on $\mathfrak{m}^2$ (by Leibniz's rule). It therefore defines a k-linear map $\mathfrak{m}/\mathfrak{m}^2 \to k$. Conversely, a k-linear map $\mathfrak{m}/\mathfrak{m}^2 \to k$ defines a derivation by composition

$$A \xrightarrow{f \mapsto df} \mathfrak{m}/\mathfrak{m}^2 \to k.$$

□

Tangent spaces and differentials

We now summarize the above discussion in the context of affine algebraic varieties.

4.30. Let V be an affine algebraic variety, and let P be a point on V. Write $\mathfrak{m}_P$ for the corresponding maximal ideal in $k[V]$ and $\mathfrak{n}_P$ for the maximal ideal $\mathfrak{m}_P\mathcal{O}_{V,P}$ in the local ring at P. There are canonical isomorphisms

$$\begin{tikzcd} T_P(V) \arrow[r] \arrow[d] & \mathrm{Der}_k(k[V],k) \arrow[r] \arrow[d] & \mathrm{Hom}_{k\text{-linear}}(\mathfrak{m}_P/\mathfrak{m}_P^2,k) \arrow[d] \\ \mathrm{Hom}_{\text{local }k\text{-algebra}}(\mathcal{O}_P,k[\varepsilon]) \arrow[r] & \mathrm{Der}_k(\mathcal{O}_P,k) \arrow[r] & \mathrm{Hom}_{k\text{-linear}}(\mathfrak{n}_P/\mathfrak{n}_P^2,k). \end{tikzcd} \tag{24}$$

In the middle term on the top row, $k[V]$ acts on k through $k[V]\to k[V]/\mathfrak{m}_P\simeq k$, and on the bottom row $\mathcal{O}_P$ acts on k through $\mathcal{O}_P\to\mathcal{O}_P/\mathfrak{n}_P\simeq k$. The maps have the following descriptions.

(a) By definition, $T_P(V)$ is the fibre of $V(k[\varepsilon])\to V(k)$ over P. To give an element of $T_P(V)$ amounts to giving a homomorphism $\alpha\colon k[V]\to k[\varepsilon]$ such that $\alpha^{-1}((\varepsilon))=\mathfrak{m}_P$.

(b) The homomorphism α in (a) can be decomposed,

$$\alpha(f)=f(\mathfrak{m})\oplus D_\alpha(f)\varepsilon,\quad f\in k[V],\ f(\mathfrak{m}_P)\in k,\ D_\alpha(f)\in k.$$

The map D_α is a k-derivation $k[V]\to k$, and D_α induces a k-linear map $\mathfrak{m}_P/\mathfrak{m}_P^2\to k$.

(c) The homomorphism $\alpha\colon k[V]\to k[\varepsilon]$ in (a) extends uniquely to a local homomorphism $\mathcal{O}_P\to k[\varepsilon]$. Similarly, a k-derivation $k[V]\to k$ extends uniquely to a k-derivation $\mathcal{O}_P\to k$.

(d) The two right hand groups are related through the isomorphism $\mathfrak{m}_P/\mathfrak{m}_P^2\to\mathfrak{n}_P/\mathfrak{n}_P^2$ of (1.15).

4.31. A regular map $\varphi\colon V\to W$ defines a map $\varphi(k[\varepsilon])\colon V(k[\varepsilon])\to W(k[\varepsilon])$. If $Q=\varphi(P)$, then φ maps the fibre over P to the fibre over Q, i.e., it defines a map

$$d\varphi\colon T_P(V)\to T_Q(W).$$

$$\begin{tikzcd} T_P(V) \arrow[r, "d\varphi"] \arrow[d] & T_Q(W) \arrow[d] \\ V(k[\varepsilon]) \arrow[r, "\varphi"] \arrow[d, "\varepsilon\mapsto 0"] & W(k[\varepsilon]) \arrow[d, "\varepsilon\mapsto 0"] \\ V(k) \arrow[r, "\varphi"] & W(k) \end{tikzcd}$$

This map of tangent spaces is called the *differential* of φ at P.

(a) When V and W are embedded as closed subvarieties of $\mathbb{A}^n$, $d\varphi$ has the description in p.87.

(b) As a map $\mathrm{Hom}(\mathcal{O}_P,k[\varepsilon])\to\mathrm{Hom}(\mathcal{O}_Q,k[\varepsilon])$, $d\varphi$ is induced by $\varphi^*\colon\mathcal{O}_Q\to\mathcal{O}_P$.

(c) As a map $\mathrm{Hom}(\mathfrak{m}_P/\mathfrak{m}_P^2,k)\to\mathrm{Hom}(\mathfrak{m}_Q/\mathfrak{m}_Q^2,k)$, $d\varphi$ is induced by the map $\mathfrak{m}_Q/\mathfrak{m}_Q^2\to\mathfrak{m}_P/\mathfrak{m}_P^2$ defined by $\varphi^*\colon k[W]\to k[V]$.

EXAMPLE 4.32. Let E be a finite dimensional vector space over k. Then

$$T_o(\mathbb{A}(E))\simeq E.$$

ASIDE 4.33. A map $\mathrm{Spm}(k[\varepsilon])\to V$ should be thought of as a curve in V but with only the first infinitesimal structure retained. Thus, the descriptions of the tangent space provided by the terms in the top row of (24) correspond to the three standard descriptions of the tangent space in differential geometry (Wikipedia TANGENT SPACE).

g Tangent cones

Let V be an algebraic subset of k^m, and let $\mathfrak{a} = I(V)$. Assume that $P = (0,\ldots,0) \in V$. Define $\mathfrak{a}_*$ to be the ideal generated by the polynomials F_* for $F \in \mathfrak{a}$, where F_* is the leading form of F (see p.81). The ***geometric tangent cone*** at P, $C_P(V)$ is $V(\mathfrak{a}_*)$, and the ***tangent cone*** is the pair $(V(\mathfrak{a}_*), k[X_1,\ldots,X_n]/\mathfrak{a}_*)$. Obviously, $C_P(V) \subset T_P(V)$.

If $\mathfrak{a}$ is principal, say $\mathfrak{a} = (F)$, then $\mathfrak{a}_* = (F_*)$, but if $\mathfrak{a} = (F_1,\ldots,F_r)$, then it need not be true that $\mathfrak{a}_* = (F_{1*},\ldots,F_{r*})$. Consider for example $\mathfrak{a} = (XY, XZ + Z(Y^2 - Z^2))$. One can show that this is an intersection of prime ideals, and hence is radical. As the polynomial

$$YZ(Y^2 - Z^2) = Y \cdot (XZ + Z(Y^2 - Z^2)) - Z \cdot (XY)$$

lies in $\mathfrak{a}$ and is homogeneous, it lies in $\mathfrak{a}_*$, but it is not in the ideal generated by XY, XZ. In fact, $\mathfrak{a}_*$ is the ideal generated by

$$XY, \quad XZ, \quad YZ(Y^2 - Z^2).$$

Let A be a local ring with maximal ideal $\mathfrak{n}$. The ***associated graded ring*** is

$$\mathrm{gr}(A) = \bigoplus\nolimits_{i \geq 0} \mathfrak{n}^i/\mathfrak{n}^{i+1}.$$

Note that if $A = B_{\mathfrak{m}}$ and $\mathfrak{n} = \mathfrak{m}A$, then $\mathrm{gr}(A) = \bigoplus \mathfrak{m}^i/\mathfrak{m}^{i+1}$ (because of (1.15)).

PROPOSITION 4.34. *The map $k[X_1,\ldots,X_n]/\mathfrak{a}_* \to \mathrm{gr}(\mathcal{O}_P)$ sending the class of X_i in $k[X_1,\ldots,X_n]/\mathfrak{a}_*$ to the class of X_i in $\mathrm{gr}(\mathcal{O}_P)$ is an isomorphism.*

PROOF. Let $\mathfrak{m}$ be the maximal ideal in $k[X_1,\ldots,X_n]/\mathfrak{a}$ corresponding to P. Then

$$\begin{aligned} \mathrm{gr}(\mathcal{O}_P) &= \sum \mathfrak{m}^i/\mathfrak{m}^{i+1} \\ &= \sum (X_1,\ldots,X_n)^i/(X_1,\ldots,X_n)^{i+1} + \mathfrak{a} \cap (X_1,\ldots,X_n)^i \\ &= \sum (X_1,\ldots,X_n)^i/(X_1,\ldots,X_n)^{i+1} + \mathfrak{a}_i \end{aligned}$$

where $\mathfrak{a}_i$ is the homogeneous piece of $\mathfrak{a}_*$ of degree i (that is, the subspace of $\mathfrak{a}_*$ consisting of homogeneous polynomials of degree i). But

$$(X_1,\ldots,X_n)^i/(X_1,\ldots,X_n)^{i+1} + \mathfrak{a}_i = i\text{th homogeneous piece of } k[X_1,\ldots,X_n]/\mathfrak{a}_*. \qquad \square$$

For an affine algebraic variety V and $P \in V$, we define the ***geometric tangent cone*** $C_P(V)$ of V at P to be $\mathrm{Spm}(\mathrm{gr}(\mathcal{O}_P)_{\mathrm{red}})$, where $\mathrm{gr}(\mathcal{O}_P)_{\mathrm{red}}$ is the quotient of $\mathrm{gr}(\mathcal{O}_P)$ by its nilradical, and we define the ***tangent cone*** to be $(C_P(V), \mathrm{gr}(\mathcal{O}_P))$.

As in the case of a curve, the dimension of the geometric tangent cone at P is the same as the dimension of V (because the Krull dimension of noetherian local ring is equal to that of its graded ring). Moreover, $\mathrm{gr}(\mathcal{O}_P)$ is a polynomial ring in $\dim V$ variables if and only if $\mathcal{O}_P$ is regular. Therefore, P is nonsingular (see below) if and only if $\mathrm{gr}(\mathcal{O}_P)$ is a polynomial ring in d variables, in which case $C_P(V) = T_P(V)$.

A regular map $\varphi\colon V \to W$ sending P to Q induces a homomorphism $\mathrm{gr}(\mathcal{O}_Q) \to \mathrm{gr}(\mathcal{O}_P)$, and hence a map $C_P(V) \to C_Q(V)$ of the geometric tangent cones.

. The map on the rings $k[X_1,\ldots,X_n]/\mathfrak{a}^*$ defined by a map of algebraic varieties is not the obvious one, i.e., it is not necessarily induced by the same map on polynomial rings as the original map. To see what it is, it is necessary to use Proposition 4.34, i.e., it is necessary to work with the rings $\mathrm{gr}(\mathcal{O}_P)$.

h Nonsingular points; the singular locus

DEFINITION 4.35. A point P on an affine algebraic variety V is said to be ***nonsingular*** or ***smooth*** if it lies on a single irreducible component W of V, and $\dim T_P(V) = \dim W$; otherwise the point is said to be ***singular***. A variety is ***nonsingular*** if all of its points are nonsingular. The set of singular points of a variety is called its ***singular locus.***

Thus, on an irreducible variety V of dimension d,

$$P \text{ is nonsingular} \iff \dim_k T_P(V) = d \iff \dim_k(\mathfrak{n}_P/\mathfrak{n}_P^2) = d.$$

PROPOSITION 4.36. *Let V be an irreducible variety of dimension d, and let P be a nonsingular point on V. Then there exist d regular functions $f_1, \ldots, f_d$ defined in an open neighbourhood U of P such that P is the only common zero of the f_i on U.*

PROOF. Suppose that P is nonsingular. Let $f_1, \ldots, f_d$ generate the maximal ideal $\mathfrak{n}_P$ in $\mathcal{O}_P$. Then $f_1, \ldots, f_d$ are all defined on some open affine neighbourhood U of P, and I claim that P is an irreducible component of the zero set $V(f_1, \ldots, f_d)$ of $f_1, \ldots, f_d$ in U. If not, there will be some irreducible component $Z \neq P$ of $V(f_1, \ldots, f_d)$ passing through P. Write $Z = V(\mathfrak{p})$ with $\mathfrak{p}$ a prime ideal in $k[U]$. Because $V(\mathfrak{p}) \subset V(f_1, \ldots, f_d)$ and because Z contains P and is not equal to it, we have

$$(f_1, \ldots, f_d) \subset \mathfrak{p} \subsetneqq \mathfrak{m}_P \qquad \text{(ideals in } k[U]).$$

On passing to the local ring $\mathcal{O}_P = k[U]_{\mathfrak{m}_P}$, we find (using 1.14) that

$$(f_1, \ldots, f_d) \subset \mathfrak{p}\mathcal{O}_P \subsetneqq \mathfrak{n}_P \qquad \text{(ideals in } \mathcal{O}_P).$$

This contradicts the assumption that the f_i generate $\mathfrak{n}_P$. Hence P is an irreducible component of $V(f_1, \ldots, f_d)$. On removing the remaining irreducible components of $V(f_1, \ldots, f_d)$ from U, we obtain an open neighbourhood of P with the required property. □

Let P be a point on an irreducible variety V, and let $f_1, \ldots, f_r$ generate the maximal ideal $\mathfrak{n}_P$ in $\mathcal{O}_P$. The proof of the proposition shows that P is an irreducible component of $V(f_1, \ldots, f_r)$, and so $r \geq d$ (see 3.45). Nakayama's lemma (1.3) shows that $f_1, \ldots, f_r$ generate $\mathfrak{n}_P$ if and only if their images in $\mathfrak{n}_P/\mathfrak{n}_P^2$ span it. Thus $\dim T_P(V) \geq \dim V$, with equality if and only if P is nonsingular.

A point P on V is nonsingular if and only if there exists an open affine neighbourhood U of P and functions $f_1, \ldots f_d$ on U such that $(f_1, \ldots, f_d)$ is the ideal of all regular functions on U zero at P.

THEOREM 4.37. *The set of nonsingular points of an affine algebraic variety is dense and open.*

PROOF. Let V be an irreducible component of the variety. It suffices to show that the singular locus of V is a proper closed subset.[1]

[1] Let $V_1, \ldots, V_r$ be the irreducible components of V. Then $V_i \cap (\bigcap_{j \neq i} V_j)$ is a proper closed subset of V_i. We show that $(V_i)_{\text{sing}}$ is a proper closed subset of V_i. It follows that $V_i \cap V_{\text{sing}}$ is the union of two proper closed subsets of V_i, and so it is proper and closed in V_i. Hence the points of V_i that are nonsingular on V form a nonempty open (hence dense) subset of V_i.

We first show that it is closed. We may suppose that $V = V(\mathfrak{a}) \subset \mathbb{A}^n$. Let $P_1, \ldots, P_r$ generate $\mathfrak{a}$. Then the singular locus is the zero set of the ideal generated by the $(n-d) \times (n-d)$ minors of the matrix

$$\mathrm{Jac}(P_1, \ldots, P_r)(\mathbf{a}) = \begin{pmatrix} \frac{\partial P_1}{\partial X_1}(\mathbf{a}) & \ldots & \frac{\partial P_1}{\partial X_m}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial P_r}{\partial X_1}(\mathbf{a}) & \ldots & \frac{\partial P_r}{\partial X_m}(\mathbf{a}) \end{pmatrix},$$

which is closed.

We now show that the singular locus is not equal to V. According to (3.36) and (3.37) some nonempty open affine subset of V is isomorphic to a nonempty open affine subset of an irreducible hypersurface in $\mathbb{A}^{d+1}$, and so we may suppose that V itself is an irreducible hypersurface in $\mathbb{A}^{d+1}$, say, equal to the zero set of the nonconstant irreducible polynomial $F(X_1, \ldots, X_{d+1})$. By (2.64), $\dim V = d$. The singular locus is the set of common zeros of the polynomials

$$F, \frac{\partial F}{\partial X_1}, \ldots, \frac{\partial F}{\partial X_{d+1}},$$

and so it will be proper unless the polynomials $\partial F/\partial X_i$ are identically zero on V. As in the proof of (4.7), if $\partial F/\partial X_i$ is identically zero on $V(F)$, then it is the zero polynomial, and so F is a polynomial in $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots X_{d+1}$ (characteristic zero) or in $X_1, \ldots, X_i^p, \ldots, X_{d+1}$ (characteristic p). Therefore, if the singular locus equals V, then F is constant (characteristic 0) or a pth power (characteristic p), which contradicts the hypothesis. □

COROLLARY 4.38. *If V is irreducible, then*

$$\dim V = \min_{P \in V} \dim T_P(V).$$

PROOF. By definition $\dim T_P(V) \geq \dim V$, with equality if and only if P is nonsingular. As there exists a nonsingular P, $\dim V$ is the minimum value of $\dim T_P(V)$. □

This formula can be useful in computing the dimension of a variety.

COROLLARY 4.39. *An irreducible algebraic variety is nonsingular if and only if the tangent spaces $T_P(V)$, $P \in V$, have constant dimension.*

PROOF. The constant dimension is the dimension of V, and so all points are nonsingular. □

COROLLARY 4.40. *Every variety on which a group acts transitively by regular maps is nonsingular.*

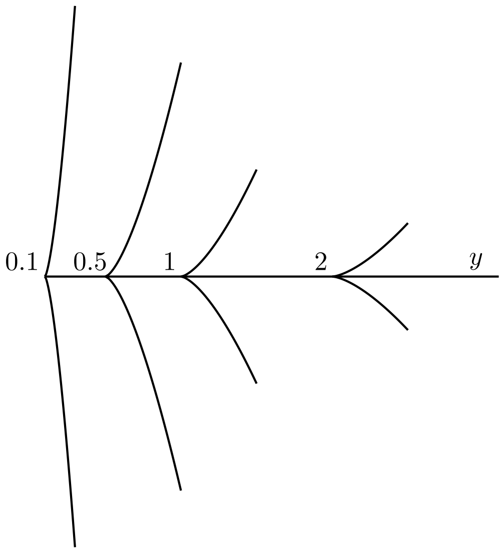
PROOF. The group must act by isomorphisms, and so the tangent spaces have constant dimension. □

In particular, every group variety (see p. g) is nonsingular.

Examples

4.41. For the surface $Z^3 = XY$, the only singular point is $(0,0,0)$. The tangent cone at $(0,0,0)$ has equation $XY = 0$, and so it is the union of two planes intersecting in the z-axis.

4.42. For the surface $V: Z^3 = X^2Y$, the singular locus is the line $X = 0 = Z$ (and the singularity at $(0,0)$ is very bad. The intersection of the surface with the surface $Y = c$ is the cuspidal curve $X^2 = Z^3/c$:



4.43. Let V be the union of the coordinate axes in $\mathbb{A}^3$, and let W be the zero set of $XY(X-Y)$ in $\mathbb{A}^2$. Each of V and W is a union of three lines meeting at the origin. Are they isomorphic as algebraic varieties? Obviously, the origin o is the only singular point on V or W. An isomorphism $V \to W$ would have to send the singular point o to the singular point o and map $T_o(V)$ isomorphically onto $T_o(W)$. But $V = V(XY, YZ, XZ)$, and so $T_o(V)$ has dimension 3, whereas T_oW has dimension 2. Therefore, V and W are not isomorphic.

i Nonsingularity and regularity

THEOREM 4.44. *Let P be a point on an irreducible variety V. Every generating set for the maximal ideal $\mathfrak{n}_P$ of $\mathcal{O}_P$ has at least d elements, and there exists a generating set with d elements if and only if P is nonsingular.*

PROOF. If $f_1, \ldots, f_r$ generate $\mathfrak{n}_P$, then the proof of (4.36) shows that P is an irreducible component of $V(f_1, \ldots, f_r)$ in some open neighbourhood U of P. Therefore (3.45) shows that $0 \geq d - r$, and so $r \geq d$. The rest of the statement has already been noted. □

COROLLARY 4.45. *A point P on an irreducible variety is nonsingular if and only if $\mathcal{O}_P$ is regular.*

PROOF. This is a restatement of the second part of the theorem. □

According to CA 22.3, a regular local ring is an integral domain. If P lies on two irreducible components of a V, then $\mathcal{O}_P$ is not an integral domain (3.14), and so $\mathcal{O}_P$ is not regular. Therefore, the corollary holds also for reducible varieties.

j Examples of tangent spaces

The description of the tangent space in terms of dual numbers is particularly convenient when our variety is given to us in terms of its points functor. For example, let M_n be the set of $n \times n$ matrices, and let I be the identity matrix. Write e for I when it is to be regarded as the identity element of GL_n.

4.46. A matrix $I + \varepsilon A$ has inverse $I - \varepsilon A$ in $M_n(k[\varepsilon])$, and so lies in $\mathrm{GL}_n(k[\varepsilon])$. Therefore,

$$\begin{aligned} T_e(\mathrm{GL}_n) &= \{I + \varepsilon A \mid A \in M_n\} \\ &\simeq M_n(k). \end{aligned}$$

4.47. Since

$$\det(I + \varepsilon A) = I + \varepsilon \mathrm{trace}(A)$$

(using that $\varepsilon^2 = 0$),

$$\begin{aligned} T_e(\mathrm{SL}_n) &= \{I + \varepsilon A \mid \mathrm{trace}(A) = 0\} \\ &\simeq \{A \in M_n(k) \mid \mathrm{trace}(A) = 0\}. \end{aligned}$$

4.48. Assume that the characteristic $\neq 2$, and let O_n be orthogonal group:

$$\mathrm{O}_n = \{A \in \mathrm{GL}_n \mid A^{\mathrm{tr}} \cdot A = I\}.$$

(A^{tr} denotes the transpose of A). This is the group of matrices preserving the quadratic form $X_1^2 + \cdots + X_n^2$. The determinant defines a surjective regular homomorphism det: $O_n \to \{\pm 1\}$, whose kernel is defined to be the special orthogonal group SO_n. For $I + \varepsilon A \in M_n(k[\varepsilon])$,

$$(I + \varepsilon A)^{\mathrm{tr}} \cdot (I + \varepsilon A) = I + \varepsilon A^{\mathrm{tr}} + \varepsilon A,$$

and so

$$\begin{aligned} T_e(\mathrm{O}_n) = T_e(\mathrm{SO}_n) &= \{I + \varepsilon A \in M_n(k[\varepsilon]) \mid A \text{ is skew-symmetric}\} \\ &\simeq \{A \in M_n(k) \mid A \text{ is skew-symmetric}\}. \end{aligned}$$

ASIDE 4.49. On the tangent space $T_e(\mathrm{GL}_n) \simeq M_n$ of GL_n, there is a bracket operation

$$[M, N] \stackrel{\mathrm{def}}{=} MN - NM$$

which makes $T_e(\mathrm{GL}_n)$ into a Lie algebra. For any closed algebraic subgroup G of GL_n, $T_e(G)$ is stable under the bracket operation on $T_e(\mathrm{GL}_n)$ and is a sub-Lie-algebra of M_n, which we denote $\mathrm{Lie}(G)$. The Lie algebra structure on $\mathrm{Lie}(G)$ is independent of the embedding of G into GL_n (in fact, it has an intrinsic definition in terms of left invariant derivations), and $G \mapsto \mathrm{Lie}(G)$ is a functor from the category of linear group varieties to that of Lie algebras.

This functor is not fully faithful, for example, every étale homomorphism $G \to G'$ defines an isomorphism $\mathrm{Lie}(G) \to \mathrm{Lie}(G')$, but it is nevertheless very useful.

Assume that k has characteristic zero. A connected algebraic group G is said to be ***semisimple*** if it has no closed connected solvable normal subgroup (except $\{e\}$). Such a group G may have a finite nontrivial centre $Z(G)$, and we call two semisimple groups G and G' ***locally isomorphic*** if $G/Z(G) \approx G'/Z(G')$. For example, SL_n is semisimple, with centre μ_n, the set of diagonal matrices $\mathrm{diag}(\zeta, \ldots, \zeta)$, $\zeta^n = 1$, and $\mathrm{SL}_n / \mu_n = \mathrm{PSL}_n$. A Lie algebra is ***semisimple*** if it has no commutative ideal (except $\{0\}$). One can prove that

$$G \text{ is semisimple} \iff \mathrm{Lie}(G) \text{ is semisimple},$$

and the map $G \mapsto \mathrm{Lie}(G)$ defines a one-to-one correspondence between the set of local isomorphism classes of semisimple algebraic groups and the set of isomorphism classes of Lie algebras. The classification of semisimple algebraic groups can be deduced from that of semisimple Lie algebras and a study of the finite coverings of semisimple algebraic groups — this is quite similar to the relation between Lie groups and Lie algebras.

Exercises

4-1. Find the singular points, and the tangent cones at the singular points, for each of

(a) $Y^3 - Y^2 + X^3 - X^2 + 3Y^2X + 3X^2Y + 2XY$;

(b) $X^4 + Y^4 - X^2Y^2$ (assume that the characteristic is not 2).

4-2. Let $V \subset \mathbb{A}^n$ be an irreducible affine variety, and let P be a nonsingular point on V. Let H be a hyperplane in $\mathbb{A}^n$ (i.e., the subvariety defined by a linear equation $\sum a_i X_i = d$ with not all a_i zero) passing through P but not containing $T_P(V)$. Show that P is a nonsingular point on each irreducible component of $V \cap H$ on which it lies. (Each irreducible component has codimension 1 in V — you may assume this.) Give an example with $H \supset T_P(V)$ and P singular on $V \cap H$. Must P be singular on $V \cap H$ if $H \supset T_P(V)$?

4-3. Given a smooth point on a variety and a tangent vector at the point, show that there is a smooth curve passing through the point with the given vector as its tangent vector (see mo111467).

4-4. Let P and Q be points on varieties V and W. Show that

$$T_{(P,Q)}(V \times W) = T_P(V) \oplus T_Q(W).$$

4-5. For each n, show that there is a curve C and a point P on C such that the tangent space to C at P has dimension n (hence C can't be embedded in $\mathbb{A}^{n-1}$).

4-6. Let I be the $n \times n$ identity matrix, and let J be the matrix $\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. The ***symplectic group*** Sp_n is the group of $2n \times 2n$ matrices A with determinant 1 such that $A^{\mathrm{tr}} \cdot J \cdot A = J$. (It is the group of matrices fixing a nondegenerate skew-symmetric form.) Find the tangent space to Sp_n at its identity element, and also the dimension of Sp_n.

4-7. Find a regular map $\alpha\colon V \to W$ which induces an isomorphism on the geometric tangent cones $C_P(V) \to C_{\alpha(P)}(W)$ but is not étale at P.

4-8. Show that the cone $X^2 + Y^2 = Z^2$ is a normal variety, even though the origin is singular (characteristic $\neq 2$). See p.172.

4-9. Let $V = V(\mathfrak{a}) \subset \mathbb{A}^n$. Suppose that $\mathfrak{a} \neq I(V)$, and for $\mathbf{a} \in V$, let $T'_{\mathbf{a}}$ be the subspace of $T_{\mathbf{a}}(\mathbb{A}^n)$ defined by the equations $(df)_{\mathbf{a}} = 0$, $f \in \mathfrak{a}$. Clearly, $T'_{\mathbf{a}} \supset T_{\mathbf{a}}(V)$, but need they always be different?

4-10. Let W be a finite-dimensional k-vector space, and let $R_W = k \oplus W$ endowed with the k-algebra structure for which $W^2 = 0$. Let V be an affine algebraic variety over k. Show that the elements of $V(R_W) \stackrel{\text{def}}{=} \mathrm{Hom}_{k\text{-algebra}}(k[V], R_W)$ are in natural one-to-one correspondence with the pairs (P, t) with $P \in V$ and $t \in W \otimes T_P(V)$ (cf. Mumford, Lectures on curves ..., 1966, p25).

CHAPTER 5

Algebraic Varieties

An algebraic variety is a ringed space that is locally isomorphic to an affine algebraic variety, just as a topological manifold is a ringed space that is locally isomorphic to an open subset of $\mathbb{R}^n$; both are required to satisfy a separation axiom.

a Algebraic prevarieties

As motivation, recall the following definitions.

DEFINITION 5.1. (a) A ***topological manifold of dimension*** n is a ringed space $(V, \mathcal{O}_V)$ such that V is Hausdorff and every point of V has an open neighbourhood U for which $(U, \mathcal{O}_V|U)$ is isomorphic to the ringed space of continuous functions on an open subset of $\mathbb{R}^n$ (cf. 3.2)).

(b) A ***differentiable manifold of dimension*** n is a ringed space such that V is Hausdorff and every point of V has an open neighbourhood U for which $(U, \mathcal{O}_V|U)$ is isomorphic to the ringed space of smooth functions on an open subset of $\mathbb{R}^n$ (cf. 3.3).

(c) A ***complex manifold of dimension*** n is a ringed space such that V is Hausdorff and every point of V has an open neighbourhood U for which $(U, \mathcal{O}_V|U)$ is isomorphic to the ringed space holomorphic functions on an open subset of $\mathbb{C}^n$ (cf. 3.4).

These definitions are easily seen to be equivalent to the more classical definitions in terms of charts and atlases.[1] Often one imposes additional conditions on V, for example, that it be connected or that it have a countable base of open subsets.

DEFINITION 5.2. An ***algebraic prevariety over*** k is a k-ringed space $(V, \mathcal{O}_V)$ such that V is quasicompact and every point of V has an open neighbourhood U for which $(U, \mathcal{O}_V|U)$ is isomorphic to the ringed space of regular functions on an algebraic set over k.

Thus, a ringed space $(V, \mathcal{O}_V)$ is an algebraic prevariety over k if there exists a finite open covering $V = \bigcup V_i$ such that $(V_i, \mathcal{O}_V|V_i)$ is an affine algebraic variety over k for all i. An algebraic variety will be defined to be an algebraic prevariety satisfying a certain separation condition.

An open subset U of an algebraic prevariety V such that $(U, \mathcal{O}_V|U)$ is an affine algebraic variety is called an ***open affine (subvariety)*** in V. Because V is a finite union of open

[1] Provided the latter are stated correctly, which is frequently not the case.

affines, and in each open affine the open affines (in fact the basic open subsets) form a base for the topology, it follows that the open affines form a base for the topology on V.

Let $(V, \mathcal{O}_V)$ be an algebraic prevariety, and let U be an open subset of V. The functions $f: U \to k$ lying in $\Gamma(U, \mathcal{O}_V)$ are called ***regular***. Note that if (U_i) is an open covering of V by affine varieties, then $f: U \to k$ is regular if and only if $f|U_i \cap U$ is regular for all i (by 3.1(c)). Thus understanding the regular functions on open subsets of V amounts to understanding the regular functions on the open affine subvarieties and how these subvarieties fit together to form V.

EXAMPLE 5.3. (Projective space). Let $\mathbb{P}^n$ denote $k^{n+1} \smallsetminus \{\text{origin}\}$ modulo the equivalence relation

$$(a_0, \ldots, a_n) \sim (b_0, \ldots, b_n) \iff (a_0, \ldots, a_n) = (cb_0, \ldots, cb_n) \text{ some } c \in k^\times.$$

Thus the equivalence classes are the lines through the origin in k^{n+1} (with the origin omitted). Write $(a_0 : \ldots : a_n)$ for the equivalence class containing $(a_0, \ldots, a_n)$. For each i, let

$$U_i = \{(a_0 : \ldots : a_i : \ldots : a_n) \in \mathbb{P}^n \mid a_i \neq 0\}.$$

Then $\mathbb{P}^n = \bigcup U_i$, and the map

$$(a_0 : \ldots : a_n) \mapsto \left(\frac{a_0}{a_i}, \ldots, \widehat{\frac{a_i}{a_i}}, \ldots, \frac{a_n}{a_i}\right) : U_i \xrightarrow{u_i} \mathbb{A}^n$$

(the term a_i/a_i is omitted) is a bijection. In Chapter 6 we shall show that there is a unique structure of a (separated) algebraic variety on $\mathbb{P}^n$ for which each U_i is an open affine subvariety of $\mathbb{P}^n$ and each map u_i is an isomorphism of algebraic varieties.

b Regular maps

In each of the examples (5.1a,b,c), a morphism of manifolds (continuous map, smooth map, holomorphic map respectively) is just a morphism of ringed spaces. This motivates the following definition.

Let $(V, \mathcal{O}_V)$ and $(W, \mathcal{O}_W)$ be algebraic prevarieties. A map $\varphi: V \to W$ is said to be ***regular*** if it is a morphism of k-ringed spaces. In other words, a continuous map $\varphi: V \to W$ is regular if $f \mapsto f \circ \varphi$ sends a regular function on an open subset U of W to a regular function on $\varphi^{-1}(U)$. A composite of regular maps is again regular (this is a general fact about morphisms of ringed spaces).

Note that we have three categories:

$$(\text{affine varieties}) \subset (\text{algebraic prevarieties}) \subset (\text{ringed spaces}).$$

Each subcategory is full, i.e., the morphisms $\mathrm{Mor}(V, W)$ are the same in the three categories.

PROPOSITION 5.4. *Let $(V, \mathcal{O}_V)$ and $(W, \mathcal{O}_W)$ be prevarieties, and let $\varphi: V \to W$ be a continuous map (of topological spaces). Let $W = \bigcup W_j$ be a covering of W by open affines, and let $\varphi^{-1}(W_j) = \bigcup V_{ji}$ be a covering of $\varphi^{-1}(W_j)$ by open affines. Then φ is regular if and only if its restrictions*

$$\varphi | V_{ji} : V_{ji} \to W_j$$

are regular for all i, j.

PROOF. We assume that φ satisfies this condition, and prove that it is regular. Let f be a regular function on an open subset U of W. Then $f|U \cap W_j$ is regular for each W_j (sheaf condition 3.1(b)), and so $f \circ \varphi|\varphi^{-1}(U) \cap V_{ji}$ is regular for each j, i (this is our assumption). It follows that $f \circ \varphi$ is regular on $\varphi^{-1}(U)$ (sheaf condition 3.1(c)). Thus φ is regular. The converse is even easier. □

ASIDE 5.5. A differentiable manifold of dimension n is locally isomorphic to an open subset of $\mathbb{R}^n$. In particular, all manifolds of the same dimension are locally isomorphic. This is not true for algebraic varieties, for two reasons:

(a) We are not assuming our varieties are nonsingular (see Chapter 5 below).

(b) The inverse function theorem fails in our context: a regular map that induces an isomorphism on the tangent space at a point P need not induce an isomorphism in a neighbourhood of P. However, see (5.55) below.

c Algebraic varieties

In the study of topological manifolds, the Hausdorff condition eliminates such bizarre possibilities as the line with the origin doubled where a sequence tending to the origin has two limits (see 5.10 below).

It is not immediately obvious how to impose a separation axiom on our algebraic varieties, because even affine algebraic varieties are not Hausdorff. The key is to restate the Hausdorff condition. Intuitively, the significance of this condition is that it prevents a sequence in the space having more than one limit. Thus a continuous map into the space should be determined by its values on a dense subset, i.e., if φ_1 and φ_2 are continuous maps $Z \to U$ that agree on a dense subset of Z then they should agree on the whole of Z. Equivalently, the set where two continuous maps $\varphi_1, \varphi_2 \colon Z \rightrightarrows U$ agree should be closed. Surprisingly, affine varieties have this property, provided φ_1 and φ_2 are required to be regular maps.

LEMMA 5.6. *Let $\varphi_1, \varphi_2 \colon Z \rightrightarrows V$ regular maps of affine algebraic varieties. The subset of Z on which φ_1 and φ_2 agree is closed.*

PROOF. There are regular functions x_i on V such that $P \mapsto (x_1(P), \ldots, x_n(P))$ identifies V with a closed subset of $\mathbb{A}^n$ (take the x_i to be any set of generators for $k[V]$ as a k-algebra). Now $x_i \circ \varphi_1$ and $x_i \circ \varphi_2$ are regular functions on Z, and the set where φ_1 and φ_2 agree is $\bigcap_{i=1}^{n} V(x_i \circ \varphi_1 - x_i \circ \varphi_2)$, which is closed. □

DEFINITION 5.7. An algebraic prevariety V is said to be ***separated*** if it satisfies the following additional condition:

> Separation axiom: for every pair of regular maps $\varphi_1, \varphi_2 \colon Z \rightrightarrows V$ with Z an affine algebraic variety, the set $\{z \in Z \mid \varphi_1(z) = \varphi_2(z)\}$ is closed in Z.

An ***algebraic variety*** over k is a separated algebraic prevariety over k.[2]

PROPOSITION 5.8. *Let φ_1 and φ_2 be regular maps $Z \rightrightarrows V$ from an algebraic prevariety Z to a separated prevariety V. The subset of Z on which φ_1 and φ_2 agree is closed.*

[2]These are sometimes called "algebraic varieties in the sense of FAC" (Serre, Jean-Pierre. Faisceaux algébriques cohérents. Ann. of Math. (2) 61, (1955). 197–278; §34). In Grothendieck's language, they are separated and reduced schemes of finite type over k (assumed to be algebraically closed), except that we omit the nonclosed points; cf. EGA IV, 10.10. Some authors use a more restrictive definition — they may require a variety to be connected, irreducible, or quasi-projective — usually because their foundations do not allow for a more flexible definition.

PROOF. Let W be the set on which φ_1 and φ_2 agree. For any open affine U of Z, $W \cap U$ is the subset of U on which $\varphi_1|U$ and $\varphi_2|U$ agree, and so $W \cap U$ is closed. This implies that W is closed because Z is a finite union of open affines. □

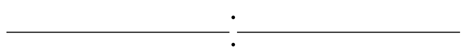
EXAMPLE 5.9. The open subspace $U = \mathbb{A}^2 \smallsetminus \{(0,0)\}$ of $\mathbb{A}^2$ becomes an algebraic variety when endowed with the sheaf $\mathcal{O}_{\mathbb{A}^2}|U$ (cf. 3.33).

A subvariety of an affine variety is said to be ***quasi-affine***. For example, $\mathbb{A}^2 \smallsetminus \{(0,0)\}$ is quasi-affine but not affine.

EXAMPLE 5.10. (The affine line with the origin doubled.)[3] Let V_1 and V_2 be copies of $\mathbb{A}^1$. Let $V^* = V_1 \sqcup V_2$ (disjoint union), and give it the obvious topology. Define an equivalence relation on V^* by

$$x \text{ (in } V_1) \sim y \text{ (in } V_2) \iff x = y \text{ and } x \neq 0.$$

Let V be the quotient space $V = V^*/{\sim}$ with the quotient topology (a set is open if and only if its inverse image in V^* is open):



Then V_1 and V_2 are open subspaces of V, $V = V_1 \cup V_2$, and $V_1 \cap V_2 = \mathbb{A}^1 - \{0\}$. Define a function on an open subset to be regular if its restriction to each V_i is regular. This makes V into a prevariety, but not a variety: it fails the separation axiom because the two maps

$$\mathbb{A}^1 = V_1 \hookrightarrow V^*, \quad \mathbb{A}^1 = V_2 \hookrightarrow V^*$$

agree exactly on $\mathbb{A}^1 - \{0\}$, which is not closed in $\mathbb{A}^1$.

Let Var_k denote the category of algebraic varieties over k and regular maps. The functor $A \rightsquigarrow \mathrm{Spm}(A)$ is a fully faithful contravariant functor $\mathsf{Aff}_k \to \mathsf{Var}_k$, and defines an equivalence of the first category with the subcategory of the second whose objects are the affine algebraic varieties.

5.11. When V is irreducible, all the rings attached to it can be identified with subrings of the field $k(V)$. For example,

$$\begin{aligned} \mathcal{O}_P &= \{g/h \in k(V) \mid h(P) \neq 0\} \\ \mathcal{O}_V(U) &= \bigcap \{\mathcal{O}_V(U') \mid U' \subset U,\ U' \text{ open affine}\} \\ &= \bigcap \{\mathcal{O}_P \mid P \in U\}. \end{aligned}$$

[3] This is the algebraic analogue of the standard example of a non Hausdorff topological space. Let $\mathbb{R}^*$ denote the real line with the origin removed but with two points $a \neq b$ added. The topology is generated by the open intervals in $\mathbb{R}$ together with the sets of the form $(u,0) \cup \{a\} \cup (0,v)$ and $(u,0) \cup \{b\} \cup (0,v)$, where $u < 0 < v$. Then X is not Hausdorff because a and b cannot be separated by disjoint open sets. Every sequence that converges to a also converges to b; for example, $1/n$ converges to both a and b.

d Maps from varieties to affine varieties

Let $(V, \mathcal{O}_V)$ be an algebraic variety, and let $\alpha: A \to \Gamma(V, \mathcal{O}_V)$ be a homomorphism from an affine k-algebra A to the k-algebra of regular functions on V. For any $P \in V$, $f \mapsto \alpha(f)(P)$ is a k-algebra homomorphism $A \to k$, and so its kernel $\varphi(P)$ is a maximal ideal in A. In this way, we get a map

$$\varphi: V \to \mathrm{spm}(A)$$

which is easily seen to be regular. Conversely, from a regular map $\varphi: V \to \mathrm{Spm}(A)$, we get a k-algebra homomorphism $f \mapsto f \circ \varphi: A \to \Gamma(V, \mathcal{O}_V)$. Since these maps are inverse, we have proved the following result.

PROPOSITION 5.12. *For an algebraic variety V and an affine k-algebra A, there is a canonical one-to-one correspondence*

$$\mathrm{Mor}(V, \mathrm{Spm}(A)) \simeq \mathrm{Hom}_{k\text{-algebra}}(A, \Gamma(V, \mathcal{O}_V)).$$

Let V be an algebraic variety such that $\Gamma(V, \mathcal{O}_V)$ is an affine k-algebra. The proposition shows that the regular map $\varphi: V \to \mathrm{Spm}(\Gamma(V, \mathcal{O}_V))$ defined by $\mathrm{id}_{\Gamma(V, \mathcal{O}_V)}$ has the following universal property: every regular map from V to an affine algebraic variety U factors uniquely through φ:

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[dr] & \mathrm{Spm}(\Gamma(V, \mathcal{O}_V)) \arrow[d, dotted, "\exists!"] \\ & U. \end{tikzcd}$$

ASIDE 5.13. For a nonaffine algebraic variety V, $\Gamma(V, \mathcal{O}_V)$ need not be finitely generated as a k-algebra.

e Subvarieties

Let $(V, \mathcal{O}_V)$ be an algebraic variety over k.

Open subvarieties

Let U be an open subset of V. Then U is a union of open affines, and it follows that $(U, \mathcal{O}_V|U)$ is a variety, called an ***open subvariety*** of V. A regular map $\varphi: W \to V$ is an ***open immersion*** if $\varphi(W)$ is open in V and φ defines an isomorphism $W \to \varphi(W)$ of varieties.

Closed subvarieties

Let Z be an closed subset of V. A function f on an open subset U of Z is regular if, for every $P \in U$, there exists a germ (U', f') of a regular function at P on V such that $f'|U' \cap Z = f|U' \cap U$. This defines a ringed structure $\mathcal{O}_Z$ on Z. To show that $(Z, \mathcal{O}_Z)$ is a variety it suffices to check that, for every open affine $U \subset V$, the ringed space $(U \cap Z, \mathcal{O}_Z|U \cap Z)$ is affine algebraic variety, but this is only an exercise (Exercise 3-2 to be precise). Such a pair $(Z, \mathcal{O}_Z)$ is called a ***closed subvariety*** of V. A regular map $\varphi: W \to V$ is a ***closed immersion*** if $\varphi(W)$ is closed in V and φ defines an isomorphism $W \to \varphi(W)$ of varieties.

Subvarieties

A subset W of a topological space V is said to be ***locally closed*** if every point P in W has an open neighbourhood U in V such that $W \cap U$ is closed in U. Equivalent conditions: W is the intersection of an open and a closed subset of V; W is open in its closure. A locally closed subset W of a variety V acquires a natural structure as a variety: write it as the intersection $W = U \cap Z$ of an open and a closed subset; Z is a variety, and W (being open in Z) therefore acquires the structure of a variety. This structure on W has the following characterization: the inclusion map $W \hookrightarrow V$ is regular, and a map $\varphi: V' \to W$ with V' a variety is regular if and only if it is regular when regarded as a map into V. With this structure, W is called a ***subvariety*** of V. A regular map $\varphi: W \to V$ is an ***immersion*** if it induces an isomorphism of W onto a subvariety of V. Every immersion is the composite of an open immersion with a closed immersion (in both orders).

Application

PROPOSITION 5.14. *A prevariety V is separated if and only if two regular maps from a prevariety to V agree on the whole prevariety whenever they agree on a dense subset of it.*

PROOF. If V is separated, then the set on which a pair of regular maps $\varphi_1, \varphi_2: Z \rightrightarrows V$ agree is closed, and so must be the whole of the Z.

Conversely, consider a pair of maps $\varphi_1, \varphi_2: Z \rightrightarrows V$, and let S be the subset of Z on which they agree. We assume that V has the property in the statement of the proposition, and show that S is closed. Let $\bar{S}$ be the closure of S in Z. According to the above discussion, $\bar{S}$ has the structure of a closed prevariety of Z and the maps $\varphi_1|\bar{S}$ and $\varphi_2|\bar{S}$ are regular. Because they agree on a dense subset of $\bar{S}$ they agree on the whole of $\bar{S}$, and so $S = \bar{S}$ is closed. □

f Prevarieties obtained by patching

PROPOSITION 5.15. *Suppose that the set V is a finite union $V = \bigcup_{i \in I} V_i$ of subsets V_i and that each V_i is equipped with ringed space structure. Assume that the following "patching" condition holds:*

for all i, j, $V_i \cap V_j$ is open in both V_i and V_j and $\mathcal{O}_{V_i}|V_i \cap V_j = \mathcal{O}_{V_j}|V_i \cap V_j$.

Then there is a unique structure of a ringed space on V for which

(a) *each inclusion $V_i \hookrightarrow V$ is a homeomorphism of V_i onto an open set, and*

(b) *for each $i \in I$, $\mathcal{O}_V|V_i = \mathcal{O}_{V_i}$.*

If every V_i is an algebraic prevariety, then so also is V, and to give a regular map from V to a prevariety W amounts to giving a family of regular maps $\varphi_i: V_i \to W$ such that $\varphi_i|V_i \cap V_j = \varphi_j|V_i \cap V_j$.

PROOF. One checks easily that the subsets $U \subset V$ such that $U \cap V_i$ is open for all i are the open subsets for a topology on V satisfying (a), and that this is the only topology to satisfy (a). Define $\mathcal{O}_V(U)$ to be the set of functions $f: U \to k$ such that $f|U \cap V_i \in \mathcal{O}_{V_i}(U \cap V_i)$ for all i. Again, one checks easily that $\mathcal{O}_V$ is a sheaf of k-algebras satisfying (b), and that it is the only such sheaf.

For the final statement, if each $(V_i, \mathcal{O}_{V_i})$ is a finite union of open affines, so also is $(V, \mathcal{O}_V)$. Moreover, to give a map $\varphi: V \to W$ amounts to giving a family of maps

$\varphi_i: V_i \to W$ such that $\varphi_i|V_i \cap V_j = \varphi_j|V_i \cap V_j$ (obviously), and φ is regular if and only $\varphi|V_i$ is regular for each i. □

Clearly, the V_i may be separated without V being separated (see, for example, 5.10). In (5.29) below, we give a condition on an open affine covering of a prevariety sufficient to ensure that the prevariety is separated.

g Products of varieties

Let V and W be objects in a category C. A triple

$$(V \times W, \quad p: V \times W \to V, \quad q: V \times W \to W)$$

is said to be the ***product*** of V and W if it has the following universal property: for every pair of morphisms $Z \to V$, $Z \to W$ in C, there exists a unique morphism $Z \to V \times W$ making the diagram

$$\begin{tikzcd} & Z \arrow[dl] \arrow[d, dotted, "\exists!"] \arrow[dr] & \\ V & V \times W \arrow[l, "p"'] \arrow[r, "q"] & W \end{tikzcd}$$

commute. In other words, the triple is a product if the map

$$\varphi \mapsto (p \circ \varphi, q \circ \varphi): \mathrm{Hom}(Z, V \times W) \to \mathrm{Hom}(Z, V) \times \mathrm{Hom}(Z, W)$$

is a bijection. The product, if it exists, is uniquely determined up to a unique isomorphism by its universal property.

For example, the product of two sets (in the category of sets) is the usual cartesian product of the sets, and the product of two topological spaces (in the category of topological spaces) is the product of the underlying sets endowed with the product topology.

We shall show that products exist in the category of algebraic varieties. Suppose, for the moment, that $V \times W$ exists. For any prevariety Z, $\mathrm{Mor}(\mathbb{A}^0, Z)$ is the underlying set of Z; more precisely, for any $z \in Z$, the map $\mathbb{A}^0 \to Z$ with image z is regular, and these are all the regular maps (cf. 3.28). Thus, from the definition of products we have

$$\begin{aligned}(\text{underlying set of } V \times W) &\simeq \mathrm{Mor}(\mathbb{A}^0, V \times W)\\ &\simeq \mathrm{Mor}(\mathbb{A}^0, V) \times \mathrm{Mor}(\mathbb{A}^0, W)\\ &\simeq (\text{underlying set of } V) \times (\text{underlying set of } W).\end{aligned}$$

Hence, our problem can be restated as follows: given two prevarieties V and W, define on the set $V \times W$ the structure of a prevariety such that

(a) the projection maps $p, q: V \times W \rightrightarrows V, W$ are regular, and

(b) a map $\varphi: T \to V \times W$ of sets (with T an algebraic prevariety) is regular if its components $p \circ \varphi, q \circ \varphi$ are regular.

There can be at most one such structure on the set $V \times W$.

Products of affine varieties

EXAMPLE 5.16. Let $\mathfrak{a}$ and $\mathfrak{b}$ be ideals in $k[X_1,\ldots,X_m]$ and $k[X_{m+1},\ldots,X_{m+n}]$ respectively, and let $(\mathfrak{a},\mathfrak{b})$ be the ideal in $k[X_1,\ldots,X_{m+n}]$ generated by the elements of $\mathfrak{a}$ and $\mathfrak{b}$. Then there is an isomorphism

$$f\otimes g\mapsto fg\colon \frac{k[X_1,\ldots,X_m]}{\mathfrak{a}}\otimes_k\frac{k[X_{m+1},\ldots,X_{m+n}]}{\mathfrak{b}}\to\frac{k[X_1,\ldots,X_{m+n}]}{(\mathfrak{a},\mathfrak{b})}.$$

Again this comes down to checking that the natural map from

$$\mathrm{Hom}_{k\text{-alg}}(k[X_1,\ldots,X_{m+n}]/(\mathfrak{a},\mathfrak{b}),R)$$

to

$$\mathrm{Hom}_{k\text{-alg}}(k[X_1,\ldots,X_m]/\mathfrak{a},R)\times\mathrm{Hom}_{k\text{-alg}}(k[X_{m+1},\ldots,X_{m+n}]/\mathfrak{b},R)$$

is a bijection. But the three sets are respectively
$V(\mathfrak{a},\mathfrak{b})=$ zero set of $(\mathfrak{a},\mathfrak{b})$ in R^{m+n},
$V(\mathfrak{a})=$ zero set of $\mathfrak{a}$ in R^m,
$V(\mathfrak{b})=$ zero set of $\mathfrak{b}$ in R^n,
and so this is obvious.

The tensor product of two k-algebras A and B has the universal property to be a product in the category of k-algebras, but with the arrows reversed. Because of the category anti-equivalence (3.25), this shows that $\mathrm{Spm}(A\otimes_k B)$ will be the product of $\mathrm{Spm}\,A$ and $\mathrm{Spm}\,B$ in the category of affine algebraic varieties once we have shown that $A\otimes_k B$ is an affine k-algebra.

PROPOSITION 5.17. *Let A and B be k-algebras with A finitely generated.*

(a) *If A and B are reduced, then so also is $A\otimes_k B$.*

(b) *If A and B are integral domains, then so also is $A\otimes_k B$.*

PROOF. Let $\alpha\in A\otimes_k B$. Then $\alpha=\sum_{i=1}^n a_i\otimes b_i$, some $a_i\in A$, $b_i\in B$. If one of the b_i is a linear combination of the remaining b , say, $b_n=\sum_{i=1}^{n-1}c_ib_i$, $c_i\in k$, then, using the bilinearity of $\otimes$, we find that

$$\alpha=\sum_{i=1}^{n-1}a_i\otimes b_i+\sum_{i=1}^{n-1}c_ia_n\otimes b_i=\sum_{i=1}^{n-1}(a_i+c_ia_n)\otimes b_i.$$

Thus we can suppose that in the original expression of α, the b_i are linearly independent over k.

Now assume A and B to be reduced, and suppose that α is nilpotent. Let $\mathfrak{m}$ be a maximal ideal of A. From $a\mapsto\bar{a}\colon A\to A/\mathfrak{m}=k$ we obtain homomorphisms

$$a\otimes b\mapsto\bar{a}\otimes b\mapsto\bar{a}b\colon A\otimes_k B\to k\otimes_k B\xrightarrow{\simeq}B$$

The image $\sum\bar{a}_ib_i$ of α under this homomorphism is a nilpotent element of B, and hence is zero (because B is reduced). As the b_i are linearly independent over k, this means that the $\bar{a}_i$ are all zero. Thus, the a_i lie in all maximal ideals $\mathfrak{m}$ of A, and so are zero (see 2.18). Hence $\alpha=0$, and we have shown that $A\otimes_k B$ is reduced.

Now assume that A and B are integral domains, and let $\alpha,\alpha'\in A\otimes_k B$ be such that $\alpha\alpha'=0$. As before, we can write $\alpha=\sum a_i\otimes b_i$ and $\alpha'=\sum a_i'\otimes b_i'$ with the sets $\{b_1,b_2,\ldots\}$

and $\{b_1', b_2', \ldots\}$ each linearly independent over k. For each maximal ideal $\mathfrak{m}$ of A, we know $(\sum \bar{a}_i b_i)(\sum \bar{a}_i' b_i') = 0$ in B, and so either $(\sum \bar{a}_i b_i) = 0$ or $(\sum \bar{a}_i' b_i') = 0$. Thus either all the $a_i \in \mathfrak{m}$ or all the $a_i' \in \mathfrak{m}$.[4] This shows that

$$\mathrm{spm}(A) = V(a_1, \ldots, a_m) \cup V(a_1', \ldots, a_n').$$

As $\mathrm{spm}(A)$ is irreducible (see 2.27), it follows that $\mathrm{spm}(A)$ equals either $V(a_1, \ldots, a_m)$ or $V(a_1', \ldots, a_n')$. In the first case $\alpha = 0$, and in the second $\alpha' = 0$. □

EXAMPLE 5.18. The proof of (5.17) fails when k is not algebraically closed, because then $A/\mathfrak{m}$ may be a finite extension of k over which the b_i become linearly dependent. The following examples show that the statement of (5.17) also fails in this case.

(a) Suppose that k is nonperfect of characteristic p, so that there exists an element α in an algebraic closure of k such that $\alpha \notin k$ but $\alpha^p \in k$. Let $k' = k[\alpha]$, and let $\alpha^p = a$. Then $(\alpha \otimes 1 - 1 \otimes \alpha) \neq 0$ in $k' \otimes_k k'$ (in fact, the elements $\alpha^i \otimes \alpha^j$, $0 \le i, j \le p-1$, form a basis for $k' \otimes_k k'$ as a k-vector space), but

$$\begin{aligned}(\alpha \otimes 1 - 1 \otimes \alpha)^p &= (a \otimes 1 - 1 \otimes a) \\ &= (1 \otimes a - 1 \otimes a) \quad \text{(because } a \in k\text{)} \\ &= 0.\end{aligned}$$

Thus $k' \otimes_k k'$ is not reduced, even though k' is a field.

(b) Let K be a finite separable extension of k and let Ω be a second field containing k. By the primitive element theorem (FT 5.1),

$$K = k[\alpha] = k[X]/(f(X)),$$

for some $\alpha \in K$ and its minimum polynomial $f(X)$. Assume that Ω is large enough to split f, say, $f(X) = \prod_i (X - \alpha_i)$ with $\alpha_i \in \Omega$. Because K/k is separable, the α_i are distinct, and so

$$\begin{aligned}\Omega \otimes_k K &\simeq \Omega[X]/(f(X)) && \text{(1.58(b))} \\ &\simeq \prod \Omega[X]/(X - \alpha_i), && \text{(1.1)}\end{aligned}$$

which is not an integral domain. For example,

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \simeq \mathbb{C}[X]/(X - i) \times \mathbb{C}[X]/(X + i) \simeq \mathbb{C} \times \mathbb{C}.$$

The proposition allows us to make the following definition.

DEFINITION 5.19. The ***product*** of the affine varieties V and W is

$$(V \times W, \mathcal{O}_{V \times W}) = \mathrm{Spm}(k[V] \otimes_k k[W])$$

with the projection maps $p, q \colon V \times W \to V, W$ defined by the homomorphisms

$$\begin{aligned}f &\mapsto f \otimes 1 \colon k[V] \to k[V] \otimes_k k[W] \\ g &\mapsto 1 \otimes g \colon k[W] \to k[V] \otimes_k k[W].\end{aligned}$$

[4] Here is where we use that k is algebraically closed. We know that $\{b_1, b_2, \ldots\}$ and $\{b_1', b_2', \ldots\}$ are linearly independent over k (emphasis on k). The elements $\bar{a}_i$ and $\bar{a}_i'$ live in $A/\mathfrak{m}$, which is a priori *only* an algebraic field extension of k. It is possible for elements of a k-algebra to be linearly independent over k but not over some extension (consider $1, i \in$ over $\mathbb{R}$). The fact that k is algebraically closed forces $A/\mathfrak{m} = k$, so we can apply the linear independence condition. (From sx599391.)

PROPOSITION 5.20. *Let V and W be affine varieties.*

(a) *The variety $(V \times W, \mathcal{O}_{V\times W})$ is the product of $(V, \mathcal{O}_V)$ and $(W, \mathcal{O}_W)$ in the category of affine algebraic varieties; in particular, the set $V \times W$ is the product of the sets V and W and p and q are the projection maps.*

(b) *If V and W are irreducible, then so also is $V \times W$.*

PROOF. (a) As noted at the start of the subsection, the first statement follows from (5.17a), and the second statement then follows by the argument on p.103.

(b) This follows from (5.17b) and (2.27). □

COROLLARY 5.21. *Let V and W be affine varieties. For every prevariety T, a map $\varphi: T \to V \times W$ is regular if $p \circ \varphi$ and $q \circ \varphi$ are regular.*

PROOF. If $p \circ \varphi$ and $q \circ \varphi$ are regular, then (5.20) implies that φ is regular when restricted to any open affine of T, which implies that it is regular on T. □

The corollary shows that $V \times W$ is the product of V and W in the category of prevarieties (hence also in the categories of varieties).

EXAMPLE 5.22. (a) It follows from (1.57) that $\mathbb{A}^{m+n}$ endowed with the projection maps

$$\mathbb{A}^m \xleftarrow{p} \mathbb{A}^{m+n} \xrightarrow{q} \mathbb{A}^n, \qquad \begin{cases} p(a_1,\ldots,a_{m+n}) = (a_1,\ldots,a_m) \\ q(a_1,\ldots,a_{m+n}) = (a_{m+1},\ldots,a_{m+n}), \end{cases}$$

is the product of $\mathbb{A}^m$ and $\mathbb{A}^n$.

(b) It follows from (5.16) that

$$V(\mathfrak{a}) \xleftarrow{p} V(\mathfrak{a},\mathfrak{b}) \xrightarrow{q} V(\mathfrak{b})$$

is the product of $V(\mathfrak{a})$ and $V(\mathfrak{b})$.

. When V and W have dimension > 0, the topology on $V \times W$ is strictly finer than product topology . For example, for the product topology on $\mathbb{A}^2 = \mathbb{A}^1 \times \mathbb{A}^1$, every proper closed subset is contained in a finite union of vertical and horizontal lines, whereas $\mathbb{A}^2$ has many more closed subsets (see 2.68).

Products in general

We now define the product of two algebraic prevarieties V and W.

Write V as a union of open affines $V = \bigcup V_i$, and note that V can be regarded as the variety obtained by patching the $(V_i, \mathcal{O}_{V_i})$; in particular, this covering satisfies the patching condition (5.15). Similarly, write W as a union of open affines $W = \bigcup W_j$. Then

$$V \times W = \bigcup V_i \times W_j$$

and the $(V_i \times W_j, \mathcal{O}_{V_i \times W_j})$ satisfy the patching condition. Therefore, we can define $(V \times W, \mathcal{O}_{V\times W})$ to be the variety obtained by patching the $(V_i \times W_j, \mathcal{O}_{V_i\times W_j})$.

PROPOSITION 5.23. *With the sheaf of k-algebras $\mathcal{O}_{V\times W}$ just defined, $V \times W$ becomes the product of V and W in the category of prevarieties. In particular, the structure of prevariety on $V \times W$ defined by the coverings $V = \bigcup V_i$ and $W = \bigcup W_j$ are independent of the coverings.*

PROOF. Let T be a prevariety, and let $\varphi: T \to V \times W$ be a map of sets such that $p \circ \varphi$ and $q \circ \varphi$ are regular. Then (5.21) implies that the restriction of φ to $\varphi^{-1}(V_i \times W_j)$ is regular. As these open sets cover T, this shows that φ is regular. □

PROPOSITION 5.24. *If V and W are separated, then so also is $V \times W$.*

PROOF. Let φ_1, φ_2 be two regular maps $U \to V \times W$. The set where φ_1, φ_2 agree is the intersection of the sets where $p \circ \varphi_1, p \circ \varphi_2$ and $q \circ \varphi_1, q \circ \varphi_2$ agree, which is closed. □

PROPOSITION 5.25. *If V and W are connected, then so also is $V \times W$.*

PROOF. For $v_0 \in V$, we have continuous maps

$$W \simeq v_0 \times W \overset{\text{closed}}{\hookrightarrow} V \times W.$$

Similarly, for $w_0 \in W$, we have continuous maps

$$V \simeq V \times w_0 \overset{\text{closed}}{\hookrightarrow} V \times W.$$

The images of V and W in $V \times W$ intersect in (v_0, w_0) and are connected, which shows that (v_0, w) and and (v, w_0) lie in the same connected component of $V \times W$ for all $v \in V$ and $w \in W$. Since v_0 and w_0 were arbitrary, this shows that any two points lie in the same connected component. □

Group varieties

A ***group variety*** is an algebraic variety G together with a group structure defined by regular maps

$$m: G \times G \to G, \quad \text{inv}: G \to G, \quad e\colon \mathbb{A}^0 \to G.$$

For example,

$$\begin{cases} \mathrm{SL}_n &= \mathrm{Spm}(k[X_{11}, X_{12}, \ldots, X_{nn}]/(\det(X_{ij}) - 1)) \\ \mathrm{SL}_n(k) &= \{M \in M_n(k) \mid \det M = 1\} \end{cases}$$

becomes a group variety when endowed with its usual group structures. Matrix multiplication

$$(a_{ij}) \cdot (b_{ij}) = (c_{ij}), \quad c_{ij} = \textstyle\sum_{l=1}^{n} a_{il} b_{lj},$$

is given by polynomials, and Cramer's rule gives an explicit expression of the entries of A^{-1} as polynomials in the entries of A. The only affine group varieties of dimension 1 are

$$\mathbb{G}_m = \mathrm{Spm}\, k[X, X^{-1}]$$

and

$$\mathbb{G}_a = \mathrm{Spm}\, k[X].$$

Every finite group N can be made into a group variety by setting

$$N = \mathrm{Spm}(A)$$

with A the k-algebra of all maps $f: N \to k$.

h The separation axiom revisited

By way of motivation, consider a topological space V and the diagonal $\Delta \subset V \times V$, $\Delta \stackrel{\text{def}}{=} \{(x,x) \mid x \in V\}$. If Δ is closed (for the product topology), then every pair of points $(x,y) \notin \Delta$ has an open neighbourhood $U \times U'$ such that $U \times U' \cap \Delta = \varnothing$. In other words, if x and y are distinct points in V, then there are open neighbourhoods U and U' of x and y respectively such that $U \cap U' = \varnothing$. Thus V is Hausdorff. Conversely, if V is Hausdorff, the reverse argument shows that Δ is closed.

For a variety V, we let $\Delta = \Delta_V$ (the diagonal) be the subset $\{(v,v) \mid v \in V\}$ of $V \times V$.

PROPOSITION 5.26. *An algebraic prevariety V is separated if and only if Δ_V is closed.*[5]

PROOF. We shall use the criterion (5.8): V is separated if and only if, for every pair of regular maps $\varphi_1, \varphi_2 \colon Z \rightrightarrows V$, the subset of Z on which φ_1 and φ_2 agree is closed.

Assume that Δ_V is closed. The map

$$(\varphi_1,\varphi_2)\colon Z \to V \times V, \quad z \mapsto (\varphi_1(z), \varphi_2(z))$$

is regular because its components φ_1 and φ_2 are regular (see p.103). In particular, it is continuous, and so $(\varphi_1,\varphi_2)^{-1}(\Delta_V)$ is closed, but this is exactly the subset on which φ_1 and φ_2 agree.

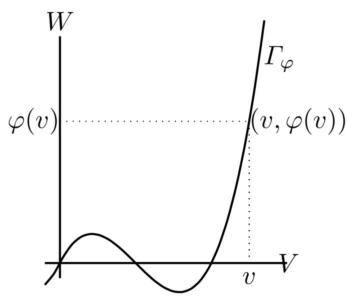
Conversely, Δ_V is the set on which the two projection maps $V \times V \to V$ agree, and so it is closed if V is separated. □

COROLLARY 5.27. *For any prevariety V, the diagonal is a locally closed subset of $V \times V$.*

PROOF. Let $P \in V$, and let U be an open affine neighbourhood of P. Then $U \times U$ is an open neighbourhood of (P,P) in $V \times V$, and $\Delta_V \cap (U \times U) = \Delta_U$, which is closed in $U \times U$ because U is separated (5.6). □

Thus Δ_V is always a subvariety of $V \times V$, and it is closed if and only if V is separated. The ***graph*** Γ_φ of a regular map $\varphi\colon V \to W$ is defined to be

$$\{(v,\varphi(v)) \in V \times W \mid v \in V\}.$$



COROLLARY 5.28. *For any morphism $\varphi\colon V \to W$ of prevarieties, the graph Γ_φ of φ is locally closed in $V \times W$, and it is closed if W is separated. The map $v \mapsto (v,\varphi(v))$ is an isomorphism of V onto Γ_φ (as algebraic prevarieties).*

PROOF. The map

$$(v,w) \mapsto (\varphi(v), w)\colon V \times W \to W \times W$$

is regular because its composites with the projections are φ and id_W which are regular. In particular, it is continuous, and as Γ_φ is the inverse image of Δ_W under this map, this proves the first statement. The second statement follows from the fact that the regular map $\Gamma_\varphi \hookrightarrow V \times W \xrightarrow{p} V$ is an inverse to $v \mapsto (v,\varphi(v))\colon V \to \Gamma_\varphi$. □

[5] Recall that the topology on $V \times V$ is *not* the product topology. Thus the statement does not contradict the fact that V is not Hausdorff.

THEOREM 5.29. *The following three conditions on a prevariety V are equivalent:*

(a) *V is separated;*

(b) *for every pair of open affines U and U' in V, $U \cap U'$ is an open affine, and the map*

$$f \otimes g \mapsto f|_{U\cap U'} \cdot g|_{U\cap U'}: k[U] \otimes_k k[U'] \to k[U \cap U']$$

is surjective;

(c) *the condition in (b) holds for the sets in some open affine covering of V.*

PROOF. Let U and U' be open affines in V. We shall prove that

(i) if Δ is closed then $U \cap U'$ affine,

(ii) when $U \cap U'$ is affine,

$$(U \times U') \cap \Delta \text{ is closed} \iff k[U] \otimes_k k[U'] \to k[U \cap U'] \text{ is surjective.}$$

Assume (a); then these statements imply (b). Assume that (b) holds for the sets in an open affine covering $(U_i)_{i \in I}$ of V. Then $(U_i \times U_j)_{(i,j)\in I\times I}$ is an open affine covering of $V \times V$, and $\Delta_V \cap (U_i \times U_j)$ is closed in $U_i \times U_j$ for each pair (i, j), which implies (a). Thus, the statements (i) and (ii) imply the theorem.

Proof of (i): The graph of the inclusion $U \cap U' \hookrightarrow V$ is the subset $(U \times U') \cap \Delta$ of $(U \cap U') \times V$. If Δ_V is closed, then $(U \times U') \cap \Delta_V$ is a closed subvariety of an affine variety, and hence is affine. Now (5.28) implies that $U \cap U'$ is affine.

Proof of (ii): Assume that $U \cap U'$ is affine. Then

$$\begin{aligned} &(U \times U') \cap \Delta_V \text{ is closed in } U \times U' \\ &\iff v \mapsto (v, v): U \cap U' \to U \times U' \text{ is a closed immersion} \\ &\iff k[U \times U'] \to k[U \cap U'] \text{ is surjective (3.34).} \end{aligned}$$

Since $k[U \times U'] = k[U] \otimes_k k[U']$, this completes the proof of (ii). □

In more down-to-earth terms, condition (b) says that $U \cap U'$ is affine and every regular function on $U \cap U'$ is a sum of functions of the form $P \mapsto f(P)g(P)$ with f and g regular functions on U and U'.

EXAMPLE 5.30. (a) Let $V = \mathbb{P}^1$, and let U_0 and U_1 be the standard open subsets (see 5.3). Then $U_0 \cap U_1 = \mathbb{A}^1 \smallsetminus \{0\}$, and the maps on rings corresponding to the inclusions $U_0 \cap U_1 \hookrightarrow U_i$ are

$$f(X) \mapsto f(X): k[X] \to k[X, X^{-1}]$$
$$f(X) \mapsto f(X^{-1}): k[X] \to k[X, X^{-1}],$$

Thus the sets U_0 and U_1 satisfy the condition in (b).

(b) Let V be $\mathbb{A}^1$ with the origin doubled (see 5.10), and let U and U' be the upper and lower copies of $\mathbb{A}^1$ in V. Then $U \cap U'$ is affine, but the maps on rings corresponding to the inclusions $U_0 \cap U_1 \hookrightarrow U_i$ are

$$X \mapsto X: k[X] \to k[X, X^{-1}]$$
$$X \mapsto X: k[X] \to k[X, X^{-1}],$$

Thus the sets U_0 and U_1 fail the condition in (b).

(c) Let V be $\mathbb{A}^2$ with the origin doubled, and let U and U' be the upper and lower copies of $\mathbb{A}^2$ in V. Then $U \cap U'$ is not affine (see 3.33).

i Fibred products

Let $\varphi\colon V \to S$ and $\psi\colon W \to S$ be regular maps of algebraic varieties. The set

$$V \times_S W \stackrel{\text{def}}{=} \{(v,w) \in V \times W \mid \varphi(v) = \psi(w)\}$$

is a closed in $V \times W$, because it is the set where $\varphi \circ p$ and $\psi \circ q$ agree, and so it has a canonical structure of an algebraic variety (see p.101). The algebraic variety $V \times_S W$ is called the ***fibred product*** of V and W over S. Note that if S consists of a single point, then $V \times_S W = V \times W$.

Write φ' for the map $(v,w) \mapsto w\colon V \times_S W \to W$ and ψ' for the map $(v,w) \mapsto v\colon V \times_S W \to V$. We then have a commutative diagram:

$$\begin{tikzcd} V \times_S W \arrow[r, "\varphi'"] \arrow[d, "\psi'"] & W \arrow[d, "\psi"] \\ V \arrow[r, "\varphi"] & S. \end{tikzcd}$$

The system $(V \times_S W, \varphi', \psi')$ has the following universal property: for any regular maps $\alpha\colon T \to V$, $\beta\colon T \to W$ such that $\varphi\alpha = \psi\beta$, there is a unique regular map $(\alpha,\beta)\colon T \to V \times_S W$ such that the following diagram

$$\begin{tikzcd} T \arrow[dr, dotted, "(\alpha,\beta)"] \arrow[drr, bend left, "\beta"] \arrow[ddr, bend right, "\alpha"'] & & \\ & V \times_S W \arrow[r, "\varphi'"] \arrow[d, "\psi'"] & W \arrow[d, "\psi"] \\ & V \arrow[r, "\varphi"] & S \end{tikzcd}$$

commutes. In other words,

$$\mathrm{Hom}(T, V \times_S W) \simeq \mathrm{Hom}(T,V) \times_{\mathrm{Hom}(T,S)} \mathrm{Hom}(T,W).$$

Indeed, there is a unique such map of sets, namely, $t \mapsto (\alpha(t), \beta(t))$, which is regular because it is as a map into $V \times W$.

The map φ' in the above diagrams is called the ***base change*** of φ with respect to ψ. For any point $P \in S$, the base change of $\varphi\colon V \to S$ with respect to $P \hookrightarrow S$ is the map $\varphi^{-1}(P) \to P$ induced by φ, which is called the ***fibre*** of V over P.

EXAMPLE 5.31. If $f\colon V \to S$ is a regular map and U is a subvariety of S, then $V \times_S U$ is the inverse image of U in V.

Notes

5.32. Since a tensor product of rings $A \otimes_R B$ has the opposite universal property to that of a fibred product, one might hope that

$$\mathrm{Spm}(A) \times_{\mathrm{Spm}(R)} \mathrm{Spm}(B) \stackrel{??}{=} \mathrm{Spm}(A \otimes_R B).$$

This is true if $A \otimes_R B$ is an affine k-algebra, but in general it may have nonzero nilpotent elements. For example, let $R = k[X]$, and consider the R-algebras

$$\begin{cases} k[X] \to k, & X \mapsto a \\ k[X] \to k[X], & X \mapsto X^p. \end{cases}$$

Then

$$A \otimes_R B \simeq k \otimes_{k[X^p]} k[X] \simeq k[X]/(X^p - a),$$

which contains the nilpotent element $x - a^{\frac{1}{p}}$ if $p = \mathrm{char}(k)$.

The correct statement is

$$\mathrm{Spm}(A) \times_{\mathrm{Spm}(R)} \mathrm{Spm}(B) \simeq \mathrm{Spm}(A \otimes_R B/\mathfrak{N}) \tag{25}$$

where $\mathfrak{N}$ is the ideal of nilpotent elements in $A \otimes_R B$. To prove this, note that for any algebraic variety T,

$$\begin{aligned} \mathrm{Mor}(T, \mathrm{Spm}(A \otimes_R B/\mathfrak{N})) &\simeq \mathrm{Hom}(A \otimes_R B/\mathfrak{N}, \mathcal{O}_T(T)) && (5.12) \\ &\simeq \mathrm{Hom}(A \otimes_R B, \mathcal{O}_T(T)) && (\text{the ring } \mathcal{O}_T(T) \text{ is reduced}) \\ &\simeq \mathrm{Hom}(A, \mathcal{O}_T(T)) \underset{\mathrm{Hom}(R, \mathcal{O}_T(T))}{\times} \mathrm{Hom}(B, \mathcal{O}_T(T)) \\ &\simeq \mathrm{Mor}(T, \mathrm{Spm}(A)) \underset{\mathrm{Mor}(T, \mathrm{Spm}(R))}{\times} \mathrm{Mor}(T, \mathrm{Spm}(B)) && (5.12). \end{aligned}$$

For the third isomorphism, we used the universal property of $A \otimes_R B$.

5.33. Fibred products may differ depending on whether we are working in the category of algebraic varieties or algebraic schemes. For example,

$$\mathrm{Spec}(A) \times_{\mathrm{Spec}(R)} \mathrm{Spec}(B) \simeq \mathrm{Spec}(A \otimes_R B)$$

in the category of schemes. Consider the map $x \mapsto x^2 \colon \mathbb{A}^1 \xrightarrow{\varphi} \mathbb{A}^1$ (see 5.49). The fibre $\varphi^{-1}(a)$ consists of two points if $a \neq 0$, and one point if $a = 0$. Thus $\varphi^{-1}(0) = \mathrm{Spm}(k[X]/(X))$. However, the scheme-theoretic fibre is $\mathrm{Spec}(k[X]/(X^2))$, which reflects the fact that 0 is "doubled" in the fibre over 0.

5.34. Fibred products exist also for prevarieties. In this case, $V \times_S W$ is only locally closed in $V \times W$.

j Dimension

Recall that, in an irreducible topological space, every nonempty open subset is dense and irreducible.

Let V be an irreducible algebraic variety V, and let U and U' be nonempty open affines in V. Then $U \cap U'$ is also a nonempty open affine (5.29), which is dense in U, and so the restriction map $\mathcal{O}_V(U) \to \mathcal{O}_V(U')$ is injective. Therefore

$$k[U] \subset k[U \cap U'] \subset k(U)$$

where $k(U)$ is the field of fractions of $k[U]$, and so $k(U)$ is also the field of fractions of $k[U \cap U']$ and of $k[U']$. Thus, attached to V there is a field $k(V)$, called ***the field of***

rational functions on V, which is the field of fractions of $k[U]$ for every open affine U in V. The ***dimension*** of V is defined to be the transcendence degree of $k(V)$ over k. Note the $\dim(V) = \dim(U)$ for any open subset U of V. In particular, $\dim(V) = \dim(U)$ for U an open affine in V. It follows that some of the results in §2 carry over — for example, if Z is a proper closed subvariety of V, then $\dim(Z) < \dim(V)$.

PROPOSITION 5.35. *Let V and W be irreducible varieties. Then*

$$\dim(V \times W) = \dim(V) + \dim(W).$$

PROOF. We may suppose V and W to be affine. Write

$$k[V] = k[x_1, \dots, x_m]$$
$$k[W] = k[y_1, \dots, y_n]$$

where the x and y have been chosen so that $\{x_1, \dots, x_d\}$ and $\{y_1, \dots, y_e\}$ are maximal algebraically independent sets of elements of $k[V]$ and $k[W]$. Then $\{x_1, \dots, x_d\}$ and $\{y_1, \dots, y_e\}$ are transcendence bases of $k(V)$ and $k(W)$ (see FT 9.12), and so $\dim(V) = d$ and $\dim(W) = e$. Then[6]

$$k[V \times W] \overset{\text{def}}{=} k[V] \otimes_k k[W] \supset k[x_1, \dots, x_d] \otimes_k k[y_1, \dots, y_e] \simeq k[x_1, \dots, x_d, y_1, \dots, y_e].$$

Therefore $\{x_1 \otimes 1, \dots, x_d \otimes 1, 1 \otimes y_1, \dots, 1 \otimes y_e\}$ will be algebraically independent in $k[V] \otimes_k k[W]$. Obviously $k[V \times W]$ is generated as a k-algebra by the elements $x_i \otimes 1$, $1 \otimes y_j$, $1 \le i \le m$, $1 \le j \le n$, and all of them are algebraic over

$$k[x_1, \dots, x_d] \otimes_k k[y_1, \dots, y_e].$$

Thus the transcendence degree of $k(V \times W)$ is $d + e$. □

We extend the definition of dimension to an arbitrary variety V as follows. An algebraic variety is a finite union of noetherian topological spaces, and so is noetherian. Consequently (see 2.31), V is a finite union $V = \bigcup V_i$ of its irreducible components, and we define $\dim(V) = \max \dim(V_i)$. When all the irreducible components of V have dimension n, V is said to be ***pure of dimension*** n (or to be of ***pure dimension*** n).

PROPOSITION 5.36. *Let V and W be closed subvarieties of $\mathbb{A}^n$; for any (nonempty) irreducible component Z of $V \cap W$,*

$$\dim(Z) \ge \dim(V) + \dim(W) - n;$$

that is,

$$\operatorname{codim}(Z) \le \operatorname{codim}(V) + \operatorname{codim}(W).$$

PROOF. In the course of the proof of (5.29), we saw that $V \cap W$ is isomorphic to $\Delta \cap (V \times W)$, and this is defined by the n equations $X_i = Y_i$ in $V \times W$. Thus the statement follows from (3.45). □

[6] In general, it is not true that if M' and N' are R-submodules of M and N, then $M' \otimes_R N'$ is an R-submodule of $M \otimes_R N$. However, this is true if R is a field, because then M' and N' will be direct summands of M and N, and tensor products preserve direct summands.

REMARK 5.37. (a) The subvariety

$$\begin{cases} X^2+Y^2 &= Z^2 \\ Z &= 0 \end{cases}$$

of $\mathbb{A}^3$ is the curve $X^2+Y^2=0$, which is the pair of lines $Y=\pm iX$ if $k=\mathbb{C}$; in particular, the codimension is 2. Note however, that real locus is $\{(0,0)\}$, which has codimension 3. Thus, Proposition 5.36 becomes false if one looks only at real points (and the pictures we draw can mislead).

(b) Proposition 5.36 becomes false if $\mathbb{A}^n$ is replaced by an arbitrary affine variety. Consider for example the affine cone V

$$X_1X_4 - X_2X_3 = 0.$$

It contains the planes,

$$Z : X_2 = 0 = X_4; \qquad Z = \{(*,0,*,0)\}$$

$$Z' : X_1 = 0 = X_3; \qquad Z' = \{(0,*,0,*)\}$$

and $Z \cap Z' = \{(0,0,0,0)\}$. Because V is a hypersurface in $\mathbb{A}^4$, it has dimension 3, and each of Z and Z' has dimension 2. Thus

$$\operatorname{codim} Z \cap Z' = 3 \not\leq 1+1 = \operatorname{codim} Z + \operatorname{codim} Z'.$$

The proof of (5.36) fails because the diagonal in $V \times V$ cannot be defined by 3 equations (it takes the same 4 that define the diagonal in $\mathbb{A}^4$) — the diagonal is not a set-theoretic complete intersection.

k Dominant maps

As in the affine case, a regular map $\varphi: V \to W$ is said to be ***dominant*** if the image of φ is dense in W. Suppose V and W are irreducible. If V' and W' are open affine subsets of V and W such that $\varphi(V') \subset W'$, then (3.34) implies that the map $f \mapsto f \circ \varphi: k[W'] \to k[V']$ is injective. Therefore it extends to a map on the fields of fractions, $k(W) \to k(V)$, and this map is independent of the choice of V' and W'.

l Rational maps; birational equivalence

Loosely speaking, a rational map from a variety V to a variety W is a regular map from a dense open subset of V to W, and a birational map is a rational map admitting a rational inverse.

Let V and W be varieties over k, and consider pairs (U, φ_U) where U is a dense open subset of V and φ_U is a regular map $U \to W$. Two such pairs (U, φ_U) and $(U', \varphi_{U'})$ are said to be ***equivalent*** if φ_U and $\varphi_{U'}$ agree on $U \cap U'$. An equivalence class of pairs is called a ***rational map*** $\varphi: V \dashrightarrow W$. A rational map φ is said to be ***defined*** at a point v of V if $v \in U$ for some $(U, \varphi_U) \in \varphi$. The set U_1 of v at which φ is defined is open, and there is a regular map $\varphi_1: U_1 \to W$ such that $(U_1, \varphi_1) \in \varphi$ — clearly, $U_1 = \bigcup_{(U,\varphi_U)\in\varphi} U$ and we can define φ_1 to be the regular map such that $\varphi_1 | U = \varphi_U$ for all $(U, \varphi_U) \in \varphi$. Hence, in the equivalence class, there is always a pair (U, φ_U) with U largest (and U is called "the open subvariety on which φ is defined").

PROPOSITION 5.38. *Let V and V' be irreducible varieties over k. A regular map $\varphi: U' \to U$ from an open subset U' of V' onto an open subset U of V defines a k-algebra homomorphism $k(V) \to k(V')$, and every such homomorphism arises in this way.*

PROOF. The first part of the statement is obvious, so let $k(V) \hookrightarrow k(V')$ be a k-algebra homomorphism. We identify $k(V)$ with a subfield of $k(V')$. Let U (resp. U') be a open affine subset of V (resp. U'). Let $k[U] = k[x_1, \ldots, x_m]$. Each $x_i \in k(V')$, which is the field of fractions of $k[U']$, and so there exists a nonzero $d \in k[U']$ such that $dx_i \in k[U']$ for all i. After inverting d, i.e., replacing U' with basic open subset, we may suppose that $k[U] \subset k[U']$. Thus, the inclusion $k(V) \hookrightarrow k(V')$ is induced by a dominant regular map $\varphi: U' \to U$. According to Proposition 9.1 below, the image of φ contains an open subset U_0 of U. Now $\varphi^{-1}(U_0) \xrightarrow{\varphi} U_0$ is the required map. □

A rational (or regular) map $\varphi: V \dashrightarrow W$ is ***birational*** if there exists a rational map $\varphi': W \dashrightarrow V$ such that $\varphi' \circ \varphi = \mathrm{id}_V$ and $\varphi \circ \varphi' = \mathrm{id}_W$ as rational maps. Two varieties V and V' are ***birationally equivalent*** if there exists a birational map from one to the other. In this case, there exist dense open subsets U and U' of V and V' respectively such that $U \approx U'$.

PROPOSITION 5.39. *Two irreducible varieties V and V' are birationally equivalent if and only if $k(V) \approx k(V')$ (as k-algebras).*

PROOF. Assume that $k(V) \approx k(V')$. We may suppose that V and W are affine, in which case the existence of $U \approx U'$ is proved in (3.36). This proves the "if" part, and the "only if" part is obvious. □

PROPOSITION 5.40. *Every irreducible algebraic variety of dimension d is birationally equivalent to a hypersurface in $\mathbb{A}^{d+1}$.*

PROOF. Let V be an irreducible variety of dimension d. According to (3.38), there exist $x_1, \ldots, x_d, x_{d+1} \in k(V)$ such that $k(V) = k(x_1, \ldots, x_d, x_{d+1})$. Let $f \in k[X_1, \ldots, X_{d+1}]$ be an irreducible polynomial satisfied by the x_i, and let H be the hypersurface $f = 0$. Then $k(V) \approx k(H)$. □

m Local study

Everything in Chapter 4, being local, extends mutatis mutandis, to general algebraic varieties.

5.41. The ***tangent space*** $T_P(V)$ at a point P on an algebraic variety V is the fibre of $V(k[\varepsilon]) \to V(k)$ over P. There are canonical isomorphisms

$$T_P(V) \simeq \mathrm{Der}_k(\mathcal{O}_P, k) \simeq \mathrm{Hom}_{k\text{-linear}}(\mathfrak{n}_P/\mathfrak{n}_P^2, k)$$

where $\mathfrak{n}_P$ is the maximal ideal of $\mathcal{O}_P$.

5.42. A point P on ***nonsingular*** (or ***smooth***) if it lies on a single irreducible component W and $\dim T_P(V) = \dim W$. A point P is nonsingular if and only if the local ring $\mathcal{O}_P$ is regular. The singular points form a proper closed subvariety, called the ***singular locus***.

5.43. A variety is ***nonsingular*** (or ***smooth***) if every point is nonsingular.

n Étale maps

DEFINITION 5.44. A regular map $\varphi\colon V \to W$ of smooth varieties is ***étale at a point*** P of V if the map $(d\varphi)_P\colon T_P(V) \to T_{\varphi(P)}(W)$ is an isomorphism; φ is ***étale*** if it is étale at all points of V.

Examples

5.45. A regular map

$$\varphi\colon \mathbb{A}^n \to \mathbb{A}^n, \quad a \mapsto (P_1(a_1,\ldots,a_n),\ldots,P_n(a_1,\ldots,a_n))$$

is étale at $\mathbf{a}$ if and only if $\operatorname{rank}\operatorname{Jac}(P_1,\ldots,P_n)(\mathbf{a}) = n$, because the map on the tangent spaces has matrix $\operatorname{Jac}(P_1,\ldots,P_n)(\mathbf{a})$). Equivalent condition: $\det\left(\frac{\partial P_i}{\partial X_j}(\mathbf{a})\right) \neq 0$

5.46. Let $V = \operatorname{Spm}(A)$ be an affine variety, and let $f = \sum c_i X^i \in A[X]$ be such that $A[X]/(f(X))$ is reduced. Let $W = \operatorname{Spm}(A[X]/(f(X))$, and consider the map $W \to V$ corresponding to the inclusion $A \hookrightarrow A[X]/(f)$. Thus

$$\begin{tikzcd} A[X]/(f) & A[X] \arrow[l, two heads] \\ & A \arrow[u] \arrow[ul] \end{tikzcd} \qquad \begin{tikzcd} W \arrow[r, hook] \arrow[dr] & V \times \mathbb{A}^1 \arrow[d] \\ & V \end{tikzcd}$$

The points of W lying over a point $\mathbf{a} \in V$ are the pairs $(\mathbf{a}, b) \in V \times \mathbb{A}^1$ such that b is a root of $\sum c_i(\mathbf{a})X^i$. I claim that the map $W \to V$ is étale at $(\mathbf{a}, b)$ if and only if b is a *simple* root of $\sum c_i(\mathbf{a})X^i$.

To see this, write $A = \operatorname{Spm} k[X_1,\ldots,X_n]/\mathfrak{a}$, $\mathfrak{a} = (f_1,\ldots,f_r)$, so that

$$A[X]/(f) = k[X_1,\ldots,X_n]/(f_1,\ldots,f_r,f).$$

The tangent spaces to W and V at $(\mathbf{a}, b)$ and $\mathbf{a}$ respectively are the null spaces of the matrices

$$\begin{pmatrix} \frac{\partial f_1}{\partial X_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial X_m}(\mathbf{a}) & 0 \\ \vdots & & \vdots & \\ \frac{\partial f_n}{\partial X_1}(\mathbf{a}) & \cdots & \frac{\partial f_n}{\partial X_m}(\mathbf{a}) & 0 \\ \frac{\partial f}{\partial X_1}(\mathbf{a}) & \cdots & \frac{\partial f}{\partial X_m}(\mathbf{a}) & \frac{\partial f}{\partial X}(\mathbf{a}, b) \end{pmatrix} \qquad \begin{pmatrix} \frac{\partial f_1}{\partial X_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial X_m}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial X_1}(\mathbf{a}) & \cdots & \frac{\partial f_n}{\partial X_m}(\mathbf{a}) \end{pmatrix}$$

and the map $T_{(\mathbf{a},b)}(W) \to T_{\mathbf{a}}(V)$ is induced by the projection map $k^{n+1} \to k^n$ omitting the last coordinate. This map is an isomorphism if and only if $\frac{\partial f}{\partial X}(\mathbf{a}, b) \neq 0$, because then every solution of the smaller set of equations extends uniquely to a solution of the larger set. But

$$\frac{\partial f}{\partial X}(\mathbf{a}, b) = \frac{d(\sum_i c_i(\mathbf{a})X^i)}{dX}(b),$$

which is zero if and only if b is a multiple root of $\sum_i c_i(\mathbf{a})X^i$. The intuitive picture is that $W \to V$ is a finite covering with $\deg(f)$ sheets, which is ramified exactly at the points where two or more sheets cross.

5.47. Consider a dominant map $\varphi: W \to V$ of smooth affine varieties, corresponding to a map $A \to B$ of rings. Suppose B can be written $B = A[Y_1, \ldots, Y_n]/(P_1, \ldots, P_n)$ (same number of polynomials as variables). A similar argument to the above shows that φ is étale if and only if $\det\left(\frac{\partial P_i}{\partial X_j}(\mathbf{a})\right)$ is never zero.

5.48. The example in (b) is typical; in fact every étale map is locally of this form, provided V is normal (in the sense defined below p.171). More precisely, let $\varphi: W \to V$ be étale at $P \in W$, and assume V to normal; then there exist a map $\varphi': W' \to V'$ with $k[W'] = k[V'][X]/(f(X))$, and a commutative diagram

$$\begin{tikzcd}
W \arrow[d, "\varphi"] & \supset & U_1 \arrow[d] & \approx & U_1' \arrow[d] & \subset & W' \arrow[d, "\varphi'"] \\
V & \supset & U_2 & \approx & U_2' & \subset & V'
\end{tikzcd}$$

with the U all open subvarieties and $P \in U_1$.

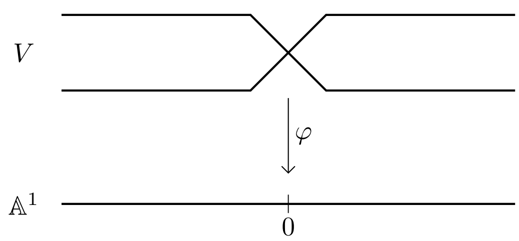
The failure of the inverse function theorem for the Zariski topology

5.49. In advanced calculus (or differential topology, or complex analysis), the inverse function theorem says that a map φ that is étale at a point $\mathbf{a}$ is a local isomorphism there, i.e., there exist open neighbourhoods U and U' of $\mathbf{a}$ and $\varphi(\mathbf{a})$ such that φ induces an isomorphism $U \to U'$. *This is not true in algebraic geometry*, at least not for the Zariski topology: a map can be étale at a point without being a local isomorphism. Consider for example the map

$$\varphi: \mathbb{A}^1 \smallsetminus \{0\} \to \mathbb{A}^1 \smallsetminus \{0\}, \quad a \mapsto a^2.$$

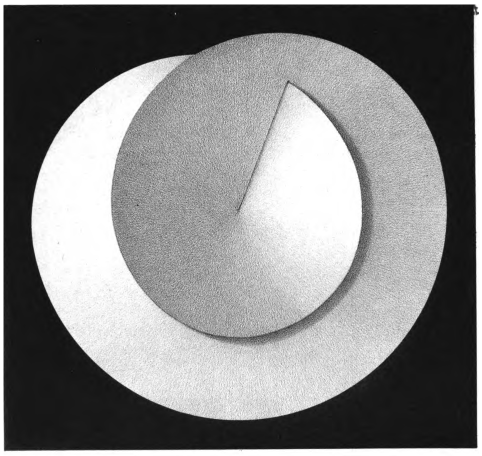
This is étale if the characteristic is $\neq 2$, because the Jacobian matrix is $(2X)$, which has rank one for all $X \neq 0$ (alternatively, it is of the form (5.46) with $f(X) = X^2 - T$, where T is the coordinate function on $\mathbb{A}^1$, and $X^2 - c$ has distinct roots for $c \neq 0$). Nevertheless, I claim that there do not exist nonempty open subsets U and U' of $\mathbb{A}^1 - \{0\}$ such that φ defines an isomorphism $U \to U'$. If there did, then φ would define an isomorphism $k[U'] \to k[U]$ and hence an isomorphism on the fields of fractions $k(\mathbb{A}^1) \to k(\mathbb{A}^1)$. But on the fields of fractions, φ defines the map $k(X) \to k(X)$, $X \mapsto X^2$, which is not an isomorphism.

5.50. Let V be the plane curve $Y^2 = X$ and φ the map $V \to \mathbb{A}^1$, $(x, y) \mapsto x$. Then φ is $2:1$ except over 0, and so we may view it schematically as



However, when viewed as a Riemann surface, $V(\mathbb{C})$ consists of two sheets joined at a single point O. As a point on the surface moves around O, it shifts from one sheet to the other. Thus the true picture is more complicated. To get a section to φ, it is necessary to remove a line in $\mathbb{C}$ from 0 to infinity, which is not closed for the Zariski topology.

It is not possible to fit the graph of the complex curve $Y^2 = X$ into 3-space, but the picture at right is an early depiction of it (from Neumann, Carl, Vorlesungen über Riemann's theorie der Abel'schen integrale, Leipzig : Teubner, 1865.



Die Riemann'sche Windungsfläche erster Ordnung.
Vergl. Seite 162-168, 213-214 und 218-221

Lith. Anst. v. M. Singer, Leipzig.

Étale maps of singular varieties

Using tangent cones, we can extend the notion of an étale morphism to singular varieties. Obviously, a regular map $\alpha \colon V \to W$ induces a homomorphism $\mathrm{gr}(\mathcal{O}_{\alpha(P)}) \to \mathrm{gr}(\mathcal{O}_P)$.

We say that α is ***étale*** at P if this is an isomorphism. Note that then there is an isomorphism of the geometric tangent cones $C_P(V) \to C_{\alpha(P)}(W)$, but this map may be an isomorphism without α being étale at P. Roughly speaking, to be étale at P, we need the map on geometric tangent cones to be an isomorphism and to preserve the "multiplicities" of the components.

It is a fairly elementary result that a local homomorphism of local rings $\alpha \colon A \to B$ induces an isomorphism on the graded rings if and only if it induces an isomorphism on the completions (Atiyah-MacDonald 1969, 10.23). Thus $\alpha \colon V \to W$ is étale at P if and only if the map $\hat{\mathcal{O}}_{\alpha(P)} \to \hat{\mathcal{O}}_P$ is an isomorphism. Hence (5.53) shows that the choice of a local system of parameters $f_1, \ldots, f_d$ at a nonsingular point P determines an isomorphism $\hat{\mathcal{O}}_P \to k[[X_1, \ldots, X_d]]$.

We can rewrite this as follows: let $t_1, \ldots, t_d$ be a local system of parameters at a nonsingular point P; then there is a canonical isomorphism $\hat{\mathcal{O}}_P \to k[[t_1, \ldots, t_d]]$. For $f \in \hat{\mathcal{O}}_P$, the image of $f \in k[[t_1, \ldots, t_d]]$ can be regarded as the Taylor series of f.

For example, let $V = \mathbb{A}^1$, and let P be the point a. Then $t = X - a$ is a local parameter at a, $\mathcal{O}_P$ consists of quotients $f(X) = g(X)/h(X)$ with $h(a) \neq 0$, and the coefficients of the Taylor expansion $\sum_{n \geq 0} a_n (X-a)^n$ of $f(X)$ can be computed as in elementary calculus courses: $a_n = f^{(n)}(a)/n!$.

PROPOSITION 5.51. *Let $\varphi \colon W \to V$ be a map of irreducible affine varieties. If $k(W)$ is a finite separable extension of $k(V)$, then φ is étale on a nonempty open subvariety of W.*

PROOF. After passing to open subvarieties, we may assume that W and V are nonsingular, and that $k[W] = k[V][X]/(f(X))$ where $f(X)$ is separable when considered as a polynomial in $k(V)$. Now the statement follows from (5.46). $\square$

ASIDE 5.52. There is an old conjecture that every étale map $\varphi\colon \mathbb{A}^n \to \mathbb{A}^n$ is an isomorphism. If we write $\varphi = (P_1, \ldots, P_n)$, then this becomes the statement:

$$\text{if } \det\left(\frac{\partial P_i}{\partial X_j}(\mathbf{a})\right) \text{ is never zero (for } \mathbf{a} \in k^n\text{), then } \varphi \text{ has a inverse.}$$

The condition, $\det\left(\frac{\partial P_i}{\partial X_j}(\mathbf{a})\right)$ never zero, implies that $\det\left(\frac{\partial P_i}{\partial X_j}\right)$ is a nonzero constant (by the Nullstellensatz 2.11 applied to the ideal generated by $\det\left(\frac{\partial P_i}{\partial X_j}\right)$). This conjecture, which is known as the Jacobian conjecture, has not been settled even for $k = \mathbb{C}$ and $n = 2$, despite the existence of several published proofs and innumerable announced proofs. It has caused many mathematicians a good deal of grief. It is probably harder than it is interesting. See the Wikipedia JACOBIAN CONJECTURE.

o Étale neighbourhoods

Recall that a regular map $\alpha\colon W \to V$ is said to be étale at a nonsingular point P of W if the map $(d\alpha)_P\colon T_P(W) \to T_{\alpha(P)}(V)$ is an isomorphism.

Let P be a nonsingular point on a variety V of dimension d. A ***local system of parameters*** at P is a family $\{f_1, \ldots, f_d\}$ of germs of regular functions at P generating the maximal ideal $\mathfrak{n}_P \subset \mathcal{O}_P$. Equivalent conditions: the images of $f_1, \ldots, f_d$ in $\mathfrak{n}_P/\mathfrak{n}_P^2$ generate it as a k-vector space (see 1.4); or $(df_1)_P, \ldots, (df_d)_P$ is a basis for dual space to $T_P(V)$.

PROPOSITION 5.53. *Let $\{f_1, \ldots, f_d\}$ be a local system of parameters at a nonsingular point P of V. Then there is a nonsingular open neighbourhood U of P such that $f_1, f_2, \ldots, f_d$ are represented by pairs $(\tilde{f}_1, U), \ldots, (\tilde{f}_d, U)$ and the map $(\tilde{f}_1, \ldots, \tilde{f}_d)\colon U \to \mathbb{A}^d$ is étale.*

PROOF. Obviously, the f_i are represented by regular functions $\tilde{f}_i$ defined on a single open neighbourhood U' of P, which, because of (4.37), we can choose to be nonsingular. The map $\alpha = (\tilde{f}_1, \ldots, \tilde{f}_d)\colon U' \to \mathbb{A}^d$ is étale at P, because the dual map to $(d\alpha)_{\mathbf{a}}$ is $(dX_i)_o \mapsto (d\tilde{f}_i)_{\mathbf{a}}$. The next lemma then shows that α is étale on an open neighbourhood U of P. □

LEMMA 5.54. *Let W and V be nonsingular varieties. If $\alpha\colon W \to V$ is étale at P, then it is étale at all points in an open neighbourhood of P.*

PROOF. The hypotheses imply that W and V have the same dimension d, and that their tangent spaces all have dimension d. We may assume W and V to be affine, say $W \subset \mathbb{A}^m$ and $V \subset \mathbb{A}^n$, and that α is given by polynomials $P_1(X_1, \ldots, X_m), \ldots, P_n(X_1, \ldots, X_m)$. Then $(d\alpha)_{\mathbf{a}}\colon T_{\mathbf{a}}(\mathbb{A}^m) \to T_{\alpha(\mathbf{a})}(\mathbb{A}^n)$ is a linear map with matrix $\left(\frac{\partial P_i}{\partial X_j}(\mathbf{a})\right)$, and α is not étale at $\mathbf{a}$ if and only if the kernel of this map contains a nonzero vector in the subspace $T_{\mathbf{a}}(V)$ of $T_{\mathbf{a}}(\mathbb{A}^n)$. Let $f_1, \ldots, f_r$ generate $I(W)$. Then α is not étale at $\mathbf{a}$ if and only if the matrix

$$\begin{pmatrix} \frac{\partial f_i}{\partial X_j}(\mathbf{a}) \\ \frac{\partial P_i}{\partial X_j}(\mathbf{a}) \end{pmatrix}$$

has rank less than m. This is a polynomial condition on $\mathbf{a}$, and so it fails on a closed subset of W, which doesn't contain P. □

Let V be a nonsingular variety, and let $P \in V$. An ***étale neighbourhood*** of a point P of V is pair $(Q, \pi\colon U \to V)$ with π an étale map from a nonsingular variety U to V and Q a point of U such that $\pi(Q) = P$.

COROLLARY 5.55. *Let V be a nonsingular variety of dimension d, and let $P \in V$. There is an open Zariski neighbourhood U of P and a map $\pi: U \to \mathbb{A}^d$ realizing (P, U) as an étale neighbourhood of $(0, \ldots, 0) \in \mathbb{A}^d$.*

PROOF. This is a restatement of the Proposition. □

ASIDE 5.56. Note the analogy with the definition of a differentiable manifold: every point P on nonsingular variety of dimension d has an open neighbourhood that is also a "neighbourhood" of the origin in $\mathbb{A}^d$. There is a "topology" on algebraic varieties for which the "open neighbourhoods" of a point are the étale neighbourhoods. Relative to this "topology", any two nonsingular varieties are locally isomorphic (this is *not* true for the Zariski topology). The "topology" is called the ***étale topology*** — see my notes Lectures on Étale Cohomology.

The inverse function theorem (for the étale topology)

THEOREM 5.57 (INVERSE FUNCTION THEOREM). *If a regular map of nonsingular varieties $\varphi: V \to W$ is étale at $P \in V$, then there exists a commutative diagram*

$$\begin{tikzcd} V \arrow[d, "\varphi"] & U_P \arrow[l, "\text{open}"'] \arrow[d, "\varphi'", "\approx"'] \\ W & U_{\varphi(P)} \arrow[l, "\text{étale}"'] \end{tikzcd}$$

with U_P an open neighbourhood U of P, $U_{f(P)}$ an étale neighbourhood $\varphi(P)$, and φ' an isomorphism.

PROOF. According to (5.54), there exists an open neighbourhood U of P such that the restriction $\varphi|U$ of φ to U is étale. To get the above diagram, we can take $U_P = U$, $U_{\varphi(P)}$ to be the étale neighbourhood $\varphi|U: U \to W$ of $\varphi(P)$, and φ' to be the identity map. □

The rank theorem

For vector spaces, the rank theorem says the following: let $\alpha: V \to W$ be a linear map of k-vector spaces of rank r; then there exist bases for V and W relative to which α has matrix $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. In other words, there is a commutative diagram

$$\begin{tikzcd} V \arrow[r, "\alpha"] \arrow[d, "\approx"] & W \arrow[d, "\approx"] \\ k^m \arrow[r, "{(x_1,\ldots,x_m)\mapsto(x_1,\ldots,x_r,0,\ldots)}"] & k^n \end{tikzcd}$$

A similar result holds locally for differentiable manifolds. In algebraic geometry, there is the following weaker analogue.

THEOREM 5.58 (RANK THEOREM). *Let $\varphi: V \to W$ be a regular map of nonsingular varieties of dimensions m and n respectively, and let $P \in V$. If $\operatorname{rank}(T_P(\varphi)) = n$, then there exists a commutative diagram*

$$\begin{tikzcd} U_P \arrow[r, "\varphi|U_P"] \arrow[d, "\text{étale}"] & W \arrow[d, "\text{étale}"] \\ \mathbb{A}^m \arrow[r, "{(x_1,\ldots,x_m)\mapsto(x_1,\ldots,x_n)}"] & \mathbb{A}^n \end{tikzcd}$$

in which U_P and $U_{\varphi(P)}$ are open neighbourhoods of P and $\varphi(P)$ respectively and the vertical maps are étale.

PROOF. Choose a local system of parameters $g_1,\ldots,g_n$ at $\varphi(P)$, and let $f_1 = g_1 \circ \varphi, \ldots, f_n = g_n \circ \varphi$. Then $df_1,\ldots,df_n$ are linearly independent forms on $T_P(V)$, and there exist $f_{n+1},\ldots,f_m$ such $df_1,\ldots,df_m$ is a basis for $T_P(V)^\vee$. Then $f_1,\ldots,f_m$ is a local system of parameters at P. According to (5.54), there exist open neighbourhoods U_P of P and $U_{\varphi(P)}$ of $\varphi(P)$ such that the maps

$$(f_1,\ldots,f_m)\colon U_P \to \mathbb{A}^m$$
$$(g_1,\ldots,g_n)\colon U_{\varphi(P)} \to \mathbb{A}^n$$

are étale. They give the vertical maps in the above diagram. □

ASIDE 5.59. Tangent vectors at a point P on a smooth manifold V can be defined to be certain equivalence classes of curves through P (Wikipedia TANGENT SPACE). For $V = \mathbb{A}^n$, there is a similar description with a curve taken to be a regular map from an open neighbourhood U of 0 in $\mathbb{A}^1$ to V. In the general case there is a map from an open neighbourhood of the point P in X onto affine space sending P to 0 and inducing an isomorphism from tangent space at P to that at 0 (5.53). Unfortunately, the maps from $U \subset \mathbb{A}^1$ to $\mathbb{A}^n$ need not lift to X, and so it is necessary to allow maps from smooth curves into X (pull-backs of the covering $X \to \mathbb{A}^n$ by the maps from U into $\mathbb{A}^n$). There is a description of the tangent vectors at a point P on a smooth algebraic variety V as certain equivalence classes of regular maps from an étale neighbourhood U of 0 in $\mathbb{A}^1$ to V.

p Smooth maps

DEFINITION 5.60. A regular map $\varphi\colon V \to W$ of nonsingular varieties is ***smooth at a point*** P of V if $(d\varphi)_P\colon T_P(V) \to T_{\varphi(P)}(W)$ is surjective; φ is ***smooth*** if it is smooth at all points of V.

THEOREM 5.61. *A map $\varphi\colon V \to W$ is smooth at $P \in V$ if and only if there exist open neighbourhoods U_P and $U_{\varphi(P)}$ of P and $\varphi(P)$ respectively such that $\varphi|U_P$ factors into*

$$U_P \xrightarrow{\text{étale}} \mathbb{A}^{\dim V - \dim W} \times U_{\varphi(P)} \xrightarrow{q} U_{\varphi(P)}.$$

PROOF. Certainly, if $\varphi|U_P$ factors in this way, it is smooth. Conversely, if φ is smooth at P, then we get a diagram as in the rank theorem. From it we get maps

$$U_P \to \mathbb{A}^m \times_{\mathbb{A}^n} U_{\varphi(P)} \to U_{\varphi(P)}.$$

The first is étale, and the second is the projection of $\mathbb{A}^{m-n} \times U_{\varphi(P)}$ onto $U_{\varphi(P)}$. □

COROLLARY 5.62. *Let V and W be nonsingular varieties. If $\varphi\colon V \to W$ is smooth at P, then it is smooth on an open neighbourhood of V.*

PROOF. In fact, it is smooth on the neighbourhood U_P in the theorem. □

Separable maps

DEFINITION 5.63. A dominant map $\varphi: W \to V$ of irreducible algebraic varieties is ***separable*** if $k(W)$ is separably generated over $k(V)$.

THEOREM 5.64. *Let $\varphi: W \to V$ be a map of irreducible varieties.*

(a) *If there exists a nonsingular point P of W such that φP is nonsingular and $(d\varphi)_P$ is surjective, then φ is dominant and separable.*

(b) *Conversely if φ is dominant and separable, then the set of $P \in W$ satisfying (a) is open and dense.*

PROOF. Replace W and V with their open subsets of nonsingular points. Then apply the rank theorem. □

q Algebraic varieties as a functors

Let R be an affine k-algebra, and let V be an algebraic variety. We define a ***point of*** V ***with coordinates in*** R to be a regular map $\mathrm{Spm}(R) \to V$. For example, if $V = V(\mathfrak{a}) \subset k^n$, then

$$V(R) = \{(a_1, \ldots, a_n) \in R^n \mid f(a_1, \ldots, a_n) = 0 \text{ all } f \in \mathfrak{a}\},$$

which is what you should expect. In particular $V(k) = V$ (as a set), i.e., V (as a set) can be identified with the set of points of V with coordinates in k. Note that

$$(V \times W)(R) = V(R) \times W(R)$$

(property of a product).

REMARK 5.65. Let V be the union of two subvarieties, $V = V_1 \cup V_2$. If V_1 and V_2 are both open, then $V(R) = V_1(R) \cup V_2(R)$, but not necessarily otherwise. For example, for any polynomial $f(X_1, \ldots, X_n)$,

$$\mathbb{A}^n = D_f \cup V(f)$$

where $D_f \simeq \mathrm{Spm}(k[X_1, \ldots, X_n, T]/(1 - Tf))$ and $V(f)$ is the zero set of f, but

$$R^n \neq \{\mathbf{a} \in R^n \mid f(\mathbf{a}) \in R^\times\} \cup \{\mathbf{a} \in R^n \mid f(\mathbf{a}) = 0\}$$

in general.

THEOREM 5.66. *A regular map $\varphi: V \to W$ of algebraic varieties defines a family of maps of sets, $\varphi(R): V(R) \to W(R)$, one for each affine k-algebra R, such that for every homomorphism $\alpha: R \to S$ of affine k-algebras,*

$$\begin{array}{ccc} R & \quad & V(R) \xrightarrow{\varphi(R)} W(R) \\ \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle V(\alpha)} \qquad\quad \downarrow{\scriptstyle V(\beta)} \\ S & & V(S) \xrightarrow{\varphi(S)} W(S) \end{array} \tag{*}$$

commutes. Every family of maps with this property arises from a unique morphism of algebraic varieties.

For a variety V, let h_V^{aff} be the functor sending an affine k-algebra R to $V(R)$. We can restate as Theorem 5.66 follows.

THEOREM 5.67. *The functor*

$$V \rightsquigarrow h_V^{\mathrm{aff}} : \mathsf{Var}_k \to \mathsf{Fun}(\mathsf{Aff}_k, \mathsf{Sets})$$

if fully faithful.

PROOF. The Yoneda lemma (q.v. Wikipedia) shows that the functor

$$V \rightsquigarrow h_V : \mathsf{Var}_k \to \mathsf{Fun}(\mathsf{Var}_k, \mathsf{Sets})$$

is fully faithful. Let φ be a morphism $h_V^{\mathrm{aff}} \to h_{V'}^{\mathrm{aff}}$, and let T be a variety. Let $(U_i)_{i \in I}$ be a finite affine covering of T. Each intersection $U_i \cap U_j$ is affine (5.29), and so φ gives rise to a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & h_V(T) \arrow[r] \arrow[d] & \prod_i h_V(U_i) \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & \prod_{i,j} h_V(U_i \cap U_j) \arrow[d] \\
0 \arrow[r] & h_{V'}(T) \arrow[r] & \prod_i h_{V'}(U_i) \arrow[r, shift left] \arrow[r, shift right] & \prod_{i,j} h_{V'}(U_i \cap U_j)
\end{tikzcd}$$

in which the pairs of maps are defined by the inclusions $U_i \cap U_j \hookrightarrow U_i, U_j$. As the rows are exact (5.15), this shows that φ_V extends uniquely to a functor $h_V \to h_{V'}$, which (by the Yoneda lemma) arises from a unique regular map $V \to V'$. □

COROLLARY 5.68. *To give an affine group variety is the same as giving a functor $G: \mathsf{Aff}_k \to \mathsf{Gp}$ such that for some n and some finite set S of polynomials in $k[X_1, X_2, \ldots, X_n]$, G is isomorphic to the functor sending R to the set of zeros of S in R^n.*

PROOF. Certainly an affine group variety defines such a functor. Conversely, the conditions imply that $G = h_V$ for an affine algebraic variety V (unique up to a unique isomorphism). The multiplication maps $G(R) \times G(R) \to G(R)$ give a morphism of functors $h_V \times h_V \to h_V$. As $h_V \times h_V \simeq h_{V \times V}$ (by definition of $V \times V$), we see that they arise from a regular map $V \times V \to V$. Similarly, the inverse map and the identity-element map are regular. □

It is not unusual for a variety to be most naturally defined in terms of its points functor. For example:

$$\begin{aligned}
\mathrm{SL}_n &: R \rightsquigarrow \{M \in M_n(R) \mid \det(M) = 1\} \\
\mathrm{GL}_n &: R \rightsquigarrow \{M \in M_n(R) \mid \det(M) \in R^\times\} \\
\mathbb{G}_a &: R \rightsquigarrow (R, +).
\end{aligned}$$

We now describe the essential image of $h \mapsto h_V : \mathsf{Var}_k \to \mathsf{Fun}(\mathsf{Aff}_k, \mathsf{Sets})$. The ***fibred product*** of two maps $\alpha_1: F_1 \to F_3$, $\alpha_2: F_2 \to F_3$ of sets is the set

$$F_1 \times_{F_3} F_2 = \{(x_1, x_2) \mid \alpha_1(x_1) = \alpha_2(x_2)\}.$$

When F_1, F_2, F_3 are functors and $\alpha_1, \alpha_2, \alpha_3$ are morphisms of functors, there is a functor $F = F_1 \times_{F_3} F_2$ such that

$$(F_1 \times_{F_3} F_2)(R) = F_1(R) \times_{F_3(R)} F_2(R)$$

for all affine k-algebras R.

To simplify the statement of the next proposition, we write U for h_U when U is an affine variety.

PROPOSITION 5.69. *A functor $F\colon \mathsf{Aff}_k \to \mathsf{Sets}$ is in the essential image of Var_k if and only if there exists an affine scheme U and a morphism $U \to F$ such that*

(a) *the functor $R \overset{\text{def}}{=} U \times_F U$ is a closed affine subvariety of $U \times U$ and the maps $R \rightrightarrows U$ defined by the projections are open immersions;*

(b) *the set $R(k)$ is an equivalence relation on $U(k)$, and the map $U(k) \to F(k)$ realizes $F(k)$ as the quotient of $U(k)$ by $R(k)$.*

PROOF. Let $F = h_V$ for V an algebraic variety. Choose a finite open affine covering $V = \bigcup U_i$ of V, and let $U = \bigsqcup U_i$. It is again an affine variety (Exercise 5-2). The functor R is $h_{U'}$ where U' is the disjoint union of the varieties $U_i \cap U_j$. These are affine (5.29), and so U' is affine. As U' is the inverse image of Δ_V in $U \times U$, it is closed (5.26). This proves (a), and (b) is obvious.

The converse is omitted for the present. $\square$

REMARK 5.70. A variety V defines a functor $R \rightsquigarrow V(R)$ from the category of all k-algebras to Sets. Again, we call the elements of $V(R)$ the ***points of V with coordinates in R***.

For example, if V is affine,

$$V(R) = \mathrm{Hom}_{k\text{-algebra}}(k[V], R).$$

More explicitly, if $V \subset k^n$ and $I(V) = (f_1, \ldots, f_m)$, then $V(R)$ is the set of solutions in R^n of the system equations

$$f_i(X_1, \ldots, X_n) = 0, \quad i = 1, \ldots, m.$$

Note that, when we allow R to have nilpotent elements, it is important to choose the f_i to generate $I(V)$ (i.e., a radical ideal) and not just an ideal $\mathfrak{a}$ such that $V(\mathfrak{a}) = V$.[7]

For a general variety V, we write V as a finite union of open affines $V = \bigcup_i V_i$, and we define $V(R)$ to be the set of families $(\alpha_i)_{i \in I} \in \prod_{i \in I} V_i(R)$ such that α_i agrees with α_j on $V_i \cap V_j$ for all $i, j \in I$. This is independent of the choice of the covering, and agrees with the previous definition when V is affine.

The functor defined by $\mathbb{A}(E)$ (see p.70) is $R \rightsquigarrow R \otimes_k E$.

A criterion for a functor to arise from an algebraic prevariety

5.71. By a functor we mean a functor from the category of affine k-algebras to sets. A subfunctor U of a functor X is ***open*** if, for all maps $\varphi\colon h^A \to X$, the subfunctor $\varphi^{-1}(U)$ of h^A is defined by an open subscheme of $\mathrm{Spm}(A)$. A family $(U_i)_{i \in I}$ of open subfunctors of X is an ***open covering*** of X if each U_i is open in X and $X = \bigcup U_i(K)$ for every field K. A functor X is ***local*** if, for all k-algebras R and all finite families $(f_i)_i$ of elements of A generating the ideal A, the sequence of sets

$$X(R) \to \prod\nolimits_i X(R_{f_i}) \rightrightarrows \prod\nolimits_{i,j} X(R_{f_i f_j})$$

[7] Let $\mathfrak{a}$ be an ideal in $k[X_1, \ldots]$. If A has no nonzero nilpotent elements, then every k-algebra homomorphism $k[X_1, \ldots] \to A$ that is zero on $\mathfrak{a}$ is also zero on $\mathrm{rad}(\mathfrak{a})$, and so

$$\mathrm{Hom}_k(k[X_1, \ldots]/\mathfrak{a}, A) \simeq \mathrm{Hom}_k(k[X_1, \ldots]/\mathrm{rad}(\mathfrak{a}), A).$$

This is not true if A has nonzero nilpotents.

is exact.

Let $\mathbb{A}^1$ denote the functor sending a k-algebra R to its underlying set. For a functor U, let $\mathcal{O}(U) = \mathrm{Hom}(U, \mathbb{A}^1)$ — it is a k-algebra.[8] A functor U is ***affine*** if $\mathcal{O}(U)$ is an affine k-algebra and the canonical map $U \to h^{\mathcal{O}(U)}$ is an isomorphism. A local functor admitting a finite covering by open affines is representable by an algebraic variety over k.

In the functorial approach to algebraic geometry, an algebraic prevariety over k is *defined* to be a functor satisfying this criterion. See, for example, I, §1, 3.11, p.13, of Demazure and Gabriel, Groupes algébriques: géométrie algébrique, généralités, groupes commutatifs. 1970.

r Rational and unirational varieties

DEFINITION 5.72. Let V be an algebraic variety over k.

(a) V is ***unirational*** if there exists a dominant rational map $\mathbb{P}^n \dashrightarrow V$.

(b) V is ***rational*** if there exists a birational map $\mathbb{P}^n \dashrightarrow V$.

In more down-to-earth terms, V is rational if $k(V)$ is a pure transcendental extension of k, and it is unirational if $k(V)$ is contained in such an extension of k.

In 1876 (over $\mathbb{C}$), Lüroth proved that every unirational curve is rational. For a proof over any field, see, for example, FT, Chapter 9. The Lüroth problem asks whether every unirational variety is rational.

Already for surfaces, this is a difficult problem. In characteristic zero, Castelnuovo and Severi proved that all unirational surfaces are rational, but in characteristic $p \neq 0$, Zariski showed that some surfaces of the form

$$Z^p = f(X, Y),$$

while obviously unirational, are not rational. Surfaces of this form are now called Zariski surfaces.

Fano attempted to find counter-examples to the Lüroth problem in dimension 3 among the so-called Fano varieties, but none of his attempted proofs satisfies modern standards. In 1971-72, three examples of nonrational unirational three-folds were found. For a description of them, and more discussion of the Lüroth problem in characteristic zero, see: Arnaud Beauville, *The Lüroth problem*, arXiv:1507.02476.

A little history

In his first proof of the Riemann hypothesis for curves over finite fields, Weil made use of the Jacobian variety of the curve, but initially he was not able to construct this as a projective variety. This led him to introduce "abstract" algebraic varieties, neither affine nor projective (in 1946). Weil first made use of the Zariski topology when he introduced fibre spaces into algebraic geometry (in 1949). For more on this, see my article: *The Riemann hypothesis over finite fields: from Weil to the present day.*

[8] Actually, one needs to be more careful to ensure that $\mathcal{O}(U)$ is a set; for example, restrict U and $\mathbb{A}^1$ to the category of k-algebras of the form $k[X_0, X_1, \ldots]/\mathfrak{a}$ for a fixed family of symbols (X_i) indexed by $\mathbb{N}$.

Exercises

5-1. Show that the only regular functions on $\mathbb{P}^1$ are the constant functions. [Thus $\mathbb{P}^1$ is not affine. When $k = \mathbb{C}$, $\mathbb{P}^1$ is the Riemann sphere (as a set), and one knows from complex analysis that the only holomorphic functions on the Riemann sphere are constant. Since regular functions are holomorphic, this proves the statement in this case. The general case is easier.]

5-2. Let V be the disjoint union of algebraic varieties $V_1, \ldots, V_n$. This set has an obvious topology and ringed space structure for which it is an algebraic variety. Show that V is affine if and only if each V_i is affine.

5-3. A monoid variety is an algebraic variety G together with a monoid structure defined by regular maps m and e. Let G be a smooth monoid variety over field of characteristic zero, and assume that $G(k)$ is a group. Show that $(x,y) \mapsto (x,xy)\colon G \times G \to G \times G$ is an isomorphism of algebraic varieties, and deduced that that inv is regular; hence (G,m) is a group variety.

5-4. Let G be an algebraic group. Show:

(a) The neutral element e of G is contained in a unique irreducible component G° of G, which is also the unique connected component of G containing e.

(b) The subvariety G° is a normal subgroup of G of finite index, and every algebraic subgroup of G of finite index contains G°.

5-5. Show that every subgroup variety of a group variety is closed.

5-6. Show that a prevariety V is separated if and only if it satisfies the following condition: a regular map $U \smallsetminus \{P\} \to V$ with U a curve and P a nonsingular point on U extends in at most one way to a regular map $U \to V$.

5-7. Prove the final statement in (5.71).

CHAPTER 6

Projective Varieties

Recall (5.3) that we defined $\mathbb{P}^n$ to be the set of equivalence classes in $k^{n+1} \smallsetminus \{\text{origin}\}$ for the relation

$$(a_0, \ldots, a_n) \sim (b_0, \ldots, b_n) \iff (a_0, \ldots, a_n) = c(b_0, \ldots, b_n) \text{ for some } c \in k^{\times}.$$

Let $(a_0 : \ldots : a_n)$ denote the equivalence class of $(a_0, \ldots, a_n)$, and let π denote the map

$$\frac{k^{n+1} \smallsetminus \{(0, \ldots, 0)\}}{\sim} \to \mathbb{P}^n.$$

Let U_i be the set of $(a_0 : \ldots : a_n) \in \mathbb{P}^n$ such that $a_i \neq 0$, and let u_i be the bijection

$$(a_0 : \ldots : a_n) \mapsto \left(\tfrac{a_0}{a_i}, \ldots, \widehat{\tfrac{a_i}{a_i}}, \ldots, \tfrac{a_n}{a_i}\right) : U_i \xrightarrow{u_i} \mathbb{A}^n \quad (\tfrac{a_i}{a_i} \text{ omitted}).$$

In this chapter, we show that $\mathbb{P}^n$ has a unique structure of an algebraic variety for which these maps become isomorphisms of affine algebraic varieties. A variety isomorphic to a closed subvariety of $\mathbb{P}^n$ is called a ***projective variety***, and a variety isomorphic to a locally closed subvariety of $\mathbb{P}^n$ is called a ***quasiprojective variety***. Every affine variety is quasiprojective, but not all algebraic varieties are quasiprojective. We study morphisms between quasiprojective varieties.

Projective varieties are important for the same reason compact manifolds are important: results are often simpler when stated for projective varieties, and the "part at infinity" often plays a role, even when we would like to ignore it. For example, a famous theorem of Bezout (see 6.37 below) says that a curve of degree m in the projective plane intersects a curve of degree n in exactly mn points (counting multiplicities). For affine curves, one has only an inequality.

a Algebraic subsets of $\mathbb{P}^n$

A polynomial $F(X_0, \ldots, X_n)$ is said to be ***homogeneous of degree*** d if it is a sum of terms $a_{i_0,\ldots,i_n} X_0^{i_0} \cdots X_n^{i_n}$ with $i_0 + \cdots + i_n = d$; equivalently,

$$F(tX_0, \ldots, tX_n) = t^d F(X_0, \ldots, X_n)$$

for all $t \in k$. The polynomials homogeneous of degree d form a subspace $k[X_0, \ldots, X_n]_d$ of $k[X_0, \ldots, X_n]$, and

$$k[X_0, \ldots, X_n] = \bigoplus_{d \geq 0} k[X_0, \ldots, X_n]_d;$$

in other words, every polynomial F can be written uniquely as a sum $F = \sum F_d$ with F_d homogeneous of degree d.

Let $P = (a_0 : \ldots : a_n) \in \mathbb{P}^n$. Then P also equals $(ca_0 : \ldots : ca_n)$ for any $c \in k^\times$, and so we can't speak of the value of a polynomial $F(X_0, \ldots, X_n)$ at P. However, if F is homogeneous, then $F(ca_0, \ldots, ca_n) = c^d F(a_0, \ldots, a_n)$, and so it does make sense to say that F is zero or not zero at P. An ***algebraic set in*** $\mathbb{P}^n$ (or ***projective algebraic set***) is the set of common zeros in $\mathbb{P}^n$ of some set of homogeneous polynomials.

EXAMPLE 6.1. Consider the projective algebraic subset of $\mathbb{P}^2$ defined by the homogeneous equation

$$E : Y^2 Z = X^3 + aXZ^2 + bZ^3. \tag{26}$$

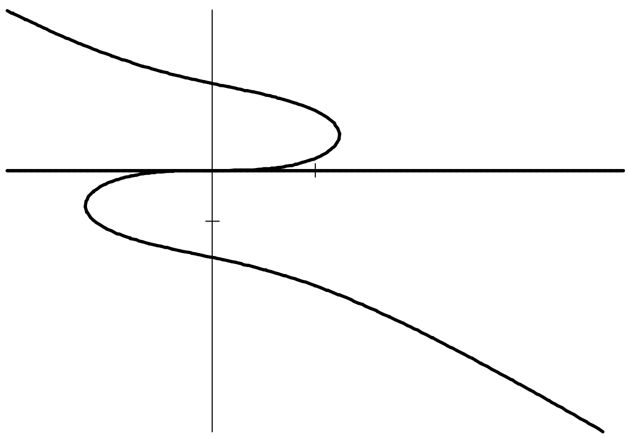
It consists of the points $(x : y : 1)$ on the affine curve $E \cap U_2$

$$Y^2 = X^3 + aX + b$$

(see 2.2) together with the point "at infinity" $(0 : 1 : 0)$. Note that $E \cap U_1$ is the affine curve

$$Z = X^3 + aXZ^2 + bZ^3,$$

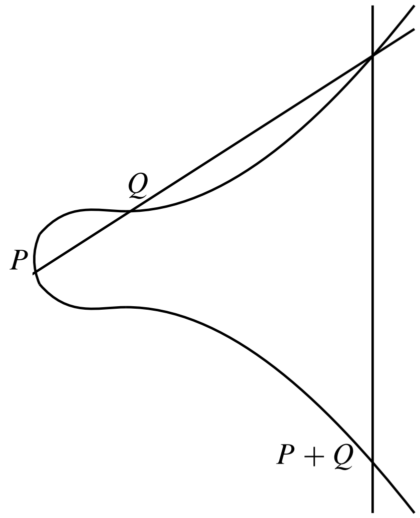
and that $(0{:}1{:}0)$ corresponds to the point $(0,0)$ on $E \cap U_1$:



$$Z = X^3 + XZ^2 + Z^3$$

As $(0,0)$ is nonsingular on $E \cap U_1$, we deduce from (4.5) that E is nonsingular unless $X^3 + aX + b$ has a multiple root. A nonsingular curve of the form (26) is called an ***elliptic curve***.

An elliptic curve has a unique structure of a group variety for which the point at infinity is the zero:



When $a, b \in \mathbb{Q}$, we can speak of the zeros of (26) with coordinates in $\mathbb{Q}$. They also form a group $E(\mathbb{Q})$, which Mordell showed to be finitely generated. It is easy to compute the torsion subgroup of $E(\mathbb{Q})$, but there is at present no known algorithm for computing the rank of $E(\mathbb{Q})$. More precisely, there is an "algorithm" which works in practice, but which has not been proved to always terminate after a finite amount of time. There is a very beautiful theory surrounding elliptic curves over $\mathbb{Q}$ and other number fields, whose origins can be traced back almost 1,800 years to Diophantus. (See my book on *Elliptic Curves* for all of this.)

An ideal $\mathfrak{a} \subset k[X_0, \ldots, X_n]$ is said to be ***graded*** or ***homogeneous*** if it contains with any polynomial F all the homogeneous components of F, i.e., if

$$F \in \mathfrak{a} \implies F_d \in \mathfrak{a}, \text{ all } d.$$

It is straightforward to check that

- ◇ an ideal is graded if and only if it is generated by (a finite set of) homogeneous polynomials;
- ◇ the radical of a graded ideal is graded;
- ◇ an intersection, product, or sum of graded ideals is graded.

For a graded ideal $\mathfrak{a}$, we let $V(\mathfrak{a})$ denote the set of common zeros of the homogeneous polynomials in $\mathfrak{a}$. Clearly

$$\mathfrak{a} \subset \mathfrak{b} \implies V(\mathfrak{a}) \supset V(\mathfrak{b}).$$

If $F_1, \ldots, F_r$ are homogeneous generators for $\mathfrak{a}$, then $V(\mathfrak{a})$ is also the set of common zeros of the F_i. Clearly every polynomial in $\mathfrak{a}$ is zero on every representative of a point in $V(\mathfrak{a})$. We write $V^{\mathrm{aff}}(\mathfrak{a})$ for the set of common zeros of $\mathfrak{a}$ in k^{n+1}. It is a ***cone*** in k^{n+1}, i.e., together with any point P it contains the line through P and the origin, and

$$V(\mathfrak{a}) = \frac{V^{\mathrm{aff}}(\mathfrak{a}) \smallsetminus \{(0, \ldots, 0)\}}{\sim}.$$

The sets $V(\mathfrak{a})$ in $\mathbb{P}^n$ have similar properties to their namesakes in $\mathbb{A}^n$.

PROPOSITION 6.2. *There are the following relations:*

(a) $V(0)=\mathbb{P}^n$; $\quad V(\mathfrak{a})=\varnothing \iff \mathrm{rad}(\mathfrak{a})\supset(X_0,\ldots,X_n)$;

(b) $V(\mathfrak{a}\mathfrak{b})=V(\mathfrak{a}\cap\mathfrak{b})=V(\mathfrak{a})\cup V(\mathfrak{b})$;

(c) $V(\sum\mathfrak{a}_i)=\bigcap V(\mathfrak{a}_i)$.

PROOF. For the second statement in (a), note that

$$\begin{aligned} V(\mathfrak{a})=\emptyset &\iff V^{\mathrm{aff}}(\mathfrak{a})\subset\{(0,\ldots,0)\}\\ &\iff \mathrm{rad}(\mathfrak{a})\supset(X_0,\ldots,X_n) \qquad \text{(strong Nullstellensatz 2.16).}\end{aligned}$$

The remaining statements can be proved directly, as in (2.10), or by using the relation between $V(\mathfrak{a})$ and $V^{\mathrm{aff}}(\mathfrak{a})$. □

Proposition 6.2 shows that the projective algebraic sets are the closed sets for a topology on $\mathbb{P}^n$. This topology is called the ***Zariski topology*** on $\mathbb{P}^n$.

If C is a cone in k^{n+1}, then $I(C)$ is a graded ideal in $k[X_0,\ldots,X_n]$: if $F(ca_0,\ldots,ca_n)=0$ for all $c\in k^{\times}$, then

$$\sum_d F_d(a_0,\ldots,a_n)\cdot c^d = F(ca_0,\ldots,ca_n)=0,$$

for infinitely many c, and so $\sum F_d(a_0,\ldots,a_n)X^d$ is the zero polynomial. For a subset S of $\mathbb{P}^n$, we define the ***affine cone over*** S in k^{n+1} to be

$$C=\pi^{-1}(S)\cup\{\text{origin}\}$$

and we set

$$I(S)=I(C).$$

Note that if S is nonempty and closed, then C is the closure of $\pi^{-1}(S)=\emptyset$, and that $I(S)$ is spanned by the homogeneous polynomials in $k[X_0,\ldots,X_n]$ that are zero on S.

PROPOSITION 6.3. *The maps V and I define inverse bijections between the set of algebraic subsets of $\mathbb{P}^n$ and the set of proper graded radical ideals of $k[X_0,\ldots,X_n]$. An algebraic set V in $\mathbb{P}^n$ is irreducible if and only if $I(V)$ is prime; in particular, $\mathbb{P}^n$ is irreducible.*

PROOF. Note that we have bijections

$$\begin{tikzcd}
\{\text{algebraic subsets of } \mathbb{P}^n\} \arrow[rr, "S\mapsto C"] & & \{\text{nonempty closed cones in } k^{n+1}\} \arrow[dl, "I"] \\
& \{\text{proper graded radical ideals in } k[X_0,\ldots,X_n]\} \arrow[ul, "V"] &
\end{tikzcd}$$

Here the top map sends S to the affine cone over S, and the maps V and I are in the sense of projective geometry and affine geometry respectively. The composite of any three of these maps is the identity map, which proves the first statement because the composite of the top map with I is I in the sense of projective geometry. Obviously, V is irreducible if and only if the closure of $\pi^{-1}(V)$ is irreducible, which is true if and only if $I(V)$ is a prime ideal. □

Note that the graded ideals $(X_0,\ldots,X_n)$ and $k[X_0,\ldots,X_n]$ are both radical, but

$$V(X_0,\ldots,X_n) = \emptyset = V(k[X_0,\ldots,X_n])$$

and so the correspondence between irreducible subsets of $\mathbb{P}^n$ and radical graded ideals is not quite one-to-one.

ASIDE 6.4. In English "homogeneous ideal" is more common than "graded ideal", but we follow Bourbaki, Alg, II, §11. A ***graded ring*** is a pair $(S,(S_d)_{d\in\mathbb{N}})$ consisting of a ring S and a family of additive subgroups S_d such that

$$\begin{cases} S = \bigoplus_{d\in\mathbb{N}} S_d \\ S_d S_e \subset S_{d+e}, \text{ all } d,e\in\mathbb{N}. \end{cases}$$

An ideal $\mathfrak{a}$ in S is ***graded*** if and only if

$$\mathfrak{a} = \bigoplus_{d\in\mathbb{N}} (\mathfrak{a}\cap S_d);$$

this means that it is a graded submodule of $(S,(S_d))$. The quotient of a graded ring S by a graded ideal $\mathfrak{a}$ is a graded ring $S/\mathfrak{a} = \bigoplus_d S_d/(\mathfrak{a}\cap S_d)$.

b The Zariski topology on $\mathbb{P}^n$

For a graded polynomial F, let

$$D(F) = \{P\in\mathbb{P}^n \mid F(P)\neq 0\}.$$

Then, just as in the affine case, $D(F)$ is open and the sets of this type form a base for the topology of $\mathbb{P}^n$. As in the opening paragraph of this chapter, we let $U_i = D(X_i)$.

To each polynomial $f(X_1,\ldots,X_n)$, we attach the homogeneous polynomial of the same degree

$$f^*(X_0,\ldots,X_n) = X_0^{\deg(f)} f\left(\frac{X_1}{X_0},\ldots,\frac{X_n}{X_0}\right),$$

and to each homogeneous polynomial $F(X_0,\ldots,X_n)$, we attach the polynomial

$$F_*(X_1,\ldots,X_n) = F(1,X_1,\ldots,X_n).$$

PROPOSITION 6.5. *Each subset U_i of $\mathbb{P}^n$ is open in the Zariski topology on $\mathbb{P}^n$, and when we endow it with the induced topology, the bijection*

$$U_i \leftrightarrow \mathbb{A}^n,\ (a_0:\ldots:1:\ldots:a_n)\leftrightarrow(a_0,\ldots,a_{i-1},a_{i+1},\ldots,a_n)$$

becomes a homeomorphism.

PROOF. It suffices to prove this with $i=0$. The set $U_0 = D(X_0)$, and so it is a basic open subset in $\mathbb{P}^n$. Clearly, for any homogeneous polynomial $F\in k[X_0,\ldots,X_n]$,

$$D(F(X_0,\ldots,X_n))\cap U_0 = D(F(1,X_1,\ldots,X_n)) = D(F_*)$$

and, for any polynomial $f\in k[X_1,\ldots,X_n]$,

$$D(f) = D(f^*)\cap U_0.$$

Thus, under the bijection $U_0\leftrightarrow\mathbb{A}^n$, the basic open subsets of $\mathbb{A}^n$ correspond to the intersections with U_i of the basic open subsets of $\mathbb{P}^n$, which proves that the bijection is a homeomorphism. □

REMARK 6.6. It is possible to use this to give a different proof that $\mathbb{P}^n$ is irreducible. We apply the criterion that a space is irreducible if and only if every nonempty open subset is dense (see p.44). Note that each U_i is irreducible, and that $U_i \cap U_j$ is open and dense in each of U_i and U_j (as a subset of U_i, it is the set of points $(a_0 : \ldots : 1 : \ldots : a_j : \ldots : a_n)$ with $a_j \neq 0$). Let U be a nonempty open subset of $\mathbb{P}^n$; then $U \cap U_i$ is open in U_i. For some i, $U \cap U_i$ is nonempty, and so must meet $U_i \cap U_j$. Therefore U meets every U_j, and so is dense in every U_j. It follows that its closure is all of $\mathbb{P}^n$.

c Closed subsets of $\mathbb{A}^n$ and $\mathbb{P}^n$

We identify $\mathbb{A}^n$ with U_0, and examine the closures in $\mathbb{P}^n$ of closed subsets of $\mathbb{A}^n$. Note that

$$\mathbb{P}^n = \mathbb{A}^n \sqcup H_\infty, \quad H_\infty = V(X_0).$$

With each ideal $\mathfrak{a}$ in $k[X_1, \ldots, X_n]$, we associate the graded ideal $\mathfrak{a}^*$ in $k[X_0, \ldots, X_n]$ generated by $\{f^* \mid f \in \mathfrak{a}\}$. For a closed subset V of $\mathbb{A}^n$, set $V^* = V(\mathfrak{a}^*)$ with $\mathfrak{a} = I(V)$.

With each graded ideal $\mathfrak{a}$ in $k[X_0, X_1, \ldots, X_n]$, we associate the ideal $\mathfrak{a}_*$ in $k[X_1, \ldots, X_n]$ generated by $\{F_* \mid F \in \mathfrak{a}\}$. When V is a closed subset of $\mathbb{P}^n$, we set $V_* = V(\mathfrak{a}_*)$ with $\mathfrak{a} = I(V)$.

PROPOSITION 6.7. *(a) Let V be a closed subset of $\mathbb{A}^n$. Then V^* is the closure of V in $\mathbb{P}^n$, and $(V^*)_* = V$. If $V = \bigcup V_i$ is the decomposition of V into its irreducible components, then $V^* = \bigcup V_i^*$ is the decomposition of V^* into its irreducible components.*

(b) Let V be a closed subset of $\mathbb{P}^n$. Then $V_* = V \cap \mathbb{A}^n$, and if no irreducible component of V lies in H_∞ or contains H_∞, then V_* is a proper subset of $\mathbb{A}^n$, and $(V_*)^* = V$.

PROOF. Straightforward. □

Examples

6.8. For

$$V : Y^2 = X^3 + aX + b,$$

we have

$$V^* : Y^2 Z = X^3 + aXZ^2 + bZ^3,$$

and $(V^*)_* = V$.

6.9. Let $V = V(f_1, \ldots, f_m)$; then the closure of V in $\mathbb{P}^n$ is the union of the irreducible components of $V(f_1^*, \ldots, f_m^*)$ not contained in H_∞. For example, let

$$V = V(X_1, X_1^2 + X_2) = \{(0,0)\};$$

then $V(X_0 X_1, X_1^2 + X_0 X_2)$ consists of the two points $(1{:}0{:}0)$ (the closure of V) and $(0{:}0{:}1)$ (which is contained in H_∞).[1]

6.10. For $V = H_\infty = V(X_0)$, we have $V_* = \emptyset = V(1)$ and $(V_*)^* = \emptyset \neq V$.

[1] Of course, in this case $\mathfrak{a} = (X_1, X_2)$, $\mathfrak{a}^* = (X_1, X_2)$, and $V^* = \{(1{:}0{:}0)\}$, and so this example doesn't contradict the proposition.

d The hyperplane at infinity

It is often convenient to think of $\mathbb{P}^n$ as being $\mathbb{A}^n = U_0$ with a hyperplane added "at infinity". More precisely, we identify the set U_0 with $\mathbb{A}^n$; the complement of U_0 in $\mathbb{P}^n$ is

$$H_\infty = \{(0:a_1:\ldots:a_n) \subset \mathbb{P}^n\},$$

which can be identified with $\mathbb{P}^{n-1}$.

For example, $\mathbb{P}^1 = \mathbb{A}^1 \sqcup H_\infty$ (disjoint union), with H_∞ consisting of a single point, and $\mathbb{P}^2 = \mathbb{A}^2 \cup H_\infty$ with H_∞ a projective line. Consider the line

$$1 + aX_1 + bX_2 = 0$$

in $\mathbb{A}^2$. Its closure in $\mathbb{P}^2$ is the line

$$X_0 + aX_1 + bX_2 = 0.$$

This line intersects the line $H_\infty = V(X_0)$ at the point $(0:-b:a)$, which equals $(0:1:-a/b)$ when $b \neq 0$. Note that $-a/b$ is the slope of the line $1 + aX_1 + bX_2 = 0$, and so the point at which a line intersects H_∞ depends only on the slope of the line: parallel lines meet in one point at infinity. We can think of the projective plane $\mathbb{P}^2$ as being the affine plane $\mathbb{A}^2$ with one point added at infinity for each "direction" in $\mathbb{A}^2$.

Similarly, we can think of $\mathbb{P}^n$ as being $\mathbb{A}^n$ with one point added at infinity for each direction in $\mathbb{A}^n$ — being parallel is an equivalence relation on the lines in $\mathbb{A}^n$, and there is one point at infinity for each equivalence class of lines.

We can replace U_0 with U_n in the above discussion, and write $\mathbb{P}^n = U_n \sqcup H_\infty$ with $H_\infty = \{(a_0:\ldots:a_{n-1}:0)$, as in Example 6.1. Note that in this example the point at infinity on the elliptic curve $Y^2 = X^3 + aX + b$ is the intersection of the closure of any vertical line with H_∞.

e $\mathbb{P}^n$ is an algebraic variety

For each i, write $\mathcal{O}_i$ for the sheaf on $U_i \subset \mathbb{P}^n$ defined by the homeomorphism $u_i : U_i \to \mathbb{A}^n$.

LEMMA 6.11. *Write $U_{ij} = U_i \cap U_j$; then $\mathcal{O}_i|U_{ij} = \mathcal{O}_j|U_{ij}$. When endowed with this sheaf, U_{ij} is an affine algebraic variety; moreover, $\Gamma(U_{ij}, \mathcal{O}_i)$ is generated as a k-algebra by the functions $(f|U_{ij})(g|U_{ij})$ with $f \in \Gamma(U_i, \mathcal{O}_i)$, $g \in \Gamma(U_j, \mathcal{O}_j)$.*

PROOF. It suffices to prove this for $(i,j) = (0,1)$. All rings occurring in the proof will be identified with subrings of the field $k(X_0, X_1, \ldots, X_n)$.

Recall that

$$U_0 = \{(a_0:a_1:\ldots:a_n) \mid a_0 \neq 0\};\ (a_0:a_1:\ldots:a_n) \leftrightarrow (\tfrac{a_1}{a_0}, \tfrac{a_2}{a_0}, \ldots, \tfrac{a_n}{a_0}) \in \mathbb{A}^n.$$

Let $k[\frac{X_1}{X_0}, \frac{X_2}{X_0}, \ldots, \frac{X_n}{X_0}]$ be the subring of $k(X_0, X_1, \ldots, X_n)$ generated by the quotients $\frac{X_i}{X_0}$ — it is the polynomial ring in the n symbols $\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}$. An element $f(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}) \in k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}]$ defines a map

$$(a_0:a_1:\ldots:a_n) \mapsto f(\tfrac{a_1}{a_0}, \ldots, \tfrac{a_n}{a_0}) : U_0 \to k,$$

and in this way $k[\frac{X_1}{X_0}, \frac{X_2}{X_0}, \ldots, \frac{X_n}{X_0}]$ becomes identified with the ring of regular functions on U_0, and U_0 with $\mathrm{Spm}\left(k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}]\right)$.

Next consider the open subset of U_0,

$$U_{01} = \{(a_0 : \ldots : a_n) \mid a_0 \neq 0, a_1 \neq 0\}.$$

It is $D(\frac{X_1}{X_0})$, and is therefore an affine subvariety of $(U_0, \mathcal{O}_0)$. The inclusion $U_{01} \hookrightarrow U_0$ corresponds to the inclusion of rings $k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}] \hookrightarrow k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}, \frac{X_0}{X_1}]$. An element $f(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}, \frac{X_0}{X_1})$ of $k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}, \frac{X_0}{X_1}]$ defines the function $(a_0 : \ldots : a_n) \mapsto f(\frac{a_1}{a_0}, \ldots, \frac{a_n}{a_0}, \frac{a_0}{a_1})$ on U_{01}.

Similarly,

$$U_1 = \{(a_0 : a_1 : \ldots : a_n) \mid a_1 \neq 0\}; \; (a_0 : a_1 : \ldots : a_n) \leftrightarrow (\tfrac{a_0}{a_1}, \ldots, \tfrac{a_n}{a_1}) \in \mathbb{A}^n,$$

and we identify U_1 with $\mathrm{Spm}\left(k[\frac{X_0}{X_1}, \frac{X_2}{X_0}, \ldots, \frac{X_n}{X_1}]\right)$. A polynomial $f(\frac{X_0}{X_1}, \ldots, \frac{X_n}{X_1})$ in $k[\frac{X_0}{X_1}, \ldots, \frac{X_n}{X_1}]$ defines the map $(a_0 : \ldots : a_n) \mapsto f(\frac{a_0}{a_1}, \ldots, \frac{a_n}{a_1}) \colon U_1 \to k$.

When regarded as an open subset of U_1, $U_{01} = D(\frac{X_0}{X_1})$, and is therefore an affine subvariety of $(U_1, \mathcal{O}_1)$, and the inclusion $U_{01} \hookrightarrow U_1$ corresponds to the inclusion of rings $k[\frac{X_0}{X_1}, \ldots, \frac{X_n}{X_1}] \hookrightarrow k[\frac{X_0}{X_1}, \ldots, \frac{X_n}{X_1}, \frac{X_1}{X_0}]$. An element $f(\frac{X_0}{X_1}, \ldots, \frac{X_n}{X_1}, \frac{X_1}{X_0})$ of $k[\frac{X_0}{X_1}, \ldots, \frac{X_n}{X_1}, \frac{X_1}{X_0}]$ defines the function $(a_0 : \ldots : a_n) \mapsto f(\frac{a_0}{a_1}, \ldots, \frac{a_n}{a_1}, \frac{a_1}{a_0})$ on U_{01}.

The two subrings $k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}, \frac{X_0}{X_1}]$ and $k[\frac{X_0}{X_1}, \ldots, \frac{X_n}{X_1}, \frac{X_1}{X_0}]$ of $k(X_0, X_1, \ldots, X_n)$ are equal, and an element of this ring defines the same function on U_{01} regardless of which of the two rings it is considered an element. Therefore, whether we regard U_{01} as a subvariety of U_0 or of U_1 it inherits the same structure as an affine algebraic variety (3.15). This proves the first two assertions, and the third is obvious: $k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}, \frac{X_0}{X_1}]$ is generated by its subrings $k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}]$ and $k[\frac{X_0}{X_1}, \frac{X_2}{X_1}, \ldots, \frac{X_n}{X_1}]$. □

PROPOSITION 6.12. *There is a unique structure of a (separated) algebraic variety on $\mathbb{P}^n$ for which each U_i is an open affine subvariety of $\mathbb{P}^n$ and each map u_i is an isomorphism of algebraic varieties.*

PROOF. Endow each U_i with the structure of an affine algebraic variety for which u_i is an isomorphism. Then $\mathbb{P}^n = \bigcup U_i$, and the lemma shows that this covering satisfies the patching condition (5.15), and so $\mathbb{P}^n$ has a unique structure of a ringed space for which $U_i \hookrightarrow \mathbb{P}^n$ is a homeomorphism onto an open subset of $\mathbb{P}^n$ and $\mathcal{O}_{\mathbb{P}^n}|U_i = \mathcal{O}_{U_i}$. Moreover, because each U_i is an algebraic variety, this structure makes $\mathbb{P}^n$ into an algebraic prevariety. Finally, the lemma shows that $\mathbb{P}^n$ satisfies the condition (5.29c) to be separated. □

EXAMPLE 6.13. Let C be the plane projective curve

$$C \colon Y^2 Z = X^3$$

and assume that $\mathrm{char}(k) \neq 2$. For each $a \in k^\times$, there is an automorphism

$$(x : y : z) \mapsto (ax : y : a^3 z) \colon C \xrightarrow{\varphi_a} C.$$

Patch two copies of $C \times \mathbb{A}^1$ together along $C \times (\mathbb{A}^1 - \{0\})$ by identifying (P, a) with $(\varphi_a(P), a^{-1})$, $P \in C$, $a \in \mathbb{A}^1 \smallsetminus \{0\}$. One obtains in this way a singular surface that is not quasiprojective (see Hartshorne 1977, Exercise 7.13). It is even complete — see below — and so if it were quasiprojective, it would be projective. In Shafarevich 1994, VI 2.3, there is an example of a nonsingular complete variety of dimension 3 that is not projective. It is known that every irreducible separated curve is quasiprojective, and every nonsingular complete surface is projective, and so these examples are minimal.

f The homogeneous coordinate ring of a projective variety

Recall (p.112) that attached to each irreducible variety V, there is a field $k(V)$ with the property that $k(V)$ is the field of fractions of $k[U]$ for any open affine $U \subset V$. We now describe this field in the case that $V = \mathbb{P}^n$. Recall that $k[U_0] = k[\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}]$. We regard this as a subring of $k(X_0, \ldots, X_n)$, and wish to identify the field of fractions of $k[U_0]$ as a subfield of $k(X_0, \ldots, X_n)$. Every nonzero $F \in k[U_0]$ can be written

$$F(\tfrac{X_1}{X_0}, \ldots, \tfrac{X_n}{X_0}) = \frac{F^*(X_0, \ldots, X_n)}{X_0^{\deg(F)}}$$

with F^* homogeneous of degree $\deg(F)$, and it follows that the field of fractions of $k[U_0]$ is

$$k(U_0) = \left\{ \frac{G(X_0, \ldots, X_n)}{H(X_0, \ldots, X_n)} \;\middle|\; G, H \text{ homogeneous of the same degree} \right\} \cup \{0\}.$$

Write $k(X_0, \ldots, X_n)_0$ for this field (the subscript 0 is short for "subfield of elements of degree 0"), so that $k(\mathbb{P}^n) = k(X_0, \ldots, X_n)_0$. Note that for $F = \frac{G}{H}$ in $k(X_0, \ldots, X_n)_0$,

$$(a_0 : \ldots : a_n) \mapsto \frac{G(a_0, \ldots, a_n)}{H(a_0, \ldots, a_n)} : D(H) \to k,$$

is a well-defined function, which is obviously regular (look at its restriction to U_i).

We now extend this discussion to any irreducible projective variety V. Such a V can be written $V = V(\mathfrak{p})$ with $\mathfrak{p}$ a graded radical ideal in $k[X_0, \ldots, X_n]$, and we define the ***homogeneous*** coordinate ring of V (with its given embedding) to be

$$k_{\text{hom}}[V] = k[X_0, \ldots, X_n]/\mathfrak{p}.$$

Note that $k_{\text{hom}}[V]$ is the ring of regular functions on the affine cone over V; therefore its dimension is $\dim(V) + 1$. It depends, not only on V, but on the embedding of V into $\mathbb{P}^n$, i.e., it is not intrinsic to V. For example,

$$(a_0 : a_1) \mapsto (a_0^2 : a_0 a_1 : a_1^2) : \mathbb{P}^1 \xrightarrow{\nu} \mathbb{P}^2$$

is an isomorphism from $\mathbb{P}^1$ onto its image $\nu(\mathbb{P}^1)$: $X_0 X_2 = X_1^2$ (see 6.23 below), but $k_{\text{hom}}[\mathbb{P}^1] = k[X_0, X_1]$, which is the affine coordinate ring of the smooth variety $\mathbb{A}^2$, whereas $k_{\text{hom}}[\nu(\mathbb{P}^1)] = k[X_0, X_1, X_2]/(X_0 X_2 - X_1^2)$, which is the affine coordinate ring of the singular variety $X_0 X_2 - X_1^2$.

We say that a nonzero $f \in k_{\text{hom}}[V]$ is ***homogeneous of degree*** d if it can be represented by a homogeneous polynomial F of degree d in $k[X_0, \ldots, X_n]$, and we say that 0 is homogeneous of degree 0.

LEMMA 6.14. *Each element of $k_{\text{hom}}[V]$ can be written uniquely in the form*

$$f = f_0 + \cdots + f_d$$

with f_i homogeneous of degree i.

PROOF. Let F represent f; then F can be written $F = F_0 + \cdots + F_d$ with F_i homogeneous of degree i; when read modulo $\mathfrak{p}$, this gives a decomposition of f of the required type. Suppose f also has a decomposition $f = \sum g_i$, with g_i represented by the homogeneous polynomial G_i of degree i. Then $F - G \in \mathfrak{p}$, and the homogeneity of $\mathfrak{p}$ implies that $F_i - G_i = (F-G)_i \in \mathfrak{p}$. Therefore $f_i = g_i$. □

It therefore makes sense to speak of homogeneous elements of $k[V]$. For such an element h, we define $D(h) = \{P \in V \mid h(P) \neq 0\}$.

Since $k_{\text{hom}}[V]$ is an integral domain, we can form its field of fractions $k_{\text{hom}}(V)$. Define

$$k_{\text{hom}}(V)_0 = \left\{ \frac{g}{h} \in k_{\text{hom}}(V) \;\middle|\; g \text{ and } h \text{ homogeneous of the same degree} \right\} \cup \{0\}.$$

PROPOSITION 6.15. *The field of rational functions on V is $k(V) \stackrel{\text{def}}{=} k_{\text{hom}}(V)_0$.*

PROOF. Consider $V_0 \stackrel{\text{def}}{=} U_0 \cap V$. As in the case of $\mathbb{P}^n$, we can identify $k[V_0]$ with a subring of $k_{\text{hom}}[V]$, and then the field of fractions of $k[V_0]$ becomes identified with $k_{\text{hom}}(V)_0$. □

g Regular functions on a projective variety

Let V be an irreducible projective variety, and let $f \in k(V)$. By definition, we can write $f = \frac{g}{h}$ with g and h homogeneous of the same degree in $k_{\text{hom}}[V]$ and $h \neq 0$. For any $P = (a_0 : \ldots : a_n)$ with $h(P) \neq 0$,

$$f(P) \stackrel{\text{def}}{=} \frac{g(a_0, \ldots, a_n)}{h(a_0, \ldots, a_n)}$$

is well-defined: if $(a_0, \ldots, a_n)$ is replaced by $(ca_0, \ldots, ca_n)$, then both the numerator and denominator are multiplied by $c^{\deg(g)} = c^{\deg(h)}$.

We can write f in the form $\frac{g}{h}$ in many different ways,[2] but if

$$f = \frac{g}{h} = \frac{g'}{h'} \quad (\text{in } k(V)_0),$$

then

$$gh' - g'h \quad (\text{in } k_{\text{hom}}[V])$$

and so

$$g(a_0, \ldots, a_n) \cdot h'(a_0, \ldots, a_n) = g'(a_0, \ldots, a_n) \cdot h(a_0, \ldots, a_n).$$

Thus, if $h'(P) \neq 0$, the two representations give the same value for $f(P)$.

PROPOSITION 6.16. *For each $f \in k(V) \stackrel{\text{def}}{=} k_{\text{hom}}(V)_0$, there is an open subset U of V where $f(P)$ is defined, and $P \mapsto f(P)$ is a regular function on U; every regular function on an open subset of V arises from a unique element of $k(V)$.*

PROOF. From the above discussion, we see that f defines a regular function on $U = \bigcup D(h)$ where h runs over the denominators of expressions $f = \frac{g}{h}$ with g and h homogeneous of the same degree in $k_{\text{hom}}[V]$.

Conversely, let f be a regular function on an open subset U of V, and let $P \in U$. Then P lies in the open affine subvariety $V \cap U_i$ for some i, and so f coincides with the function

[2] Unless $k_{\text{hom}}[V]$ is a unique factorization domain, there will be no preferred representation $f = \frac{g}{h}$.

defined by some $f_P \in k(V \cap U_i) = k(V)$ on an open neighbourhood of P. If f coincides with the function defined by $f_Q \in k(V)$ in a neighbourhood of a second point Q of U, then f_P and f_Q define the same function on some open affine U', and so $f_P = f_Q$ as elements of $k[U'] \subset k(V)$. This shows that f is the function defined by f_P on the whole of U. □

REMARK 6.17. (a) The elements of $k(V) = k_{\mathrm{hom}}(V)_0$ should be regarded as the algebraic analogues of meromorphic functions on a complex manifold; the regular functions on an open subset U of V are the "meromorphic functions without poles" on U. [In fact, when $k = \mathbb{C}$, this is more than an analogy: a nonsingular projective algebraic variety over $\mathbb{C}$ defines a complex manifold, and the meromorphic functions on the manifold are precisely the rational functions on the variety. For example, the meromorphic functions on the Riemann sphere are the rational functions in z.]

(b) We shall see presently (6.24) that, for any nonzero homogeneous $h \in k_{\mathrm{hom}}[V]$, $D(h)$ is an open affine subset of V. The ring of regular functions on it is

$$k[D(h)] = \{g/h^m \mid g \text{ homogeneous of degree } m \deg(h)\} \cup \{0\}.$$

We shall also see that the ring of regular functions on V itself is just k, i.e., any regular function on an irreducible (connected will do) projective variety is constant. However, if U is an open nonaffine subset of V, then the ring $\Gamma(U, \mathcal{O}_V)$ of regular functions can be almost anything — it needn't even be a finitely generated k-algebra!

h Maps from projective varieties

We describe the morphisms from a projective variety to another variety.

PROPOSITION 6.18. *The map*

$$\pi : \mathbb{A}^{n+1} \smallsetminus \{\text{origin}\} \to \mathbb{P}^n,\ (a_0, \ldots, a_n) \mapsto (a_0 : \ldots : a_n)$$

is an open morphism of algebraic varieties. A map $\alpha : \mathbb{P}^n \to V$ with V a prevariety is regular if and only if $\alpha \circ \pi$ is regular.

PROOF. The restriction of π to $D(X_i)$ is the projection

$$(a_0, \ldots, a_n) \mapsto (\tfrac{a_0}{a_i} : \ldots : \tfrac{a_n}{a_i}) : k^{n+1} \smallsetminus V(X_i) \to U_i,$$

which is the regular map of affine varieties corresponding to the map of k-algebras

$$k\left[\frac{X_0}{X_i}, \ldots, \frac{X_n}{X_i}\right] \to k[X_0, \ldots, X_n][X_i^{-1}].$$

(In the first algebra $\frac{X_j}{X_i}$ is to be thought of as a single symbol.) It now follows from (5.4) that π is regular.

Let U be an open subset of $k^{n+1} \smallsetminus \{\text{origin}\}$, and let U' be the union of all the lines through the origin that meet U, that is, $U' = \pi^{-1}\pi(U)$. Then U' is again open in $k^{n+1} \smallsetminus \{\text{origin}\}$, because $U' = \bigcup cU$, $c \in k^\times$, and $x \mapsto cx$ is an automorphism of $k^{n+1} \smallsetminus \{\text{origin}\}$. The complement Z of U' in $k^{n+1} \smallsetminus \{\text{origin}\}$ is a closed cone, and the proof of (6.3) shows that its image is closed in $\mathbb{P}^n$; but $\pi(U)$ is the complement of $\pi(Z)$. Thus π sends open sets to open sets.

The rest of the proof is straightforward. □

Thus, the regular maps $\mathbb{P}^n \to V$ are just the regular maps $\mathbb{A}^{n+1} \smallsetminus \{\text{origin}\} \to V$ factoring through $\mathbb{P}^n$ (as maps of sets).

REMARK 6.19. Consider polynomials $F_0(X_0,\ldots,X_m),\ldots,F_n(X_0,\ldots,X_m)$ of the same degree. The map

$$(a_0:\ldots:a_m) \mapsto (F_0(a_0,\ldots,a_m):\ldots:F_n(a_0,\ldots,a_m))$$

obviously defines a regular map to $\mathbb{P}^n$ on the open subset of $\mathbb{P}^m$ where not all F_i vanish, that is, on the set $\bigcup D(F_i) = \mathbb{P}^n \smallsetminus V(F_1,\ldots,F_n)$. Its restriction to any subvariety V of $\mathbb{P}^m$ will also be regular. It may be possible to extend the map to a larger set by representing it by different polynomials. Conversely, every such map arises in this way, at least locally. More precisely, there is the following result.

PROPOSITION 6.20. *Let $V = V(\mathfrak{a}) \subset \mathbb{P}^m$ and $W = V(\mathfrak{b}) \subset \mathbb{P}^n$. A map $\varphi\colon V \to W$ is regular if and only if, for every $P \in V$, there exist polynomials*

$$F_0(X_0,\ldots,X_m),\ldots,F_n(X_0,\ldots,X_m),$$

homogeneous of the same degree, such that

$$\varphi((b_0:\ldots:b_n)) = (F_0(b_0,\ldots,b_m):\ldots:F_n(b_0,\ldots,b_m))$$

for all points $(b_0:\ldots:b_m)$ in some neighbourhood of P in $V(\mathfrak{a})$.

PROOF. Straightforward. □

EXAMPLE 6.21. We prove that the circle $X^2+Y^2=Z^2$ is isomorphic to $\mathbb{P}^1$. This equation can be rewritten $(X+iY)(X-iY)=Z^2$, and so, after a change of variables, the equation of the circle becomes $C: XZ=Y^2$. Define

$$\varphi\colon \mathbb{P}^1 \to C,\ (a:b) \mapsto (a^2:ab:b^2).$$

For the inverse, define

$$\psi\colon C \to \mathbb{P}^1 \quad \text{by} \begin{cases} (a:b:c) \mapsto (a:b) & \text{if } a \neq 0 \\ (a:b:c) \mapsto (b:c) & \text{if } b \neq 0 \end{cases}.$$

Note that,

$$a \neq 0 \neq b, \quad ac = b^2 \implies \frac{c}{b} = \frac{b}{a}$$

and so the two maps agree on the set where they are both defined. Clearly, both φ and ψ are regular, and one checks directly that they are inverse.

i Some classical maps of projective varieties

We list some of the classic maps.

HYPERPLANE SECTIONS AND COMPLEMENTS

6.22. Let $L = \sum c_i X_i$ be a nonzero linear form in $n+1$ variables. Then the map

$$(a_0 : \ldots : a_n) \mapsto \left(\frac{a_0}{L(\mathbf{a})}, \ldots, \frac{a_n}{L(\mathbf{a})}\right)$$

is a bijection of $D(L) \subset \mathbb{P}^n$ onto the hyperplane $L(X_0, X_1, \ldots, X_n) = 1$ of $\mathbb{A}^{n+1}$, with inverse

$$(a_0, \ldots, a_n) \mapsto (a_0 : \ldots : a_n).$$

Both maps are regular — for example, the components of the first map are the regular functions $\frac{X_j}{\sum c_i X_i}$. As $V(L-1)$ is affine, so also is $D(L)$, and its ring of regular functions is $k[\frac{X_0}{\sum c_i X_i}, \ldots, \frac{X_n}{\sum c_i X_i}]$. In this ring, each quotient $\frac{X_j}{\sum c_i X_i}$ is to be thought of as a single symbol, and $\sum c_j \frac{X_j}{\sum c_i X_i} = 1$; thus it is a polynomial ring in n symbols; any one symbol $\frac{X_j}{\sum c_i X_i}$ for which $c_j \neq 0$ can be omitted.

For a fixed $P = (a_0 : \ldots : a_n) \in \mathbb{P}^n$, the set of $\mathbf{c} = (c_0 : \ldots : c_n)$ such that

$$L_{\mathbf{c}}(P) \stackrel{\text{def}}{=} \sum c_i a_i \neq 0$$

is a nonempty open subset of $\mathbb{P}^n$ ($n > 0$). Therefore, for any finite set S of points of $\mathbb{P}^n$,

$$\{\mathbf{c} \in \mathbb{P}^n \mid S \subset D(L_{\mathbf{c}})\}$$

is a nonempty open subset of $\mathbb{P}^n$ (because $\mathbb{P}^n$ is irreducible). In particular, S is contained in an open affine subset $D(L_{\mathbf{c}})$ of $\mathbb{P}^n$. Moreover, if $S \subset V$ where V is a closed subvariety of $\mathbb{P}^n$, then $S \subset V \cap D(L_{\mathbf{c}})$: any finite set of points of a projective variety is contained in an open affine subvariety.

THE VERONESE MAP; HYPERSURFACE SECTIONS

6.23. Let

$$I = \{(i_0, \ldots, i_n) \in \mathbb{N}^{n+1} \mid \sum i_j = m\}.$$

Note that I indexes the monomials of degree m in $n+1$ variables. It has $\binom{m+n}{m}$ elements[3]. Write $\nu_{n,m} = \binom{m+n}{m} - 1$, and consider the projective space $\mathbb{P}^{\nu_{n,m}}$ whose coordinates are indexed by I; thus a point of $\mathbb{P}^{\nu_{n,m}}$ can be written $(\ldots : b_{i_0 \ldots i_n} : \ldots)$. The Veronese mapping is defined to be

$$v : \mathbb{P}^n \to \mathbb{P}^{\nu_{n,m}},\ (a_0 : \ldots : a_n) \mapsto (\ldots : b_{i_0 \ldots i_n} : \ldots), \quad b_{i_0 \ldots i_n} = a_0^{i_0} \ldots a_n^{i_n}.$$

[3]This can be proved by induction on $m+n$. If $m = 0 = n$, then $\binom{0}{0} = 1$, which is correct. A general homogeneous polynomial of degree m can be written uniquely as

$$F(X_0, X_1, \ldots, X_n) = F_1(X_1, \ldots, X_n) + X_0 F_2(X_0, X_1, \ldots, X_n)$$

with F_1 homogeneous of degree m and F_2 homogeneous of degree $m-1$. But

$$\binom{m+n}{n} = \binom{m+n-1}{m} + \binom{m+n-1}{m-1}$$

because they are the coefficients of X^m in

$$(X+1)^{m+n} = (X+1)(X+1)^{m+n-1},$$

and this proves the induction.

In other words, the Veronese mapping sends an $n+1$-tuple $(a_0{:}\dots:a_n)$ to the set of monomials in the a_i of degree m. For example, when $n=1$ and $m=2$, the Veronese map is

$$\mathbb{P}^1 \to \mathbb{P}^2,\ (a_0:a_1) \mapsto (a_0^2 : a_0a_1 : a_1^2).$$

Its image is the curve $\nu(\mathbb{P}^1) : X_0X_2 = X_1^2$, and the map

$$(b_{2,0}:b_{1,1}:b_{0,2}) \mapsto \begin{cases} (b_{2,0}:b_{1,1}) \text{ if } b_{2,0} \neq 1 \\ (b_{1,1}:b_{0,2}) \text{ if } b_{0,2} \neq 0. \end{cases}$$

is an inverse $\nu(\mathbb{P}^1) \to \mathbb{P}^1$. (Cf. Example 6.22.)

When $n=1$ and m is general, the Veronese map is

$$\mathbb{P}^1 \to \mathbb{P}^m,\ (a_0:a_1) \mapsto (a_0^m : a_0^{m-1}a_1 : \dots : a_1^m).$$

I claim that, in the general case, the image of ν is a closed subset of $\mathbb{P}^{\nu_{n,m}}$ and that ν defines an isomorphism of projective varieties $\nu{:}\mathbb{P}^n \to \nu(\mathbb{P}^n)$.

First note that the map has the following interpretation: if we regard the coordinates a_i of a point P of $\mathbb{P}^n$ as being the coefficients of a linear form $L = \sum a_iX_i$ (well-defined up to multiplication by nonzero scalar), then the coordinates of $\nu(P)$ are the coefficients of the homogeneous polynomial L^m with the binomial coefficients omitted.

As $L \neq 0 \Rightarrow L^m \neq 0$, the map ν is defined on the whole of $\mathbb{P}^n$, that is,

$$(a_0,\dots,a_n) \neq (0,\dots,0) \Rightarrow (\dots,b_{i_0\dots i_n},\dots) \neq (0,\dots,0).$$

Moreover, $L_1 \neq cL_2 \Rightarrow L_1^m \neq cL_2^m$, because $k[X_0,\dots,X_n]$ is a unique factorization domain, and so ν is injective. It is clear from its definition that ν is regular.

We shall see in the next chapter that the image of any projective variety under a regular map is closed, but in this case we can prove directly that $\nu(\mathbb{P}^n)$ is defined by the system of equations:

$$b_{i_0\dots i_n}b_{j_0\dots j_n} = b_{k_0\dots k_n}b_{\ell_0\dots \ell_n}, \qquad i_h + j_h = k_h + \ell_h, \text{ all } h. \tag{*}$$

Obviously $\mathbb{P}^n$ maps into the algebraic set defined by these equations. Conversely, let

$$V_i = \{(\dots : b_{i_0\dots i_n} : \dots) \mid b_{0\dots 0m0\dots 0} \neq 0\}.$$

Then $\nu(U_i) \subset V_i$ and $\nu^{-1}(V_i) = U_i$. It is possible to write down a regular map $V_i \to U_i$ inverse to $\nu|U_i$: for example, define $V_0 \to \mathbb{P}^n$ to be

$$(\dots : b_{i_0\dots i_n} : \dots) \mapsto (b_{m,0,\dots,0} : b_{m-1,1,0,\dots,0} : b_{m-1,0,1,0,\dots,0} : \dots : b_{m-1,0,\dots,0,1}).$$

Finally, one checks that $\nu(\mathbb{P}^n) \subset \bigcup V_i$.

For any closed variety $W \subset \mathbb{P}^n$, $\nu|W$ is an isomorphism of W onto a closed subvariety $\nu(W)$ of $\nu(\mathbb{P}^n) \subset \mathbb{P}^{\nu_{n,m}}$.

6.24. The Veronese mapping has a very important property. If F is a nonzero homogeneous form of degree $m \geq 1$, then $V(F) \subset \mathbb{P}^n$ is called a ***hypersurface of degree*** m and $V(F) \cap W$ is called a ***hypersurface section*** of the projective variety W. When $m=1$, "surface" is replaced by "plane".

Now let H be the hypersurface in $\mathbb{P}^n$ of degree m

$$\sum a_{i_0\dots i_n} X_0^{i_0} \cdots X_n^{i_n} = 0,$$

and let L be the hyperplane in $\mathbb{P}^{\nu_{n,m}}$ defined by

$$\sum a_{i_0\ldots i_n} X_{i_0\ldots i_n}.$$

Then $\nu(H) = \nu(\mathbb{P}^n) \cap L$, i.e.,

$$H(\mathbf{a}) = 0 \iff L(\nu(\mathbf{a})) = 0.$$

Thus for any closed subvariety W of $\mathbb{P}^n$, ν defines an isomorphism of the hypersurface section $W \cap H$ of V onto the hyperplane section $\nu(W) \cap L$ of $\nu(W)$. This observation often allows one to reduce questions about hypersurface sections to questions about hyperplane sections.

As one example of this, note that ν maps the complement of a hypersurface section of W isomorphically onto the complement of a hyperplane section of $\nu(W)$, which we know to be affine. Thus the complement of any hypersurface section of a projective variety is an affine variety.

AUTOMORPHISMS OF $\mathbb{P}^n$

6.25. An element $A = (a_{ij})$ of GL_{n+1} defines an automorphism of $\mathbb{P}^n$:

$$(x_0 : \ldots : x_n) \mapsto (\ldots : \sum a_{ij} x_j : \ldots);$$

clearly it is a regular map, and the inverse matrix gives the inverse map. Scalar matrices act as the identity map.

Let $\mathrm{PGL}_{n+1} = \mathrm{GL}_{n+1}/k^\times I$, where I is the identity matrix, that is, PGL_{n+1} is the quotient of GL_{n+1} by its centre. Then PGL_{n+1} is the complement in $\mathbb{P}^{(n+1)^2-1}$ of the hypersurface $\det(X_{ij}) = 0$, and so it is an affine variety with ring of regular functions

$$k[\mathrm{PGL}_{n+1}] = \{F(\ldots, X_{ij}, \ldots)/\det(X_{ij})^m \mid \deg(F) = m \cdot (n+1)\} \cup \{0\}.$$

It is an affine group variety.

The homomorphism $\mathrm{PGL}_{n+1} \to \mathrm{Aut}(\mathbb{P}^n)$ is obviously injective. We sketch a proof that it is surjective.[4] Consider a hypersurface

$$H : F(X_0, \ldots, X_n) = 0$$

in $\mathbb{P}^n$ and a line

$$L = \{(ta_0 : \ldots : ta_n) \mid t \in k\}$$

in $\mathbb{P}^n$. The points of $H \cap L$ are given by the solutions of

$$F(ta_0, \ldots, ta_n) = 0,$$

which is a polynomial of degree $\leq \deg(F)$ in t unless $L \subset H$. Therefore, $H \cap L$ contains $\leq \deg(F)$ points, and it is not hard to show that for a fixed H and most L it will contain exactly $\deg(F)$ points. Thus, the hyperplanes are exactly the closed subvarieties H of $\mathbb{P}^n$ such that

(a) $\dim(H) = n - 1$,

[4]This is related to the fundamental theorem of projective geometry — see E. Artin, Geometric Algebra, Interscience, 1957, Theorem 2.26.

(b) $\#(H \cap L) = 1$ for all lines L not contained in H.

These are geometric conditions, and so any automorphism of $\mathbb{P}^n$ must map hyperplanes to hyperplanes. But on an open subset of $\mathbb{P}^n$, such an automorphism takes the form

$$(b_0 : \ldots : b_n) \mapsto (F_0(b_0, \ldots, b_n) : \ldots : F_n(b_0, \ldots, b_n))$$

where the F_i are homogeneous of the same degree d (see 6.20). Such a map will take hyperplanes to hyperplanes if only if $d = 1$.

THE SEGRE MAP

6.26. This is the mapping

$$((a_0 : \ldots : a_m), (b_0 : \ldots : b_n)) \mapsto ((\ldots : a_i b_j : \ldots)) \colon \mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^{mn+m+n}.$$

The index set for $\mathbb{P}^{mn+m+n}$ is $\{(i,j) \mid 0 \le i \le m, \quad 0 \le j \le n\}$. Note that if we interpret the tuples on the left as being the coefficients of two linear forms $L_1 = \sum a_i X_i$ and $L_2 = \sum b_j Y_j$, then the image of the pair is the set of coefficients of the homogeneous form of degree 2, $L_1 L_2$. From this observation, it is obvious that the map is defined on the whole of $\mathbb{P}^m \times \mathbb{P}^n$ ($L_1 \neq 0 \neq L_2 \Rightarrow L_1 L_2 \neq 0$) and is injective. On any subset of the form $U_i \times U_j$ it is defined by polynomials, and so it is regular. Again one can show that it is an isomorphism onto its image, which is the closed subset of $\mathbb{P}^{mn+m+n}$ defined by the equations

$$w_{ij} w_{kl} - w_{il} w_{kj} = 0$$

– see Shafarevich 1994, I 5.1. For example, the map

$$((a_0 : a_1), (b_0 : b_1)) \mapsto (a_0 b_0 : a_0 b_1 : a_1 b_0 : a_1 b_1) \colon \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3$$

has image the hypersurface

$$H : \quad WZ = XY.$$

The map

$$(w : x : y : z) \mapsto ((w : y), (w : x))$$

is an inverse on the set where it is defined. [Incidentally, $\mathbb{P}^1 \times \mathbb{P}^1$ is not isomorphic to $\mathbb{P}^2$, because in the first variety there are closed curves, e.g., two vertical lines, that don't intersect.]

If V and W are closed subvarieties of $\mathbb{P}^m$ and $\mathbb{P}^n$, then the Segre map sends $V \times W$ isomorphically onto a closed subvariety of $\mathbb{P}^{mn+m+n}$. Thus products of projective varieties are projective.

The product $\mathbb{P}^1 \times \mathbb{P}^n$ contains many disjoint copies of $\mathbb{P}^n$ as closed subvarieties. Therefore a finite disjoint union of copies of $\mathbb{P}^n$ is projective, which shows that a finite disjoint union of projective varieties is projective.

There is an explicit description of the topology on $\mathbb{P}^m \times \mathbb{P}^n$: the closed sets are the sets of common solutions of families of equations

$$F(X_0, \ldots, X_m; Y_0, \ldots, Y_n) = 0$$

with F separately homogeneous in the X and in the Y .

PROJECTIONS WITH GIVEN CENTRE

6.27. Let $L_1, \ldots, L_{n-d}$ be linearly independent linear forms in $n+1$ variables. Their zero set E in k^{n+1} has dimension $d+1$, and so their zero set in $\mathbb{P}^n$ is a d-dimensional linear space. Define $\pi\colon \mathbb{P}^n - E \to \mathbb{P}^{n-d-1}$ by $\pi(a) = (L_1(a) : \ldots : L_{n-d}(a))$; such a map is called a ***projection with centre*** E. If V is a closed subvariety disjoint from E, then π defines a regular map $V \to \mathbb{P}^{n-d-1}$. More generally, if $F_1, \ldots, F_r$ are homogeneous forms of the same degree, and $Z = V(F_1, \ldots, F_r)$, then $a \mapsto (F_1(a) : \ldots : F_r(a))$ is a morphism $\mathbb{P}^n - Z \to \mathbb{P}^{r-1}$.

By carefully choosing the centre E, it is possible to linearly project any smooth curve in $\mathbb{P}^n$ isomorphically onto a curve in $\mathbb{P}^3$, and nonisomorphically (but bijectively on an open subset) onto a curve in $\mathbb{P}^2$ with only nodes as singularities.[5] For example, suppose we have a nonsingular curve C in $\mathbb{P}^3$. To project to $\mathbb{P}^2$ we need three linear forms L_0, L_1, L_2 and the centre of the projection is the point P_0 where all forms are zero. We can think of the map as projecting from the centre P_0 onto some (projective) plane by sending the point P to the point where P_0P intersects the plane. To project C to a curve with only ordinary nodes as singularities, one needs to choose P_0 so that it doesn't lie on any tangent to C, any trisecant (line crossing the curve in 3 points), or any chord at whose extremities the tangents are coplanar. See for example Samuel, P., Lectures on Old and New Results on Algebraic Curves, Tata Notes, 1966.

Projecting a nonsingular variety in $\mathbb{P}^n$ to a lower dimensional projective space usually introduces singularities. Hironaka proved that every singular variety arises in this way in characteristic zero. See Chapter 8.

APPLICATION

PROPOSITION 6.28. *Every finite set S of points of a quasiprojective variety V is contained in an open affine subset of V.*

PROOF. Regard V as a subvariety of $\mathbb{P}^n$, let $\bar{V}$ be the closure of V in $\mathbb{P}^n$, and let $Z = \bar{V} \smallsetminus V$. Because $S \cap Z = \emptyset$, for each $P \in S$ there exists a homogeneous polynomial $F_P \in I(Z)$ such that $F_P(P) \neq 0$. We may suppose that the F_P have the same degree. An elementary argument shows that some linear combination F of the F_P, $P \in S$, is nonzero at each P. Then F is zero on Z, and so $\bar{V} \cap D(F)$ is an open affine of V, but F is nonzero at each P, and so $\bar{V} \cap D(F)$ contains S. □

j Maps to projective space

Under construction.

k Projective space without coordinates

Let E be a vector space over k of dimension n. The set $\mathbb{P}(E)$ of lines through zero in E has a natural structure of an algebraic variety: the choice of a basis for E defines a bijection $\mathbb{P}(E) \to \mathbb{P}^n$, and the inherited structure of an algebraic variety on $\mathbb{P}(E)$ is independent of

[5] A nonsingular curve of degree d in $\mathbb{P}^2$ has genus $\frac{(d-1)(d-2)}{2}$. Thus, if g is not of this form, a curve of genus g can't be realized as a nonsingular curve in $\mathbb{P}^2$.

the choice of the basis (because the bijections defined by two different bases differ by an automorphism of $\mathbb{P}^n$). Note that in contrast to $\mathbb{P}^n$, which has $n+1$ distinguished hyperplanes, namely, $X_0 = 0, \ldots, X_n = 0$, no hyperplane in $\mathbb{P}(E)$ is distinguished.

l The functor defined by projective space

Let R be a k-algebra. A submodule M of an R-module N is said to be a direct summand of N if there exists another submodule M' of M (a complement of M) such that $N = M \oplus M'$. Let M be a direct summand of a finitely generated projective R-module N. Then M is also finitely generated and projective, and so $M_{\mathfrak{m}}$ is a free $R_{\mathfrak{m}}$-module of finite rank for every maximal ideal $\mathfrak{m}$ in R. If $M_{\mathfrak{m}}$ is of constant rank r, then we say that M has rank r. See CA §12.

Let

$$P^n(R) = \{\text{direct summands of rank 1 of } R^{n+1}\}.$$

Then P^n is a functor from k-algebras to sets. When K is a field, every K-subspace of K^{n+1} is a direct summand, and so $\mathbb{P}^n(K)$ consists of the lines through the origin in K^{n+1}.

Let H_i be the hyperplane $X_i = 0$ in k^{n+1}, and let

$$P_i(R) = \{L \in P^n(R) \mid L \oplus H_{iR} = R^{n+1}\}.$$

Let $L \in P_i(R)$; then

$$e_i = \ell + \sum_{j \neq i} a_j e_j.$$

Now

$$L \mapsto (a_j)_{j \neq i} \colon P_i(R) \to U_i(R) \simeq R^n$$

is a bijection. These combine to give an isomorphism $P^n(R) \to \mathbb{P}^n(R)$:

$$\begin{tikzcd}
P^n(R) \arrow[r] \arrow[d, dashed] & \prod_{0 \le i \le n} P_i(R) \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & \prod_{0 \le i, j \le n} P_i(R) \cap P_j(R) \arrow[d] \\
\mathbb{P}^n(R) \arrow[r] & \prod_{0 \le i \le n} U_i(R) \arrow[r, shift left] \arrow[r, shift right] & \prod_{0 \le i, j \le n} U_i(R) \cap U_j(R).
\end{tikzcd}$$

More generally, to give a regular map from a variety V to $\mathbb{P}^n$ is the same as giving an isomorphism class of pairs $(L, (s_0, \ldots, s_n))$ where L is an invertible sheaf on V and $s_0, \ldots, s_n$ are sections of L that generate it.

m Grassmann varieties

Let E be a vector space over k of dimension n, and let $G_d(E)$ be the set of d-dimensional subspaces of E. When $d = 0$ or n, $G_d(E)$ has a single element, and so from now on we assume that $0 < d < n$. Fix a basis for E, and let $S \in G_d(E)$. The choice of a basis for S then determines a $d \times n$ matrix $A(S)$ whose rows are the coordinates of the basis elements. Changing the basis for S multiplies $A(S)$ on the left by an invertible $d \times d$ matrix. Thus, the family of $d \times d$ minors of $A(S)$ is determined up to multiplication by a nonzero constant, and so defines a point $P(S)$ in $\mathbb{P}^{\binom{n}{d}-1}$.

PROPOSITION 6.29. *The map $S \mapsto P(S)\colon G_d(E) \to \mathbb{P}^{\binom{n}{d}-1}$ is injective, with image a closed subset of $\mathbb{P}^{\binom{n}{d}-1}$.*

We give the proof below. The maps P defined by different bases of E differ by an automorphism of $\mathbb{P}^{\binom{n}{d}-1}$, and so the statement is independent of the choice of the basis — later (6.34) we shall give a "coordinate-free description" of the map. The map realizes $G_d(E)$ as a projective algebraic variety called the ***Grassmann variety*** of d-dimensional subspaces of E.

EXAMPLE 6.30. The affine cone over a line in $\mathbb{P}^3$ is a two-dimensional subspace of k^4. Thus, $G_2(k^4)$ can be identified with the set of lines in $\mathbb{P}^3$. Let L be a line in $\mathbb{P}^3$, and let $\mathbf{x} = (x_0 : x_1 : x_2 : x_3)$ and $\mathbf{y} = (y_0 : y_1 : y_2 : y_3)$ be distinct points on L. Then

$$P(L) = (p_{01} : p_{02} : p_{03} : p_{12} : p_{13} : p_{23}) \in \mathbb{P}^5, \quad p_{ij} \stackrel{\text{def}}{=} \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix},$$

depends only on L. The map $L \mapsto P(L)$ is a bijection from $G_2(k^4)$ onto the quadric

$$\Pi : X_{01}X_{23} - X_{02}X_{13} + X_{03}X_{12} = 0$$

in $\mathbb{P}^5$. For a direct elementary proof of this, see (9.41, 9.42) below.

REMARK 6.31. Let S' be a subspace of E of complementary dimension $n-d$, and let $G_d(E)_{S'}$ be the set of $S \in G_d(V)$ such that $S \cap S' = \{0\}$. Fix an $S_0 \in G_d(E)_{S'}$, so that $E = S_0 \oplus S'$. For any $S \in G_d(V)_{S'}$, the projection $S \to S_0$ given by this decomposition is an isomorphism, and so S is the graph of a homomorphism $S_0 \to S'$:

$$s \mapsto s' \iff (s, s') \in S.$$

Conversely, the graph of any homomorphism $S_0 \to S'$ lies in $G_d(V)_{S'}$. Thus,

$$G_d(V)_{S'} \approx \mathrm{Hom}(S_0, S') \approx \mathrm{Hom}(E/S', S'). \tag{27}$$

The isomorphism $G_d(V)_{S'} \approx \mathrm{Hom}(E/S', S')$ depends on the choice of S_0 — it is the element of $G_d(V)_{S'}$ corresponding to $0 \in \mathrm{Hom}(E/S', S')$. The decomposition $E = S_0 \oplus S'$ gives a decomposition $\mathrm{End}(E) = \begin{pmatrix} \mathrm{End}(S_0) & \mathrm{Hom}(S', S_0) \\ \mathrm{Hom}(S_0, S') & \mathrm{End}(S') \end{pmatrix}$, and the bijections (27) show that the group $\begin{pmatrix} 1 & 0 \\ \mathrm{Hom}(S_0, S') & 1 \end{pmatrix}$ acts simply transitively on $G_d(E)_{S'}$.

REMARK 6.32. The bijection (27) identifies $G_d(E)_{S'}$ with the affine variety $\mathbb{A}(\mathrm{Hom}(S_0, S'))$ defined by the vector space $\mathrm{Hom}(S_0, S')$ (cf. p.70). Therefore, the tangent space to $G_d(E)$ at S_0,

$$T_{S_0}(G_d(E)) \simeq \mathrm{Hom}(S_0, S') \simeq \mathrm{Hom}(S_0, E/S_0). \tag{28}$$

Since the dimension of this space doesn't depend on the choice of S_0, this shows that $G_d(E)$ is nonsingular (4.39).

REMARK 6.33. Let B be the set of all bases of E. The choice of a basis for E identifies B with GL_n, which is the principal open subset of $\mathbb{A}^{n^2}$ where $\det \neq 0$. In particular, B has a natural structure as an irreducible algebraic variety. The map $(e_1, \ldots, e_n) \mapsto \langle e_1, \ldots, e_d \rangle \colon B \to G_d(E)$ is a surjective regular map, and so $G_d(E)$ is also irreducible.

REMARK 6.34. The exterior algebra $\bigwedge E = \bigoplus_{d \geq 0} \bigwedge^d E$ of E is the quotient of the tensor algebra by the ideal generated by all vectors $e \otimes e$, $e \in E$. The elements of $\bigwedge^d E$ are called ***(exterior) d-vectors***.The exterior algebra of E is a finite-dimensional graded algebra over k with $\bigwedge^0 E = k$, $\bigwedge^1 E = E$; if $e_1, \ldots, e_n$ form an ordered basis for V, then the $\binom{n}{d}$ wedge products $e_{i_1} \wedge \ldots \wedge e_{i_d}$ $(i_1 < \cdots < i_d)$ form an ordered basis for $\bigwedge^d E$. In particular, $\bigwedge^n E$ has dimension 1. For a subspace S of E of dimension d, $\bigwedge^d S$ is the one-dimensional subspace of $\bigwedge^d E$ spanned by $e_1 \wedge \ldots \wedge e_d$ for any basis $e_1, \ldots, e_d$ of S. Thus, there is a well-defined map

$$S \mapsto \bigwedge^d S : G_d(E) \to \mathbb{P}(\bigwedge^d E) \tag{29}$$

which the choice of a basis for E identifies with $S \mapsto P(S)$. Note that the subspace spanned by $e_1, \ldots, e_n$ can be recovered from the line through $e_1 \wedge \ldots \wedge e_d$ as the space of vectors v such that $v \wedge e_1 \wedge \ldots \wedge e_d = 0$ (cf. 6.35 below).

FIRST PROOF OF PROPOSITION 6.29.

Fix a basis $e_1, \ldots, e_n$ of E, and let $S_0 = \langle e_1, \ldots, e_d \rangle$ and $S' = \langle e_{d+1}, \ldots, e_n \rangle$. Order the coordinates in $\mathbb{P}^{\binom{n}{d}-1}$ so that

$$P(S) = (a_0 : \ldots : a_{ij} : \ldots : \ldots)$$

where a_0 is the left-most $d \times d$ minor of $A(S)$, and a_{ij}, $1 \leq i \leq d$, $d < j \leq n$, is the minor obtained from the left-most $d \times d$ minor by replacing the ith column with the jth column. Let U_0 be the ("typical") standard open subset of $\mathbb{P}^{\binom{n}{d}-1}$ consisting of the points with nonzero zeroth coordinate. Clearly,[6] $P(S) \in U_0$ if and only if $S \in G_d(E)_{S'}$. We shall prove the proposition by showing that $P : G_d(E)_{S'} \to U_0$ is injective with closed image.

For $S \in G_d(E)_{S'}$, the projection $S \to S_0$ is bijective. For each i, $1 \leq i \leq d$, let

$$e'_i = e_i + \textstyle\sum_{d < j \leq n} a_{ij} e_j \tag{30}$$

denote the unique element of S projecting to e_i. Then $e'_1, \ldots, e'_d$ is a basis for S. Conversely, for any $(a_{ij}) \in k^{d(n-d)}$, the e'_i defined by (30) span an $S \in G_d(E)_{S'}$ and project to the e_i . Therefore, $S \leftrightarrow (a_{ij})$ gives a one-to-one correspondence $G_d(E)_{S'} \leftrightarrow k^{d(n-d)}$ (this is a restatement of (27) in terms of matrices).

Now, if $S \leftrightarrow (a_{ij})$, then

$$P(S) = (1 : \ldots : a_{ij} : \ldots : \ldots : f_k(a_{ij}) : \ldots)$$

where $f_k(a_{ij})$ is a polynomial in the a_{ij} whose coefficients are independent of S. Thus, $P(S)$ determines (a_{ij}) and hence also S. Moreover, the image of $P : G_d(E)_{S'} \to U_0$ is the graph of the regular map

$$(\ldots, a_{ij}, \ldots) \mapsto (\ldots, f_k(a_{ij}), \ldots) : \mathbb{A}^{d(n-d)} \to \mathbb{A}^{\binom{n}{d}-d(n-d)-1},$$

which is closed (5.28).

[6] If $e \in S' \cap S$ is nonzero, we may choose it to be part of the basis for S, and then the left-most $d \times d$ submatrix of $A(S)$ has a row of zeros. Conversely, if the left-most $d \times d$ submatrix is singular, we can change the basis for S so that it has a row of zeros; then the basis element corresponding to the zero row lies in $S' \cap S$.

SECOND PROOF OF PROPOSITION 6.29.

An exterior d-vector v is said to be ***pure*** (or ***decomposable***) if there exist vectors $e_1, \ldots, e_d \in V$ such that $v = e_1 \wedge \ldots \wedge e_d$. According to (6.34), the image of $G_d(E)$ in $\mathbb{P}(\bigwedge^d E)$ consists of the lines through the pure d-vectors.

LEMMA 6.35. *Let w be a nonzero d-vector and let*

$$M(w) = \{v \in E \mid v \wedge w = 0\};$$

then $\dim_k M(w) \leq d$, with equality if and only if w is pure.

PROOF. Let $e_1, \ldots, e_m$ be a basis of $M(w)$, and extend it to a basis $e_1, \ldots, e_m, \ldots, e_n$ of V. Write

$$w = \sum_{1 \leq i_1 < \ldots < i_d} a_{i_1 \ldots i_d} e_{i_1} \wedge \ldots \wedge e_{i_d}, \quad a_{i_1 \ldots i_d} \in k.$$

If there is a nonzero term in this sum in which e_j does not occur, then $e_j \wedge w \neq 0$. Therefore, each nonzero term in the sum is of the form $ae_1 \wedge \ldots \wedge e_m \wedge \ldots$. It follows that $m \leq d$, and $m = d$ if and only if $w = ae_1 \wedge \ldots \wedge e_d$ with $a \neq 0$. □

For a nonzero d-vector w, let $[w]$ denote the line through w. The lemma shows that $[w] \in G_d(E)$ if and only if the linear map $v \mapsto v \wedge w \colon E \mapsto \bigwedge^{d+1} E$ has rank $\leq n-d$ (in which case the rank is $n-d$). Thus $G_d(E)$ is defined by the vanishing of the minors of order $n-d+1$ of this map. [7]

Flag varieties

The discussion in the last subsection extends easily to chains of subspaces. Let $\mathbf{d} = (d_1, \ldots, d_r)$ be a sequence of integers with $0 < d_1 < \cdots < d_r < n$, and let $G_{\mathbf{d}}(E)$ be the set of flags

$$F: \quad E \supset E^1 \supset \cdots \supset E^r \supset 0 \tag{31}$$

with E^i a subspace of E of dimension d_i. The map

$$G_{\mathbf{d}}(E) \xrightarrow{F \mapsto (V^i)} \prod_i G_{d_i}(E) \subset \prod_i \mathbb{P}(\bigwedge^{d_i} E)$$

[7] In more detail, the map

$$w \mapsto (v \mapsto v \wedge w) \colon \bigwedge^d E \to \operatorname{Hom}_k(E, \bigwedge^{d+1} E)$$

is injective and linear, and so defines an injective regular map

$$\mathbb{P}(\bigwedge^d E) \hookrightarrow \mathbb{P}(\operatorname{Hom}_k(E, \bigwedge^{d+1} E)).$$

The condition rank $\leq n-d$ defines a closed subset W of $\mathbb{P}(\operatorname{Hom}_k(E, \bigwedge^{d+1} E))$ (once a basis has been chosen for E, the condition becomes the vanishing of the minors of order $n-d+1$ of a linear map $E \to \bigwedge^{d+1} E$), and

$$G_d(E) = \mathbb{P}(\bigwedge^d E) \cap W.$$

realizes $G_{\mathbf{d}}(E)$ as a closed subset[8] $\prod_i G_{d_i}(E)$, and so it is a projective variety, called a ***flag variety***. The tangent space to $G_{\mathbf{d}}(E)$ at the flag F consists of the families of homomorphisms

$$\varphi^i : E^i \to V/E^i, \quad 1 \le i \le r, \tag{32}$$

that are compatible in the sense that

$$\varphi^i | E^{i+1} \equiv \varphi^{i+1} \bmod E^{i+1}.$$

ASIDE 6.36. A basis $e_1, \ldots, e_n$ for E is ***adapted to*** the flag F if it contains a basis $e_1, \ldots, e_{j_i}$ for each E^i. Clearly, every flag admits such a basis, and the basis then determines the flag. As in (6.33), this implies that $G_{\mathbf{d}}(E)$ is irreducible. Because $\mathrm{GL}(E)$ acts transitively on the set of bases for E, it acts transitively on $G_{\mathbf{d}}(E)$. For a flag F, the subgroup $P(F)$ stabilizing F is an algebraic subgroup of $\mathrm{GL}(E)$, and the map

$$g \mapsto gF_0 : \mathrm{GL}(E)/P(F_0) \to G_{\mathbf{d}}(E)$$

is an isomorphism of algebraic varieties. Because $G_{\mathbf{d}}(E)$ is projective, this shows that $P(F_0)$ is a parabolic subgroup of $\mathrm{GL}(V)$.

n Bezout's theorem

Let V be a hypersurface in $\mathbb{P}^n$ (that is, a closed subvariety of dimension $n-1$). For such a variety, $I(V) = (F(X_0, \ldots, X_n))$ with F a homogenous polynomial without repeated factors. We define the ***degree*** of V to be the degree of F.

The next theorem is one of the oldest, and most famous, in algebraic geometry.

THEOREM 6.37. *Let C and D be curves in $\mathbb{P}^2$ of degrees m and n respectively. If C and D have no irreducible component in common, then they intersect in exactly mn points, counted with appropriate multiplicities.*

PROOF. Decompose C and D into their irreducible components. Clearly it suffices to prove the theorem for each irreducible component of C and each irreducible component of D. We can therefore assume that C and D are themselves irreducible.

We know from (2.63) that $C \cap D$ is of dimension zero, and so is finite. After a change of variables, we can assume that $a \neq 0$ for all points $(a:b:c) \in C \cap D$.

Let $F(X,Y,Z)$ and $G(X,Y,Z)$ be the polynomials defining C and D, and write

$$F = s_0 Z^m + s_1 Z^{m-1} + \cdots + s_m, \qquad G = t_0 Z^n + t_1 Z^{n-1} + \cdots + t_n$$

with s_i and t_j polynomials in X and Y of degrees i and j respectively. Clearly $s_m \neq 0 \neq t_n$, for otherwise F and G would have Z as a common factor. Let R be the resultant of F and G, regarded as polynomials in Z. It is a homogeneous polynomial of degree mn in X and Y, or else it is identically zero. If the latter occurs, then for every $(a,b) \in k^2$, $F(a,b,Z)$ and $G(a,b,Z)$ have a common zero, which contradicts the finiteness of $C \cap D$. Thus R is a nonzero polynomial of degree mn. Write $R(X,Y) = X^{mn} R_*(\frac{Y}{X})$ where $R_*(T)$ is a polynomial of degree $\le mn$ in $T = \frac{Y}{X}$.

[8] For example, if u_i is a pure d_i-vector and u_{i+1} is a pure d_{i+1}-vector, then it follows from (6.35) that $M(u_i) \subset M(u_{i+1})$ if and only if the map

$$v \mapsto (v \wedge u_i, v \wedge u_{i+1}) : V \to \bigwedge^{d_i+1} V \oplus \bigwedge^{d_{i+1}+1} V$$

has rank $\le n - d_i$ (in which case it has rank $n - d_i$). Thus, $G_{\mathbf{d}}(V)$ is defined by the vanishing of many minors.

Suppose first that $\deg R_* = mn$, and let $\alpha_1, \ldots, \alpha_{mn}$ be the roots of R_* (some of them may be multiple). Each such root can be written $\alpha_i = \frac{b_i}{a_i}$, and $R(a_i, b_i) = 0$. According to (7.28) this means that the polynomials $F(a_i, b_i, Z)$ and $G(a_i, b_i, Z)$ have a common root c_i. Thus $(a_i : b_i : c_i)$ is a point on $C \cap D$, and conversely, if $(a : b : c)$ is a point on $C \cap D$ (so $a \neq 0$), then $\frac{b}{a}$ is a root of $R_*(T)$. Thus we see in this case, that $C \cap D$ has precisely mn points, provided we take the multiplicity of $(a : b : c)$ to be the multiplicity of $\frac{b}{a}$ as a root of R_*.

Now suppose that R_* has degree $r < mn$. Then $R(X, Y) = X^{mn-r} P(X, Y)$ where $P(X, Y)$ is a homogeneous polynomial of degree r not divisible by X. Obviously $R(0, 1) = 0$, and so there is a point $(0 : 1 : c)$ in $C \cap D$, in contradiction with our assumption. □

REMARK 6.38. The above proof has the defect that the notion of multiplicity has been too obviously chosen to make the theorem come out right. It is possible to show that the theorem holds with the following more natural definition of multiplicity. Let P be an isolated point of $C \cap D$. There will be an affine neighbourhood U of P and regular functions f and g on U such that $C \cap U = V(f)$ and $D \cap U = V(g)$. We can regard f and g as elements of the local ring $\mathcal{O}_P$, and clearly $\operatorname{rad}(f, g) = \mathfrak{m}$, the maximal ideal in $\mathcal{O}_P$. It follows that $\mathcal{O}_P/(f, g)$ is finite-dimensional over k, and we define the multiplicity of P in $C \cap D$ to be $\dim_k(\mathcal{O}_P/(f, g))$. For example, if C and D cross transversely at P, then f and g will form a system of local parameters at P — $(f, g) = \mathfrak{m}$ — and so the multiplicity is one.

The attempt to find good notions of multiplicities in very general situations motivated much of the most interesting work in commutative algebra in the second half of the twentieth century.

o Hilbert polynomials (sketch)

Recall that for a projective variety $V \subset \mathbb{P}^n$,

$$k_{\text{hom}}[V] = k[X_0, \ldots, X_n]/\mathfrak{b} = k[x_0, \ldots, x_n],$$

where $\mathfrak{b} = I(V)$. We observed that $\mathfrak{b}$ is graded, and therefore $k_{\text{hom}}[V]$ is a graded ring:

$$k_{\text{hom}}[V] = \bigoplus\nolimits_{m \geq 0} k_{\text{hom}}[V]_m,$$

where $k_{\text{hom}}[V]_m$ is the subspace generated by the monomials in the x_i of degree m. Clearly $k_{\text{hom}}[V]_m$ is a finite-dimensional k-vector space.

THEOREM 6.39. *There is a unique polynomial $P(V, T)$ such that $P(V, m) = \dim_k k[V]_m$ for all m sufficiently large.*

PROOF. Omitted. □

EXAMPLE 6.40. For $V = \mathbb{P}^n$, $k_{\text{hom}}[V] = k[X_0, \ldots, X_n]$, and (see the footnote on page 139), $\dim k_{\text{hom}}[V]_m = \binom{m+n}{n} = \frac{(m+n)\cdots(m+1)}{n!}$, and so

$$P(\mathbb{P}^n, T) = \binom{T+n}{n} = \frac{(T+n)\cdots(T+1)}{n!}.$$

The polynomial $P(V, T)$ in the theorem is called the ***Hilbert polynomial*** of V. Despite the notation, it depends not just on V but also on its embedding in projective space.

THEOREM 6.41. *Let V be a projective variety of dimension d and degree δ; then*

$$P(V,T) = \frac{\delta}{d!}T^d + \text{terms of lower degree.}$$

PROOF. Omitted. □

The ***degree*** of a projective variety is the number of points in the intersection of the variety and of a general linear variety of complementary dimension (see later).

EXAMPLE 6.42. Let V be the image of the Veronese map

$$(a_0 : a_1) \mapsto (a_0^d : a_0^{d-1}a_1 : \ldots : a_1^d) \colon \mathbb{P}^1 \to \mathbb{P}^d.$$

Then $k_{\text{hom}}[V]_m$ can be identified with the set of homogeneous polynomials of degree $m \cdot d$ in two variables (look at the map $\mathbb{A}^2 \to \mathbb{A}^{d+1}$ given by the same equations), which is a space of dimension $dm + 1$, and so

$$P(V,T) = dT + 1.$$

Thus V has dimension 1 (which we certainly knew) and degree d.

Macaulay knows how to compute Hilbert polynomials.

REFERENCES: Hartshorne 1977, I.7; Atiyah and Macdonald 1969, Chapter 11; Harris 1992, Lecture 13.

p Dimensions

The results for affine varieties extend to projective varieties with one important simplification: if V and W are projective varieties of dimensions r and s in $\mathbb{P}^n$ and $r + s \geq n$, then $V \cap W \neq \emptyset$.

THEOREM 6.43. *Let $V = V(\mathfrak{a}) \subset \mathbb{P}^n$ be a projective variety of dimension ≥ 1, and let $f \in k[X_0, \ldots, X_n]$ be homogeneous, nonconstant, and $\notin \mathfrak{a}$; then $V \cap V(f)$ is nonempty and of pure codimension 1.*

PROOF. Since the dimension of a variety is equal to the dimension of any dense open affine subset, the only part that doesn't follow immediately from (3.42) is the fact that $V \cap V(f)$ is nonempty. Let $V^{\text{aff}}(\mathfrak{a})$ be the zero set of $\mathfrak{a}$ in $\mathbb{A}^{n+1}$ (that is, the affine cone over V). Then $V^{\text{aff}}(\mathfrak{a}) \cap V^{\text{aff}}(f)$ is nonempty (it contains $(0,\ldots,0)$), and so it has codimension 1 in $V^{\text{aff}}(\mathfrak{a})$. Clearly $V^{\text{aff}}(\mathfrak{a})$ has dimension ≥ 2, and so $V^{\text{aff}}(\mathfrak{a}) \cap V^{\text{aff}}(f)$ has dimension ≥ 1. This implies that the polynomials in $\mathfrak{a}$ have a zero in common with f other than the origin, and so $V(\mathfrak{a}) \cap V(f) \neq \emptyset$. □

COROLLARY 6.44. *Let $f_1, \cdots, f_r$ be homogeneous nonconstant elements of $k[X_0, \ldots, X_n]$; and let Z be an irreducible component of $V \cap V(f_1, \ldots f_r)$. Then codim$(Z) \leq r$, and if $\dim(V) \geq r$, then $V \cap V(f_1, \ldots f_r)$ is nonempty.*

PROOF. Induction on r, as before. □

PROPOSITION 6.45. *Let Z be a irreducible closed subvariety of V; if codim$(Z) = r$, then there exist homogeneous polynomials $f_1, \ldots, f_r$ in $k[X_0, \ldots, X_n]$ such that Z is an irreducible component of $V \cap V(f_1, \ldots, f_r)$.*

PROOF. Use the same argument as in the proof (3.47). □

PROPOSITION 6.46. *Every pure closed subvariety Z of $\mathbb{P}^n$ of codimension one is principal, i.e., $I(Z) = (f)$ for some f homogeneous element of $k[X_0, \ldots, X_n]$.*

PROOF. Follows from the affine case. □

COROLLARY 6.47. *Let V and W be closed subvarieties of $\mathbb{P}^n$; if $\dim(V) + \dim(W) \geq n$, then $V \cap W \neq \emptyset$, and every irreducible component of it has $\mathrm{codim}(Z) \leq \mathrm{codim}(V) + \mathrm{codim}(W)$.*

PROOF. Write $V = V(\mathfrak{a})$ and $W = V(\mathfrak{b})$, and consider the affine cones $V' = V(\mathfrak{a})$ and $W' = V(\mathfrak{b})$ over them. Then

$$\dim(V') + \dim(W') = \dim(V) + 1 + \dim(W) + 1 \geq n + 2.$$

As $V' \cap W' \neq \emptyset$, $V' \cap W'$ has dimension ≥ 1, and so it contains a point other than the origin. Therefore $V \cap W \neq \emptyset$. The rest of the statement follows from the affine case. □

PROPOSITION 6.48. *Let V be a closed subvariety of $\mathbb{P}^n$ of dimension $r < n$; then there is a linear projective variety E of dimension $n - r - 1$ (that is, E is defined by $r + 1$ independent linear forms) such that $E \cap V = \emptyset$.*

PROOF. Induction on r. If $r = 0$, then V is a finite set, and the lemma below shows that there is a hyperplane in k^{n+1} not meeting V.

Suppose $r > 0$, and let $V_1, \ldots, V_s$ be the irreducible components of V. By assumption, they all have dimension $\leq r$. The intersection E_i of all the linear projective varieties containing V_i is the smallest such variety. The lemma below shows that there is a hyperplane H containing none of the nonzero E_i; consequently, H contains none of the irreducible components V_i of V, and so each $V_i \cap H$ is a pure variety of dimension $\leq r - 1$ (or is empty). By induction, there is an linear subvariety E' not meeting $V \cap H$. Take $E = E' \cap H$. □

LEMMA 6.49. *Let W be a vector space of dimension d over an infinite field k, and let $E_1, \ldots, E_r$ be a finite set of nonzero subspaces of W. Then there is a hyperplane H in W containing none of the E_i.*

PROOF. Pass to the dual space V of W. The problem becomes that of showing V is not a finite union of proper subspaces $E_i^\vee$. Replace each $E_i^\vee$ by a hyperplane H_i containing it. Then H_i is defined by a nonzero linear form L_i. We have to show that $\prod L_j$ is not identically zero on V. But this follows from the statement that a polynomial in n variables, with coefficients not all zero, can not be identically zero on k^n (Exercise 1-1). □

Let V and E be as in Proposition 6.48. If E is defined by the linear forms $L_0, \ldots, L_r$ then the projection $a \mapsto (L_0(a) : \cdots : L_r(a))$ defines a map $V \to \mathbb{P}^r$. We shall see later that this map is finite, and so it can be regarded as a projective version of the Noether normalization theorem.

In general, a regular map from a variety V to $\mathbb{P}^n$ corresponds to a line bundle on V and a set of global sections of the line bundle. All line bundles on $\mathbb{A}^n \smallsetminus \{\text{origin}\}$ are trivial (see, for example, Hartshorne II 7.1 and II 6.2), from which it follows that all regular maps $\mathbb{A}^{n+1} \smallsetminus \{\text{origin}\} \to \mathbb{P}^m$ are given by a family of homogeneous polynomials. Assuming this, it is possible to prove the following result.

COROLLARY 6.50. *Let $\alpha : \mathbb{P}^n \to \mathbb{P}^m$ be regular; if $m < n$, then α is constant.*

PROOF. Let $\pi: \mathbb{A}^{n+1} - \{\text{origin}\} \to \mathbb{P}^n$ be the map $(a_0, \ldots, a_n) \mapsto (a_0 : \ldots : a_n)$. Then $\alpha \circ \pi$ is regular, and there exist polynomials $F_0, \ldots, F_m \in k[X_0, \ldots, X_n]$ such that $\alpha \circ \pi$ is the map

$$(a_0, \ldots, a_n) \mapsto (F_0(a) : \ldots : F_m(a)).$$

As $\alpha \circ \pi$ factors through $\mathbb{P}^n$, the F_i must be homogeneous of the same degree. Note that

$$\alpha(a_0 : \ldots : a_n) = (F_0(a) : \ldots : F_m(a)).$$

If $m < n$ and the F_i are nonconstant, then (6.43) shows they have a common zero and so α is not defined on all of $\mathbb{P}^n$. Hence the F_i must be constant. □

q Products

It is useful to have an explicit description of the topology on some product varieties.

The topology on $\mathbb{P}^m \times \mathbb{P}^n$.

Suppose we have a collection of polynomials $F_i(X_0, \ldots, X_m; Y_0, \ldots, Y_n)$, $i \in I$, each of which is separately homogeneous in the X and Y. Then the equations

$$F_i(X_0, \ldots, X_m; Y_0, \ldots, Y_n) = 0, \quad i \in I,$$

define a closed subset of $\mathbb{P}^m \times \mathbb{P}^n$, and every closed subset of $\mathbb{P}^m \times \mathbb{P}^n$ arises in this way from a (finite) set of polynomials.

The topology on $\mathbb{A}^m \times \mathbb{P}^n$

The closed subsets of $\mathbb{A}^m \times \mathbb{P}^n$ are exactly those defined by sets of equations

$$F_i(X_1, \ldots, X_m; Y_0, \ldots, Y_n) = 0, \quad i \in I,$$

with each F_i homogeneous in the Y.

The topology on $V \times \mathbb{P}^n$

Let V be an irreducible affine algebraic variety. We look more closely at the topology on $V \times \mathbb{P}^n$ in terms of ideals. Let $A = k[V]$, and let $B = A[X_0, \ldots, X_n]$. Note that $B = A \otimes_k k[X_0, \ldots, X_n]$, and so we can view it as the ring of regular functions on $V \times \mathbb{A}^{n+1}$: for $f \in A$ and $g \in k[X_0, \ldots, X_n]$, $f \otimes g$ is the function

$$(v, \mathbf{a}) \mapsto f(v) \cdot g(\mathbf{a}) : V \times \mathbb{A}^{n+1} \to k.$$

The ring B has an obvious grading — a monomial $aX_0^{i_0} \ldots X_n^{i_n}$, $a \in A$, has degree $\sum i_j$ — and so we have the notion of a graded ideal $\mathfrak{b} \subset B$. It makes sense to speak of the zero set $V(\mathfrak{b}) \subset V \times \mathbb{P}^n$ of such an ideal. For any ideal $\mathfrak{a} \subset A$, $\mathfrak{a}B$ is graded, and $V(\mathfrak{a}B) = V(\mathfrak{a}) \times \mathbb{P}^n$.

LEMMA 6.51. *(a) For each graded ideal $\mathfrak{b} \subset B$, the set $V(\mathfrak{b})$ is closed, and every closed subset of $V \times \mathbb{P}^n$ is of this form.*

(b) The set $V(\mathfrak{b})$ is empty if and only if $\operatorname{rad}(\mathfrak{b}) \supset (X_0, \ldots, X_n)$.

(c) If V is irreducible, then $V = V(\mathfrak{b})$ for some graded prime ideal $\mathfrak{b}$.

PROOF. (a) In the case that $A = k$, we proved this in (6.1) and (6.2), and similar arguments apply in the present more general situation. For example, to see that $V(\mathfrak{b})$ is closed, cover $\mathbb{P}^n$ with the standard open affines U_i and show that $V(\mathfrak{b}) \cap U_i$ is closed for all i.

The set $V(\mathfrak{b})$ is empty if and only if the cone $V^{\text{aff}}(\mathfrak{b}) \subset V \times \mathbb{A}^{n+1}$ defined by $\mathfrak{b}$ is contained in $V \times \{\text{origin}\}$. But

$$\sum a_{i_0 \ldots i_n} X_0^{i_0} \ldots X_n^{i_n}, \quad a_{i_0 \ldots i_n} \in k[V],$$

is zero on $V \times \{\text{origin}\}$ if and only if its constant term is zero, and so

$$I^{\text{aff}}(V \times \{\text{origin}\}) = (X_0, X_1, \ldots, X_n).$$

Thus, the Nullstellensatz shows that $V(\mathfrak{b}) = \emptyset \Rightarrow \text{rad}(\mathfrak{b}) = (X_0, \ldots, X_n)$. Conversely, if $X_i^N \in \mathfrak{b}$ for all i, then obviously $V(\mathfrak{b})$ is empty.

For (c), note that if $V(\mathfrak{b})$ is irreducible, then the closure of its inverse image in $V \times \mathbb{A}^{n+1}$ is also irreducible, and so $IV(\mathfrak{b})$ is prime. □

Exercises

6-1. Show that a point P on a projective curve $F(X,Y,Z) = 0$ is singular if and only if $\partial F/\partial X$, $\partial F/\partial Y$, and $\partial F/\partial Z$ are all zero at P. If P is nonsingular, show that the tangent line at P has the (homogeneous) equation

$$(\partial F/\partial X)_P X + (\partial F/\partial Y)_P Y + (\partial F/\partial Z)_P Z = 0.$$

Verify that $Y^2 Z = X^3 + aXZ^2 + bZ^3$ is nonsingular if $X^3 + aX + b$ has no repeated root, and find the tangent line at the point at infinity on the curve.

6-2. Let L be a line in $\mathbb{P}^2$ and let C be a nonsingular conic in $\mathbb{P}^2$ (i.e., a curve in $\mathbb{P}^2$ defined by a homogeneous polynomial of degree 2). Show that either

(a) L intersects C in exactly 2 points, or

(b) L intersects C in exactly 1 point, and it is the tangent at that point.

6-3. Let $V = V(Y - X^2, Z - X^3) \subset \mathbb{A}^3$. Prove

(a) $I(V) = (Y - X^2, Z - X^3)$,

(b) $ZW - XY \in I(V)^* \subset k[W,X,Y,Z]$, but $ZW - XY \notin ((Y - X^2)^*, (Z - X^3)^*)$. (Thus, if $F_1, \ldots, F_r$ generate $\mathfrak{a}$, it does not follow that $F_1^*, \ldots, F_r^*$ generate $\mathfrak{a}^*$, even if $\mathfrak{a}^*$is radical.)

6-4. Let $P_0, \ldots, P_r$ be points in $\mathbb{P}^n$. Show that there is a hyperplane H in $\mathbb{P}^n$ passing through P_0 but *not* passing through any of $P_1, \ldots, P_r$.

6-5. Is the subset

$$\{(a:b:c) \mid a \neq 0, \quad b \neq 0\} \cup \{(1:0:0)\}$$

of $\mathbb{P}^2$ locally closed?

6-6. Show that the image of the Segre map $\mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^{mn+m+n}$ (see 6.26) is not contained in any hyperplane of $\mathbb{P}^{mn+m+n}$.

6-7. Write 0, 1, ∞ for the points $(0{:}1)$, $(1{:}1)$, and $(1{:}0)$ on $\mathbb{P}^1$.

(a) Let α be an automorphism of $\mathbb{P}^1$ such that

$$\alpha(0) = 0, \quad \alpha(1) = 1, \quad \alpha(\infty) = \infty.$$

Show that α is the identity map.

(b) Let P_0, P_1, P_2 be distinct points on $\mathbb{P}^1$. Show that there exists an $\alpha \in \mathrm{PGL}_2(k)$ such that

$$\alpha(0) = P_0, \quad \alpha(1) = P_1, \quad \alpha(\infty) = P_2.$$

(c) Deduce that $\mathrm{Aut}(\mathbb{P}^1) \simeq \mathrm{PGL}_2(k)$.

6-8. Show that the functor

$$R \rightsquigarrow P^n(R) = \{\text{direct summands of rank 1 of } R^{n+1}\}$$

satisfies the criterion (5.71) to arise from an algebraic prevariety. (This gives an alternative definition of $\mathbb{P}^n$.)

CHAPTER **7**

Complete Varieties

Complete varieties are the analogues in the category of algebraic varieties of compact topological spaces in the category of Hausdorff topological spaces. Recall that the image of a compact space under a continuous map is compact, and hence is closed if the image space is Hausdorff. Moreover, a Hausdorff space V is compact if and only if, for all topological spaces T, the projection map $q: V \times T \to T$ is closed, i.e., maps closed sets to closed sets (see Bourbaki, N., General Topology, I, 10.2, Corollary 1 to Theorem 1).

a Definition and basic properties

Definition

DEFINITION 7.1. An algebraic variety V is ***complete*** if for all algebraic varieties T, the projection map $q: V \times T \to T$ is closed.

Note that a complete variety is required to be separated — we really mean it to be a variety and not a prevariety. We shall see (7.22) that projective varieties are complete.

EXAMPLE 7.2. Consider the projection map

$$(x, y) \mapsto y: \mathbb{A}^1 \times \mathbb{A}^1 \to \mathbb{A}^1$$

This is not closed; for example, the variety $V : XY = 1$ is closed in $\mathbb{A}^2$ but its image in $\mathbb{A}^1$ omits the origin. However, when we replace V with its closure in $\mathbb{P}^1 \times \mathbb{A}^1$, its projection becomes the whole of $\mathbb{A}^1$. To see this, note that

$$\bar{V} = \{((x:z), y) \in \mathbb{P}^1 \times \mathbb{A}^1 \mid xy = z^2\}$$

contains V as an open dense subset, and so must be its closure in $\mathbb{P}^1 \times \mathbb{A}^1$. The point $((x:0), 0)$ of $\bar{V}$ maps to 0.

Properties

7.3. *Closed subvarieties of complete varieties are complete.*

Let Z be a closed subvariety of a complete variety V. For any variety T, $Z \times T$ is closed in $V \times T$, and so the restriction of the closed map $q: V \times T \to T$ to $Z \times T$ is also closed.

7.4. *A variety is complete if and only if its irreducible components are complete.*

Each irreducible component is closed, and hence complete if the variety is complete (7.3). Conversely, suppose that the irreducible components V_i of a variety V are complete. If Z is closed in $V \times T$, then $Z_i \stackrel{\text{def}}{=} Z \cap (V_i \times T)$ is closed in $V_i \times T$. Therefore, $q(Z_i)$ is closed in T, and so $q(Z) = \bigcup q(Z_i)$ is also closed.

7.5. *Products of complete varieties are complete.*

Let $V_1, \ldots, V_n$ be complete varieties, and let T be a variety. The projection $\left(\prod_i V_i\right) \times T \to T$ is the composite of the projections

$$V_1 \times \cdots \times V_n \times T \to V_2 \times \cdots \times V_n \times T \to \cdots \to V_n \times T \to T,$$

all of which are closed.

7.6. *If $\varphi: W \to V$ is surjective and W is complete, then V is complete.*

Let T be a variety, and let Z be a closed subset of $V \times T$. Let Z' be the inverse image of Z in $W \times T$. Then Z' is closed, and its image in T equals that of Z.

7.7. *Let $\varphi: W \to V$ be a regular map of varieties. If W is complete, then $\varphi(W)$ is a complete closed subvariety of V. In particular, every complete subvariety of a variety is closed.*

Let $\Gamma_\varphi = \{(w, \varphi(w))\} \subset W \times V$ be the graph of φ. It is a closed subset of $W \times V$ (because V is a variety, see 5.28), and $\varphi(W)$ is the projection of Γ_φ into V. Therefore $\varphi(W)$ is closed, and (7.6) shows that it is complete. The second statement follows from the first applied to the identity map.

7.8. *A regular map $V \to \mathbb{P}^1$ from a complete connected variety V is either constant or surjective.*

The only proper closed subsets of $\mathbb{P}^1$ are the finite sets, and such a set is connected if and only if it consists of a single point. Because $\varphi(V)$ is connected and closed, it must either be a single point (and φ is constant) or $\mathbb{P}^1$ (and φ is onto).

7.9. *The only regular functions on a complete connected variety are the constant functions.*

A regular function on a variety V is a regular map $f: V \to \mathbb{A}^1 \subset \mathbb{P}^1$, to which we can apply (7.8).

7.10. *A regular map $\varphi: V \to W$ from a complete connected variety to an affine variety has image equal to a point. In particular, every complete connected affine variety is a point.*

Embed W as a closed subvariety of $\mathbb{A}^n$, and write $\varphi = (\varphi_1, \ldots, \varphi_n)$ where φ_i is the composite of φ with the coordinate function $x_i: \mathbb{A}^n \to \mathbb{A}^1$. Each φ_i is a regular function on V, and hence is constant. (Alternatively, apply 5.12.) This proves the first statement, and the second follows from the first applied to the identity map.

7.11. *In order to show that a variety V is complete, it suffices to check that $q: V \times T \to T$ is a closed mapping when T is affine (or even an affine space $\mathbb{A}^n$).*

Every variety T can be written as a finite union of open affine subvarieties $T = \bigcup T_i$. If Z is closed in $V \times T$, then $Z_i \overset{\text{def}}{=} Z \cap (V \times T_i)$ is closed in $V \times T_i$. Therefore, $q(Z_i)$ is closed in T_i for all i. As $q(Z_i) = q(Z) \cap T_i$, this shows that $q(Z)$ is closed. This shows that it suffices to check that $V \times T \to T$ is closed for all affine varieties T. But T can be realized as a closed subvariety of $\mathbb{A}^n$, and then $V \times T \to T$ is closed if $V \times \mathbb{A}^n \to \mathbb{A}^n$ is closed.

Remarks

7.12. The statement that a complete variety V is closed in every larger variety W perhaps explains the name: if V is complete, W is connected, and $\dim V = \dim W$, then $V = W$. Contrast $\mathbb{A}^n \subset \mathbb{P}^n$.

7.13. Here is another criterion: a variety V is complete if and only if every regular map $C \smallsetminus \{P\} \to V$ extends uniquely to a regular map $C \to V$; here P is a nonsingular point on a curve C. Intuitively, this says that all Cauchy sequences have limits in V and that the limits are unique.

b Proper maps

DEFINITION 7.14. A regular map $\varphi\colon V \to S$ of varieties is said to be ***proper*** if it is "universally closed", that is, if for all regular maps $T \to S$, the base change $\varphi'\colon V \times_S T \to T$ of φ is closed.

7.15. For example, a variety V is complete if and only if the map $V \to \{\text{point}\}$ is proper.

7.16. From its very definition, it is clear that the base change of a proper map is proper. In particular,

(a) if V is complete, then $V \times S \to S$ is proper,

(b) if $\varphi\colon V \to S$ is proper, then the fibre $\varphi^{-1}(P)$ over a point P of S is complete.

7.17. If $\varphi\colon V \to S$ is proper, and W is a closed subvariety of V, then $W \xrightarrow{\varphi} S$ is proper.

PROPOSITION 7.18. *A composite of proper maps is proper.*

PROOF. Let $V_3 \to V_2 \to V_1$ be proper maps, and let T be a variety. Consider the diagram

$$\begin{array}{ccc}
V_3 & \longleftarrow & V_3 \times_{V_2} (V_2 \times_{V_1} T) \simeq V_3 \times_{V_1} T \\
\downarrow & & \downarrow{\scriptstyle\text{closed}} \\
V_2 & \longleftarrow & V_2 \times_{V_1} T \\
\downarrow & & \downarrow{\scriptstyle\text{closed}} \\
V_1 & \longleftarrow & T.
\end{array}$$

Both smaller squares are cartesian, and hence so also is the outer square. The statement is now obvious from the fact that a composite of closed maps is closed. □

COROLLARY 7.19. *If $V \to S$ is proper and S is complete, then V is complete.*

PROOF. Special case of the proposition. □

COROLLARY 7.20. *The inverse image of a complete variety under a proper map is complete.*

PROOF. Let $\varphi: V \to S$ be proper, and let Z be a complete subvariety of S. Then $V \times_S Z \to Z$ is proper, and $V \times_S Z \simeq \varphi^{-1}(Z)$. □

EXAMPLE 7.21. Let $f \in k[T_1, \ldots, T_n, X, Y]$ be homogeneous of degree m in X and Y, and let H be the subvariety of $\mathbb{A}^n \times \mathbb{P}^1$ defined by

$$f(T_1, \ldots, T_n, X, Y) = 0.$$

The projection map $\mathbb{A}^n \times \mathbb{P}^1 \to \mathbb{A}^n$ defines a regular map $H \to \mathbb{A}^n$, which is proper (7.22, 7.15). The fibre over a point $(t_1, \ldots, t_n) \in \mathbb{A}^n$ is the subvariety of $\mathbb{P}^1$ defined by the polynomial

$$f(t_1, \ldots, t_n, X, Y) = a_0 X^m + a_1 X^{m-1} Y + \cdots + a_m Y^m, \quad a_i \in k.$$

Assume that not all a_i are zero. Then this is a homogeneous of degree m and so the fibre always has m points counting multiplicities. The points that "disappeared off to infinity" when $\mathbb{P}^1$ was taken to be $\mathbb{A}^1$ (see p.49) have literally become the point at infinity on $\mathbb{P}^1$.

c Projective varieties are complete

The reader may skip this section since the main theorem is given a more explicit proof in Theorem 7.31 below.

THEOREM 7.22. *A projective variety is complete.*

PROOF. After (7.3), it suffices to prove the Theorem for projective space $\mathbb{P}^n$ itself; thus we have to prove that the projection map $\mathbb{P}^n \times W \to W$ is a closed mapping in the case that W is an irreducible affine variety (7.11).

Write p for the projection $W \times \mathbb{P}^n \to W$. We have to show that Z closed in $W \times \mathbb{P}^n$ implies that $p(Z)$ closed in W. If Z is empty, this is true, and so we can assume it to be nonempty. Then Z is a finite union of irreducible closed subsets Z_i of $W \times \mathbb{P}^n$, and it suffices to show that each $p(Z_i)$ is closed. Thus we may assume that Z is irreducible, and hence that $Z = V(\mathfrak{b})$ with $\mathfrak{b}$ a graded prime ideal in $B = A[X_0, \ldots, X_n]$ (6.51).

If $p(Z)$ is contained in some closed subvariety W' of W, then Z is contained in $W' \times \mathbb{P}^n$, and we can replace W with W'. This allows us to assume that $p(Z)$ is dense in W, and we now have to show that $p(Z) = W$.

Because $p(Z)$ is dense in W, the image of the cone $V^{\text{aff}}(\mathfrak{b})$ under the projection $W \times \mathbb{A}^{n+1} \to W$ is also dense in W, and so (see 3.34a) the map $A \to B/\mathfrak{b}$ is injective.

Let $w \in W$: we shall show that if $w \notin p(Z)$, i.e., if there does not exist a $P \in \mathbb{P}^n$ such that $(w, P) \in Z$, then $p(Z)$ is empty, which is a contradiction.

Let $\mathfrak{m} \subset A$ be the maximal ideal corresponding to w. Then $\mathfrak{m}B + \mathfrak{b}$ is a graded ideal, and $V(\mathfrak{m}B + \mathfrak{b}) = V(\mathfrak{m}B) \cap V(\mathfrak{b}) = (w \times \mathbb{P}^n) \cap V(\mathfrak{b})$, and so w will be in the image of Z unless $V(\mathfrak{m}B + \mathfrak{b}) \neq \emptyset$. But if $V(\mathfrak{m}B + \mathfrak{b}) = \emptyset$, then $\mathfrak{m}B + \mathfrak{b} \supset (X_0, \ldots, X_n)^N$ for some N (by 6.51b), and so $\mathfrak{m}B + \mathfrak{b}$ contains the set B_N of homogeneous polynomials of degree N. Because $\mathfrak{m}B$ and $\mathfrak{b}$ are graded ideals,

$$B_N \subset \mathfrak{m}B + \mathfrak{b} \implies B_N = \mathfrak{m}B_N + B_N \cap \mathfrak{b}.$$

In detail: the first inclusion says that an $f \in B_N$ can be written $f = g + h$ with $g \in \mathfrak{m}B$ and $h \in \mathfrak{b}$. On equating homogeneous components, we find that $f_N = g_N + h_N$. Moreover: $f_N = f$; if $g = \sum m_i b_i$, $m_i \in \mathfrak{m}$, $b_i \in B$, then $g_N = \sum m_i b_{iN}$; and $h_N \in \mathfrak{b}$ because $\mathfrak{b}$ is homogeneous. Together these show $f \in \mathfrak{m}B_N + B_N \cap \mathfrak{b}$.

Let $M = B_N / B_N \cap \mathfrak{b}$, regarded as an A-module. The displayed equation says that $M = \mathfrak{m}M$. The argument in the proof of Nakayama's lemma (1.3) shows that $(1+m)M = 0$ for some $m \in \mathfrak{m}$. Because $A \to B/\mathfrak{b}$ is injective, the image of $1+m$ in $B/\mathfrak{b}$ is nonzero. But $M = B_N/B_N \cap \mathfrak{b} \subset B/\mathfrak{b}$, which is an integral domain, and so the equation $(1+m)M = 0$ implies that $M = 0$. Hence $B_N \subset \mathfrak{b}$, and so $X_i^N \in \mathfrak{b}$ for all i, which contradicts the assumption that $Z = V(\mathfrak{b})$ is nonempty. □

Remarks

7.23. Every complete curve is projective.

7.24. Every nonsingular complete surface is projective (Zariski), but there exist singular complete surfaces that are not projective (Nagata).

7.25. There exist nonsingular complete three-dimensional varieties that are not projective (Nagata, Hironaka).

7.26. A nonsingular complete irreducible variety V is projective if and only if every finite set of points of V is contained in an open affine subset of V (Conjecture of Chevalley; proved by Kleiman[1]; see 6.22 for the necessity).

d Elimination theory

When given a system of polynomial equations to solve, we first use some of the equations to eliminate some of the variables; we then find the solutions of the reduced system, and go back to find the solutions of the original system. Elimination theory does this more systematically.

Note that the fact that $\mathbb{P}^n$ is complete has the following explicit restatement: for each system of polynomial equations

$$(*)\left\{\begin{array}{c} P_1(X_1,\ldots,X_m;Y_0,\ldots,Y_n) = 0 \\ \vdots \\ P_r(X_1,\ldots,X_m;Y_0,\ldots,Y_n) = 0 \end{array}\right.$$

such that each P_i is homogeneous in the Y , there exists a system of polynomial equations

$$(**)\left\{\begin{array}{c} R_1(X_1,\ldots,X_m) = 0 \\ \vdots \\ R_s(X_1,\ldots,X_m) = 0 \end{array}\right.$$

with the following property; an m-tuple $(a_1,\ldots,a_m)$ is a solution of (**) if and only if there exists a nonzero n-tuple $(b_0,\ldots,b_n)$ such that $(a_1,\ldots,a_m,b_0,\ldots,b_n)$ is a solution of (*). In

[1] Kleiman, Steven L., Toward a numerical theory of ampleness. Ann. of Math. (2) 84 1966 293–344 (Theorem 3, p.327, et seq.). See also, Hartshorne, Robin, Ample subvarieties of algebraic varieties. Lecture Notes in Mathematics, Vol. 156 Springer, 1970, I §9 p45.

other words, the polynomials $P_i(a_1,\ldots,a_m;Y_0,\ldots,Y_n)$ have a common zero if and only if $R_j(a_1,\ldots,a_m)=0$ for all j. The polynomials R_j are said to have been obtained from the polynomials P_i by elimination of the variables Y_i.

Unfortunately, the proof we gave of the completeness of $\mathbb{P}^n$, while short and elegant, gives no indication of how to construct (**) from (*). The purpose of elimination theory is to provide an algorithm for doing this.

Elimination theory: special case

Let $P = s_0X^m + s_1X^{m-1} + \cdots + s_m$ and $Q = t_0X^n + t_1X^{n-1} + \cdots + t_n$ be polynomials. The ***resultant*** of P and Q is defined to be the determinant

$$\begin{array}{cl} \begin{vmatrix} s_0 & s_1 & \ldots & s_m & & \\ & s_0 & \ldots & & s_m & \\ & & \ldots & & & \ldots \\ t_0 & t_1 & \ldots & t_n & & \\ & t_0 & \ldots & & t_n & \\ & & \ldots & & & \ldots \end{vmatrix} & \begin{array}{l} n \text{ rows} \\ \\ \\ \\ m \text{ rows} \\ \end{array} \end{array}$$

There are n rows of s's and m rows of t's, so that the matrix is $(m+n)\times(m+n)$; all blank spaces are to be filled with zeros. The resultant is a polynomial in the coefficients of P and Q.

PROPOSITION 7.27. *The resultant* $\mathrm{Res}(P,Q)=0$ *if and only if*

(a) *both s_0 and t_0 are zero; or*

(b) *the two polynomials have a common root.*

PROOF. If (a) holds, then $\mathrm{Res}(P,Q)=0$ because the first column is zero. Suppose that α is a common root of P and Q, so that there exist polynomials P_1 and Q_1 of degrees $m-1$ and $n-1$ respectively such that

$$P(X) = (X-\alpha)P_1(X), \qquad Q(X) = (X-\alpha)Q_1(X).$$

Using these equalities, we find that

$$P(X)Q_1(X) - Q(X)P_1(X) = 0. \tag{33}$$

On equating the coefficients of $X^{m+n-1},\ldots,X,1$ in (33) to zero, we find that the coefficients of P_1 and Q_1 are the solutions of a system of $m+n$ linear equations in $m+n$ unknowns. The matrix of coefficients of the system is the transpose of the matrix

$$\begin{pmatrix} s_0 & s_1 & \ldots & s_m & & \\ & s_0 & \ldots & & s_m & \\ & & \ldots & & & \ldots \\ t_0 & t_1 & \ldots & t_n & & \\ & t_0 & \ldots & & t_n & \\ & & \ldots & & & \ldots \end{pmatrix}$$

The existence of the solution shows that this matrix has determinant zero, which implies that $\mathrm{Res}(P,Q)=0$.

Conversely, suppose that $\operatorname{Res}(P,Q)=0$ but neither s_0 nor t_0 is zero. Because the above matrix has determinant zero, we can solve the linear equations to find polynomials P_1 and Q_1 satisfying (33). A root α of P must be also be a root of P_1 or of Q. If the former, cancel $X-\alpha$ from the left hand side of (33), and consider a root β of $P_1/(X-\alpha)$. As $\deg P_1 < \deg P$, this argument eventually leads to a root of P that is not a root of P_1, and so must be a root of Q. □

The proposition can be restated in projective terms. We define the resultant of two homogeneous polynomials

$$P(X,Y)=s_0X^m+s_1X^{m-1}Y+\cdots+s_mY^m,\quad Q(X,Y)=t_0X^n+\cdots+t_nY^n,$$

exactly as in the nonhomogeneous case.

PROPOSITION 7.28. *The resultant* $\operatorname{Res}(P,Q)=0$ *if and only if* P *and* Q *have a common zero in* $\mathbb{P}^1$.

PROOF. The zeros of $P(X,Y)$ in $\mathbb{P}^1$ are of the form:

(a) $(1:0)$ in the case that $s_0=0$;

(b) $(a:1)$ with a a root of $P(X,1)$.

Since a similar statement is true for $Q(X,Y)$, (7.28) is a restatement of (7.27). □

Now regard the coefficients of P and Q as indeterminates. The pairs of polynomials (P,Q) are parametrized by the space $\mathbb{A}^{m+1}\times\mathbb{A}^{n+1}=\mathbb{A}^{m+n+2}$. Consider the closed subset $V(P,Q)$ in $\mathbb{A}^{m+n+2}\times\mathbb{P}^1$. The proposition shows that its projection on $\mathbb{A}^{m+n+2}$ is the set defined by $\operatorname{Res}(P,Q)=0$. Thus, not only have we shown that the projection of $V(P,Q)$ is closed, but we have given an algorithm for passing from the polynomials defining the closed set to those defining its projection.

Elimination theory does this in general. Given a family of polynomials

$$P_i(T_1,\ldots,T_m;X_0,\ldots,X_n),$$

homogeneous in the X_i, elimination theory gives an algorithm for finding polynomials $R_j(T_1,\ldots,T_n)$ such that the $P_i(a_1,\ldots,a_m;X_0,\ldots,X_n)$ have a common zero if and only if $R_j(a_1,\ldots,a_n)=0$ for all j. (Theorem 7.22 shows only that the R_j exist.)

Maple can find the resultant of two polynomials in one variable: for example, entering "resultant$((x+a)^5,(x+b)^5,x)$" gives the answer $(-a+b)^{25}$. Explanation: the polynomials have a common root if and only if $a=b$, and this can happen in 25 ways. Macaulay doesn't seem to know how to do more.

Elimination theory: general case

In this subsection, we give a proof of Theorem 7.22, following Cartier and Tate 1978[2], which is a more explicit than that given above. Throughout, k is a field (not necessarily algebraically closed) and K is an algebraically closed field containing k.

THEOREM 7.29. *For any graded ideal* $\mathfrak{a}$ *in* $k[X_0,\ldots,X_n]$, *exactly one of the following statements is true:*

[2]Cartier, P., Tate, J., A simple proof of the main theorem of elimination theory in algebraic geometry. Enseign. Math. (2) 24 (1978), no. 3-4, 311–317.

(a) *there exists an integer $d_0 \geq 0$ such that $\mathfrak{a}$ contains every homogeneous polynomial of degree $d \geq d_0$;*

(b) *the ideal $\mathfrak{a}$ has a nontrivial zero in K^{n+1}.*

PROOF. Statement (a) says that the radical of $\mathfrak{a}$ contains $(X_0, \ldots, X_n)$, and so the theorem is a restatement of (6.2c), which we deduced from the strong Nullstellensatz. For a direct proof of it, see the article of Cartier and Tate. □

THEOREM 7.30. *Let $R = \bigoplus_{d \in \mathbb{N}} R_d$ be a graded k-algebra such that $R_0 = k$, R is generated as a k-algebra by R_1, and R_d is finite-dimensional for all d. Then exactly one of the following statements is true:*

(a) *there exists an integer $d_0 \geq 0$ such that $R_d = 0$ for all $d \geq d_0$;*

(b) *no $R_d = 0$, and there exists a k-algebra homomorphism $R \to K$ whose kernel is not equal to $R^+ \stackrel{\text{def}}{=} \bigoplus_{d \geq 1} R_d$.*

PROOF. The hypotheses on R say that it is a quotient of $k[X_0, \ldots, X_n]$ by a graded ideal. Therefore (7.30) is a restatement of (7.29). □

Let $P_1, \ldots, P_r$ be polynomials in $k[T_1, \ldots, T_m; X_0, \ldots, X_n]$ with P_j homogeneous of degree d_j in the variables $X_0, \ldots, X_n$. Let J be the ideal $(P_1, \ldots, P_r)$ in $k[T_1, \ldots, T_m; X_0, \ldots, X_n]$, and let $\mathfrak{A}$ be the ideal of polynomials f in $k[T_1, \ldots, T_m]$ with the following property: there exists an integer $N \geq 1$ such that $fX_0^N, \ldots, fX_n^N$ all lie in J.

THEOREM 7.31. *Let V be the zero set of J in $\mathbb{A}^n(K) \times \mathbb{P}^n(K)$. The projection of V into $\mathbb{A}^n(K)$ is the zero set of $\mathfrak{A}$.*

Consider the ring $B = k[T_1, \ldots, T_m; X_0, \ldots, X_n]$ and its subring $B_0 = k[T_1, \ldots, T_m]$. Then B is a graded B_0-algebra with B_d the B_0-submodule generated by the monomials of degree d in $X_0, \ldots, X_n$, and J is a homogeneous (graded) ideal in B. Let $A = \bigoplus_{d \in \mathbb{N}} A_d$ be the quotient graded ring $B/J = \bigoplus_{d \in \mathbb{N}} B_d/(B_d \cap J)$. Let $\mathfrak{S}$ be the ideal of elements a of A_0 such that $aA_d = 0$ for all sufficiently large d.

THEOREM 7.32. *A ring homomorphism $\varphi: A_0 \to K$ extends to a ring homomorphism $\Psi: A \to K$ not annihilating the ideal $A^+ \stackrel{\text{def}}{=} \bigoplus_{d \geq 1} A_d$ if and only if $\varphi(\mathfrak{S}) = 0$.*

Following Cartier and Tate, we leave it to reader to check that (7.32) is equivalent to (7.31).

Proof of Theorem 7.32

We shall prove (7.32) for any graded ring $A = \bigoplus_{d \geq 0} A_d$ satisfying the following two conditions:

(a) as an A_0-algebra, A is generated by A_1;

(b) for every $d \geq 0$, A_d is finitely generated as an A_0-module.

In the statement of the theorem, K is any algebraically closed field.

The proof proceeds by replacing A with other graded rings with the properties (a) and (b) and also having the property that no A_d is zero.

Let $\varphi: A_0 \to K$ be a homomorphism such that $\varphi(\mathfrak{S}) = 0$, and let $\mathfrak{P} = \mathrm{Ker}(\varphi)$. Then $\mathfrak{P}$ is a prime ideal of A_0 containing $\mathfrak{S}$.

Step 1. Let J be the ideal of elements a of A for which there exists an $s \in A_0 \smallsetminus \mathfrak{P}$ such that $sa = 0$. For every $d \geq 0$, the annihilator of the A_0-module A_d is contained in $\mathfrak{S}$, hence in $\mathfrak{P}$, and so $J \cap A_d \neq A_d$. The ideal J is graded, and the quotient ring $A' = A/J$ has the required properties.

Step 2. Let A'' be the ring of fractions of A' whose denominators are in $\Sigma \stackrel{\text{def}}{=} A'_0 \smallsetminus \mathfrak{P}$. Let A''_d be the set of fractions with numerator in A'_d and denominator in Σ. Then $A'' = \bigoplus_{d \geq 0} A''_d$ is a graded ring with the required properties, and A''_0 is a local ring with maximal ideal $\mathfrak{P}'' \stackrel{\text{def}}{=} \mathfrak{P}' \cdot A'_0$.

Step 3. Let R be the quotient of A'' by the graded ideal $\mathfrak{P}'' \cdot A''$. As A''_d is a nonzero finitely generated module over the local ring A''_0, Nakayama's lemma shows that $A''_d \neq \mathfrak{P}'' A''_d$. Therefore R is graded ring with the required properties, and $k = R_0 \stackrel{\text{def}}{=} A''_0/\mathfrak{P}''$ is a field.

Step 4. At this point R satisfies the hypotheses of Theorem 7.30. Let ε be the composite of the natural maps

$$A \to A' \to A'' \to R.$$

In degree 0, this is nothing but the natural map from A_0 to k with kernel $\mathfrak{P}$. As φ has the same kernel, it factors through ε_0, making K into an algebraically closed extension of k. Now, by Theorem 7.30, there exists a k-algebra homomorphism $f: R \to K$ such that $f(R^+) \neq 0$. The composite map $\Psi = f \circ \varepsilon$ has the required properties. □

See Cox, Little, and O'Shea. Ideals, varieties, and algorithms. 1992, Chapter 8, Section 5, for more on elimination theory.

ASIDE 7.33. Elimination theory became unfashionable several decades ago—one prominent algebraic geometer went so far as to announce that Theorem 7.22 eliminated elimination theory from mathematics,[3] provoking Abhyankar, who prefers equations to abstractions, to start the chant "eliminate the eliminators of elimination theory". With the rise of computers, it has become fashionable again.

e The rigidity theorem; abelian varieties

The paucity of maps between complete varieties has some interesting consequences. First an observation: for any point $w \in W$, the projection map $V \times W \to V$ defines an isomorphism $V \times \{w\} \to V$ with inverse $v \mapsto (v, w): V \to V \times W$ (this map is regular because its components are).

THEOREM 7.34 (RIGIDITY THEOREM). *Let $\varphi: V \times W \to T$ be a regular map, and assume that V is complete, V and W are irreducible, and T is separated. If $\varphi(v, w_0)$ is independent of v for one $w_0 \in W$, then $\varphi(v, w) = g(w)$ with g a regular map $g: W \to T$.*

$$\begin{tikzcd} V \times W \arrow[d, "\varphi"'] \arrow[dr, "q"] & \\ T & W \arrow[l, dotted, "g"] \end{tikzcd}$$

PROOF. Choose a $v_0 \in V$, and consider the regular map

$$g: W \to T, \quad w \mapsto \varphi(v_0, w).$$

[3] Weil 1946/1962, p.31: "The device that follows, which, it may be hoped, finally eliminates from algebraic geometry the last traces of elimination-theory, is borrowed from C. Chevalley's Princeton lectures." Demazure 2012 quotes Dieudonné as saying: "Il faut éliminer la théorie de l'élimination."

We shall show that $\varphi = g \circ q$. Because V is complete, the projection map $q: V \times W \to W$ is closed. Let U be an open affine neighbourhood U of $f(v_0, w_0)$; then $T \smallsetminus U$ is closed in T, $\varphi^{-1}(T \smallsetminus U)$ is closed in $V \times W$, and

$$C \stackrel{\text{def}}{=} q(\varphi^{-1}(T \smallsetminus U))$$

is closed in W. By definition, C consists of the $w \in W$ such that $\varphi(v, w) \notin U$ for some $v \in V$, and so

$$W \smallsetminus C = \{w \in W \mid \varphi(V \times \{w\}) \subset U\}.$$

As $f(V, w_0) = f(v_0, w_0)$, we see that $w_0 \in W \smallsetminus C$. Therefore $W \smallsetminus C$ is nonempty, and so it is dense in W. As $V \times \{w\}$ is complete and U is affine, $\varphi(V \times \{w\})$ must be a point whenever $w \in W \smallsetminus C$ (see 7.10); in fact

$$\varphi(V \times \{w\}) = \varphi(v_0, w) = g(w).$$

We have shown that φ and $g \circ q$ agree on the dense subset $V \times (W \smallsetminus C)$ of $V \times W$, and therefore on the whole of $V \times W$. □

COROLLARY 7.35. *Let $\varphi: V \times W \to T$ be a regular map, and assume that V is complete, that V and W are irreducible, and that T is separated. If there exist points $v_0 \in V$, $w_0 \in W$, $t_0 \in T$ such that*

$$\varphi(V \times \{w_0\}) = \{t_0\} = \varphi(\{v_0\} \times W),$$

then $\varphi(V \times W) = \{t_0\}$.

PROOF. With g as in the proof of the theorem,

$$\varphi(v, w) = g(w) = \varphi(v_0, w) = t_0.$$

□

In more colloquial terms, the corollary says that if φ collapses a vertical and a horizontal slice to a point, then it collapses the whole of $V \times W$ to a point, which must therefore be "rigid".

DEFINITION 7.36. An ***abelian variety*** is a complete connected group variety.

THEOREM 7.37. *Every regular map $\alpha: A \to B$ of abelian varieties is the composite of a homomorphism with a translation; in particular, a regular map $\alpha: A \to B$ such that $\alpha(0) = 0$ is a homomorphism.*

PROOF. After composing α with a translation, we may suppose that $\alpha(0) = 0$. Consider the map

$$\varphi: A \times A \to B, \qquad \varphi(a, a') = \alpha(a + a') - \alpha(a) - \alpha(a').$$

Then $\varphi(A \times 0) = 0 = \varphi(0 \times A)$ and so $\varphi = 0$. This means that α is a homomorphism. □

COROLLARY 7.38. *The group law on an abelian variety is commutative.*

PROOF. Commutative groups are distinguished among all groups by the fact that the map taking an element to its inverse is a homomorphism: if $(gh)^{-1} = g^{-1}h^{-1}$, then, on taking inverses, we find that $gh = hg$. Since the negative map, $a \mapsto -a: A \to A$, takes the identity element to itself, the preceding corollary shows that it is a homomorphism. □

f Chow's Lemma

The next theorem is a useful tool in extending results from projective varieties to complete varieties. It shows that a complete variety is not far from a projective variety.

THEOREM 7.39 (CHOW'S LEMMA). *For every complete irreducible variety V, there exists a surjective regular map $f: V' \to V$ from a projective algebraic variety V' to V such that, for some dense open subset U of V, f induces an isomorphism $f^{-1}(U) \to U$ (in particular, f is birational).*

Write V as a finite union of nonempty open affines, $V = U_1 \cup \ldots \cup U_n$, and let $U = \bigcap U_i$. Because V is irreducible, U is a dense in V. Realize each U_i as a dense open subset of a projective variety P_i. Then $P \stackrel{\text{def}}{=} \prod_i P_i$ is a projective variety (6.26). We shall construct an algebraic variety V' and regular maps $f: V' \to V$ and $g: V' \to P$ such that

(a) f is surjective and induces an isomorphism $f^{-1}(U) \to U$;

(b) g is a closed immersion (hence V' is projective).

Let φ_0 (resp. φ_i) denote the inclusion of U into V (resp. into P_i), and let

$$\varphi = (\varphi_0, \varphi_1, \ldots, \varphi_n): U \to V \times P_1 \times \cdots \times P_n,$$

be the diagonal map. We set $U' = \varphi(U)$ and V' equal to the closure of U' in $V \times P_1 \times \cdots \times P_n$. The projection maps $p: V \times P \to V$ and $q: V \times P \to P$ restrict to regular maps $f: V' \to V$ and $g: V' \to P$. Thus, we have a commutative diagram

$$\begin{tikzcd}
& & V \\
U \arrow[r, "\varphi"] \arrow[urr, bend left, "\varphi_0"] & V' \arrow[r, hook] \arrow[ur, "f"] \arrow[dr, "g"'] & V \times P \arrow[u, "p"'] \arrow[d, "q"] \\
& & P
\end{tikzcd} \tag{34}$$

PROOF OF (a)

In the upper-left triangle of the diagram (34), the maps φ and φ_0 are isomorphisms from U onto its images U' and U. Therefore f restricts to an isomorphism $U' \to U$. The set U' is the graph of the map $(\varphi_1, \ldots, \varphi_n): U \to P$, which is closed in $U \times P$ (5.28), and so

$$U' = V' \cap (U \times P) = f^{-1}(U).$$

The map f is dominant, and $f(V') = p(V)$, which is closed because P is complete. Hence f is surjective.

PROOF OF (b)

We first show that g is an immersion. As this is a local condition, it suffices to find open subsets $V_i \subset P$ such that $\bigcup q^{-1}(V_i) \supset V'$ and each map $V' \cap q^{-1}(V_i) \xrightarrow{g} V_i$ is an immersion.

We set

$$V_i = p_i^{-1}(U_i) = P_1 \times \cdots \times U_i \times \cdots \times P_n$$

where p_i is the projection map $P \to P_i$.

We first show that the sets $q^{-1}(V_i)$ cover V'. The sets U_i cover V, hence the sets $f^{-1}(U_i)$ cover V', and so it suffices to show that

$$q^{-1}(V_i) \supset f^{-1}(U_i)$$

for all i. Consider the diagrams

$$\begin{tikzcd} q^{-1}(V_i) \arrow[r] \arrow[d, hook] & U_i \arrow[d, hook, "\varphi_i"] \\ V \times P \arrow[r, "p_i \circ q"] & P_i \end{tikzcd} \qquad \begin{tikzcd} f^{-1}(U_i) \arrow[r, "f"] \arrow[d, hook] & U_i \arrow[d, hook, "\varphi_i"] \\ V \times P \arrow[r, "p_i \circ q"] & P_i. \end{tikzcd} \qquad \begin{tikzcd} U \arrow[r, "\varphi_0"] \arrow[d, hook, "\varphi"] & U_i \arrow[d, hook, "\varphi_i"] \\ V \times P \arrow[r, "p_i \circ q"] & P_i \end{tikzcd}$$

The diagram at left is cartesian, and so it suffices to show that the middle diagram commutes. But U' is dense in V', hence in $f^{-1}(U_i)$, and so it suffices to prove that the middle diagram commutes with $f^{-1}(U_i)$ replaced by U'. But then it becomes the diagram at right, which obviously commutes.

We next show that

$$V' \cap q^{-1}(V_i) \xrightarrow{g} V_i$$

is an immersion for each i. Recall that

$$V_i = U_i \times P^i, \quad \text{where } P^i = \prod\nolimits_{j \neq i} P_j.$$

and so

$$q^{-1}(V_i) = V \times U_i \times P^i \subset V \times P.$$

Let Γ_i denote the graph of the map

$$\left(U_i \times P^i \xrightarrow{p_i} U_i \hookrightarrow V\right).$$

Being a graph, Γ_i is closed in $V \times (U_i \times P^i)$ and the projection map $V \times (U_i \times P^i) \to U_i \times P^i$ restricts to an isomorphism $\Gamma_i \to U_i \times P^i$. In other words, Γ_i is closed in $q^{-1}(V_i)$, and the projection map $q^{-1}(V_i) \to V_i$ restricts to an isomorphism $\Gamma_i \to V_i$. As Γ_i is closed in $q^{-1}(V_i)$ and contains U', it contains $V' \cap q^{-1}(V_i)$, and so the projection map $q^{-1}(V_i) \to V_i$ restricts to an immersion $V' \cap q^{-1}(V_i) \to V_i$.

Finally, $V \times P$ is complete because V and P are, and so V' is complete (7.21). Hence $g(V)$ is closed (7.21), and so g is a closed immersion.

Notes

7.40. Let V be a complete variety, and let $V_1, \ldots, V_s$ be the irreducible components of V. Each V_i is complete (7.4), and so there exists a surjective birational regular map $V_i' \to V_i$ with V_i' projective (7.39). Now $\bigsqcup V_i'$ is projective (6.26), and the composite

$$\bigsqcup V_i' \to \bigsqcup V_i \to V$$

is surjective and birational.

7.41. There is the following more general statement:

> For every algebraic variety V, there exists a projective algebraic variety V' and a birational regular map φ from an open dense subset U of V' onto V whose graph is closed in $V' \times V$; the subset U equals V' if and only if V is complete.

Cf. Serre 1955-56, p.12.[4]

7.42. Chow (1956, Lemma 1)[5] proved essentially the statement (7.41) by essentially the above argument. He used the lemma to prove that all homogeneous spaces are quasiprojective. See also EGA II, 5.6.1.

g Analytic spaces; Chow's theorem

We summarize a little of Serre 1955-56.

A subset V of $\mathbb{C}^n$ is ***analytic*** if every $v \in V$ admits an open neighbourhood U in $\mathbb{C}^n$ such that $V \cap U$ is the zero set of a finite collection of holomorphic functions on U. An analytic subset is locally closed.

Let V' be an open subset of an analytic set V. A function $f: V' \to \mathbb{C}$ is ***holomorphic*** if, for every $v \in V'$, there exists an open neighbourhood U of v in $\mathbb{C}^n$ and a holomorphic function h on U such that $f = h|V' \cap U$. The holomorphic functions on open subsets of V define on V the structure of a $\mathbb{C}$-ringed space.

DEFINITION 7.43. An ***analytic space*** is a $\mathbb{C}$-ringed space $(V, \mathcal{O}_V)$ satisfying the following two conditions:

(a) there exists an open covering $V = \bigcup V_i$ of V such that, for each i, the $\mathbb{C}$-ringed space $(V_i, \mathcal{O}_V|V_i)$ is isomorphic to an analytic set equipped with its sheaf of holomorphic functions;

(b) the topological space V is Hausdorff.

PROPOSITION 7.44. *An algebraic variety V is complete if and only if $V(\mathbb{C})$ is compact in for the complex topology.*

PROOF. The proof uses Chow's lemma (ibid. Proposition 6, p.12). □

There is a natural functor $V \rightsquigarrow V^{\mathrm{an}}$ from algebraic varieties over $\mathbb{C}$ to complex analytic spaces (ibid. §2).

We omit the definition of a coherent sheaf of $\mathcal{O}_V$-modules.

THEOREM 7.45. *Let V be a projective variety over $\mathbb{C}$. Then the functor $\mathcal{F} \mapsto \mathcal{F}^{\mathrm{an}}$ is an equivalence from the category of coherent $\mathcal{O}_{V^{\mathrm{an}}}$-modules to the category of coherent $\mathcal{O}_V$-modules, under which locally free modules correspond. In particular, $\Gamma(V^{\mathrm{an}}, \mathcal{O}_{V^{\mathrm{an}}}) \simeq \Gamma(V, \mathcal{O}_V)$.*

PROOF. This summarizes the main results of Serre 1955-56 (Théoréme 2,3, p.19, p.20). □

THEOREM 7.46 (CHOW'S THEOREM). *Every closed analytic subset of a projective variety is algebraic.*

[4] Serre, Jean-Pierre. Géométrie algébrique et géométrie analytique. Ann. Inst. Fourier, Grenoble 6 (1955–1956), 1–42.

[5] Chow, Wei-Liang. On the projective embedding of homogeneous varieties. Algebraic geometry and topology. A symposium in honor of S. Lefschetz, pp. 122–128. Princeton University Press, Princeton, N. J., 1957.

PROOF. Let V be a projective space, and let Z be a closed analytic subset of V^{an}. A theorem of Henri Cartan states that $\mathcal{O}_{Z^{\mathrm{an}}}$ is a coherent analytic sheaf on V^{an}, and so there exists a coherent algebraic sheaf $\mathcal{F}$ on V such that $\mathcal{F}^{\mathrm{an}} = \mathcal{O}_{Z^{\mathrm{an}}}$. The support of $\mathcal{F}$ is Zariski closed, and equals Z (ibid. p.29). □

THEOREM 7.47. *Every compact analytic subset of an algebraic variety is algebraic.*

PROOF. Let V be an algebraic variety, and let Z be a compact analytic subset. By Chow's lemma (7.41), there exists a projective variety V', a dense open subset U of V', and a surjective regular map $\varphi\colon U \to V$ whose graph Γ is closed in $V \times V'$. Let $\Gamma' = \Gamma \cap (Z \times V')$. As Z and V' are compact and Γ is closed, Γ' is compact, and so its projection V'' on V' is also compact. On the other hand, $V'' = f^{-1}(Z)$, which shows that it is an analytic subset of U, and therefore also of V'. According to Chow's theorem, it is a Zariski closed subset of V' (hence an algebraic variety). Now $Z = f(V'')$ is constructible (Zariski sense; see 9.7 below), and therefore its Zariski closure coincides with its closure for the complex topology, but (by assumption) it is closed. □

COROLLARY 7.48. *Let V and W be algebraic varieties over $\mathbb{C}$. If V is complete, then every holomorphic map $f\colon V^{\mathrm{an}} \to W^{\mathrm{an}}$ is algebraic.*

PROOF. Apply the preceding theorem to the graph of f. □

EXAMPLE 7.49. The graph of $z \mapsto e^z\colon \mathbb{C} \to \mathbb{C} \times \mathbb{C}$ is closed in $\mathbb{C} \times \mathbb{C}$ but it is not Zariski closed.

h Nagata's Embedding Theorem

A necessary condition for a prevariety to be an open subvariety of a complete variety is that it be separated. An important theorem of Nagata says that this condition is also sufficient.

THEOREM 7.50. *Every variety V admits an open immersion $V \hookrightarrow W$ into a complete variety W.*

If V is affine, then one can embed $V \hookrightarrow \mathbb{A}^n \hookrightarrow \mathbb{P}^n$, and take W to be the closure of V in $\mathbb{P}^n$. The proof in the general case is quite difficult. See:

> Nagata, Masayoshi. Imbedding of an abstract variety in a complete variety. J. Math. Kyoto Univ. 2 1962 1–10; A generalization of the imbedding problem of an abstract variety in a complete variety. J. Math. Kyoto Univ. 3 1963 89–102.

For a modern exposition, see:

> Lütkebohmert, W. On compactification of schemes. Manuscripta Math. 80 (1993), no. 1, 95–111.

In the 1970s, Deligne translated Nagata's work into the language of schemes. His personal notes are available in three versions.

> Deligne, P., Le théorème de plongement de Nagata, Kyoto J. Math. 50, Number 4 (2010), 661-670.
>
> Conrad, B., Deligne's notes on Nagata compactifications. J. Ramanujan Math. Soc. 22 (2007), no. 3, 205–257.
>
> Vojta, P., Nagata's embedding theorem, 19pp., 2007, arXiv:0706.1907.

See also:

Temkin, Michael. Relative Riemann-Zariski spaces. Israel J. Math. 185 (2011), 1–42.

A little history

Weil introduced the term "complete variety" to denote the algebraic geometers analogue of a compact manifold when he defined abstract algebraic varieties.

Exercises

7-1. Identify the set of homogeneous polynomials $F(X,Y) = \sum a_{ij} X^i Y^j$, $0 \leq i, j \leq m$, with an affine space. Show that the subset of reducible polynomials is closed.

7-2. Let V and W be complete irreducible varieties, and let A be an abelian variety. Let P and Q be points of V and W. Show that any regular map $h\colon V \times W \to A$ such that $h(P,Q) = 0$ can be written $h = f \circ p + g \circ q$ where $f\colon V \to A$ and $g\colon W \to A$ are regular maps carrying P and Q to 0 and p and q are the projections $V \times W \to V, W$.

CHAPTER 8

Normal Varieties; (Quasi-)finite maps; Zariski's Main Theorem

We begin by studying normal varieties. These varieties have some of the good properties of nonsingular varieties, and it is easy to show that every variety is birationally equivalent to a normal variety. After studying finite and quasi-finite maps, we discuss the celebrated Zariski's Main Theorem (ZMT), which says that every quasi-finite map of algebraic varieties can be obtained from a finite map by removing a closed subset from the source variety. In its original form, the theorem says that a birational regular map to a normal algebraic variety fails to be a local isomorphism only at points where the fibre has dimension > 0.

a Normal varieties

Recall (1.42) that a normal domain is an integral domain that is integrally closed in its field of fractions. Moreover, that an integral domain A is normal if and only if $A_{\mathfrak{m}}$ is normal for every maximal ideal $\mathfrak{m}$ in A (see 1.49).

DEFINITION 8.1. A point P on an algebraic variety V is ***normal*** if $\mathcal{O}_{V,P}$ is a normal domain. An algebraic variety is said to be ***normal*** if all of its points are normal.

Since the local ring at a point lying on two irreducible components can't be an integral domain (see 3.14), a normal variety is a disjoint union of its irreducible components, which are therefore its connected components.

PROPOSITION 8.2. *The following conditions on an irreducible variety V are equivalent.*

(a) *The variety V is normal.*

(b) *For all open affine subsets U of V, the ring $\mathcal{O}_V(U)$ is a normal domain.*

(c) *For all open subsets U of V, a rational function on V that satisfies a monic polynomial equation on U whose coefficients are regular on U is itself regular on U.*

PROOF. The equivalence of (a) and (b) follows from (1.49).

(a) $\Longrightarrow$ (c). Let U be an open subset of V, and let $f \in k(V)$ satisfy

$$f^n + a_1 f^{n-1} + \cdots + a_n = 0, \quad a_i \in \mathcal{O}_V(U),$$

(equality in $k(V)$). Then $a_i \in \mathcal{O}_V(U) \subset \mathcal{O}_P$ for all $P \in U$, and so $f \in \mathcal{O}_P$ for all $P \in U$. This implies that $f \in \mathcal{O}_V(U)$ (5.11).

(c) $\implies$ (b). The condition applied to an open affine subset U of V implies that $\mathcal{O}_V(U)$ is integrally closed in $k(V)$. □

A regular local noetherian ring is normal — this is a difficult result that we don't prove here (see CA 22.5 for references). Conversely, a normal local domain *of dimension one* is regular. Thus nonsingular varieties are normal, and normal curves are nonsingular. However, a normal surface need not be nonsingular: the cone

$$X^2 + Y^2 - Z^2 = 0$$

is normal, but it is singular at the origin — the tangent space at the origin is k^3.

The singular locus of a normal variety V must have dimension $\leq \dim V - 2$ (see 8.12 below). For example, a normal surface can only have isolated singularities — the singular locus can't contain a curve. In particular, the surface $Z^3 = X^2Y$ (see 4.42) is not normal.

The normalization of an algebraic variety

Let $E \supset F$ be a finite extension of fields. The extension E/F is said to be normal if the minimum polynomial of every element of E splits in E. Let F^{al} be an algebraic closure of F containing E. The composite in F^{al} of the fields σE, $\sigma \in \mathrm{Aut}(E/F)$, is normal over F (and is called the normal closure of F in F^{al}). If E is normal over F, then E is Galois over $E^{\mathrm{Aut}(E/F)}$ (FT, 3.10), and $E^{\mathrm{Aut}(E/F)}$ is purely inseparable over F (because $\mathrm{Hom}_F(E^{\mathrm{Aut}(E/F)}, F^{\mathrm{al}})$ consists of a single element).

PROPOSITION 8.3. *Let A be a finitely generated k-algebra. Assume that A is an integral domain, and let E be a finite field extension of its field of fractions F. Then the integral closure A' of A in E is a finite A-algebra (hence a finitely generated k-algebra).*

PROOF. According to the Noether normalization theorem 2.45, A contains a polynomial subalgebra A_0 and is finite over A_0. Now E is a finite extension of $F(A_0)$ and A' is the integral closure of A_0 in E, and so we only need to consider the case that A is a polynomial ring $k[X_1, \ldots, X_d]$.

Let $\tilde{E}$ denote the normal closure of E in some algebraic closure of F containing E, and let $\tilde{A}$ denote the integral closure of A in $\tilde{E}$. If $\tilde{A}$ is finitely generated as an A-module, then so is its submodule A' (because A is noetherian). Therefore we only need to consider the case that E is normal over F.

According to the above discussion, $E \supset E_1 \supset F$ with E Galois over E_1 and E_1 purely inseparable over F. Let A_1 denote the integral closure of A in E_1. Then A' is a finite A_1-algebra (1.51), and so it suffices to show that A_1 is a finite A-algebra. Therefore we only need to consider the case that E is purely inseparable over F.

In this case, k has characteristic $p \neq 0$, and, for each $x \in E$, there is a power $q(x)$ of p such that $x^{q(x)} \in F$. As E is finitely generated over F, there is a single power q of p such that $x^q \in F$ for all $x \in E$. Let F^{al} denote an algebraic closure of F containing E. For each i, there is a unique $Y_i \in F^{\mathrm{al}}$ such that $Y_i^q = X_i$. Now

$$F = k(X_1, \ldots, X_d) \subset E \subset k(Y_1, \ldots, Y_d)$$

and

$$A = k[X_1, \ldots, X_d] \subset A' \subset k[Y_1, \ldots, Y_d]$$

because $k[Y_1, \ldots, Y_d]$ contains A and is integrally closed (1.32, 1.43). Obviously $k[Y_1, \ldots, Y_d]$ is a finite A-algebra, and this implies, as before, that A' is a finite A-algebra. □

COROLLARY 8.4. *Let A be as in (8.3). If $A_{\mathfrak{m}}$ is normal for some maximal ideal $\mathfrak{m}$ in A, then A_h is normal for some $h \in A \smallsetminus \mathfrak{m}$.*

PROOF. Let A' be the integral closure of A in its field of fractions. Then $A' = A[f_1,\ldots,f_m]$ for some $f_i \in A'$. Now $(A')_{\mathfrak{m}} \overset{1.47}{=} (A_{\mathfrak{m}})' = A_{\mathfrak{m}}$, and so there exists an $h \in A \smallsetminus \mathfrak{m}$ such that, for all i, $hf_i \in A$. Now $A'_h = A_h$, and so A_h is normal. □

The proposition shows that if A is an integral domain finitely generated over k, then the integral closure A' of A in a finite extension E of $F(A)$ has the same properties. Therefore, $\mathrm{Spm}(A')$ is an irreducible algebraic variety, called the ***normalization*** of $\mathrm{Spm}(A)$ in E. This construction extends without difficulty to nonaffine varieties.

PROPOSITION 8.5. *Let V be an irreducible algebraic variety, and let K be a finite field extension of $k(V)$. Then there exists an irreducible algebraic variety W with $k(W) = K$ and a regular map $\varphi: W \to V$ such that, for all open affines U in V, $\varphi^{-1}(U)$ is affine and $k[\varphi^{-1}(U)]$ is the integral closure of $k[U]$ in K.*

The map φ (or just W) is called the ***normalization*** of V in K.

PROOF. For each $v \in V$, let $W(v)$ be the set of maximal ideals in the integral closure of $\mathcal{O}_v$ in K. Let $W = \bigsqcup_{v\in V} W(v)$, and let $\varphi: W \to V$ be the map sending the points of $W(v)$ to v. For an open affine subset U of V,

$$\varphi^{-1}(U) \simeq \mathrm{spm}(k[U]')$$

where $k[U]'$ is the integral closure of $k[U]$ in K. We endow W with the k-ringed space structure for which

$$(\varphi^{-1}(U), \mathcal{O}_W|\varphi^{-1}(U)) \simeq \mathrm{Spm}(k[U]').$$

A routine argument shows that $(W, \mathcal{O}_W)$ is an algebraic variety with the required properties. □

EXAMPLE 8.6. (a) The normalization of the cuspidal cubic $V: Y^2 = X^3$ in $k(V)$ is the map $\mathbb{A}^1 \to V$, $t \mapsto (t^2, t^3)$ (see 3.29).

(b) The normalization of the nodal cubic $V: Y^2 = X^3 + X^2$ (4.10) in $k(V)$ is the map $\mathbb{A}^1 \to V$, $t \mapsto (t^2 - 1, t^3 - t)$.

PROPOSITION 8.7. *The normal points in an irreducible algebraic variety form a dense open subset.*

PROOF. Corollary 8.4 shows that the set of normal points is open, and it remains to show that it is nonempty. Let V be an irreducible algebraic variety. According to (3.37, 3.38), V is birationally equivalent to a hypersurface H in $\mathbb{A}^{d+1}$, $d = \dim V$,

$$H: \quad a_0X^m + a_1X^{m-1} + \cdots + a_m, \quad a_i \in k[T_1,\ldots,T_d], \quad a_0 \neq 0, \quad m \in \mathbb{N};$$

moreover, $T_1,\ldots,T_d$ can be chosen to be a separating transcendence basis for $k(V)$ over k. Therefore the discriminant D of the polynomial $a_0X^m + \cdots + a_m$ is nonzero (it is an element of $k[T_1,\ldots,T_d]$).

Let $A = k[T_1,\ldots,T_d]$; then $k[H] = A[X]/(a_0X^m + \cdots + a_m) = A[x]$. Let

$$y = c_0 + \cdots + c_{m-1}x^{m-1}, \quad c_i \in k(T_1,\ldots,T_d), \tag{35}$$

be an element of $k(H)$ integral over A. For each $j \in \mathbb{N}$, $\mathrm{Tr}_{k(H)/F(A)}(yx^j)$ is a sum of conjugates of yx^j, and hence is integral over A (cf. the proof of 1.44). As it lies in $F(A)$, it is an element of A. On multiplying (35) with x^j and taking traces, we get a system of linear equations

$$c_0 \cdot \mathrm{Tr}(x^j) + c_1 \cdot \mathrm{Tr}(x^{1+j}) + \cdots + c_{m-1} \cdot \mathrm{Tr}(x^{m-1+j}) = \mathrm{Tr}(yx^j), \quad j = 0, \ldots, m-1.$$

By Cramer's rule (p.24),

$$\det(\mathrm{Tr}(x^{i+j})) \cdot c_l \in A, \quad l = 0, \ldots, m-1.$$

But $\det(\mathrm{Tr}(x^{i+j})) = D$,[1] and so $c_l \in A[D^{-1}]$. Hence $k[H]$ becomes normal once we invert the nonzero element D. We have shown that H contains a dense open normal subvariety, which implies that V does also. □

PROPOSITION 8.8. *For every irreducible algebraic variety V, there exists a surjective regular map $\varphi\colon V' \to V$ from a normal algebraic variety V' to V such that, for some dense open subset U of V, φ induces an isomorphism $\varphi^{-1}(U) \to U$ (in particular φ is birational).*

PROOF. Proposition 8.7 shows that the normalization of V in $k(V)$ has this property. □

8.9. More generally, for a dominant map $\varphi\colon W \to V$ of irreducible algebraic varieties, there exists a ***normalization*** of V in W. For each open affine U in V we have

$$k[U] \subset \Gamma(\varphi^{-1}(U), \mathcal{O}_W) \subset k(W).$$

The integral closure $k[U]'$ of $\Gamma(U, \mathcal{O}_V)$ in $\Gamma(\varphi^{-1}(U), \mathcal{O}_W)$ is a finite $k[U]$-algebra (because it is a $k[U]$-submodule of the integral closure of $k[U]$ in $k(W)$). The normalization of V in W is a regular map $\varphi'\colon V' \to V$ such that, for every open affine U in V,

$$(\varphi'^{-1}(U), \mathcal{O}_{V'}) = \mathrm{Spm}(k[U]').$$

In particular, φ' is an affine map. For example, if W and V are affine, then $V' = \mathrm{Spm}(k[V]')$ where $k[V]'$ is the integral closure of $k[V]$ in $k[W]$. There is a commutative triangle

$$\begin{tikzcd} W \arrow[rr, "j"] \arrow[dr, "\varphi"'] & & V' \arrow[dl, "\varphi'"] \\ & V. & \end{tikzcd}$$

b Regular functions on normal varieties

DEFINITION 8.10. An algebraic variety V is ***factorial at a point*** P if $\mathcal{O}_P$ is a factorial domain. The variety V is ***factorial*** if it is factorial at all points P.

When V is factorial, it *does not follow* that $\mathcal{O}_V(U)$ is factorial for all open affines U in V.

A ***prime divisor*** Z on a variety V is a closed irreducible subvariety of codimension 1. Let Z be a prime divisor on V, and let $P \in V$; we say that Z is ***locally principal*** at P if there exists an open affine neighbourhood U of P and an $f \in k[U]$ such that $I(Z \cap U) = (f)$; the regular function f is then called a ***local equation*** for Z at P. If $P \notin Z$, then Z is locally principal at P because then we can choose U so that $Z \cap U = \emptyset$, and $I(Z \cap U) = (1)$.

[1] See, for example, (2.34) of my notes *Algebraic Number Theory*.

PROPOSITION 8.11. *An irreducible variety V is factorial at a point P if and only if every prime divisor on V is locally principal at P.*

PROOF. Recall that an integral domain is factorial if and only if every prime ideal of height 1 is principal (1.24, 3.52). □

PROPOSITION 8.12. *The codimension of the singular locus in a normal variety is at least* 2.

PROOF. Let V be a normal algebraic variety of dimension d, and suppose that its singular locus has an irreducible component W of codimension 1. After replacing V with an open subvariety, we may suppose that it is affine and that W is principal, say, $W=(f)$ (see 8.11). There exists a nonsingular point P on W (4.37). Let $(U,f_1),\ldots,(U,f_{d-1})$ be germs of functions at P (on V) whose restrictions to W generate the maximal ideal in $\mathcal{O}_{W,P}$ (cf. 4.36). Then $(U,f_1),\ldots,(U,f_{d-1}),(U,f)$ generate the maximal ideal in $\mathcal{O}_{V,P}$, and so P is nonsingular on V. This contradicts the definition of W. □

SUMMARY 8.13. For an algebraic variety V,

$$\text{nonsingular} \implies \text{factorial} \implies \text{normal} \implies \text{singular locus has codimension} \geq 2.$$

- ◇ The variety $X_1^2+\cdots+X_5^2$ is factorial but singular.
- ◇ The cone $Z^2=XY$ in $\mathbb{A}^3$ is normal but not factorial (see 9.39 below).
- ◇ The variety $\mathrm{Spm}(k[X,XY,Y^2,Y^3])$ is a surface in $\mathbb{A}^4$ with exactly one singular point, namely, the origin. Its singular locus has codimension 2, but the variety is not normal (the normalization $k[X,XY,Y^2,Y^3]$ is $k[X,Y]$).
- ◇ Every singular curve has singular locus of codimension 1 (hence fails all conditions).

ZEROS AND POLES OF RATIONAL FUNCTIONS ON NORMAL VARIETIES

Let V be a normal irreducible variety. A ***divisor*** on V is an element of the free abelian group $\mathrm{Div}(V)$ generated by the prime divisors. Thus a divisor D can be written uniquely as a finite (formal) sum

$$D=\sum n_i Z_i,\quad n_i\in\mathbb{Z},\quad Z_i \text{ a prime divisor on } V.$$

The ***support*** $|D|$ of D is the union of the Z_i corresponding to nonzero n_i. A divisor is said to be ***effective*** (or ***positive***) if $n_i\geq 0$ for all i. We get a partial ordering on the divisors by defining $D\geq D'$ to mean $D-D'\geq 0$.

Because V is normal, there is associated with every prime divisor Z on V a discrete valuation ring $\mathcal{O}_Z$. This can be defined, for example, by choosing an open affine subvariety U of V such that $U\cap Z\neq\emptyset$; then $U\cap Z$ is a maximal proper closed subset of U, and so the ideal $\mathfrak{p}$ corresponding to it is minimal among the nonzero ideals of $R=\Gamma(U,\mathcal{O})$; so $R_\mathfrak{p}$ is a normal domain with exactly one nonzero prime ideal $\mathfrak{p}R$ — it is therefore a discrete valuation ring (4.20), which is defined to be $\mathcal{O}_Z$. More intrinsically we can define $\mathcal{O}_Z$ to be the set of rational functions on V that are defined an open subset U of V meeting Z.

Let ord_Z be the valuation $k(V)^\times\xrightarrow{\text{onto}}\mathbb{Z}$ with valuation ring $\mathcal{O}_Z$; thus, if π is a prime element of $\mathcal{O}_Z$, then

$$a=\text{unit}\times\pi^{\mathrm{ord}_Z(a)}.$$

The divisor of a nonzero element f of $k(V)$ is defined to be

$$\mathrm{div}(f)=\sum \mathrm{ord}_Z(f)\cdot Z.$$

The sum is over all the prime divisors of V, but in fact $\mathrm{ord}_Z(f) = 0$ for all but finitely many Z. In proving this, we can assume that V is affine (because it is a finite union of affines), say $V = \mathrm{Spm}(R)$. Then $k(V)$ is the field of fractions of R, and so we can write $f = g/h$ with $g, h \in R$, and $\mathrm{div}(f) = \mathrm{div}(g) - \mathrm{div}(h)$. Therefore, we can assume $f \in R$. The zero set of f, $V(f)$ either is empty or is a finite union of prime divisors, $V = \bigcup Z_i$ (see 3.42) and $\mathrm{ord}_Z(f) = 0$ unless Z is one of the Z_i.

The map

$$f \mapsto \mathrm{div}(f) \colon k(V)^\times \to \mathrm{Div}(V)$$

is a homomorphism. A divisor of the form $\mathrm{div}(f)$ is said to be ***principal***, and two divisors are said to be ***linearly equivalent***, denoted $D \sim D'$, if they differ by a principal divisor.

When V is nonsingular, the ***Picard group*** $\mathrm{Pic}(V)$ of V is defined to be the group of divisors on V modulo principal divisors. (The definition of the Picard group of a general algebraic variety agrees with this definition only for nonsingular varieties; it may differ for normal varieties.)

THEOREM 8.14. *Let V be a normal variety, and let f be rational function on V. If f has no zeros or poles on an open subset U of V, then f is regular on U.*

PROOF. We may assume that V is connected, hence irreducible. Now apply the following statement (proof omitted):

> a noetherian domain is normal if and only if $A_{\mathfrak{p}}$ is a discrete valuation ring for all prime ideals $\mathfrak{p}$ of height 1 and $A = \bigcap_{\mathrm{ht}(\mathfrak{p})=1} A_{\mathfrak{p}}$.

COROLLARY 8.15. *A rational function on a normal variety, regular outside a subset of codimension ≥ 2, is regular everywhere.*

PROOF. This is a restatement of the theorem. □

COROLLARY 8.16. *Let V and W be affine varieties with V normal, and let $\varphi \colon V \smallsetminus Z \to W$ be a regular map defined on the complement of a closed subset Z of V. If $\mathrm{codim}(Z) \geq 2$, then φ extends to a regular map on the whole of V.*

PROOF. We may suppose that W is affine, and embed it as a closed subvariety of $\mathbb{A}^n$. The map $V \smallsetminus Z \to W \hookrightarrow \mathbb{A}^n$ is given by n regular functions on $V \smallsetminus Z$, each of which extends to V. Therefore $V \smallsetminus Z \to \mathbb{A}^n$ extends to $\mathbb{A}^n$, and its image is contained in W. □

c Finite and quasi-finite maps

Finite maps

DEFINITION 8.17. A regular map $\varphi \colon W \to V$ of algebraic varieties is ***finite*** if there exists a finite covering $V = \bigcup_i U_i$ of V by open affines such that, for each i, the set $\varphi^{-1}(U_i)$ is affine and $k[\varphi^{-1}(U_i)]$ is a finite $k[U_i]$-algebra.

EXAMPLE 8.18. Let V be an irreducible algebraic variety, and let $\varphi \colon W \to V$ be the normalization of V in a finite extension of $k(V)$. Then φ is finite. This follows from the definition (8.5) and (8.3).

The next lemma shows that, for maps of affine algebraic varieties, the above definition agrees with the definition (2.39).

LEMMA 8.19. *A regular map $\varphi\colon W \to V$ of affine algebraic varieties is finite if and only if $k[W]$ is a finite $k[V]$-algebra.*

PROOF. The necessity being obvious, we prove the sufficiency. For simplicity, we shall assume in the proof that V and W are irreducible . Let $(U_i)_i$ be a finite family of open affines covering V and such that, for each i, the set $\varphi^{-1}(U_i)$ is affine and $k[\varphi^{-1}(U_i)]$ is a finite $k[U_i]$-algebra.

Each U_i is a finite union of basic open subsets of V. These are also basic open subsets of U_i, because $D(f)\cap U_i = D(f|U_i)$, and so we may assume that the original U_i are basic open subsets of V, say, $U_i = D(f_i)$ with $f_i \in A$.

Let $A = k[V]$ and $B = k[W]$. We are given that $(f_1,\ldots,f_n) = A$ and that B_{f_i} is a finite A_{f_i}-algebra for each i. We have to show that B is a finite A-algebra.

Let $\{b_{i1},\ldots,b_{im_i}\}$ generate B_{f_i} as an A_{f_i}-module. After multiplying through by a power of f_i, we may assume that the b_{ij} lie in B. We shall show that the family of all b_{ij} generate B as an A-module. Let $b \in B$. Then $b/1 \in B_{f_i}$, and so

$$b = \frac{a_{i1}}{f_i^{r_i}}b_{i1} + \cdots + \frac{a_{im_i}}{f_i^{r_i}}b_{im_i},\ \text{some } a_{ij} \in A \text{ and } r_i \in \mathbb{N}.$$

The ideal $(f_1^{r_1},\ldots,f_n^{r_n}) = A$ because any maximal ideal containing $(f_1^{r_1},\ldots,f_n^{r_n})$ would have to contain $(f_1,\ldots,f_n) = A$. Therefore,

$$1 = h_1 f_1^{r_1} + \cdots + h_n f_n^{r_n},\ \text{some } h_i \in A.$$

Now

$$\begin{aligned} b = b\cdot 1 &= h_1\cdot b f_1^{r_1} + \cdots + h_n \cdot b f_n^{r_n} \\ &= h_1(a_{11}b_{11} + \cdots + a_{1m_1}b_{1m_1}) + \cdots + h_n(a_{n1}b_{n1} + \cdots + a_{nm_n}b_{nm_n}), \end{aligned}$$

as required. □

LEMMA 8.20. *Let $\varphi\colon W \to V$ be a regular map with V affine, and let U be an open affine in V. There is a canonical isomorphism of k-algebras*

$$\Gamma(W,\mathcal{O}_W)\otimes_{k[V]} k[U] \to \Gamma(\varphi^{-1}(U),\mathcal{O}_W).$$

PROOF. Let $U' = \varphi^{-1}(U)$. The map is defined by the $k[V]$-bilinear pairing

$$(f,g)\mapsto (f|_{U'}, g\circ\varphi|_{U'})\colon \Gamma(W,\mathcal{O}_W)\times k[U] \to \Gamma(U',\mathcal{O}_W).$$

When W is also affine, it is an isomorphism (see 5.31, 5.32).

Let $W = \bigcup W_i$ be a finite open affine covering of W, and consider the commutative diagram:

$$\begin{tikzcd} 0 \arrow[r] & \Gamma(W,\mathcal{O}_W)\otimes_{k[V]} k[U] \arrow[r] \arrow[d] & \prod_i \Gamma(W_i,\mathcal{O}_W)\otimes_{k[V]} k[U] \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & \prod_{i,j} \Gamma(W_{ij},\mathcal{O}_W)\otimes_{k[V]} k[U] \arrow[d] \\ 0 \arrow[r] & \Gamma(U',\mathcal{O}_W) \arrow[r] & \prod_i \Gamma(U'\cap W_i,\mathcal{O}_W) \arrow[r, shift left] \arrow[r, shift right] & \Gamma(U\cap W_{ij},\mathcal{O}_W) \end{tikzcd}$$

Here $W_{ij} = W_i \cap W_j$. The bottom row is exact because $\mathcal{O}_W$ is a sheaf, and the top row is exact because $\mathcal{O}_W$ is a sheaf and $k[U]$ is flat over $k[V]$.[2] The varieties W_i and $W_i \cap W_j$ are all affine, and so the two vertical arrows at right are products of isomorphisms. This implies that the first is also an isomorphism. □

PROPOSITION 8.21. *Let $\varphi\colon W \to V$ be a regular map of algebraic varieties. If φ is finite, then, for every open affine U in V, $\varphi^{-1}(U)$ is affine and $k[\varphi^{-1}(U)]$ is a finite $k[U]$-algebra.*

PROOF. Let V_i be an open affine covering of V (which we may suppose to be finite) such that $W_i \overset{\text{def}}{=} \varphi^{-1}(V_i)$ is an affine subvariety of W for all i and $k[W_i]$ is a finite $k[V_i]$-algebra. Let U be an open affine in V, and let $U' = \varphi^{-1}(U)$. Then $\Gamma(U', \mathcal{O}_W)$ is a subalgebra of $\prod_i \Gamma(U' \cap W_i, \mathcal{O}_W)$, and so it is an affine k-algebra finite over $k[U]$.[3] We have a morphism of varieties over V

$$\begin{tikzcd} U' \arrow[rr, "\text{canonical}"] \arrow[dr] & & \operatorname{Spm}(\Gamma(U', \mathcal{O}_W)) \arrow[dl] \\ & V & \end{tikzcd} \tag{36}$$

which we shall show to be an isomorphism. We know that each of the maps

$$U' \cap W_i \to \operatorname{Spm}(\Gamma(U' \cap W_i, \mathcal{O}_W))$$

is an isomorphism. But (8.21) shows that $\operatorname{Spm}(\Gamma(U' \cap W_i, \mathcal{O}_W))$ is the inverse image of V_i in $\operatorname{Spm}(\Gamma(U', \mathcal{O}_W))$. Therefore the canonical morphism is an isomorphism over each V_i, and so it is an isomorphism. □

SUMMARY 8.22. Let $\varphi\colon W \to V$ be a regular map, and consider the following condition on an open affine subset U of V:

(*) $\quad \varphi^{-1}(U)$ is affine and $k[\varphi^{-1}(U)]$ is a finite over $k[U]$.

The map φ is finite if (*) holds for the open affines in some covering of V, in which case (*) holds for all open affines of V.

PROPOSITION 8.23. (a) *Closed immersions are finite.*

(b) *The composite of two finite morphisms is finite.*

(c) *The product of two finite morphisms is finite.*

PROOF. (a) Let Z be a closed subvariety of a variety V, and let U be an open affine subvariety of V. Then $Z \cap U$ is a closed subvariety of U. It is therefore affine, and the map $Z \cap U \to U$ corresponds to a map $A \to A/\mathfrak{a}$ of rings, which is obviously finite.

This proves (a). As to be finite is a local condition, it suffices to prove (a) and (b) for maps of affine varieties. Then the statements become statements in commutative algebra.

[2] A sequence $0 \to M' \to M \to M''$ is exact if and only if $0 \to A_\mathfrak{m} \otimes_A M' \to A_\mathfrak{m} \otimes_A M \to A_\mathfrak{m} \otimes_A M''$ is exact for all maximal ideals $\mathfrak{m}$ of A (1.21). This implies the claim because $k[U]_{\mathfrak{m}_P} \simeq \mathcal{O}_{U,P} \simeq \mathcal{O}_{V,P} \simeq k[V]_{\mathfrak{m}_P}$ for all $P \in U$.

[3] Recall that a module over a noetherian ring is noetherian if and only if it is finitely generated, and that a submodule of a noetherian module is noetherian. Therefore, a submodule of a finitely generated module is finitely generated.

(b) If B is a finite A-algebra and C is a finite B-algebra, then C is a finite A-algebra. To see this, note that if $\{b_i\}$ is a set of generators for B as an A-module, and $\{c_j\}$ is a set of generators for C as a B-module, then $\{b_i c_j\}$ is a set of generators for C as an A-module.

(c) If B and B' are respectively finite A and A'-algebras, then $B \otimes_k B'$ is a finite $A \otimes_k A'$-algebra. To see this, note that if $\{b_i\}$ is a set of generators for B as an A-module, and $\{b'_j\}$ is a set of generators for B' as an A-module, the $\{b_i \otimes b'_j\}$ is a set of generators for $B \otimes_A B'$ as an A-module. □

By way of contrast, open immersions are rarely finite. For example, the inclusion $\mathbb{A}^1 \smallsetminus \{0\} \hookrightarrow \mathbb{A}^1$ is not finite because the ring $k[T, T^{-1}]$ is not finitely generated as a $k[T]$-module (any finitely generated $k[T]$-submodule of $k[T, T^{-1}]$ is contained in $T^{-n}k[T]$ for some n).

THEOREM 8.24. *Finite maps of algebraic varieties are closed.*

PROOF. It suffices to prove this for affine varieties. Let $\varphi\colon W \to V$ be a finite map of affine varieties, and let Z be a closed subset of W. The restriction of φ to Z is finite (by 8.23a and b), and so we can replace W with Z; we then have to show that $\mathrm{Im}(\varphi)$ is closed. The map corresponds to a finite map of rings $A \to B$. This will factors as $A \to A/\mathfrak{a} \hookrightarrow B$, from which we obtain maps

$$\mathrm{Spm}(B) \to \mathrm{Spm}(A/\mathfrak{a}) \hookrightarrow \mathrm{Spm}(A).$$

The second map identifies $\mathrm{Spm}(A/\mathfrak{a})$ with the closed subvariety $V(\mathfrak{a})$ of $\mathrm{Spm}(A)$, and so it remains to show that the first map is surjective. This is a consequence of the going-up theorem (1.53). □

The base change of a finite map

Recall that the base change of a regular map $\varphi\colon V \to S$ is the map φ' in the diagram:

$$\begin{array}{ccc}
V \times_S W & \xrightarrow{\psi'} & V \\
\downarrow{\scriptstyle \varphi'} & & \downarrow{\scriptstyle \varphi} \\
W & \xrightarrow{\psi} & S.
\end{array}$$

PROPOSITION 8.25. *The base change of a finite map is finite.*

PROOF. We may assume that all the varieties concerned are affine. Then the statement becomes: if A is a finite R-algebra, then $A \otimes_R B/\mathfrak{N}$ is a finite B-algebra, which is obvious. □

PROPOSITION 8.26. *Finite maps of algebraic varieties are proper.*

PROOF. The base change of a finite map is finite, and hence closed. □

COROLLARY 8.27. *Let $\varphi\colon V \to S$ be finite; if S is complete, then so also is V.*

PROOF. Combine (7.19) and (8.26). □

Quasi-finite maps

Recall that the fibres of a regular map $\varphi: W \to V$ are the closed subvarieties $\varphi^{-1}(P)$ of W for $P \in V$. As for affine varieties (2.39), we say that a regular map of algebraic varieties is ***quasi-finite*** if all of its fibres are finite.

PROPOSITION 8.28. *A finite map $\varphi: W \to V$ is quasi-finite.*

PROOF. Let $P \in V$; we wish to show $\varphi^{-1}(P)$ is finite. After replacing V with an affine neighbourhood of P, we can suppose that it is affine, and then W will be affine also. The map φ then corresponds to a map $\alpha: A \to B$ of affine k-algebras, and a point Q of W maps to P if and only $\alpha^{-1}(\mathfrak{m}_Q) = \mathfrak{m}_P$. But this holds if and only if $\mathfrak{m}_Q \supset \alpha(\mathfrak{m}_P)$, and so the points of W mapping to P are in one-to-one correspondence with the maximal ideals of $B/\alpha(\mathfrak{m})B$. Clearly $B/\alpha(\mathfrak{m})B$ is generated as a k-vector space by the image of any generating set for B as an A-module, and so it is a finite k-algebra. The next lemma shows that it has only finitely many maximal ideals. □

LEMMA 8.29. *A finite k-algebra A has only finitely many maximal ideals.*

PROOF. Let $\mathfrak{m}_1, \ldots, \mathfrak{m}_n$ be maximal ideals in A. They are obviously coprime in pairs, and so the Chinese Remainder Theorem (1.1) shows that the map

$$A \to A/\mathfrak{m}_1 \times \cdots \times A/\mathfrak{m}_n, \qquad a \mapsto (\ldots, a_i \bmod \mathfrak{m}_i, \ldots),$$

is surjective. It follows that

$$\dim_k A \geq \sum \dim_k(A/\mathfrak{m}_i) \geq n$$

— here $\dim_k$ means dimension as a k-vector space. □

Finite and ***quasi-finite*** maps of prevarieties are defined as for varieties.

Examples

8.30. The projection from the curve $XY = 1$ onto the X axis (see p.69) is quasi-finite but not finite — its image is not closed in $\mathbb{A}^1$, and $k[X, X^{-1}]$ is not finite over $k[X]$.

8.31. The map

$$t \mapsto (t^2, t^3): \mathbb{A}^1 \to V(Y^2 - X^3) \subset \mathbb{A}^2$$

from the line to the cuspidal cubic is finite because the image of $k[X, Y]$ in $k[T]$ is $k[T^2, T^3]$, and $\{1, T\}$ is a set of generators for $k[T]$ as a $k[T^2, T^3]$-module (see 3.29).

8.32. The map $\mathbb{A}^1 \to \mathbb{A}^1$, $a \mapsto a^m$ is finite.

8.33. The obvious map

$$(\mathbb{A}^1 \text{ with the origin doubled }) \to \mathbb{A}^1$$

is quasi-finite but not finite (the inverse image of $\mathbb{A}^1$ is not affine).

8.34. The map $\mathbb{A}^2 \smallsetminus \{\text{origin}\} \hookrightarrow \mathbb{A}^2$ is quasi-finite but not finite, because the inverse image of $\mathbb{A}^2$ is not affine (see 3.33). The map

$$\mathbb{A}^2 \smallsetminus \{(0,0)\} \sqcup \{O\} \to \mathbb{A}^2$$

sending O to $(0,0)$ is bijective but not finite (here $\{O\} = \text{Spm}(k) = \mathbb{A}^0$).

8.35. The map (8.31), and the Frobenius map

$$(t_1, \ldots, t_n) \mapsto (t_1^p, \ldots, t_n^p) \colon \mathbb{A}^n \to \mathbb{A}^n$$

in characteristic $p \neq 0$, are examples of finite bijective regular maps that are not isomorphisms.

8.36. Let $V = \mathbb{A}^2 = \text{Spec}(k[X,Y])$ and let f be the map defined on the ring level by

$$\begin{aligned} X &\mapsto X = A \\ Y &\mapsto XY^2 + Y + 1 = B. \end{aligned}$$

Then f is (obviously) quasi-finite, but it is not finite. For this we have to show that $k[X,Y]$ is not integral over its subring $k[A,B]$. The minimum polynomial of Y over $k[A,B]$ is

$$AY^2 + Y + 1 - B = 0,$$

which shows that it is not integral over $k[A,B]$ (see 1.44). Alternatively, one can show directly that Y can never satisfy an equation

$$Y^s + g_1(A,B)Y^{s-1} + \cdots + g_s(A,B) = 0, \qquad g_i(A,B) \in k[A,B],$$

by multiplying the equation by A.

8.37. Let V be the hyperplane

$$X^n + T_1 X^{n-1} + \cdots + T_n = 0$$

in $\mathbb{A}^{n+1}$, and consider the projection map

$$(a_1, \ldots, a_n, x) \mapsto (a_1, \ldots, a_n) \colon V \to \mathbb{A}^n.$$

The fibre over a point $(a_1, \ldots, a_n) \in \mathbb{A}^n$ is the set of solutions of

$$X^n + a_1 X^{n-1} + \cdots + a_n = 0,$$

and so it has exactly n points, counted with multiplicities. The map is certainly quasi-finite; it is also finite because it corresponds to the finite map of k-algebras,

$$k[T_1, \ldots, T_n] \to k[T_1, \ldots, T_n, X]/(X^n + T_1 X^{n-1} + \cdots + T_n).$$

See also the more general example p.49.

8.38. Let V be the hyperplane

$$T_0X^n + T_1X^{n-1} + \cdots + T_n = 0$$

in $\mathbb{A}^{n+2}$. The projection map

$$(a_0,\ldots,a_n,x) \mapsto (a_1,\ldots,a_n)\colon V \xrightarrow{\varphi} \mathbb{A}^{n+1}$$

has finite fibres except for the fibre above $o = (0,\ldots,0)$, which is $\mathbb{A}^1$. Its restriction to $V \smallsetminus \varphi^{-1}(o)$ is quasi-finite, but not finite. Above points of the form $(0,\ldots,0,*,\ldots,*)$ some of the roots "vanish off to ∞". (Example (8.30) is a special case of this.) See also the more general example p.49.

8.39. Let

$$P(X,Y) = T_0X^n + T_1X^{n-1}Y + \cdots + T_nY^n,$$

and let V be its zero set in $\mathbb{P}^1 \times (\mathbb{A}^{n+1} \smallsetminus \{o\})$. In this case, the projection map $V \to \mathbb{A}^{n+1} \smallsetminus \{o\}$ is finite.

d The fibres of finite maps

Let $\varphi\colon W \to V$ be a finite dominant morphism of irreducible varieties. Then $\dim(W) = \dim(V)$, and so $k(W)$ is a finite field extension of $k(V)$. Its degree is called the ***degree*** of the map φ. The map φ is said to be ***separable*** if the field $k(W)$ is separable over $k(V)$. Recall that $|S|$ denotes the number of elements in a finite set S.

THEOREM 8.40. *Let $\varphi\colon W \to V$ be a finite surjective regular map of irreducible varieties, and assume that V is normal.*

(a) *For all $P \in V$, $\left|\varphi^{-1}(P)\right| \le \deg(\varphi)$.*

(b) *The set of points P of V such that $\left|\varphi^{-1}(P)\right| = \deg(\varphi)$ is an open subset of V, and it is nonempty if φ is separable.*

Before proving the theorem, we give examples to show that we need W to be separated and V to be normal in (a), and that we need $k(W)$ to be separable over $k(V)$ for the second part of (b).

EXAMPLE 8.41. (a) The map

$$\{\mathbb{A}^1 \text{ with origin doubled}\} \to \mathbb{A}^1$$

has degree one and is one-to-one except over the origin where it is two-to-one.

(b) Let C be the curve $Y^2 = X^3 + X^2$, and consider the map

$$t \mapsto (t^2-1, t(t^2-1))\colon \mathbb{A}^1 \to C.$$

It is one-to-one except that the points $t = \pm 1$ both map to 0. On coordinate rings, it corresponds to the inclusion

$$k[x,y] \hookrightarrow k[T], \begin{cases} x \mapsto T^2-1 \\ y \mapsto T(T^2-1) \end{cases},$$

and so is of degree one. The ring $k[x,y]$ is not integrally closed — in fact $k[T]$ is the integral closure of $k[x,y]$ in its field of fractions $k(x,y) = k(T)$.

(c) The Frobenius map

$$(a_1,\ldots,a_n) \mapsto (a_1^p,\ldots,a_n^p)\colon \mathbb{A}^n \to \mathbb{A}^n$$

in characteristic $p \neq 0$ is bijective on points, but has degree p^n. The field extension corresponding to the map is

$$k(X_1,\ldots,X_n) \supset k(X_1^p,\ldots,X_n^p)$$

which is purely inseparable.

LEMMA 8.42. *Let $Q_1,\ldots,Q_r$ be distinct points on an affine variety V. Then there is a regular function f on V taking distinct values at the Q_i.*

PROOF. We can embed V as closed subvariety of $\mathbb{A}^n$, and then it suffices to prove the statement with $V = \mathbb{A}^n$ — almost any linear form will do. □

PROOF (OF 8.40). In proving (a) of the theorem, we may assume that V and W are affine, and so the map corresponds to a finite map of k-algebras, $k[V] \to k[W]$. Let $\varphi^{-1}(P) = \{Q_1,\ldots,Q_r\}$. According to the lemma, there exists an $f \in k[W]$ taking distinct values at the Q_i. Let

$$F(T) = T^m + a_1T^{m-1} + \cdots + a_m$$

be the minimum polynomial of f over $k(V)$. It has degree $m \leq [k(W):k(V)] = \deg\varphi$, and it has coefficients in $k[V]$ because V is normal (see 1.44). Now $F(f) = 0$ implies $F(f(Q_i)) = 0$, i.e.,

$$f(Q_i)^m + a_1(P)\cdot f(Q_i)^{m-1} + \cdots + a_m(P) = 0.$$

Therefore the $f(Q_i)$ are all roots of a single polynomial of degree m, and so $r \leq m \leq \deg(\varphi)$.

In order to prove the first part of (b), we show that, if there is a point $P \in V$ such that $\varphi^{-1}(P)$ has $\deg(\varphi)$ elements, then the same is true for all points in an open neighbourhood of P. Choose f as in the last paragraph corresponding to such a P. Then the polynomial

$$T^m + a_1(P)\cdot T^{m-1} + \cdots + a_m(P) = 0 \tag{*}$$

has $r = \deg\varphi$ distinct roots, and so $m = r$. Consider the discriminant disc F of F. Because (*) has distinct roots, $\mathrm{disc}(F)(P) \neq 0$, and so $\mathrm{disc}(F)$ is nonzero on an open neighbourhood U of P. The factorization

$$k[V] \to k[V][T]/(F) \overset{T\mapsto f}{\to} k[W]$$

gives a factorization

$$W \to \mathrm{Spm}(k[V][T]/(F)) \to V.$$

Each point $P' \in U$ has exactly m inverse images under the second map, and the first map is finite and dominant, and therefore surjective (recall that a finite map is closed). This proves that $\varphi^{-1}(P')$ has at least $\deg(\varphi)$ points for $P' \in U$, and part (a) of the theorem then implies that it has exactly $\deg(\varphi)$ points.

We now show that if the field extension is separable, then there exists a point such that $\varphi^{-1}(P)$ has $\deg\varphi$ elements. Because $k(W)$ is separable over $k(V)$, there exists a

$f \in k[W]$ such that $k(V)[f] = k(W)$. Its minimum polynomial F has degree $\deg(\varphi)$ and its discriminant is a nonzero element of $k[V]$. The diagram

$$W \to \operatorname{Spm}(A[T]/(F)) \to V$$

shows that $\#\varphi^{-1}(P) \geq \deg(\varphi)$ for P a point such that $\operatorname{disc}(f)(P) \neq 0$. □

Let $E \supset F$ be a finite extension of fields. The elements of E separable over F form a subfield F^{sep} of E, and the separable degree of E over F is defined to be the degree of F^{sep} over F. The ***separable degree*** of a finite surjective map $\varphi\colon W \to V$ of irreducible varieties is the separable degree of $k(W)$ over $k(V)$.

THEOREM 8.43. *Let $\varphi\colon W \to V$ be a finite surjective regular map of irreducible varieties, and assume that V is normal.*

(a) For all $P \in V$, $\left|\varphi^{-1}(P)\right| \leq \operatorname{sepdeg}(\varphi)$, with equality holding on a dense open subset U.

(b) For all i,

$$V_i = \{P \in V \mid \left|\varphi^{-1}(P)\right| \leq i\}$$

is closed in V.

PROOF. If φ is separable, this was proved in (8.40). If φ is purely inseparable, then φ is one-to-one because, for some q, the Frobenius map $V^{(q^{-1})} \xrightarrow{F} V$ factors through φ. To prove the general case, factor φ as the composite of a purely inseparable map with a separable map. □

ASIDE 8.44. A finite map from a variety onto a normal variety is open (hence both open and closed). For an elementary proof, see Theorem 63.12 of Musili, C. Algebraic geometry for beginners. Texts and Readings in Mathematics, 20. Hindustan Book Agency, New Delhi, 2001.

e Zariski's main theorem

In this section, we explain a fundamental theorem of Zariski.

Statement and proof

One obvious way to construct a nonfinite quasi-finite map is to take a finite map $W \to V$ and remove a closed subset of W. Zariski's Main Theorem (ZMT) shows that, for algebraic varieties, every quasi-finite map arises in this way.

THEOREM 8.45 (ZARISKI'S MAIN THEOREM). *Every quasi-finite map of algebraic varieties $\varphi\colon W \to V$ factors into $W \overset{j}{\hookrightarrow} V' \xrightarrow{\varphi'} V$ with φ' finite and j an open immersion:*

$$\begin{tikzcd} W \arrow[rr, hook, "\text{open immersion}"] \arrow[dr, "\text{quasi-finite}"'] & & V' \arrow[dl, "\text{finite}"] \\ & V & \end{tikzcd}$$

When φ is a dominant map of irreducible varieties, the statement is true with $\varphi'\colon V' \to V$ equal to the normalization of V in W (in the sense of 8.9).

The key result needed to prove (8.45) is the following statement from commutative algebra. For a ring A and a prime ideal $\mathfrak{p}$ in A, $\kappa(\mathfrak{p})$ denotes the field of fractions of $A/\mathfrak{p}$.

THEOREM 8.46 (LOCAL VERSION OF ZMT). *Let A be a commutative ring, and let $i\colon A \to B$ be a finitely generated A-algebra. Let $\mathfrak{q}$ be a prime ideal of B, and let $\mathfrak{p} = i^{-1}(\mathfrak{q})$. Finally, let A' denote the integral closure of A in B. If $B_{\mathfrak{q}}/\mathfrak{p}B_{\mathfrak{q}}$ is a finite $\kappa(\mathfrak{p})$-algebra, then there exists an $f \in A'$ not in $\mathfrak{q}$ such that the map $A'_f \to B_f$ is an isomorphism.*

PROOF. The proof is quite elementary, but intricate — see §17 of my notes CA. □

Recall that a point v in a topological space V is isolated if $\{v\}$ is an open subset of V. The isolated points v of an algebraic variety V are those such that $\{v\}$ is both open and closed. Thus they are the irreducible components of V of dimension 0.

Let $\varphi\colon W \to V$ be a continuous map of topological spaces. We say that $w \in W$ is ***isolated in its fibre*** if it is isolated in the subspace $\varphi^{-1}(\varphi(w))$ of W. Let $\varphi\colon A \to B$ be a homomorphism of finitely generated k-algebras, and consider $\mathrm{spm}(\varphi)\colon \mathrm{spm}(B) \to \mathrm{spm}(A)$; then $\mathfrak{n} \in \mathrm{spm}(B)$ is isolated in its fibre if and only if $B_{\mathfrak{n}}/\mathfrak{m}B_{\mathfrak{n}}$ is a finite k-algebra; here $\mathfrak{m} = \varphi^{-1}(\mathfrak{n})$.

PROPOSITION 8.47. *Let $\varphi\colon W \to V$ be a regular map of algebraic varieties. The set W' of points of W isolated in their fibres is open in W.*

PROOF. Let $w \in W'$. Let W_w and V_v be open affine neighbourhoods of w and $v = \varphi(w)$ such that $\varphi(W_w) \subset V_v$, and let $A = k[V_v]$ and $B = k[W_w]$. Let $\mathfrak{n} = \{f \in B \mid f(w) = 0\}$ — it is the maximal ideal in B corresponding to w.

Let A' be the integral closure of A in B. Theorem 8.46 shows that there exists an $f \in A'$ not in $\mathfrak{m}$ such that $A'_f \simeq B_f$. Write A' as the union of the finitely generated A-subalgebras A_i of A' containing f:

$$A' = \bigcup_i A_i.$$

Because A' is integral over A, each A_i is finite over A (see 1.35). We have

$$B_f \simeq A'_f = \bigcup_i A_{if}.$$

Because B_f is a finitely generated A-algebra, $B_f = A_{if}$ for all sufficiently large A_i. As the A_i are finite over A, B_f is quasi-finite over A, and $\mathrm{spm}(B_f)$ is an open neighbourhood of w consisting of quasi-finite points. □

PROPOSITION 8.48. *Every quasi-finite map of affine algebraic varieties $\varphi\colon W \to V$ factors into $W \xrightarrow{j} V' \xrightarrow{\varphi'} V$ with j a dominant open immersion and φ' finite.*

PROOF. Let $A = k[V]$ and $B = k[W]$. Because φ is quasi-finite, Theorem 8.46 shows that there exist $f_i \in A'$ such that the sets $\mathrm{spm}(B_{f_i})$ form an open covering of W and $A'_{f_i} \simeq B_{f_i}$ for all i. As W quasi-compact, finitely many sets $\mathrm{spm}(B_{f_i})$ suffice to cover W. The argument in the proof of (8.47) shows that there exists an A-subalgebra A'' of A', finite over A, which contains $f_1, \ldots, f_n$ and is such that $B_{f_i} \simeq A''_{f_i}$ for all i. Now the map $W = \mathrm{Spm}(B) \to \mathrm{Spm}(A'')$ is an open immersion because it is when restricted to $\mathrm{Spm}(B_{f_i})$ for each i. As $\mathrm{Spm}(A'') \to \mathrm{Spm}(A) = V$ is finite, we can take $V' = \mathrm{Spm}(A'')$. □

Recall (Exercise 8-3) that a regular map $\varphi\colon W \to V$ is affine if $\varphi^{-1}(U)$ is affine whenever U is an open affine subset of V.

PROPOSITION 8.49. *Let $\varphi: W \to V$ be an affine map of irreducible algebraic varieties. Then the map $j: W \to V'$ from W into the normalization V' of V in W (8.9) is an open immersion.*

PROOF. Let U be an open affine in V. Let $A = k[U]$ and $B = k[\varphi^{-1}(U)]$. In this case, the normalization A' of A in B is finite over A (because it is contained in the normalization of A in $k(W)$, which is finite over A (8.3)). Thus, in the proof of (8.48) we can take $A'' = A'$, and then $\varphi^{-1}(U) \to \mathrm{Spm}(A')$ is an open immersion. As $\mathrm{Spm}(A')$ is an open subvariety of V' and the sets $\varphi^{-1}(U)$ cover W, this implies that $j: W \to V'$ is an open immersion. □

As $V' \to V$ is finite, this proves Theorem 8.45 in the case that φ is an affine map of irreducible varieties. To deduce the general case of Theorem 8.45 from (8.44) requires an additional argument. See Theorem 12.83 of Görtz, U. and Wedhorn, T., Algebraic geometry I. Vieweg + Teubner, Wiesbaden, 2010.

NOTES

8.50. Let $\varphi: W \to V$ be a quasi-finite map of algebraic varieties. In (8.45), we may replace V' with the closure of the image of j. Thus, there is a factorization $\varphi = \varphi' \circ j$ with φ' finite and j a dominant open immersion.

8.51. Theorem 8.45 is false for prevarieties (see 8.33). However, it is true for *separated* maps of prevarieties. (A regular map $\varphi: V \to S$ of algebraic prevarieties is ***separated*** if the image $\Delta_{V/S}$ of the map $v \mapsto (v,v): V \to V \times_S V$ is closed; the map φ is separated if V is separated.)

8.52. Assume that V is normal in (8.45). Then φ' is open (8.44), and so φ is open. Thus, every quasi-finite map from an algebraic variety to a normal algebraic variety is open.

Applications to finite maps

Zariski's main theorem allows us to give a geometric criteria for a regular map to be finite.

PROPOSITION 8.53. *Every quasi-finite regular map $\varphi: W \to V$ of algebraic varieties with W complete is finite.*

PROOF. The map $j: W \hookrightarrow V'$ in (8.45) is an isomorphism of W onto its image $j(W)$ in V'. If W is complete (7.21), then $j(W)$ is closed, and so the restriction of φ' to $j(W)$ is finite. □

PROPOSITION 8.54. *Every proper quasi-finite map $\varphi: W \to V$ of algebraic varieties is finite.*

PROOF. Factor φ into $W \overset{j}{\hookrightarrow} W' \overset{\alpha}{\to} W$ with α finite and j an open immersion. Factor j into

$$W \xrightarrow{w \mapsto (w, jw)} W \times_V W' \xrightarrow{(w,w') \mapsto w'} W'.$$

The image of the first map is Γ_j, which is closed because W' is a variety (see 5.28; W' is separated because it is finite over a variety — exercise). Because φ is proper, the second map is closed. Hence j is an open immersion with closed image. It follows that its image is a connected component of W', and that W is isomorphic to that connected component. □

NOTES

8.55. When W and V are curves, every surjective map $W \to V$ is closed. Thus it is easy to give examples of closed surjective quasi-finite, but nonfinite, maps. Consider, for example, the map

$$\left(\mathbb{A}^1 \smallsetminus \{0\}\right) \sqcup \mathbb{A}^0 \to \mathbb{A}^1,$$

sending each $a \in \mathbb{A}^1 \smallsetminus \{0\}$ to a and $O \in \mathbb{A}^0$ to 0. This doesn't violate the Proposition 8.54, because the map is only closed, not universally closed.

Applications to birational maps

Recall (p.114) that a regular map $\varphi\colon W \to V$ of irreducible varieties is said to be ***birational*** if it induces an isomorphism $k(V) \to k(W)$ on the fields of rational functions.

8.56. One may ask how a birational regular map $\varphi\colon W \to V$ can fail to be an isomorphism. Here are three examples.

(a) The inclusion of an open subset into a variety is birational.

(b) The map (8.31) from $\mathbb{A}^1$ to the cuspidal cubic,

$$\mathbb{A}^1 \to C, \quad t \mapsto (t^2, t^3),$$

is birational. Here C is the cubic $Y^2 = X^3$, and the map $k[C] \to k[\mathbb{A}^1] = k[T]$ identifies $k[C]$ with the subring $k[T^2, T^3]$ of $k[T]$. Both rings have $k(T)$ as their fields of fractions.

(c) For any smooth variety V and point $P \in V$, there is a regular birational map $\varphi\colon V' \to V$ such that the restriction of φ to $V' \smallsetminus \varphi^{-1}(P)$ is an isomorphism onto $V \smallsetminus P$, but $\varphi^{-1}(P)$ is the projective space attached to the vector space $T_P(V)$. See the section on blow-ups below.

The next result says that, if we require the target variety to be normal (thereby excluding example (b)), and we require the map to be quasi-finite (thereby excluding example (c)), then we are left with (a).

PROPOSITION 8.57. *Let $\varphi\colon W \to V$ be a birational regular map of irreducible varieties. If V is normal and the map φ is quasi-finite, then φ is an isomorphism from W onto an open subvariety of V.*

PROOF. Factor φ as in the Theorem 8.45. For each open affine subset U of V, $k[\varphi'^{-1}(U)]$ is the integral closure of $k[U]$ in $k(W)$. But $k(W) = k(V)$ (because φ is birational), and $k[U]$ is integrally closed in $k(V)$ (because V is normal), and so $U = \varphi'^{-1}(U)$ (as varieties). It follows that $V' = V$. □

8.58. In topology, a continuous bijective map $\varphi\colon W \to V$ need not be a homeomorphism, but it is if W is compact and V is Hausdorff. Similarly, a bijective regular map of algebraic varieties need not be an isomorphism. Here are three examples:

(a) In characteristic p, the Frobenius map

$$(x_1, \ldots, x_n) \mapsto (x_1^p, \ldots, x_n^p)\colon \mathbb{A}^n \to \mathbb{A}^n$$

is bijective and regular, but it is not an isomorphism even though $\mathbb{A}^n$ is normal.

(b) The map $t \mapsto (t^2, t^3)$ from $\mathbb{A}^1$ to the cuspidal cubic (see 8.56b) is bijective, but not an isomorphism.

(c) Consider the regular map $\mathbb{A}^1 \to \mathbb{A}^1$ sending x to $1/x$ for $x \neq 0$ and 0 to 0. Its graph Γ is the union of $(0,0)$ and the hyperbola $xy = 1$, which is a closed subvariety of $\mathbb{A}^1 \times \mathbb{A}^1$. The projection $(x, y) \mapsto x : \Gamma \to \mathbb{A}^1$ is a bijective, regular, birational map, but it is not an isomorphism even though $\mathbb{A}^1$ is normal.

If we require the map to be birational (thereby excluding example (a)), V to be normal (thereby excluding example (b)), and the varieties to be irreducible (thereby excluding example (c)), then the map is an isomorphism.

PROPOSITION 8.59. *Let $\varphi: W \to V$ be a bijective regular map of irreducible algebraic varieties. If the map φ is birational and V is normal, then φ is an isomorphism.*

PROOF. The hypotheses imply that φ is an isomorphism of W onto an open subset of V (8.57). Because φ is bijective, the open subset must be the whole of V. □

In fact, example (a) can be excluded by requiring that φ be generically separable (instead of birational).

PROPOSITION 8.60. *Let $\varphi: W \to V$ be a bijective regular map of irreducible varieties. If V is normal and $k(W)$ is separably generated over $k(V)$, then φ is an isomorphism.*

PROOF. Because φ is bijective, $\dim(W) = \dim(V)$ (see Theorem 9.9 below) and the separable degree of $k(W)$ over $k(V)$ is 1 (apply 8.40 to the variety V' in 8.45). Hence φ is birational, and we may apply (8.59). □

8.61. In functional analysis, the closed graph theorem states that, if a linear map $\varphi: W \to V$ between two Banach spaces has a closed graph $\Gamma \overset{\text{def}}{=} \{(w, \varphi w) \mid w \in W\}$, then φ is continuous (q.v. Wikipedia). One can ask (cf. mo113858) whether a similar statement is true in algebraic geometry. Specifically, if $\varphi: W \to V$ is a map (in the set-theoretic sense) of algebraic varieties V, W whose graph is closed (for the Zariski topology), then is φ a regular map? The answer is no in general. For example, even in characteristic zero, the map $(t^2, t^3) \to t : C \to \mathbb{A}^1$ inverse to that in (8.56b) has closed graph but is not regular. In characteristic p, the inverse of the Frobenius map $x \mapsto x^p$ provides another counterexample. For a third counterexample, see (8.58c). The projection π from Γ to W is a bijective regular map, and so φ will be regular if π is an isomorphism. According to (8.60), π is an isomorphism if the varieties are irreducible, W is normal, and π is generically separable. In particular, a map between irreducible normal algebraic varieties in characteristic zero is regular if its graph is closed.

Variants of Zariski's main theorem

Mumford, 1966,[4] III, §9, lists the following variants of ZMT.

Original form (8.57) Let $\varphi: W \to V$ be a birational regular map of irreducible varieties. If V is normal and φ is quasi-finite, then φ is an isomorphism of W onto an open subvariety of V.

[4] Introduction to Algebraic Geometry, Harvard notes. Reprinted as "The Red Book of Varieties and Schemes" (with the introduction of misprints) by Springer 1999.

Topological form Let V be a normal variety over $\mathbb{C}$, and let $v \in V$. Let S be the singular locus of V. Then the complex neighbourhoods U of v such that $U \smallsetminus U \cap S$ is connected form a base for the system of complex neighbourhoods of v.

Power series form Let V be a normal variety, and let $\mathcal{O}_{V,Z}$ be the local ring attached to an irreducible closed subset of V (cf. p.175). If $\mathcal{O}_{V,Z}$ is an integrally closed integral domain, then so also is its completion.

Grothendieck's form (8.45) Every quasi-finite map of algebraic varieties factors as the composite of an open immersion with a finite map.

Connectedness theorem Let $\varphi\colon W \to V$ be a proper birational map, and let v be a (closed) normal point of V. The $\varphi^{-1}(v)$ is a connected set (in the Zariski topology).

The original form of the theorem was proved by Zariski using a fairly direct argument whose method doesn't seem to generalize.[5] The power series form was also proved by Zariski, who showed that it implied the original form. The last two forms are much deeper and were proved by Grothendieck. See the discussion in Mumford 1966.

NOTES. The original form of the theorem (8.57) is the "Main theorem" of Zariski, O., Foundations of a general theory of birational correspondences. Trans. Amer. Math. Soc. 53, (1943). 490–542.

f Stein factorization

The following important theorem shows that the fibres of a proper map are disconnected only because the fibres of finite maps are disconnected.

THEOREM 8.62 (STEIN FACTORIZATION). *Every proper map $\varphi\colon W \to V$ of algebraic varieties factors into $W \overset{\varphi_1}{\to} W' \overset{\varphi_2}{\to} V$ with φ_1 proper with connected fibres and φ_2 finite.*

When V is affine, this is the factorization

$$W \to \operatorname{Spm}(\mathcal{O}_W(W)) \to V.$$

The first major step in the proof of the theorem is to show that $\varphi_*\mathcal{O}_W$ is a coherent sheaf on V. Here $\varphi_*\mathcal{O}_W$ is the sheaf of $\mathcal{O}_V$-algebras on V,

$$U \rightsquigarrow \mathcal{O}_W(\varphi^{-1}(U)).$$

To say that $\varphi_*\mathcal{O}_W$ is coherent means that, on every open affine subset U of V, it is the sheaf of $\mathcal{O}_U$-algebras defined by a finite $k[U]$-algebra. This, in turn, means that there exists a regular map $\varphi_2\colon \operatorname{Spm}(\varphi_*\mathcal{O}_W) \to V$ that, over every open affine subset U of V, is the map attached by Spm to the map of k-algebras $k[U] \to \mathcal{O}_W(\varphi^{-1}(U))$.

The Stein factorization is then

$$W \xrightarrow{\varphi_1} W' \overset{\text{def}}{=} \operatorname{Spm}(\varphi_*\mathcal{O}_W) \xrightarrow{\varphi_2} V.$$

By construction, φ_2 is finite and $\varphi_1\colon W \to W'$ has the property that $\mathcal{O}_{W'} \to \varphi_{1*}\mathcal{O}_W$ is an isomorphism. That its fibres are connected is a consequence of the following extension of Zariski's connectedness theorem to non birational maps.

[5] See Lang, S., Introduction to Algebraic Geometry, 1958, V 2, for Zariski's original statement and proof of this theorem. See Springer, T.A., Linear Algebraic Groups, 1998, 5.2.8, for a direct proof of (8.59).

THEOREM 8.63. *Let $\varphi: W \to V$ be a proper map such that the map $\mathcal{O}_V \to \varphi_*\mathcal{O}_W$ is an isomorphism. Then the fibres of φ are connected.*

See Hartshorne 1977, III, §11.

NOTES. The Stein factorization was originally proved by Stein for complex spaces (q.v. Wikipedia).

g Blow-ups

Under construction.

Let P be a nonsingular point on an algebraic variety V, and let $T_p(V)$ be the tangent space at P. The blow-up of V at P is a regular map $\tilde{V} \to V$ that replaces P with the projective space $\mathbb{P}(T_P(V))$. More generally, the blow-up at P replaces P with $\mathbb{P}(C_P(V))$ where $C_P(V)$ is the geometric tangent cone at P.

Blowing up the origin in $\mathbb{A}^n$

Let O be the origin in $\mathbb{A}^n$, and let $\pi: \mathbb{A}^n \smallsetminus \{O\} \to \mathbb{P}^{n-1}$ be the map $(a_1,\ldots,a_n) \mapsto (a_1:\ldots:a_n)$. Let Γ_π be the graph of π, and let $\widetilde{\mathbb{A}^n}$ be the closure of Γ_π in $\mathbb{A}^n \times \mathbb{P}^{n-1}$. The map $\sigma: \widetilde{\mathbb{A}^n} \to \mathbb{A}^n$ defined by the projection map $\mathbb{A}^n \times \mathbb{P}^{n-1} \to \mathbb{A}^n$ is the blow-up of $\mathbb{A}^n$ at O.

Blowing up a point on a variety

Examples

8.64. The nodal cubic

8.65. The cuspidal cubic

h Resolution of singularities

Let V be an algebraic variety. A ***desingularization*** of V is birational regular map $\pi: W \to V$ such that W is nonsingular and π is proper; if V is projective, then W should also be projective, and π should induce an isomorphism

$$W \smallsetminus \pi^{-1}(\mathrm{Sing}(V)) \to V \smallsetminus \mathrm{Sing}(V).$$

In other words, the nonsingular variety W is the same as V except over the singular locus of V. When a variety admits a desingularization, then we say that ***resolution of singularities*** holds for V.

Note that with "nonsingular" replaced by "normalization", the normalization of V (see 8.5) provides such a map (resolution of abnormalities).

Nagata's embedding theorem 7.50 shows that it suffices to prove resolution of singularities for complete varieties, and Chow's lemma 7.39 then shows that it suffices to prove resolution of singularities for projective varieties. From now on, we shall consider only projective varieties.

Resolution of singularities for curves was first obtained using blow-ups (see Chapter 7 of Fulton's book, Algebraic Curves). Zariski introduced the notion of the normalization of a variety, and observed that the normalization $\pi: \tilde{V} \to V$ of a curve V in $k(V)$ is a desingularization of V.

There were several proofs of resolution of singularities for surfaces over $\mathbb{C}$, but the first to be accepted as rigorous is that of Walker (patching Jung's local arguments; 1935). For a surface V, normalization gives a surface with only point singularities (8.12), which can then be blown up. Zariski showed that the desingularization of a surface in characteristic zero can be obtained by alternating normalizations and blow-ups.

The resolution of singularities for three-folds in characteristic zero is much more difficult, and was first achieved by Zariski (Ann. of Math. 1944). His result was extended to nonzero characteristic by his student Abhyankar and to all varieties in characteristic zero by his student Hironaka.

The resolution of singularities for higher dimensional varieties in nonzero characteristic is one of the most important outstanding problems in algebraic geometry. In 1996, de Jong proved a weaker result in which, instead of the map π being birational, $k(W)$ is allowed to be a finite extension of $k(V)$.

A little history

Normal varieties were introduced by Zariski in a paper, Amer. J. Math. 61, 1939, p.249–194. There he noted that the singular locus of a normal variety has codimension at least 2 and that the system of hyperplane sections of a normal variety relative to a projective embedding is complete (i.e., is a complete rational equivalence class). Zariski's introduction of the notion of a normal variety and of the normalization of a variety was an important insertion of commutative algebra into algebraic geometry. It is not easy to give a geometric intuition for "normal". One criterion is that a variety is normal if and only if every surjective finite birational map onto it is an isomorphism (8.57). See mo109395 for a discussion of this question.

Exercises

8-1. Prove that a finite map is an isomorphism if and only if it is bijective and étale. (Cf. Harris 1992, 14.9.)

8-2. Give an example of a surjective quasi-finite regular map that is not finite (different from any in the notes).

8-3. Let $\varphi: V \to W$ be a regular map with the property that $\varphi^{-1}(U)$ is an open affine subset of W whenever U is an open affine subset of V (such a map is said to be ***affine***). Show that if V is separated, then so also is W.

8-4. For every $n \geq 1$, find a finite map $\varphi: W \to V$ with the following property: for all $1 \leq i \leq n$,

$$V_i \stackrel{\text{def}}{=} \{P \in V \mid \varphi^{-1}(P) \text{ has } \leq i \text{ points}\}$$

is a nonempty closed subvariety of dimension i.

CHAPTER **9**

Regular Maps and Their Fibres

Consider again the regular map $\varphi: \mathbb{A}^2 \to \mathbb{A}^2$, $(x,y) \mapsto (x,xy)$ (Exercise 3-3). The line $X = c$ maps to the line $Y = cX$. As c runs over the elements of k, this line sweeps out the whole x,y-plane except for the y-axis, and so the image of φ is

$$C = (\mathbb{A}^2 \smallsetminus \{y\text{-axis}\}) \cup \{(0,0)\},$$

which is neither open nor closed, and, in fact, is not even locally closed. The fibre

$$\varphi^{-1}(a,b) = \begin{cases} \text{point } (a, b/a) & \text{if } a \neq 0 \\ Y\text{-axis} & \text{if } (a,b) = (0,0) \\ \emptyset & \text{if } a = 0,\, b \neq 0. \end{cases}$$

From this unpromising example, it would appear that it is not possible to say anything about the image of a regular map or its fibres. However, it turns out that almost everything that can go wrong already goes wrong for in this example. We shall show:

(a) the image of a regular map is a finite union of locally closed sets;

(b) the dimensions of the fibres can jump only over closed subsets;

(c) the number of elements (if finite) in the fibres can drop only on closed subsets, provided the map is finite, the target variety is normal, and k has characteristic zero.

a The constructibility theorem

THEOREM 9.1. *Let $\varphi: W \to V$ be a dominant regular map of irreducible affine algebraic varieties. Then $\varphi(W)$ contains an open dense subset of V.*

PROOF. Because φ is dominant, the map $f \mapsto f \circ \varphi: k[V] \to k[W]$ is injective. According to Lemma 9.4 below, there exists a nonzero $a \in k[V]$ such that every homomorphism $\alpha: k[V] \to k$ such that $\alpha(a) \neq 0$ extends to a homomorphism $\beta: k[W] \to k$ with $\beta(1) \neq 0$. In particular, for $P \in D(a)$, the homomorphism $g \mapsto g(P): k[V] \to k$ extends to a nonzero homomorphism $\beta: k[W] \to k$. The kernel of β is a maximal ideal of $k[W]$ whose zero set is a point Q of W such that $\varphi(Q) = P$. □

Before beginning the proof of Lemma 9.4, we should look at an example.

EXAMPLE 9.2. Let A be an affine k-algebra, and let $B = A[T]/(f)$ with $f = a_m T^m + \cdots + a_0$. When does a homomorphism $\alpha\colon A \to k$ extend to B? The extensions of α correspond to roots of the polynomial $\alpha(a_m) T^m + \cdots + \alpha(a_0)$ in k, and so there exists an extension unless this is a nonzero constant polynomial. In particular, α extends if $\alpha(a_m) \neq 0$.

LEMMA 9.3. *Let $A \subset B$ be finitely generated k-algebras. Assume that A and B are integral domains, and that B is generated by a single element, say, $B = A[t] \simeq A[T]/\mathfrak{a}$. Let $\mathfrak{c} \subset A$ be the set of leading coefficients of the polynomials in $\mathfrak{a}$. Then every homomorphism $\alpha\colon A \to k$ such that $\alpha(\mathfrak{c}) \neq 0$ extends to a homomorphism $B \to k$.*

PROOF. Note that $\mathfrak{c}$ is an ideal in A. If $\mathfrak{a} = 0$, then every homomorphism α extends. Thus we may assume that $\mathfrak{a} \neq 0$. Let $f = a_m T^m + \cdots + a_0$ be a nonzero polynomial of minimum degree in $\mathfrak{a}$ such that $\alpha(a_m) \neq 0$. Because $B \neq 0$, we have that $m \geq 1$.

Extend α to a homomorphism $\tilde{\alpha}\colon A[T] \to k[T]$ by sending T to T. The k-submodule of $k[T]$ generated by $\tilde{\alpha}(\mathfrak{a})$ is an ideal (because $T \cdot \sum c_i \tilde{\alpha}(g_i) = \sum c_i \tilde{\alpha}(g_i T)$).

Unless $\tilde{\alpha}(\mathfrak{a})$ contains a nonzero constant, it generates a proper ideal in $k[T]$, which will have a zero c in k (2.11). The homomorphism

$$A[T] \xrightarrow{\tilde{\alpha}} k[T] \xrightarrow{h \mapsto h(c)} k, \qquad T \mapsto T \mapsto c$$

then factors through $A[T]/\mathfrak{a} = B$ and extends α.

In the contrary case, $\mathfrak{a}$ contains a polynomial

$$g(T) = b_n T^n + \cdots + b_0, \quad \alpha(b_i) = 0 \quad (i > 0), \quad \alpha(b_0) \neq 0.$$

On dividing $f(T)$ into $g(T)$, we find that

$$a_m^d g(T) = q(T) f(T) + r(T), \quad d \in \mathbb{N}, \quad q, r \in A[T], \quad \deg r < m.$$

On applying $\tilde{\alpha}$ to this equation, we obtain

$$\alpha(a_m)^d \alpha(b_0) = \tilde{\alpha}(q)\tilde{\alpha}(f) + \tilde{\alpha}(r).$$

Because $\tilde{\alpha}(f)$ has degree $m > 0$, we must have $\tilde{\alpha}(q) = 0$, and so $\tilde{\alpha}(r)$ is a nonzero constant. After replacing $g(T)$ with $r(T)$, we may assume $n < m$. If $m = 1$, such a $g(T)$ can't exist, and so we may suppose $m > 1$ and (by induction) that the lemma holds for smaller values of m.

For $h(T) = c_r T^r + c_{r-1} T^{r-1} + \cdots + c_0$, let $h'(T) = c_r + \cdots + c_0 T^r$. Then the A-module generated by the polynomials $T^s h'(T)$, $s \geq 0$, $h \in \mathfrak{a}$, is an ideal $\mathfrak{a}'$ in $A[T]$. Moreover, $\mathfrak{a}'$ contains a nonzero constant if and only if $\mathfrak{a}$ contains a nonzero polynomial cT^r, which implies $t = 0$ and $A = B$ (since B is an integral domain).

If $\mathfrak{a}'$ does not contain nonzero constants, then set $B' = A[T]/\mathfrak{a}' = A[t']$. Then $\mathfrak{a}'$ contains the polynomial $g' = b_n + \cdots + b_0 T^n$, and $\alpha(b_0) \neq 0$. Because $\deg g' < m$, the induction hypothesis implies that α extends to a homomorphism $B' \to k$. Therefore, there is a $c \in k$ such that, for all $h(T) = c_r T^r + c_{r-1} T^{r-1} + \cdots + c_0 \in \mathfrak{a}$,

$$h'(c) = \alpha(c_r) + \alpha(c_{r-1})c + \cdots + c_0 c^r = 0.$$

On taking $h = g$, we see that $c = 0$, and on taking $h = f$, we obtain the contradiction $\alpha(a_m) = 0$. $\square$

LEMMA 9.4. *Let $A \subset B$ be finitely generated k-algebras. Assume that A and B are integral domains, and let b be a nonzero element of B. Then there exists a nonzero $a \in A$ with the following property: every homomorphism $\alpha: A \to k$ from A into k such that $\alpha(a) \neq 0$ extends to a homomorphism $\beta: B \to k$ such that $\beta(b) \neq 0$.*

PROOF Suppose that we know the proposition in the case that B is generated by a single element, and write $B = A[x_1, \ldots, x_n]$. Then there exists an element $b_{n-1} \in A[x_1, \ldots, x_{n-1}]$ with the following property: every homomorphism $\alpha: A[x_1, \ldots, x_{n-1}] \to k$ such that $\alpha(b_{n-1}) \neq 0$ extends to a homomorphism $\beta: B \to k$ such that $\beta(b) \neq 0$. Then there exists a $b_{n-2} \in A[x_1, \ldots, x_{n-2}]$ etc. Continuing in this fashion, we obtain an element $a \in A$ with the required property.

Thus we may assume $B = A[x]$. Let $\mathfrak{a}$ be the kernel of the homomorphism $X \mapsto x$, $A[X] \to A[x]$.

Case (i). The ideal $\mathfrak{a} = (0)$. Write

$$b = f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n, \quad a_i \in A,$$

and take $a = a_0$. If $\alpha: A \to k$ is such that $\alpha(a_0) \neq 0$, then there exists a $c \in k$ such that $f(c) \neq 0$, and we can take β to be the homomorphism $\sum d_i x^i \mapsto \sum \alpha(d_i) c^i$.

Case (ii). The ideal $\mathfrak{a} \neq (0)$. Let

$$f(T) = a_m T^m + \cdots a_0, \quad a_m \neq 0,$$

be an element of $\mathfrak{a}$ of minimum degree. Let $h(T) \in A[T]$ represent b. As b is nonzero, $h \notin \mathfrak{a}$. Because f is irreducible over the field of fractions of A, it and h are coprime over that field. Hence there exist $u, v \in A[T]$ and $c \in A \smallsetminus \{0\}$ such that

$$uh + vf = c.$$

It follows now that ca_m satisfies our requirements, for if $\alpha(ca_m) \neq 0$, then α can be extended to $\beta: B \to k$ by the preceding lemma, and $\beta(u(x) \cdot b) = \beta(c) \neq 0$, and so $\beta(b) \neq 0$. □

ASIDE 9.5. It is also possible to deduce Theorem 9.1 from the generic freeness theorem (CA 21.11).

In order to generalize (9.1) to arbitrary maps of arbitrary varieties, we need the notion of a constructible set. Let W be a topological space. A subset C of W is said to ***constructible*** if it is a finite union of sets of the form $U \cap Z$ with U open and Z closed. Obviously, if C is constructible and $V \subset W$, then $C \cap V$ is constructible. A constructible set in $\mathbb{A}^n$ is definable by a finite number of polynomials; more precisely, it is defined by a finite number of statements of the form

$$f(X_1, \cdots, X_n) = 0, \qquad g(X_1, \cdots, X_n) \neq 0$$

combined using only "and" and "or" (or, better, statements of the form $f = 0$ combined using "and", "or", and "not"). The next proposition shows that a constructible set C that is dense in an irreducible variety V must contain a nonempty open subset of V. Contrast $\mathbb{Q}$, which is dense in $\mathbb{R}$ (real topology), but does not contain an open subset of $\mathbb{R}$, or any infinite subset of $\mathbb{A}^1$ that omits an infinite set.

PROPOSITION 9.6. *Let C be a constructible set whose closure $\bar{C}$ is irreducible. Then C contains a nonempty open subset of $\bar{C}$.*

PROOF. We are given that $C = \bigcup(U_i \cap Z_i)$ with each U_i open and each Z_i closed. We may assume that each set $U_i \cap Z_i$ in this decomposition is nonempty. Clearly $\bar{C} \subset \bigcup Z_i$, and as $\bar{C}$ is irreducible, it must be contained in one of the Z_i. For this i

$$C \supset U_i \cap Z_i \supset U_i \cap \bar{C} \supset U_i \cap C \supset U_i \cap (U_i \cap Z_i) = U_i \cap Z_i.$$

Thus $U_i \cap Z_i = U_i \cap \bar{C}$ is a nonempty open subset of $\bar{C}$ contained in C. □

THEOREM 9.7. *A regular map $\varphi: W \to V$ sends constructible sets to constructible sets. In particular, if U is a nonempty open subset of W, then $\varphi(U)$ contains a nonempty open subset of its closure in V.*

PROOF We first prove the "in particular" statement of the Theorem. By considering suitable open affine coverings of W and V, we see that it suffices to prove this in the case that both W and V are affine. If $W_1, \ldots, W_r$ are the irreducible components of W, then the closure of $\varphi(W)$ in V, $\varphi(W)^- = \varphi(W_1)^- \cup \ldots \cup \varphi(W_r)^-$, and so it suffices to prove the statement in the case that W is irreducible. We may also replace V with $\varphi(W)^-$, and so assume that both W and V are irreducible. We are in the situation of (9.1).

We now prove the theorem. Let W_i be the irreducible components of W. Then $C \cap W_i$ is constructible in W_i, and $\varphi(W)$ is the union of the $\varphi(C \cap W_i)$; it is therefore constructible if the $\varphi(C \cap W_i)$ are. Hence we may assume that W is irreducible. Moreover, C is a finite union of its irreducible components, and these are closed in C; they are therefore constructible. We may therefore assume that C also is irreducible; $\bar{C}$ is then an irreducible closed subvariety of W.

We shall prove the theorem by induction on the dimension of W. If $\dim(W) = 0$, then the statement is obvious because W is a point. If $\bar{C} \neq W$, then $\dim(\bar{C}) < \dim(W)$, and because C is constructible in $\bar{C}$, we see that $\varphi(C)$ is constructible (by induction). We may therefore assume that $\bar{C} = W$. But then $\bar{C}$ contains a nonempty open subset of W, and so the case just proved shows that $\varphi(C)$ contains an nonempty open subset U of its closure. Replace V be the closure of $\varphi(C)$, and write

$$\varphi(C) = U \cup \varphi(C \cap \varphi^{-1}(V - U)).$$

Then $\varphi^{-1}(V - U)$ is a proper closed subset of W (the complement of $V - U$ is dense in V and φ is dominant). As $C \cap \varphi^{-1}(V - U)$ is constructible in $\varphi^{-1}(V - U)$, the set $\varphi(C \cap \varphi^{-1}(V - U))$ is constructible in V by induction, which completes the proof. □

ASIDE 9.8. Let X be a subset of $\mathbb{C}^n$. If X is constructible for the Zariski topology on $\mathbb{C}^n$, then the closure of X for the Zariski topology is equal to its closure for the complex topology.

b The fibres of morphisms

We wish to examine the fibres of a regular map $\varphi: W \to V$. We can replace V by the closure of $\varphi(W)$ in V and so assume that φ is dominant.

THEOREM 9.9. *Let $\varphi: W \to V$ be a dominant regular map of irreducible varieties. Then*

(a) $\dim(W) \geq \dim(V)$;

(b) *if $P \in \varphi(W)$, then*

$$\dim(\varphi^{-1}(P)) \geq \dim(W) - \dim(V)$$

for every $P \in V$, with equality holding exactly on a nonempty open subset U of V.

(c) *The sets*

$$V_i = \{P \in V \mid \dim(\varphi^{-1}(P)) \geq i\}$$

are closed in $\varphi(W)$.

In other words, for P on a dense open subset U of V, the fibre $\varphi^{-1}(P)$ has the expected dimension $\dim(W) - \dim(V)$. On the closed complement of U (possibly empty), the dimension of the fibre is $> \dim(W) - \dim(V)$, and it may jump further on closed subsets.

Before proving the theorem, we should look at an example.

EXAMPLE 9.10. Consider the subvariety $W \subset V \times \mathbb{A}^m$ defined by r linear equations

$$\sum_{j=1}^{m} a_{ij} X_j = 0, \quad a_{ij} \in k[V], \quad i = 1, \ldots, r,$$

and let φ be the projection $W \to V$. For $P \in V$, $\varphi^{-1}(P)$ is the set of solutions of system of equations

$$\sum_{j=1}^{m} a_{ij}(P) X_j = 0, \quad a_{ij}(P) \in k, \quad i = 1, \ldots, r,$$

and so its dimension is $m - \operatorname{rank}(a_{ij}(P))$. Since the rank of the matrix $(a_{ij}(P))$ drops on closed subsets, the dimension of the fibre jumps on closed subsets. More precisely, for each $r \in \mathbb{N}$,

$$\{P \in V \mid \operatorname{rank}(a_{ij}(P)) \leq r\}$$

is a closed subset of V (see Exercise 2-2); hence, for each $r' \in \mathbb{N}$,

$$\{P \in V \mid \dim \varphi^{-1}(P) \geq r'\}$$

is closed in V.

PROOF. (a) Because the map is dominant, there is a homomorphism $k(V) \hookrightarrow k(W)$, and obviously $\operatorname{tr\,deg}_k k(V) \leq \operatorname{tr\,deg}_k k(W)$ (an algebraically independent subset of $k(V)$ remains algebraically independent in $k(W)$).

(b) In proving the first part of (b), we may replace V by any open neighbourhood of P. In particular, we can assume V to be affine. Let m be the dimension of V. From (3.47) we know that there exist regular functions $f_1, \ldots, f_m$ such that P is an irreducible component of $V(f_1, \ldots, f_m)$. After replacing V by a smaller neighbourhood of P, we can suppose that $P = V(f_1, \ldots, f_m)$. Then $\varphi^{-1}(P)$ is the zero set of the regular functions $f_1 \circ \varphi, \ldots, f_m \circ \varphi$, and so (if nonempty) has codimension $\leq m$ in W (see 3.45). Hence

$$\dim \varphi^{-1}(P) \geq \dim W - m = \dim(W) - \dim(V).$$

In proving the second part of (b), we can replace both W and V with open affine subsets. Since φ is dominant, $k[V] \to k[W]$ is injective, and we may regard it as an inclusion (we identify a function x on V with $x \circ \varphi$ on W). Then $k(V) \subset k(W)$. Write $k[V] = k[x_1, \ldots, x_M]$ and $k[W] = k[y_1, \ldots, y_N]$, and suppose V and W have dimensions m and n respectively. Then $k(W)$ has transcendence degree $n - m$ over $k(V)$, and we may suppose that $y_1, \ldots, y_{n-m}$ are algebraically independent over $k[x_1, \ldots, x_m]$, and that the remaining y_i are algebraic over $k[x_1, \ldots, x_m, y_1, \ldots, y_{n-m}]$. There are therefore relations

$$F_i(x_1, \ldots, x_m, y_1, \ldots, y_{n-m}, y_i) = 0, \quad i = n-m+1, \ldots, N. \tag{37}$$

with $F_i(X_1,\ldots,X_m,Y_1,\ldots,Y_{n-m},Y_i)$ a nonzero polynomial. We write $\bar{y}_i$ for the restriction of y_i to $\varphi^{-1}(P)$. Then

$$k[\varphi^{-1}(P)] = k[\bar{y}_1,\ldots,\bar{y}_N].$$

The equations (37) give an algebraic relation among the functions $x_1,\ldots,y_i$ on W. When we restrict them to $\varphi^{-1}(P)$, they become equations:

$$F_i(x_1(P),\ldots,x_m(P),\bar{y}_1,\ldots,\bar{y}_{n-m},\bar{y}_i) = 0, \quad i = n-m+1,\ldots,N.$$

If these are nontrivial algebraic relations, i.e., if none of the polynomials

$$F_i(x_1(P),\ldots,x_m(P),Y_1,\ldots,Y_{n-m},Y_i)$$

is identically zero, then the transcendence degree of $k(\bar{y}_1,\ldots,\bar{y}_N)$ over k will be $\leq n-m$.

Thus, regard $F_i(x_1,\ldots,x_m,Y_1,\ldots,Y_{n-m},Y_i)$ as a polynomial in the Y's with coefficients polynomials in the x's. Let V_i be the closed subvariety of V defined by the simultaneous vanishing of the coefficients of this polynomial—it is a proper closed subset of V. Let $U = V \smallsetminus \bigcup V_i$—it is a nonempty open subset of V. If $P \in U$, then none of the polynomials $F_i(x_1(P),\ldots,x_m(P),Y_1,\ldots,Y_{n-m},Y_i)$ is identically zero, and so for $P \in U$, the dimension of $\varphi^{-1}(P)$ is $\leq n-m$, and hence $= n-m$ by (a).

Finally, if for a particular point P, $\dim \varphi^{-1}(P) = n-m$, then we can modify the above argument to show that the same is true for all points in an open neighbourhood of P.

(c) We prove this by induction on the dimension of V—it is obviously true if $\dim V = 0$. We know from (b) that there is an open subset U of V such that

$$\dim \varphi^{-1}(P) = n-m \iff P \in U.$$

Let Z be the complement of U in V; thus $Z = V_{n-m+1}$. Let $Z_1,\ldots,Z_r$ be the irreducible components of Z. On applying the induction to the restriction of φ to the map $\varphi^{-1}(Z_j) \to Z_j$ for each j, we obtain the result. □

Recall that a regular map $\varphi\colon W \to V$ of algebraic varieties is closed if, for example, W is complete (7.7).

PROPOSITION 9.11. *Let $\varphi\colon W \to V$ be a regular surjective closed map of varieties, and let $n \in \mathbb{N}$. If V is irreducible and all fibres $\varphi^{-1}(P)$ of φ are irreducible of dimension n, then W is irreducible of dimension* $\dim(V)+n$.

PROOF. Let Z be a irreducible closed subset of W, and consider the map $\varphi|Z\colon Z \to V$; it has fibres $(\varphi|Z)^{-1}(P) = \varphi^{-1}(P) \cap Z$. There are three possibilities.

(a) $\varphi(Z) \neq V$. Then $\varphi(Z)$ is a proper closed subset of V.

(b) $\varphi(Z) = V$, $\dim(Z) < n + \dim(V)$. Then (b) of (9.9) shows that there is a nonempty open subset U of V such that for $P \in U$,

$$\dim(\varphi^{-1}(P) \cap Z) = \dim(Z) - \dim(V) < n.$$

Thus, for $P \in U$, the fibre $\varphi^{-1}(P)$ is not contained in Z.

(c) $\varphi(Z) = V$, $\dim(Z) \geq n + \dim(V)$. Then (b) of (9.9) shows that

$$\dim(\varphi^{-1}(P) \cap Z) \geq \dim(Z) - \dim(V) \geq n$$

for all P; thus $\varphi^{-1}(P) \subset Z$ for all $P \in V$, and so $Z = W$; moreover $\dim Z = n$.

Now let $Z_1,\ldots,Z_r$ be the irreducible components of W. I claim that (c) holds for at least one of the Z_i. Otherwise, there will be an open subset U of V such that for P in U, $\varphi^{-1}(P)$ is contained in *none* of the Z_i; but $\varphi^{-1}(P)$ is irreducible and $\varphi^{-1}(P) = \bigcup(\varphi^{-1}(P) \cap Z_i)$, and so this is impossible. □

c Flat maps and their fibres

Flat maps

Let A be a ring, and let B be an A-algebra. If the sequence of A-modules

$$0 \to N' \xrightarrow{\alpha} N \xrightarrow{\beta} N'' \to 0$$

is exact, then the sequence of B-modules

$$B \otimes_A N' \xrightarrow{1\otimes\alpha} B \otimes_A N \xrightarrow{1\otimes\beta} B \otimes_A N'' \to 0$$

is exact,[1] but $B \otimes_A N' \to B \otimes_A N$ need not be injective. For example, when we tensor the exact sequence of $k[X]$-modules

$$0 \to k[X] \xrightarrow{f \mapsto X\cdot f} k[X] \xrightarrow{f \mapsto f \bmod (X)} k[X]/(X) \to 0$$

with k, we get the sequence

$$k \xrightarrow{0} k \xrightarrow{\text{id}} k \to 0.$$

DEFINITION 9.12. An A-algebra B is ***flat*** if

$$M \to N \text{ injective} \implies B \otimes_A M \to B \otimes_A N \text{ injective.}$$

It is ***faithfully flat*** if, in addition,

$$B \otimes_A M = 0 \implies M = 0.$$

Therefore, an A-algebra B is flat if and only if the functor $M \rightsquigarrow B \otimes_A M$ from A-modules to B-modules is exact.

EXAMPLE 9.13. (a) Let S be a multiplicative subset of A. Then $S^{-1}A$ is a flat A-algebra (1.18). (b) Every open immersion is flat (obvious). (c) The composite of two flat maps is flat (obvious).

PROPOSITION 9.14. *Let $A \to A'$ be a homomorphism of rings. If $A \to B$ is flat, then so also is $A' \to B \otimes_A A'$.*

PROOF. For any A'-module M,

$$(B \otimes_A A') \otimes_{A'} M \simeq B \otimes_A (A' \otimes_{A'} M) \simeq B \otimes_A M.$$

In other words, tensoring an A'-module M with $B \otimes_A A'$ is the same as tensoring M (regarded as an A-module) with B. Therefore it preserves exact sequences. □

[1] The surjectivity of $1 \otimes \beta$ is obvious. Let $B \otimes_A N \xrightarrow{\phi} Q$ be the cokernel of $1 \otimes \alpha$. Because

$$(1 \otimes \beta) \circ (1 \otimes \alpha) = 1 \otimes (\beta \circ \alpha) = 0,$$

there is a unique A-linear map $f : Q \to B \otimes_A N''$ such that $f \circ \phi = 1 \otimes \beta$. We shall construct an inverse g to f. Let $b \in B$, and let $n \in N$. If $\beta(n) = 0$, then $n = \alpha(n')$ for some $n' \in N'$; hence $b \otimes n = b \otimes \alpha(n')$, and so $\phi(b \otimes n) = 0$. It follows by linearity that $\phi(b \otimes n_1) = \phi(b \otimes n_2)$ if $\beta(n_1) = \beta(n_2)$, and so the A-bilinear map

$$B \times N \to Q, \quad (b, n) \mapsto \phi(b \otimes n)$$

factors through $B \times N''$. It therefore defines an A-linear map g: $B \otimes_A N'' \to Q$. To show that f and g are inverse, it suffices to check that $g \circ f = \text{id}_Q$ on elements of the form $\phi(b \otimes n)$ and that $f \circ g = \text{id}_{B \otimes_A N''}$ on elements of the form $b \otimes \beta(n)$ — both are obvious.

PROPOSITION 9.15. *A homomorphism $\alpha: A \to B$ of rings is flat if and only if, for all maximal ideals $\mathfrak{n}$ in B, the map $A_{\alpha^{-1}(\mathfrak{n})} \to B_{\mathfrak{n}}$ is flat.*

PROOF. Let $\mathfrak{n}$ be a prime ideal of B, and let $\mathfrak{m} = \alpha^{-1}(\mathfrak{n})$ — it is a prime ideal in A.

If $A \to B$ is flat, then so is $A_{\mathfrak{m}} \to A_{\mathfrak{m}} \otimes_A B \simeq S_{\mathfrak{m}}^{-1}B$ (9.14). The map $S_{\mathfrak{m}}^{-1}B \to S_{\mathfrak{n}}^{-1}B = B_{\mathfrak{n}}$ is flat (9.13a), and so the composite $A_{\mathfrak{m}} \to B_{\mathfrak{n}}$ is flat (9.13c).

For the converse, let $N' \to N$ be an injective homomorphism of A-modules, and let $\mathfrak{n}$ be a maximal ideal of B. Then $A_{\mathfrak{m}} \otimes_A (N' \to N)$ is injective (9.13). Therefore, the map

$$B_{\mathfrak{n}} \otimes_A (N' \to N) \simeq B_{\mathfrak{n}} \otimes_{A_{\mathfrak{m}}} (A_{\mathfrak{m}} \otimes_A (N' \to N))$$

is injective, and so the kernel M of $B \otimes_A (N' \to N)$ has the property that $M_{\mathfrak{n}} = 0$. Let $x \in M$, and let $\mathfrak{a} = \{b \in B \mid bx = 0\}$. For each maximal ideal $\mathfrak{n}$ of B, x maps to zero in $M_{\mathfrak{n}}$, and so $\mathfrak{a}$ contains an element not in $\mathfrak{n}$. Hence $\mathfrak{a} = B$, and so $x = 0$. □

PROPOSITION 9.16. *A flat homomorphism $\varphi: A \to B$ is faithfully flat if and only if every maximal ideal $\mathfrak{m}$ of A is of the form $\varphi^{-1}(\mathfrak{n})$ for some maximal ideal $\mathfrak{n}$ of B.*

PROOF. $\Rightarrow$: Let $\mathfrak{m}$ be a maximal ideal of A, and let $M = A/\mathfrak{m}$; then

$$B \otimes_A M \simeq B/\varphi(\mathfrak{m})B.$$

As $B \otimes_A M \neq 0$, we see that $\varphi(\mathfrak{m})B \neq B$. Therefore $\varphi(\mathfrak{m})$ is contained in a maximal ideal $\mathfrak{n}$ of B. Now $\varphi^{-1}(\mathfrak{n})$ is a proper ideal in A containing $\mathfrak{m}$, and hence equals $\mathfrak{m}$.

$\Leftarrow$: Let M be a nonzero A-module. Let x be a nonzero element of M, and let $\mathfrak{a} = \mathrm{ann}(x) \overset{\text{def}}{=} \{a \in A \mid ax = 0\}$. Then $\mathfrak{a}$ is an ideal in A, and $M' \overset{\text{def}}{=} Ax \simeq A/\mathfrak{a}$. Moreover, $B \otimes_A M' \simeq B/\varphi(\mathfrak{a}) \cdot B$ and, because $A \to B$ is flat, $B \otimes_A M'$ is a submodule of $B \otimes_A M$. Because $\mathfrak{a}$ is proper, it is contained in a maximal ideal $\mathfrak{m}$ of A, and therefore

$$\varphi(\mathfrak{a}) \subset \varphi(\mathfrak{m}) \subset \mathfrak{n}$$

for some maximal ideal $\mathfrak{n}$ of A. Hence $\varphi(\mathfrak{a}) \cdot B \subset \mathfrak{n} \neq B$, and so $B \otimes_A M \supset B \otimes_A M' \neq 0$. □

COROLLARY 9.17. *A flat local homomorphism $A \to B$ of local rings is faithfully flat.*

PROOF. Let $\mathfrak{m}$ and $\mathfrak{n}$ be the (unique) maximal ideals of A and B. By hypothesis, $\mathfrak{n}^c = \mathfrak{m}$, and so the statement follows from the proposition. □

Properties of flat maps

LEMMA 9.18. *Let B be an A-algebra, and let $\mathfrak{p}$ be a prime ideal of A. The prime ideals of B contracting to $\mathfrak{p}$ are in natural one-to-one correspondence with the prime ideals of $B \otimes_A \kappa(\mathfrak{p})$.*

PROOF. Let $S = A \smallsetminus \mathfrak{p}$. Then $\kappa(\mathfrak{p}) = S^{-1}(A/\mathfrak{p})$. Therefore we obtain $B \otimes_A \kappa(\mathfrak{p})$ from B by first passing to $B/\mathfrak{p}B$ and then making the elements of A not in $\mathfrak{p}$ act invertibly. After the first step, we are left with the prime ideals $\mathfrak{q}$ of B such that $\mathfrak{q}^c \supset \mathfrak{p}$, and after the second step only with those such $\mathfrak{q}^c \cap S = \emptyset$, i.e., such that $\mathfrak{q}^c = \mathfrak{p}$. □

PROPOSITION 9.19. *Let B be a faithfully flat A-algebra. Every prime ideal $\mathfrak{p}$ of A is of the form $\mathfrak{q}^c$ for some prime ideal $\mathfrak{q}$ of B.*

PROOF. The ring $B \otimes_A \kappa(\mathfrak{p})$ is nonzero, because $\kappa(\mathfrak{p}) \neq 0$ and $A \to B$ is faithfully flat, and so it has a prime (even maximal) ideal $\mathfrak{q}$. For this ideal, $\mathfrak{q}^c = \mathfrak{p}$. □

SUMMARY 9.20. A flat homomorphism $\varphi: A \to B$ is faithfully flat if the image of

$$\operatorname{spec}(\varphi): \operatorname{spec}(B) \to \operatorname{spec}(A)$$

includes all maximal ideals of A, in which case it includes all prime ideals of A.

PROPOSITION 9.21 (GOING-DOWN THEOREM FOR FLAT MAPS). *Let $A \to B$ be a flat homomorphism. Let $\mathfrak{p} \supset \mathfrak{p}'$ be prime ideals in A, and let $\mathfrak{q}$ be a prime ideal in B such that $\mathfrak{q}^c = \mathfrak{p}$. Then $\mathfrak{q}$ contains a prime ideal $\mathfrak{q}'$ such that $\mathfrak{q}'^c = \mathfrak{p}'$:*

$$\begin{array}{ccccc} B & \quad & \mathfrak{q} & \supset & \mathfrak{q}' \\ | & & | & & | \\ A & & \mathfrak{p} & \supset & \mathfrak{p}'. \end{array}$$

PROOF. Because $A \to B$ is flat, the homomorphism $A_{\mathfrak{p}} \to B_{\mathfrak{q}}$ is flat, and because $\mathfrak{p}A_{\mathfrak{p}} = (\mathfrak{q}B_{\mathfrak{q}})^c$, it is faithfully flat (9.16). The ideal $\mathfrak{p}'A_{\mathfrak{p}}$ is prime (1.14), and so there exists a prime ideal of $B_{\mathfrak{q}}$ lying over $\mathfrak{p}'A_{\mathfrak{p}}$ (by 9.19). The contraction of this ideal to B is contained in $\mathfrak{q}$ and contracts to $\mathfrak{p}'$ in A. □

DEFINITION 9.22. A regular map $\varphi: W \to V$ of algebraic varieties is ***flat*** if, for all $P \in W$, the map $\mathcal{O}_{V,\varphi Q} \to \mathcal{O}_{W,P}$ is flat, and it is ***faithfully flat*** if it is flat and surjective.

PROPOSITION 9.23. *A regular map $\varphi: W \to V$ of affine algebraic varieties is flat (resp. faithfully flat) if and only if the map $f \mapsto f \circ \varphi: k[V] \to k[W]$ is flat (resp. faithfully flat).*

PROOF. Apply (9.15) and (9.16). □

PROPOSITION 9.24. *Let $\varphi: W \to V$ be a flat map of affine algebraic varieties. Let $S \subset S'$ be closed irreducible subsets of V, and let T be a closed irreducible subset of W such that $\varphi(T)$ is a dense subset of S. Then there exists a closed irreducible subset T' of W containing T and such that $\varphi(T')$ is a dense subset of S'.*

PROOF. Let $\mathfrak{p} = I(S)$, $\mathfrak{p}' = I(S')$, and $\mathfrak{q} = I(T)$. Then $\mathfrak{p} \supset \mathfrak{p}'$ because $S \subset S'$. Moreover $\mathfrak{q}^c = \mathfrak{p}$ because $T \xrightarrow{\varphi} S$ is dominant and so the map $k[S] = k[V]/\mathfrak{p} \to k[T]/\mathfrak{q}$ is injective. According to (9.21), there exists a prime ideal $\mathfrak{q}'$ in $k[W]$ contained in $\mathfrak{q}$ and such that $\mathfrak{q}'^c = \mathfrak{p}'$. Now $V(\mathfrak{q}')$ has the required properties. □

THEOREM 9.25 (GENERIC FLATNESS). *For every regular map $\varphi: W \to V$ of irreducible algebraic varieties, there exists a nonempty open subset U of V such that $\varphi^{-1}(U) \xrightarrow{\varphi} U$ is faithfully flat.*

PROOF. We may assume that W and V are affine, say, $V = \operatorname{Spm}(A)$ and $W = \operatorname{Spm}(B)$. Let F be the field of fractions of A. We regard B as a subring of $F \otimes_A B$.

As $F \otimes_A B$ is a finitely generated F-algebra, the Noether normalization theorem (2.45) shows that there exist elements $x_1, \ldots, x_m$ of $F \otimes_A B$ such that $F[x_1, \ldots, x_m]$ is a polynomial ring over F and $F \otimes_A B$ is a finite $F[x_1, \ldots, x_m]$-algebra. After multiplying each x_i by an element of A, we may suppose that it lies in B. Let $b_1, \ldots, b_n$ generate B as an A-algebra. Each b_i satisfies a monic polynomial equation with coefficients in $F[x_1, \ldots, x_m]$.

Let $a \in A$ be a common denominator for the coefficients of these polynomials. Then each b_i is integral over A_a. As the b_i generate B_a as an A_a-algebra, this shows that B_a is a finite $A_a[x_1,\ldots,x_m]$-algebra (1.36). Therefore, after replacing A with A_a and B with B_a, we may suppose that B is a finite $A[x_1,\ldots,x_m]$-algebra.

$$\begin{tikzcd}
B \arrow[r, "\text{injective}"] & F \otimes_A B \arrow[r] & E \otimes_{A[x_1,\ldots,x_m]} B \\
A[x_1,\ldots,x_m] \arrow[r] \arrow[u, "\text{finite}"'] & F[x_1,\ldots,x_m] \arrow[r] \arrow[u, "\text{finite}"'] & E \stackrel{\text{def}}{=} F(x_1,\ldots,x_n) \arrow[u, "\text{finite}"'] \\
A \arrow[r] \arrow[u] & F. \arrow[u] &
\end{tikzcd}$$

Let $E = F(x_1,\ldots,x_m)$ be the field of fractions of $A[x_1,\ldots,x_m]$, and let $b_1,\ldots,b_r$ be elements of B that form a basis for $E \otimes_{A[x_1,\ldots,x_m]} B$ as an E-vector space. Each element of B can be expressed as a linear combination of the b_i with coefficients in E. Let q be a common denominator for the coefficients arising from a set of generators for B as an $A[x_1,\ldots,x_m]$-module. Then $b_1,\ldots,b_r$ generate B_q as an $A[x_1,\ldots,x_m]_q$-module. In other words, the map

$$(c_1,\ldots,c_r) \mapsto \sum c_i b_i \colon A[x_1,\ldots,x_m]_q^r \to B_q \tag{*}$$

is surjective. This map becomes an isomorphism when tensored with E over $A[x_1,\ldots,x_m]_q$, which implies that each element of its kernel is killed by a nonzero element of $A[x_1,\ldots,x_m]_q$ and so is zero (because $A[x_1,\ldots,x_n]_q$ is an integral domain). Hence the map (*) is an isomorphism, and so B_q is free of finite rank over $A[x_1,\ldots,x_m]_q$. Let a be some nonzero coefficient of the polynomial q, and consider the maps

$$A_a \to A_a[x_1,\ldots,x_m] \to A_a[x_1,\ldots,x_m]_q \to B_{aq}.$$

The first and third arrows realize their targets as nonzero free modules over their sources, and so are faithfully flat. The middle arrow is flat by (9.13). Let $\mathfrak{m}$ be a maximal ideal in A_a. Then $\mathfrak{m}A_a[x_1,\ldots,x_m]$ does not contain the polynomial q because the coefficient a of q is invertible in A_a. Hence $\mathfrak{m}A_a[x_1,\ldots,x_m]_q$ is a proper ideal of $A_a[x_1,\ldots,x_m]_q$, and so the map $A_a \to A_a[x_1,\ldots,x_m]_q$ is faithfully flat (apply 9.16). This completes the proof. □

LEMMA 9.26. *Let V be an algebraic variety. A constructible subset C of V is closed if it has the following property: let Z be a closed irreducible subset of V; if $Z \cap C$ contains a dense open subset of Z, then $Z \subset C$.*

PROOF. Let Z be an irreducible component of $\bar{C}$. Then $Z \cap C$ is constructible and it is dense in Z, and so it contains a nonempty open subset U of Z (9.6). Hence $Z \subset C$. □

THEOREM 9.27. *A flat map $\varphi\colon W \to V$ of algebraic varieties is open.*

PROOF. Let U be an open subset of W. Then $\varphi(U)$ is constructible (9.7) and the going-down theorem (9.21) implies that $V \smallsetminus \varphi(U)$ satisfies the hypotheses of the lemma. Therefore $V \smallsetminus \varphi(U)$ is closed. □

COROLLARY 9.28. *Let $\varphi\colon W \to V$ be a regular map of irreducible algebraic varieties. Then there exists a dense open subset U of W such that $\varphi(U)$ is open, $U = \varphi^{-1}(\varphi U)$, and $U \xrightarrow{\varphi} \varphi(U)$ is flat.*

PROOF. According to (9.25), there exists a dense open subset U of V such that $\varphi^{-1}(U) \xrightarrow{\varphi} U$ is flat. In particular, $\varphi(\varphi^{-1}(U))$ is open in V (9.27). Note that $\varphi^{-1}(\varphi(\varphi^{-1}(U)) = \varphi^{-1}(U)$. Let $U' = \varphi^{-1}(U)$. Then U' is a dense open subset of W, $\varphi(U')$ is open, $U' = \varphi^{-1}(\varphi U')$, and $U' \xrightarrow{\varphi} \varphi(U')$ is flat. □

Fibres and flatness

The notion of flatness allows us to sharpen our earlier results.

PROPOSITION 9.29. *Let $\varphi\colon W \to V$ be a dominant map of irreducible algebraic varieties. Let $P \in \varphi(V)$. Then*

$$\dim\left(\varphi^{-1}(P)\right) \geq \dim(W) - \dim(V), \tag{38}$$

and equality holds if φ is flat.

PROOF. The inequality was proved in (9.9). If φ is flat, then we shall prove (more precisely) that, if Z is an irreducible component of $\varphi^{-1}(P)$, then

$$\dim(Z) = \dim(W) - \dim(V).$$

After replacing V with an open neighbourhood of P and W with an open subset intersecting V, we may suppose that both V and W are affine. Let

$$V \supset V_1 \supset \cdots \supset V_m = \{P\}$$

be a maximal chain of distinct irreducible closed subsets of V (so $m = \dim(V)$). Now $\varphi(Z) = \{P\}$, and so (see 9.24) there exists a chain of irreducible closed subsets

$$W \supset W_1 \supset \cdots \supset W_m = Z$$

such that $\varphi(W_i)$ is a dense subset of V_i. Let

$$Z \supset Z_1 \supset \cdots \supset Z_n$$

be a maximal chain of distinct irreducible closed subsets of V (so $n = \dim(Z)$). The existence of the chain

$$W \supset W_1 \supset \cdots \supset W_m \supset Z_1 \supset \cdots \supset Z_n$$

shows that

$$\dim(W) \geq m + n = \dim(V) + \dim(Z).$$

Together with (38), this implies that we have equality. □

PROPOSITION 9.30. *Let $\varphi\colon W \to V$ be a dominant map of irreducible algebraic varieties. Let $P \in \varphi(V)$. Then*

$$\dim\left(\varphi^{-1}(P)\right) \geq \dim(W) - \dim(V).$$

There exists a dense open subset U of W such that $\varphi(U)$ is open in V, $U = \varphi^{-1}(\varphi(U))$, and equality holds for all $P \in \varphi(U)$.

PROOF. Let U be an open subset of W as in (9.28). □

PROPOSITION 9.31. *Let $\varphi: W \to V$ be a dominant map of irreducible varieties. Let S be a closed irreducible subset of V, and let T be an irreducible component of $\varphi^{-1}(S)$ such that $\varphi(T)$ is dense in S. Then*

$$\dim(T) \geq \dim(S) + \dim(W) - \dim(V),$$

and equality holds if φ is flat.

PROOF. The inequality can be proved by a similar argument to that in (9.9) — see, for example, Hochschild 1981,[2] X, Theorem 2.1. The equality can be deduced by the same argument as in (9.29). □

PROPOSITION 9.32. *Let $\varphi: W \to V$ be a dominant map of irreducible varieties. There exists a nonempty open subset U of W such that $\varphi(U)$ is open, $U = \varphi^{-1}(\varphi U)$, and $U \xrightarrow{\varphi} \varphi(U)$ is flat. If S is a closed irreducible subset of V meeting $\varphi(U)$, and T is an irreducible component of $\varphi^{-1}(S)$ meeting U, then*

$$\dim(T) = \dim(S) + \dim(W) - \dim(V).$$

PROOF. Let U be an open subset of W as in (9.28). □

FINITE MAPS

PROPOSITION 9.33. *Let V be an irreducible algebraic variety. A finite map $\varphi: W \to V$ is flat if and only if*

$$\sum_{Q \mapsto P} \dim_k \mathcal{O}_Q/\mathfrak{m}_P\mathcal{O}_Q$$

is independent of $P \in W$.

PROOF. It suffices to prove this with V affine, in which case it follows from CA 12.6 (equivalence of (d) and (e)). □

The integer $\dim_k \mathcal{O}_Q/\mathfrak{m}_P\mathcal{O}_Q$ is the ***multiplicity*** of Q in its fibre. The theorem says that a finite map is flat if and only if the number of points in each fibre (counting multiplicities) is constant.

For example, let V be the subvariety of $\mathbb{A}^{n+1}$ defined by an equation

$$X^m + a_1X^{m-1} + \cdots + a_m = 0, \quad a_i \in k[T_1, \ldots, T_n]$$

and let $\varphi: V \to \mathbb{A}^n$ be the projection map (see p.49). The fibre over a point P of $\mathbb{A}^n$ is the set of points (P, c) with c a root of the polynomial

$$X^m + a_1(P)X^{m-1} + \cdots + a_m(P) = 0.$$

The multiplicity of (P, c) in its fibre is the multiplicity of c as a root of the polynomial. Therefore $\sum_{Q \mapsto P} \dim_k \mathcal{O}_Q/\mathfrak{m}_P\mathcal{O}_Q = m$ for every P, and so the map φ is flat.

[2] Hochschild, Gerhard P., Basic theory of algebraic groups and Lie algebras. Springer, 1981.

Criteria for flatness

THEOREM 9.34. *Let $\varphi\colon A \to B$ be a local homomorphism of noetherian local rings, and let $\mathfrak{m}$ be the maximal ideal of A. If A is regular, B is Cohen-Macaulay, and*

$$\dim(B) = \dim(A) + \dim(B/\mathfrak{m}B),$$

then φ is flat.

PROOF. Matsumura, H., Commutative Ring Theory, CUP 1986, 23.1. □

9.35. We don't define notion of being Cohen-Macaulay here (see ibid. p.134), but merely list some of its properties.

(a) A noetherian ring A is Cohen-Macaulay if and only if $A_{\mathfrak{m}}$ is Cohen-Macaulay for every maximal ideal $\mathfrak{m}$ of A (this is part of the definition).

(b) Zero-dimensional and reduced one-dimensional noetherian rings are Cohen-Macaulay (ibid. p.139).

(c) Regular noetherian rings are Cohen-Macaulay (ibid. p.137).

(d) Let $\varphi\colon A \to B$ be a flat local homomorphism of noetherian local rings, and let $\mathfrak{m}$ be the maximal ideal of A. Then B is Cohen-Macaulay if and only if both A and $B/\mathfrak{m}B$ are Cohen-Macaulay (ibid. p.181).

PROPOSITION 9.36. *Let $\varphi\colon A \to B$ be a finite homomorphism noetherian rings with A regular. Then φ is flat if and only if B is Cohen-Macaulay.*

PROOF. Note that $\dim(B/\mathfrak{m}B)$ is zero-dimensional,[3] hence Cohen-Macaulay, for every maximal ideal $\mathfrak{m}$ of A (9.35b), and that $\mathrm{ht}(\mathfrak{n}) = \mathrm{ht}(\mathfrak{n}^c)$ for every maximal ideal $\mathfrak{n}$ of B. If φ is flat, then B is Cohen-Macaulay by (9.35d). Conversely, if B is Cohen-Macaulay, then φ is flat by (9.34). □

EXAMPLE 9.37. Let A be a finite $k[X_1,\ldots,X_n]$-algebra (cf. 2.45). The map $k[X_1,\ldots,X_n] \to A$ is flat if and only if A is Cohen-Macaulay.

An algebraic variety V is said to be ***Cohen-Macaulay*** if $\mathcal{O}_{V,P}$ is Cohen-Macaulay for all $P \in V$. An affine algebraic variety V is Cohen-Macaulay if and only if $k[V]$ is Cohen-Macaulay (9.35a). A nonsingular variety is Cohen-Macaulay (9.35c).

THEOREM 9.38. *Let V and W be algebraic varieties with V nonsingular and W Cohen-Macaulay. A regular map $\varphi\colon W \to V$ is flat if and only if*

$$\dim \varphi^{-1}(P) = \dim W - \dim V \tag{39}$$

for all $P \in V$.

PROOF. Immediate consequence of (9.34). □

[3]Note that $C \overset{\text{def}}{=} B/\mathfrak{m}B = B \otimes_A A/\mathfrak{m}$ is a finite k-algebra. Therefore it has only finitely many maximal ideals. Every prime ideal in C is an intersection of maximal ideals (2.18), but a prime ideal can equal a finite intersection of ideals only if it equals one of the ideals.

ASIDE 9.39. The theorem fails with "nonsingular" weakened to "normal". Let $\mathbb{Z}/2\mathbb{Z}$ act on $W \stackrel{\text{def}}{=} \mathbb{A}^2$ by $(x,y) \mapsto (-x,-y)$. The quotient of W by this action is the quadric cone $V \subset \mathbb{A}^3$ defined by $TV = U^2$. The quotient map $\varphi\colon W \to V$ is $(x,y) \mapsto (t,u,v) = (x^2, xy, y^2)$. The variety W is nonsingular, and V is normal because $k[V] = k[X,Y]^G$ (cf. CA 23.12). Moreover φ is finite, and so its fibres have constant dimension 0, but it is not flat because

$$\sum_{Q \mapsto P} \dim_k \mathcal{O}_Q/\mathfrak{m}_P\mathcal{O}_Q = \begin{cases} 3 & \text{if } P = (0,0,0) \\ 2 & \text{otherwise} \end{cases}$$

(see 9.33). See mo117043.

d Lines on surfaces

As an application of some of the above results, we consider the problem of describing the set of lines on a surface of degree m in $\mathbb{P}^3$. To avoid possible problems, we assume for the rest of this chapter that k has characteristic zero.

We first need a way of describing lines in $\mathbb{P}^3$. Recall that we can associate with each projective variety $V \subset \mathbb{P}^n$ an affine cone over $\tilde{V}$ in k^{n+1}. This allows us to think of points in $\mathbb{P}^3$ as being one-dimensional subspaces in k^4, and lines in $\mathbb{P}^3$ as being two-dimensional subspaces in k^4. To such a subspace $W \subset k^4$, we can attach a one-dimensional subspace $\bigwedge^2 W$ in $\bigwedge^2 k^4 \approx k^6$, that is, to each line L in $\mathbb{P}^3$, we can attach point $p(L)$ in $\mathbb{P}^5$. Not every point in $\mathbb{P}^5$ should be of the form $p(L)$—heuristically, the lines in $\mathbb{P}^3$ should form a four-dimensional set. (Fix two planes in $\mathbb{P}^3$; giving a line in $\mathbb{P}^3$ corresponds to choosing a point on each of the planes.) We shall show that there is natural one-to-one correspondence between the set of lines in $\mathbb{P}^3$ and the set of points on a certain hyperspace $\Pi \subset \mathbb{P}^5$. Rather than using exterior algebras, I shall usually give the old-fashioned proofs.

Let L be a line in $\mathbb{P}^3$ and let $\mathbf{x} = (x_0 : x_1 : x_2 : x_3)$ and $\mathbf{y} = (y_0 : y_1 : y_2 : y_3)$ be distinct points on L. Then

$$p(L) = (p_{01} : p_{02} : p_{03} : p_{12} : p_{13} : p_{23}) \in \mathbb{P}^5, \quad p_{ij} \stackrel{\text{def}}{=} \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix},$$

depends only on L. The p_{ij} are called the Plücker coordinates of L, after Plücker (1801-1868).

In terms of exterior algebras, write e_0, e_1, e_2, e_3 for the canonical basis for k^4, so that $\mathbf{x}$, regarded as a point of k^4 is $\sum x_i e_i$, and $\mathbf{y} = \sum y_i e_i$; then $\bigwedge^2 k^4$ is a 6-dimensional vector space with basis $e_i \wedge e_j$, $0 \le i < j \le 3$, and $x \wedge y = \sum p_{ij} e_i \wedge e_j$ with p_{ij} given by the above formula.

We define p_{ij} for all i,j, $0 \le i,j \le 3$ by the same formula — thus $p_{ij} = -p_{ji}$.

LEMMA 9.40. *The line L can be recovered from $p(L)$ as follows:*

$$L = \{(\textstyle\sum_j a_j p_{0j} : \sum_j a_j p_{1j} : \sum_j a_j p_{2j} : \sum_j a_j p_{3j}) \mid (a_0 : a_1 : a_2 : a_3) \in \mathbb{P}^3\}.$$

PROOF. Let $\tilde{L}$ be the cone over L in k^4—it is a two-dimensional subspace of k^4—and let $\mathbf{x} = (x_0, x_1, x_2, x_3)$ and $\mathbf{y} = (y_0, y_1, y_2, y_3)$ be two linearly independent vectors in $\tilde{L}$. Then

$$\tilde{L} = \{f(\mathbf{y})\mathbf{x} - f(\mathbf{x})\mathbf{y} \mid f\colon k^4 \to k \text{ linear}\}.$$

Write $f = \sum a_j X_j$; then

$$f(\mathbf{y})\mathbf{x} - f(\mathbf{x})\mathbf{y} = (\textstyle\sum a_j p_{0j}, \sum a_j p_{1j}, \sum a_j p_{2j}, \sum a_j p_{3j}). \qquad \square$$

LEMMA 9.41. *The point $p(L)$ lies on the quadric $\Pi \subset \mathbb{P}^5$ defined by the equation*

$$X_{01}X_{23} - X_{02}X_{13} + X_{03}X_{12} = 0.$$

PROOF. This can be verified by direct calculation, or by using that

$$0 = \begin{vmatrix} x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \\ x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \end{vmatrix} = 2(p_{01}p_{23} - p_{02}p_{13} + p_{03}p_{12})$$

(expansion in terms of 2×2 minors). □

LEMMA 9.42. *Every point of Π is of the form $p(L)$ for a unique line L.*

PROOF. Assume $p_{03} \neq 0$; then the line through the points $(0 : p_{01} : p_{02} : p_{03})$ and $(p_{03} : p_{13} : p_{23} : 0)$ has Plücker coordinates

$$\begin{aligned}(-p_{01}p_{03} : -p_{02}p_{03} : -p_{03}^2 : \underbrace{p_{01}p_{23} - p_{02}p_{13}}_{-p_{03}p_{12}} : -p_{03}p_{13} : -p_{03}p_{23}) \\ = (p_{01} : p_{02} : p_{03} : p_{12} : p_{13} : p_{23}).\end{aligned}$$

A similar construction works when one of the other coordinates is nonzero, and this way we get inverse maps. □

Thus we have a canonical one-to-one correspondence

$$\{\text{lines in } \mathbb{P}^3\} \leftrightarrow \{\text{points on } \Pi\};$$

that is, we have identified the set of lines in $\mathbb{P}^3$ with the points of an algebraic variety. We may now use the methods of algebraic geometry to study the set. (This is a special case of the Grassmannians discussed in §6.)

We next consider the set of homogeneous polynomials of degree m in 4 variables,

$$F(X_0, X_1, X_2, X_3) = \sum_{i_0+i_1+i_2+i_3=m} a_{i_0i_1i_2i_3} X_0^{i_0} \dots X_3^{i_3}.$$

LEMMA 9.43. *The set of homogeneous polynomials of degree m in 4 variables is a vector space of dimension $\binom{3+m}{m}$*

PROOF. See the footnote p.139. □

Let $\nu = \binom{3+m}{m} - 1 = \frac{(m+1)(m+2)(m+3)}{6} - 1$, and regard $\mathbb{P}^\nu$ as the projective space attached to the vector space of homogeneous polynomials of degree m in 4 variables (p.143). Then we have a surjective map

$$\mathbb{P}^\nu \to \{\text{surfaces of degree } m \text{ in } \mathbb{P}^3\},$$

$$(\dots : a_{i_0i_1i_2i_3} : \dots) \mapsto V(F), \qquad F = \sum a_{i_0i_1i_2i_3} X_0^{i_0} X_1^{i_1} X_2^{i_2} X_3^{i_3}.$$

The map is not quite injective—for example, X^2Y and XY^2 define the same surface—but nevertheless, we can (somewhat loosely) think of the points of $\mathbb{P}^\nu$ as being (possibly degenerate) surfaces of degree m in $\mathbb{P}^3$.

Let $\Gamma_m \subset \Pi \times \mathbb{P}^\nu \subset \mathbb{P}^5 \times \mathbb{P}^\nu$ be the set of pairs (L, F) consisting of a line L in $\mathbb{P}^3$ lying on the surface $F(X_0, X_1, X_2, X_3) = 0$.

THEOREM 9.44. *The set Γ_m is a irreducible closed subset of $\Pi \times \mathbb{P}^\nu$; it is therefore a projective variety. The dimension of Γ_m is* $\frac{m(m+1)(m+5)}{6} + 3$.

EXAMPLE 9.45. For $m = 1$, Γ_m is the set of pairs consisting of a plane in $\mathbb{P}^3$ and a line on the plane. The theorem says that the dimension of Γ_1 is 5. Since there are ∞^3 planes in $\mathbb{P}^3$, and each has ∞^2 lines on it, this seems to be correct.

PROOF. We first show that Γ_m is closed. Let

$$p(L) = (p_{01} : p_{02} : \ldots) \qquad F = \sum a_{i_0 i_1 i_2 i_3} X_0^{i_0} \cdots X_3^{i_3}.$$

From (9.40) we see that L lies on the surface $F(X_0, X_1, X_2, X_3) = 0$ if and only if

$$F(\sum b_j p_{0j} : \sum b_j p_{1j} : \sum b_j p_{2j} : \sum b_j p_{3j}) = 0, \text{ all } (b_0, \ldots, b_3) \in k^4.$$

Expand this out as a polynomial in the b_j's with coefficients polynomials in the $a_{i_0 i_1 i_2 i_3}$ and p_{ij}'s. Then $F(...) = 0$ for all $\mathbf{b} \in k^4$ if and only if the coefficients of the polynomial are all zero. But each coefficient is of the form

$$P(\ldots, a_{i_0 i_1 i_2 i_3}, \ldots; p_{01}, p_{02} : \ldots)$$

with P homogeneous separately in the a's and p's, and so the set is closed in $\Pi \times \mathbb{P}^\nu$ (cf. the discussion in 6.51).

It remains to compute the dimension of Γ_m. We shall apply Proposition 9.11 to the projection map

$$\begin{array}{ccc} (L, F) & & \Gamma_m \subset \Pi \times \mathbb{P}^\nu \\ \downarrow & & \downarrow \varphi \\ L & & \Pi \end{array}$$

For $L \in \Pi$, $\varphi^{-1}(L)$ consists of the homogeneous polynomials of degree m such that $L \subset V(F)$ (taken up to nonzero scalars). After a change of coordinates, we can assume that L is the line

$$\begin{cases} X_0 = 0 \\ X_1 = 0, \end{cases}$$

i.e., $L = \{(0, 0, *, *)\}$. Then L lies on $F(X_0, X_1, X_2, X_3) = 0$ if and only if X_0 or X_1 occurs in each nonzero monomial term in F, i.e.,

$$F \in \varphi^{-1}(L) \iff a_{i_0 i_1 i_2 i_3} = 0 \text{ whenever } i_0 = 0 = i_1.$$

Thus $\varphi^{-1}(L)$ is a linear subspace of $\mathbb{P}^\nu$; in particular, it is irreducible. We now compute its dimension. Recall that F has $\nu + 1$ coefficients altogether; the number with $i_0 = 0 = i_1$ is $m + 1$, and so $\varphi^{-1}(L)$ has dimension

$$\frac{(m+1)(m+2)(m+3)}{6} - 1 - (m+1) = \frac{m(m+1)(m+5)}{6} - 1.$$

We can now deduce from (9.11) that Γ_m is irreducible and that

$$\dim(\Gamma_m) = \dim(\Pi) + \dim(\varphi^{-1}(L)) = \frac{m(m+1)(m+5)}{6} + 3,$$

as claimed. $\square$

Now consider the other projection By definition

$$\psi^{-1}(F) = \{L \mid L \text{ lies on } V(F)\}.$$

EXAMPLE 9.46. Let $m = 1$. Then $\nu = 3$ and $\dim \Gamma_1 = 5$. The projection $\psi\colon \Gamma_1 \to \mathbb{P}^3$ is surjective (every plane contains at least one line), and (9.9) tells us that $\dim \psi^{-1}(F) \geq 2$. In fact of course, the lines on any plane form a 2-dimensional family, and so $\psi^{-1}(F) = 2$ for all F.

THEOREM 9.47. *When $m > 3$, the surfaces of degree m containing no line correspond to an open subset of $\mathbb{P}^\nu$.*

PROOF. We have

$$\dim \Gamma_m - \dim \mathbb{P}^\nu = \frac{m(m+1)(m+5)}{6} + 3 - \frac{(m+1)(m+2)(m+3)}{6} + 1 = 4 - (m+1).$$

Therefore, if $m > 3$, then $\dim \Gamma_m < \dim \mathbb{P}^\nu$, and so $\psi(\Gamma_m)$ is a proper closed subvariety of $\mathbb{P}^\nu$. This proves the claim. □

We now look at the case $m = 2$. Here $\dim \Gamma_m = 10$, and $\nu = 9$, which suggests that ψ should be surjective and that its fibres should all have dimension ≥ 1. We shall see that this is correct.

A quadric is said to be ***nondegenerate*** if it is defined by an irreducible polynomial of degree 2. After a change of variables, any nondegenerate quadric will be defined by an equation

$$XW = YZ.$$

This is just the image of the Segre mapping (see 6.26)

$$(a_0 : a_1), (b_0 : b_1) \mapsto (a_0b_0 : a_0b_1 : a_1b_0 : a_1b_1) : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3.$$

There are two obvious families of lines on $\mathbb{P}^1 \times \mathbb{P}^1$, namely, the horizontal family and the vertical family; each is parametrized by $\mathbb{P}^1$, and so is called a ***pencil of lines***. They map to two families of lines on the quadric:

$$\left\{\begin{array}{l} t_0X = t_1Z \\ t_0Y = t_1W \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{l} t_0X = t_1Y \\ t_0Z = t_1W. \end{array}\right.$$

Since a degenerate quadric is a surface or a union of two surfaces, we see that every quadric surface contains a line, that is, that $\psi\colon \Gamma_2 \to \mathbb{P}^9$ is surjective. Thus (9.9) tells us that all the fibres have dimension ≥ 1, and the set where the dimension is > 1 is a proper closed subset. In fact the dimension of the fibre is > 1 exactly on the set of reducible F's, which we know to be closed (this was a homework problem in the original course).

It follows from the above discussion that if F is nondegenerate, then $\psi^{-1}(F)$ is isomorphic to the disjoint union of two lines, $\psi^{-1}(F) \approx \mathbb{P}^1 \cup \mathbb{P}^1$. Classically, one defines a ***regulus*** to be a nondegenerate quadric surface together with a choice of a pencil of lines. One can show that the set of reguli is, in a natural way, an algebraic variety R, and that, over the set of nondegenerate quadrics, ψ factors into the composite of two regular maps:

$$\begin{array}{ccl} \Gamma_2 - \psi^{-1}(S) & = & \text{pairs, } (F, L) \text{ with } L \text{ on } F; \\ \downarrow & & \\ R & = & \text{set of reguli;} \\ \downarrow & & \\ \mathbb{P}^9 - S & = & \text{set of nondegenerate quadrics.} \end{array}$$

The fibres of the top map are connected, and of dimension 1 (they are all isomorphic to $\mathbb{P}^1$), and the second map is finite and two-to-one. Factorizations of this type occur quite generally (see the Stein factorization theorem (9.38) below).

We now look at the case $m = 3$. Here $\dim \Gamma_3 = 19$; $\nu = 19$: we have a map

$$\psi \colon \Gamma_3 \to \mathbb{P}^{19}.$$

THEOREM 9.48. *The set of cubic surfaces containing exactly* 27 *lines corresponds to an open subset of* $\mathbb{P}^{19}$*; the remaining surfaces either contain an infinite number of lines or a nonzero finite number* ≤ 27.

EXAMPLE 9.49. (a) Consider the Fermat surface

$$X_0^3 + X_1^3 + X_2^3 + X_3^3 = 0.$$

Let ζ be a primitive cube root of one. There are the following lines on the surface, $0 \leq i, j \leq 2$:

$$\begin{cases} X_0 + \zeta^i X_1 = 0 \\ X_2 + \zeta^j X_3 = 0 \end{cases} \qquad \begin{cases} X_0 + \zeta^i X_2 = 0 \\ X_1 + \zeta^j X_3 = 0 \end{cases} \qquad \begin{cases} X_0 + \zeta^i X_3 = 0 \\ X_1 + \zeta^j X_2 = 0 \end{cases}$$

There are three sets, each with nine lines, for a total of 27 lines.

(b) Consider the surface

$$X_1 X_2 X_3 = X_0^3.$$

In this case, there are exactly three lines. To see this, look first in the affine space where $X_0 \neq 0$—here we can take the equation to be $X_1 X_2 X_3 = 1$. A line in $\mathbb{A}^3$ can be written in parametric form $X_i = a_i t + b_i$, but a direct inspection shows that no such line lies on the surface. Now look where $X_0 = 0$, that is, in the plane at infinity. The intersection of the surface with this plane is given by $X_1 X_2 X_3 = 0$ (homogeneous coordinates), which is the union of three lines, namely,

$$X_1 = 0;\ X_2 = 0;\ X_3 = 0.$$

Therefore, the surface contains exactly three lines.

(c) Consider the surface

$$X_1^3 + X_2^3 = 0.$$

Here there is a pencil of lines:

$$\begin{cases} t_0 X_1 = t_1 X_0 \\ t_0 X_2 = -t_1 X_0. \end{cases}$$

(In the affine space where $X_0 \neq 0$, the equation is $X^3 + Y^3 = 0$, which contains the line $X = t$, $Y = -t$, all t.)

We now discuss the proof of Theorem 9.48). If $\psi \colon \Gamma_3 \to \mathbb{P}^{19}$ were not surjective, then $\psi(\Gamma_3)$ would be a proper closed subvariety of $\mathbb{P}^{19}$, and the nonempty fibres would *all* have dimension ≥ 1 (by 9.9), which contradicts two of the above examples. Therefore the map is surjective[4], and there is an open subset U of $\mathbb{P}^{19}$ where the fibres have dimension 0; outside U, the fibres have dimension > 0.

Given that every cubic surface has at least one line, it is not hard to show that there is an open subset U' where the cubics have exactly 27 lines (see Reid, 1988, pp106–110); in fact, U' can be taken to be the set of nonsingular cubics. According to (8.26), the restriction of ψ to $\psi^{-1}(U)$ is finite, and so we can apply (8.40) to see that all cubics in $U - U'$ have fewer than 27 lines.

[4] According to Miles Reid (1988, p126) every adult algebraic geometer knows the proof that every cubic contains a line.

REMARK 9.50. The twenty-seven lines on a cubic surface were discovered in 1849 by Salmon and Cayley, and have been much studied—see A. Henderson, The Twenty-Seven Lines Upon the Cubic Surface, Cambridge University Press, 1911. For example, it is known that the group of permutations of the set of 27 lines preserving intersections (that is, such that $L \cap L' \neq \emptyset \iff \sigma(L) \cap \sigma(L') \neq \emptyset$) is isomorphic to the Weyl group of the root system of a simple Lie algebra of type E_6, and hence has 25920 elements.

It is known that there is a set of 6 skew lines on a nonsingular cubic surface V. Let L and L' be two skew lines. Then "in general" a line joining a point on L to a point on L' will meet the surface in exactly one further point. In this way one obtains an invertible regular map from an open subset of $\mathbb{P}^1 \times \mathbb{P}^1$ to an open subset of V, and hence V is birationally equivalent to $\mathbb{P}^2$.

e Bertini's theorem

Let $X \subset \mathbb{P}^n$ be a nonsingular projective variety. The hyperplanes H in $\mathbb{P}^n$ form a projective space $\mathbb{P}^{n\vee}$ (the "dual" projective space). The set of hyperplanes H not containing X and such that $X \cap H$ is nonsingular, form an open subset of $\mathbb{P}^{n\vee}$. If $\dim(X) \geq 2$, then the intersections $X \cap H$ are connected.

f Birational classification

Recall that two varieties V and W are birationally equivalent if $k(V) \approx k(W)$. This means that the varieties themselves become isomorphic once a proper closed subset has been removed from each (3.36).

The main problem of birational algebraic geometry is to classify algebraic varieties up to birational equivalence by finding a particularly good representative in each equivalence class.

For curves this is easy: in each birational equivalence class there is exactly one nonsingular projective curve (up to isomorphism). More precisely, the functor $V \rightsquigarrow k(V)$ is a contravariant equivalence from the category of nonsingular projective algebraic curves over k and dominant maps to the category of fields finitely generated and of transcendence degree 1 over k.

For surfaces, the problem is already much more difficult because many surfaces, even projective and nonsingular, will have the same function field. For example, every blow-up of a point on a surface produces a birationally equivalent surface.

A nonsingular projective surface is said to be ***minimal*** if it can not be obtained from another such surface by blowing up. The main theorem for surfaces (Enriques 1914, Kodaira 1966) says that a birational equivalence class contains either

(a) a unique minimal surface, or

(b) a surface of the form $C \times \mathbb{P}^1$ for a unique nonsingular projective curve C.

In higher dimensions, the problem becomes very involved, although much progress has been made — see Wikipedia MINIMAL MODEL PROGRAM.

Exercises

9-1. Let G be a connected group variety, and consider an action of G on a variety V, i.e., a regular map $G \times V \to V$ such that $(gg')v = g(g'v)$ for all $g, g' \in G$ and $v \in V$. Show that each orbit $O = Gv$ of G is open in its closure $\bar{O}$, and that $\bar{O} \smallsetminus O$ is a union of orbits of strictly lower dimension. Deduce that each orbit is a nonsingular subvariety of V, and that there exists at least one closed orbit.

9-2. Let $G = \mathrm{GL}_2 = V$, and let G act on V by conjugation. According to the theory of Jordan canonical forms, the orbits are of three types:

(a) Characteristic polynomial $X^2 + aX + b$; distinct roots.

(b) Characteristic polynomial $X^2 + aX + b$; minimal polynomial the same; repeated roots.

(c) Characteristic polynomial $X^2 + aX + b = (X - \alpha)^2$; minimal polynomial $X - \alpha$.

For each type, find the dimension of the orbit, the equations defining it (as a subvariety of V), the closure of the orbit, and which other orbits are contained in the closure.

(You may assume, if you wish, that the characteristic is zero. Also, you may assume the following (fairly difficult) result: for any closed subgroup H of an group variety G, G/H has a natural structure of an algebraic variety with the following properties: $G \to G/H$ is regular, and a map $G/H \to V$ is regular if the composite $G \to G/H \to V$ is regular; $\dim G/H = \dim G - \dim H$.)

[The enthusiasts may wish to carry out the analysis for GL_n.]

9-3. Find $3d^2$ lines on the Fermat projective surface

$$X_0^d + X_1^d + X_2^d + X_3^d = 0, \quad d \geq 3, \quad (p, d) = 1, \quad p \text{ the characteristic.}$$

9-4. (a) Let $\varphi: W \to V$ be a quasi-finite dominant regular map of irreducible varieties. Show that there are open subsets U' and U of W and V such that $\varphi(U') \subset U$ and $\varphi: U' \to U$ is finite.

(b) Let G be a group variety acting transitively on irreducible varieties W and V, and let $\varphi: W \to V$ be G-equivariant regular map satisfying the hypotheses in (a). Then φ is finite, and hence proper.

Solutions to the exercises

1-1 Use induction on n. For $n = 1$, use that a nonzero polynomial in one variable has only finitely many roots (which follows from unique factorization, for example). Now suppose $n > 1$ and write $f = \sum g_i X_n^i$ with each $g_i \in k[X_1,\ldots,X_{n-1}]$. If f is not the zero polynomial, then some g_i is not the zero polynomial. Therefore, by induction, there exist $(a_1,\ldots,a_{n-1}) \in k^{n-1}$ such that $f(a_1,\ldots,a_{n-1},X_n)$ is not the zero polynomial. Now, by the degree-one case, there exists a b such that $f(a_1,\ldots,a_{n-1},b) \neq 0$.

1-2 $(X+2Y,Z)$; Gaussian elimination (to reduce the matrix of coefficients to row echelon form); (1), unless the characteristic of k is 2, in which case the ideal is $(X+1,Z+1)$.

2-1 $W = Y$-axis, and so $I(W) = (X)$. Clearly,

$$(X^2,XY^2) \subset (X) \subset \operatorname{rad}(X^2,XY^2)$$

and $\operatorname{rad}((X)) = (X)$. On taking radicals, we find that $(X) = \operatorname{rad}(X^2,XY^2)$.

2-2 The $d \times d$ minors of a matrix are polynomials in the entries of the matrix, and the set of matrices with rank $\leq r$ is the set where all $(r+1)\times(r+1)$ minors are zero.

2-3 Clearly $V = V(X_n - X_1^n,\ldots,X_2 - X_1^2)$. The map

$$X_i \mapsto T^i : k[X_1,\ldots,X_n] \to k[T]$$

induces an isomorphism $k[V] \to \mathbb{A}^1$. [Hence $t \mapsto (t,\ldots,t^n)$ is an isomorphism of affine varieties $\mathbb{A}^1 \to V$.]

2-4 We use that the prime ideals are in one-to-one correspondence with the irreducible closed subsets Z of $\mathbb{A}^2$. For such a set, $0 \leq \dim Z \leq 2$.

Case $\dim Z = 2$. Then $Z = \mathbb{A}^2$, and the corresponding ideal is (0).

Case $\dim Z = 1$. Then $Z \neq \mathbb{A}^2$, and so $I(Z)$ contains a nonzero polynomial $f(X,Y)$. If $I(Z) \neq (f)$, then $\dim Z = 0$ by (2.64, 2.63). Hence $I(Z) = (f)$.

Case $\dim Z = 0$. Then Z is a point (a,b) (see 2.62), and so $I(Z) = (X-a,Y-b)$.

2-6 The statement $\mathrm{Hom}_{k\text{-algebras}}(A \otimes_{\mathbb{Q}} k, B \otimes_{\mathbb{Q}} k) \neq \emptyset$ can be interpreted as saying that a certain set of polynomials has a zero in k. If the polynomials have a common zero in $\mathbb{C}$, then the ideal they generate in $\mathbb{C}[X_1,\ldots]$ does not contain 1. *A fortiori* the ideal they generate in $\mathbb{Q}[X_1,\ldots]$ does not contain 1, and so the Nullstellensatz (2.11) implies that the polynomials have a common zero in k.

2-7 Regard $\mathrm{Hom}_A(M,N)$ as an affine space over k; the elements not isomorphisms are the zeros of a polynomial; because M and N become isomorphic over k^{al}, the polynomial is not identically zero; therefore it has a nonzero in k (Exercise 1-1).

3-1 A map $\alpha : \mathbb{A}^1 \to \mathbb{A}^1$ is continuous for the Zariski topology if the inverse images of finite sets are finite, whereas it is regular only if it is given by a polynomial $P \in k[T]$, so it is easy to give examples, e.g., any map α such that $\alpha^{-1}(\text{point})$ is finite but arbitrarily large.

3-3 The image omits the points on the Y-axis except for the origin. The complement of the image is not dense, and so it is not open, but any polynomial zero on it is also zero at $(0,0)$, and so it not closed.

3-4 Let i be an element of k with square -1. The map $(x,y) \mapsto (x+iy, x-iy)$ from the circle to the hyperbola has inverse $(x,y) \mapsto ((x+y)/2, (x-y)/2i)$. The k-algebra $k[X,Y]/(XY-1) \simeq k[X,X^{-1}]$, which is not isomorphic to $k[X]$ (too many units).

3-5 No, because both $+1$ and -1 map to $(0,0)$. The map on rings is

$$k[x,y] \to k[T], \quad x \mapsto T^2-1, \quad y \mapsto T(T^2-1),$$

which is not surjective (T is not in the image).

5-1 Let f be regular on $\mathbb{P}^1$. Then $f|U_0 = P(X) \in k[X]$, where X is the regular function $(a_0{:}a_1) \mapsto a_1/a_0 {:} U_0 \to k$, and $f|U_1 = Q(Y) \in k[Y]$, where Y is $(a_0{:}a_1) \mapsto a_0/a_1$. On $U_0 \cap U_1$, X and Y are reciprocal functions. Thus $P(X)$ and $Q(1/X)$ define the same function on $U_0 \cap U_1 = \mathbb{A}^1 \smallsetminus \{0\}$. This implies that they are equal in $k(X)$, and must both be constant.

5-2 Note that $\Gamma(V, \mathcal{O}_V) = \prod \Gamma(V_i, \mathcal{O}_{V_i})$ — to give a regular function on $\bigsqcup V_i$ is the same as to give a regular function on each V_i (this is the "obvious" ringed space structure). Thus, if V is affine, it must equal $\operatorname{Specm}(\prod A_i)$, where $A_i = \Gamma(V_i, \mathcal{O}_{V_i})$, and so $V = \bigsqcup \operatorname{Specm}(A_i)$ (use the description of the ideals in $A \times B$ on in Section 1a). Etc..

5-5 Let H be an algebraic subgroup of G. By definition, H is locally closed, i.e., open in its Zariski closure $\bar{H}$. Assume first that H is connected. Then $\bar{H}$ is a connected algebraic group, and it is a disjoint union of the cosets of H. It follows that $H = \bar{H}$. In the general case, H is a finite disjoint union of its connected components; as one component is closed, they all are.

4-1 (b) The singular points are the common solutions to

$$\begin{cases} 4X^3 - 2XY^2 = 0 & \implies \quad X = 0 \text{ or } Y^2 = 2X^2 \\ 4Y^3 - 2X^2Y = 0 & \implies \quad Y = 0 \text{ or } X^2 = 2Y^2 \\ X^4 + Y^4 - X^2Y^2 = 0. \end{cases}$$

Thus, only $(0,0)$ is singular, and the variety is its own tangent cone.

4-2 Directly from the definition of the tangent space, we have that

$$T_{\mathbf{a}}(V \cap H) \subset T_{\mathbf{a}}(V) \cap T_{\mathbf{a}}(H).$$

As

$$\dim T_{\mathbf{a}}(V \cap H) \geq \dim V \cap H = \dim V - 1 = \dim T_{\mathbf{a}}(V) \cap T_{\mathbf{a}}(H),$$

we must have equalities everywhere, which proves that $\mathbf{a}$ is nonsingular on $V \cap H$. (In particular, it can't lie on more than one irreducible component.)

The surface $Y^2 = X^2 + Z$ is smooth, but its intersection with the X-Y plane is singular.

No, P needn't be singular on $V \cap H$ if $H \supset T_P(V)$ — for example, we could have $H \supset V$ or H could be the tangent line to a curve.

4-4 We can assume V and W to affine, say

$$I(V) = \mathfrak{a} \subset k[X_1, \ldots, X_m]$$
$$I(W) = \mathfrak{b} \subset k[X_{m+1}, \ldots, X_{m+n}].$$

If $\mathfrak{a} = (f_1,\ldots,f_r)$ and $\mathfrak{b} = (g_1,\ldots,g_s)$, then $I(V\times W) = (f_1,\ldots,f_r,g_1,\ldots,g_s)$. Thus, $T_{(\mathbf{a},\mathbf{b})}(V\times W)$ is defined by the equations

$$(df_1)_{\mathbf{a}} = 0,\ldots,(df_r)_{\mathbf{a}} = 0,(dg_1)_{\mathbf{b}} = 0,\ldots,(dg_s)_{\mathbf{b}} = 0,$$

which can obviously be identified with $T_{\mathbf{a}}(V)\times T_{\mathbf{b}}(W)$.

4-5 Take C to be the union of the coordinate axes in $\mathbb{A}^n$. (Of course, if you want C to be irreducible, then this is more difficult...)

4-6 A matrix A satisfies the equations

$$(I+\varepsilon A)^{\mathrm{tr}}\cdot J\cdot(I+\varepsilon A) = I$$

if and only if

$$A^{\mathrm{tr}}\cdot J + J\cdot A = 0.$$

Such an A is of the form $\begin{pmatrix} M & N \\ P & Q\end{pmatrix}$ with M,N,P,Q $n\times n$-matrices satisfying

$$N^{\mathrm{tr}} = N,\quad P^{\mathrm{tr}} = P,\quad M^{\mathrm{tr}} = -Q.$$

The dimension of the space of A's is therefore

$$\frac{n(n+1)}{2}\text{ (for }N) + \frac{n(n+1)}{2}\text{ (for }P) + n^2\text{ (for }M,Q) = 2n^2+n.$$

4-7 Let C be the curve $Y^2 = X^3$, and consider the map $\mathbb{A}^1\to C$, $t\mapsto(t^2,t^3)$. The corresponding map on rings $k[X,Y]/(Y^2)\to k[T]$ is not an isomorphism, but the map on the geometric tangent cones is an isomorphism.

4-8 The singular locus V_{sing} has codimension ≥ 2 in V, and this implies that V is normal. [Idea of the proof: let $f\in k(V)$ be integral over $k[V]$, $f\notin k[V]$, $f = g/h$, $g,h\in k[V]$; for any $P\in V(h)\smallsetminus V(g)$, $\mathcal{O}_P$ is not integrally closed, and so P is singular.]

4-9 No! Let $\mathfrak{a} = (X^2Y)$. Then $V(\mathfrak{a})$ is the union of the X and Y axes, and $IV(\mathfrak{a}) = (XY)$. For $\mathbf{a} = (a,b)$,

$$\begin{aligned}(dX^2Y)_{\mathbf{a}} &= 2ab(X-a) + a^2(Y-b)\\ (dXY)_{\mathbf{a}} &= b(X-a) + a(Y-b).\end{aligned}$$

If $a\neq 0$ and $b = 0$, then the equations

$$\begin{aligned}(dX^2Y)_{\mathbf{a}} &= a^2Y = 0\\ (dXY)_{\mathbf{a}} &= aY = 0\end{aligned}$$

have the same solutions.

6-1 Let $P = (a:b:c)$, and assume $c\neq 0$. Then the tangent line at $P = (\frac{a}{c}:\frac{b}{c}:1)$ is

$$\left(\frac{\partial F}{\partial X}\right)_P X + \left(\frac{\partial F}{\partial Y}\right)_P Y - \left(\left(\frac{\partial F}{\partial X}\right)_P\left(\frac{a}{c}\right) + \left(\frac{\partial F}{\partial Y}\right)_P\left(\frac{b}{c}\right)\right)Z = 0.$$

Now use that, because F is homogeneous,

$$F(a,b,c) = 0 \implies \left(\frac{\partial F}{\partial X}\right)_P a + \left(\frac{\partial F}{\partial Y}\right)_P + \left(\frac{\partial F}{\partial Z}\right)_P c = 0.$$

(This just says that the tangent plane at (a,b,c) to the affine cone $F(X,Y,Z)=0$ passes through the origin.) The point at ∞ is $(0:1:0)$, and the tangent line is $Z=0$, the line at ∞. [The line at ∞ meets the cubic curve at only one point instead of the expected 3, and so the line at ∞ "touches" the curve, and the point at ∞ is a point of inflexion.]

6-2 The equation defining the conic must be irreducible (otherwise the conic is singular). After a linear change of variables, the equation will be of the form $X^2+Y^2=Z^2$ (this is proved in calculus courses). The equation of the line in $aX+bY=cZ$, and the rest is easy. [Note that this is a special case of Bezout's theorem (6.37) because the multiplicity is 2 in case (b).]

6-3 (a) The ring

$$k[X,Y,Z]/(Y-X^2,Z-X^3)=k[x,y,z]=k[x]\simeq k[X],$$

which is an integral domain. Therefore, $(Y-X^2,Z-X^3)$ is a radical ideal.

(b) The polynomial $F=Z-XY=(Z-X^3)-X(Y-X^2)\in I(V)$ and $F^*=ZW-XY$. If

$$ZW-XY=(YW-X^2)f+(ZW^2-X^3)g,$$

then, on equating terms of degree 2, we would find

$$ZW-XY=a(YW-X^2),$$

which is false.

6-4 Let $P=(a_0:\ldots:a_n)$ and $Q=(b_0:\ldots:b_n)$ be two points of $\mathbb{P}^n$, $n\geq 2$. The condition that the hyperplane $L_{\mathbf{c}}$: $\sum c_iX_i=0$ pass through P and not through Q is that

$$\sum a_ic_i=0,\quad \sum b_ic_i\neq 0.$$

The $(n+1)$-tuples $(c_0,\ldots,c_n)$ satisfying these conditions form a nonempty open subset of the hyperplane $H\colon\sum a_iX_i=0$ in $\mathbb{A}^{n+1}$. On applying this remark to the pairs (P_0,P_i), we find that the $(n+1)$-tuples $\mathbf{c}=(c_0,\ldots,c_n)$ such that P_0 lies on the hyperplane $L_{\mathbf{c}}$ but not $P_1,\ldots,P_r$ form a nonempty open subset of H.

6-5 The subset

$$C=\{(a:b:c)\mid a\neq 0,\quad b\neq 0\}\cup\{(1:0:0)\}$$

of $\mathbb{P}^2$ is not locally closed. Let $P=(1:0:0)$. If the set C were locally closed, then P would have an open neighbourhood U in $\mathbb{P}^2$ such that $U\cap C$ is closed. When we look in U_0, P becomes the origin, and

$$C\cap U_0=(\mathbb{A}^2\smallsetminus\{X\text{-axis}\})\cup\{\text{origin}\}.$$

The open neighbourhoods U of P are obtained by removing from $\mathbb{A}^2$ a finite number of curves not passing through P. It is not possible to do this in such a way that $U\cap C$ is closed in U ($U\cap C$ has dimension 2, and so it can't be a proper closed subset of U; we can't have $U\cap C=U$ because any curve containing all nonzero points on X-axis also contains the origin).

6-6 Let $\sum c_{ij}X_{ij}=0$ be a hyperplane containing the image of the Segre map. We then have

$$\sum c_{ij}a_ib_j=0$$

for all $\mathbf{a} = (a_0, \ldots, a_m) \in k^{m+1}$ and $\mathbf{b} = (b_0, \ldots, b_n) \in k^{n+1}$. In other words,

$$\mathbf{a}C\mathbf{b}^t = 0$$

for all $\mathbf{a} \in k^{m+1}$ and $\mathbf{b} \in k^{n+1}$, where C is the matrix (c_{ij}). This equation shows that $\mathbf{a}C = 0$ for all $\mathbf{a}$, and this implies that $C = 0$.

7-2 Define $f(v) = h(v, Q)$ and $g(w) = h(P, w)$, and let $\varphi = h - (f \circ p + g \circ q)$. Then $\varphi(v, Q) = 0 = \varphi(P, w)$, and so the rigidity theorem (7.35) implies that φ is identically zero.

8-2 For example, consider

$$(\mathbb{A}^1 \smallsetminus \{1\}) \to \mathbb{A}^1 \overset{x \mapsto x^n}{\to} \mathbb{A}^1$$

for $n > 1$ an integer prime to the characteristic. The map is obviously quasi-finite, but it is not finite because it corresponds to the map of k-algebras

$$X \mapsto X^n : k[X] \to k[X, (X-1)^{-1}]$$

which is not finite (the elements $1/(X-1)^i$, $i \geq 1$, are linearly independent over $k[X]$, and so also over $k[X^n]$).

8-3 Assume that V is separated, and consider two regular maps $f, g: Z \rightrightarrows W$. We have to show that the set on which f and g agree is closed in Z. The set where $\varphi \circ f$ and $\varphi \circ g$ agree is closed in Z, and it contains the set where f and g agree. Replace Z with the set where $\varphi \circ f$ and $\varphi \circ g$ agree. Let U be an open affine subset of V, and let $Z' = (\varphi \circ f)^{-1}(U) = (\varphi \circ g)^{-1}(U)$. Then $f(Z')$ and $g(Z')$ are contained in $\varphi^{-1}(U)$, which is an open affine subset of W, and is therefore separated. Hence, the subset of Z' on which f and g agree is closed. This proves the result.

[Note that the problem implies the following statement: if $\varphi: W \to V$ is a finite regular map and V is separated, then W is separated.]

8-4 Let $V = \mathbb{A}^n$, and let W be the subvariety of $\mathbb{A}^n \times \mathbb{A}^1$ defined by the polynomial

$$\textstyle\prod_{i=1}^n (X - T_i) = 0.$$

The fibre over $(t_1, \ldots, t_n) \in \mathbb{A}^n$ is the set of roots of $\prod(X - t_i)$. Thus, $V_n = \mathbb{A}^n$; V_{n-1} is the union of the linear subspaces defined by the equations

$$T_i = T_j, \quad 1 \leq i, j \leq n, \quad i \neq j;$$

V_{n-2} is the union of the linear subspaces defined by the equations

$$T_i = T_j = T_k, \quad 1 \leq i, j, k \leq n, \quad i, j, k \text{ distinct},$$

and so on.

9-1 Consider an orbit $O = Gv$. The map $g \mapsto gv: G \to O$ is regular, and so O contains an open subset U of $\bar{O}$ (9.7). If $u \in U$, then $gu \in gU$, and gU is also a subset of O which is open in $\bar{O}$ (because $P \mapsto gP: V \to V$ is an isomorphism). Thus O, regarded as a topological subspace of $\bar{O}$, contains an open neighbourhood of each of its points, and so must be open in $\bar{O}$.

We have shown that O is locally closed in V, and so has the structure of a subvariety. From (4.37), we know that it contains at least one nonsingular point P. But then gP is nonsingular, and every point of O is of this form.

From set theory, it is clear that $\bar{O} \smallsetminus O$ is a union of orbits. Since $\bar{O} \smallsetminus O$ is a proper closed subset of $\bar{O}$, all of its subvarieties must have dimension $< \dim \bar{O} = \dim O$.

Let O be an orbit of lowest dimension. The last statement implies that $O = \bar{O}$.

9-2 An orbit of type (a) is closed, because it is defined by the equations

$$\mathrm{Tr}(A) = -a, \quad \det(A) = b,$$

(as a subvariety of V). It is of dimension 2, because the centralizer of $\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$, $\alpha \neq \beta$, is $\left\{ \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} \right\}$, which has dimension 2.

An orbit of type (b) is of dimension 2, but is not closed: it is defined by the equations

$$\mathrm{Tr}(A) = -a, \quad \det(A) = b, \quad A \neq \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}, \quad \alpha = \text{root of } X^2 + aX + b.$$

An orbit of type (c) is closed of dimension 0: it is defined by the equation $A = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$. An orbit of type (b) contains an orbit of type (c) in its closure.

9-3 Let ζ be a primitive dth root of 1. Then, for each i, j, $1 \leq i, j \leq d$, the following equations define lines on the surface

$$\begin{cases} X_0 + \zeta^i X_1 = 0 \\ X_2 + \zeta^j X_3 = 0 \end{cases} \quad \begin{cases} X_0 + \zeta^i X_2 = 0 \\ X_1 + \zeta^j X_3 = 0 \end{cases} \quad \begin{cases} X_0 + \zeta^i X_3 = 0 \\ X_1 + \zeta^j X_2 = 0. \end{cases}$$

There are three sets of lines, each with d^2 lines, for a total of $3d^2$ lines.

9-4 (a) Compare the proof of Theorem 9.9.

(b) Use the transitivity, and apply Proposition 8.26.

Index

algebra
  affine, 63
  finite, 11
  finitely generated, 11
algebraically dependent, 33
algebraically independent, 33
$\mathbb{A}^n$, 35
analytic space, 167
axiom
  separation, 99

base change, 110
basis
  transcendence, 33
birationally equivalent, 71, 114
boundary, 53

codimension, 54
  pure, 54
complete intersection
  ideal-theoretic, 77
  local, 77
  set-theoretic, 76
component
  of a function, 48
cone, 129
  affine over a set, 130
content of a polynomial, 22
convergent, 58
Cramer's rule, 24
curve, 53
  elliptic, 36, 128, 133

degree
  of a hypersurface, 148
  of a map, 182
  of a projective variety, 150
derivation, 89
desingularization, 190
differential, 85
dimension, 72, 112
  of a topological space, 52
  of an algebraic set, 53
  pure, 53, 112
direct limit, 20
direct system, 20
directed set, 20
discrete valuation ring, 84
divisor, 175
  effective, 175
  locally principal, 174
  positive, 175
  prime, 174
  principal, 176
  support of, 175
domain
  factorial, 21
  normal, 26
  unique factorization, 21

element
  integral over a ring, 24
  irreducible, 21
  prime, 21

$F(A)$, 16
faithfully flat, 199
fibre, 110
field of rational functions, 48, 111
flat, 199, 201
form
  leading, 81
function
  holomorphic, 167
  rational, 61
  regular, 47, 59, 98

generate, 11
germ
  of a function, 58

graph
of a regular map, 108
group
symplectic, 96
group variety, 107

homogeneous, 135
homomorphism
finite, 11
of algebras, 11
hypersurface, 48, 140
hypersurface section, 140

ideal, 11
generated by a subset, 12
graded, 129, 131
homogeneous, 129
maximal, 12
prime, 12
radical, 40
immersion, 102
closed, 70, 101
open, 101
integral closure, 25
integral domain, 11
integrally closed, 26
irreducible components, 45
isolated in its fibre, 185
isomorphic
locally, 95

$\kappa(\mathfrak{p})$, 185

lemma
Gauss's, 22
Nakayama's, 14
prime avoidance, 75
Zariski's, 39
linearly equivalent, 176
local equation, 174
local ring
regular, 14
local system of parameters, 118

manifold
complex, 97
differentiable, 97
topological, 97
map
affine, 191
bilinear, 30
birational, 114, 187
dominant, 49, 70, 113
étale, 115, 117
finite, 49, 73, 176, 180
Frobenius, 67
proper, 157
quasi-finite, 49, 180
rational, 113
regular, 48
Segre, 142
separable, 121, 182
Veronese, 139
minimal surface, 211
morphism
of affine algebraic varieties, 63
of ringed spaces, 62
$\mathfrak{m}_P$, 40
multiplicity, 204
of a point, 82

n-fold, 53
neighbourhood
étale, 118
nilpotent, 40
node, 82
nondegenerate quadric, 209
normalization, 173, 174

open affine, 69
open subset
basic, 48
principal, 48

pencil of lines, 209
Picard group, 176
point
factorial, 174
multiple, 85
nonsingular, 80, 85
normal, 171
ordinary multiple, 82
singular, 85
smooth, 80, 85
with coordinates in a ring, 121
polynomial
Hilbert, 149
homogeneous, 127
primitive, 22
prevariety

algebraic, 97
separated, 99
product
fibred, 110
of algebraic varieties, 106
of objects, 103
tensor, 31
projection with centre, 143

radical
of an ideal, 40
rational map, 113
real locus, 36
regular map, 98
of affine algebraic varieties, 63
of algebraic sets, 48
regulus, 209
resolution of singularities, 190
resultant, 160
ring
associated graded, 91
coordinate, 47
graded, 131
local, 14
noetherian, 14
normal, 34
of dual numbers, 87
reduced, 40
ringed space, 58

section of a sheaf, 58
semisimple
group, 95
Lie algebra, 95
separable degree, 184
set
(projective) algebraic, 128
constructible, 195
sheaf
of algebras, 57
singular locus, 81
Spm(A), 64
spm(A), 64
stalk, 58
subring, 11
subset
algebraic, 35
analytic, 167
multiplicative, 15
subspace
locally closed, 102
subvariety, 102
closed, 68
open affine, 97
surface, 53

T_1 space, 44
tangent cone, 81, 91
geometric, 81, 91
tangent space, 80, 85
tensor product
of modules, 30
theorem
Bezout's , 148
Chinese Remainder, 13
going-up, 29
Hilbert basis, 37
Hilbert Nullstellensatz, 38
Noether normalization, 51
Stein factorization, 189
strong Hilbert Nullstellensatz, 41
Zariski's main, 184
topological space
irreducible , 44
noetherian, 44
quasicompact, 44
topology
étale, 119
Zariski, 38, 130

variety
abelian, 164
affine algebraic, 63
algebraic, 99
Cohen-Macaulay, 205
complete, 155
factorial, 174
flag, 148
Grassmann, 145
group, 107
monoid, 125
normal, 171
projective, 127
quasi-affine, 100
quasi-projective, 127
rational, 124
unirational, 124

zero set, 35

MATH 216: FOUNDATIONS OF ALGEBRAIC GEOMETRY

math216.wordpress.com

June 11, 2013 draft

Note to reader: the figures, index, and formatting have yet to be properly dealt with. There remain a few other issues still to be dealt with in the main part of the notes.

Contents

Preface 11
0.1. For the reader 12
0.2. For the expert 15
0.3. Background and conventions 17
0.4. $\star\star$ The goals of this book 18

Part I. Preliminaries 21

Chapter 1. Some category theory 23
1.1. Motivation 23
1.2. Categories and functors 25
1.3. Universal properties determine an object up to unique isomorphism 31
1.4. Limits and colimits 39
1.5. Adjoints 43
1.6. An introduction to abelian categories 46
1.7. $\star$ Spectral sequences 56

Chapter 2. Sheaves 69
2.1. Motivating example: The sheaf of differentiable functions. 69
2.2. Definition of sheaf and presheaf 71
2.3. Morphisms of presheaves and sheaves 76
2.4. Properties determined at the level of stalks, and sheafification 80
2.5. Sheaves of abelian groups, and $\mathscr{O}_X$-modules, form abelian categories 84
2.6. The inverse image sheaf 87
2.7. Recovering sheaves from a "sheaf on a base" 90

Part II. Schemes 95

Chapter 3. Toward affine schemes: the underlying set, and topological space 97
3.1. Toward schemes 97
3.2. The underlying set of affine schemes 99
3.3. Visualizing schemes I: generic points 111
3.4. The underlying topological space of an affine scheme 112
3.5. A base of the Zariski topology on Spec A: Distinguished open sets 115
3.6. Topological (and Noetherian) properties 116
3.7. The function $I(\cdot)$, taking subsets of Spec A to ideals of A 124

Chapter 4. The structure sheaf, and the definition of schemes in general 127
4.1. The structure sheaf of an affine scheme 127
4.2. Visualizing schemes II: nilpotents 130

4.3. Definition of schemes 133
4.4. Three examples 137
4.5. Projective schemes, and the Proj construction 143

Chapter 5. Some properties of schemes 151
5.1. Topological properties 151
5.2. Reducedness and integrality 153
5.3. Properties of schemes that can be checked "affine-locally" 155
5.4. Normality and factoriality 159
5.5. Where functions are supported: Associated points of schemes 164

Part III. Morphisms 173

Chapter 6. Morphisms of schemes 175
6.1. Introduction 175
6.2. Morphisms of ringed spaces 176
6.3. From locally ringed spaces to morphisms of schemes 178
6.4. Maps of graded rings and maps of projective schemes 184
6.5. Rational maps from reduced schemes 186
6.6. ⋆ Representable functors and group schemes 192
6.7. ⋆⋆ The Grassmannian (initial construction) 197

Chapter 7. Useful classes of morphisms of schemes 199
7.1. An example of a reasonable class of morphisms: Open embeddings 199
7.2. Algebraic interlude: Lying Over and Nakayama 200
7.3. A gazillion finiteness conditions on morphisms 205
7.4. Images of morphisms: Chevalley's theorem and elimination theory 214

Chapter 8. Closed embeddings and related notions 221
8.1. Closed embeddings and closed subschemes 221
8.2. More projective geometry 226
8.3. Smallest closed subschemes such that ... 232
8.4. Effective Cartier divisors, regular sequences and regular embeddings 236

Chapter 9. Fibered products of schemes, and base change 241
9.1. They exist 241
9.2. Computing fibered products in practice 247
9.3. Interpretations: Pulling back families, and fibers of morphisms 250
9.4. Properties preserved by base change 255
9.5. ⋆ Properties not preserved by base change, and how to fix them 257
9.6. Products of projective schemes: The Segre embedding 264
9.7. Normalization 267

Chapter 10. Separated and proper morphisms, and (finally!) varieties 273
10.1. Separated morphisms (and quasiseparatedness done properly) 273
10.2. Rational maps to separated schemes 283
10.3. Proper morphisms 287

Part IV. "Geometric" properties: Dimension and smoothness 293

Chapter 11. Dimension 295

11.1. Dimension and codimension 295
11.2. Dimension, transcendence degree, and Noether normalization 299
11.3. Codimension one miracles: Krull's and Hartogs's Theorems 307
11.4. Dimensions of fibers of morphisms of varieties 313
11.5. ⋆⋆ Proof of Krull's Principal Ideal and Height Theorems 318

Chapter 12. Regularity and smoothness 321
12.1. The Zariski tangent space 321
12.2. Regularity, and smoothness over a field 326
12.3. Examples 332
12.4. Bertini's Theorem 335
12.5. Discrete valuation rings: Dimension 1 Noetherian regular local rings 338
12.6. Smooth (and étale) morphisms (first definition) 343
12.7. ⋆ Valuative criteria for separatedness and properness 347
12.8. ⋆ More sophisticated facts about regular local rings 351
12.9. ⋆ Filtered rings and modules, and the Artin-Rees Lemma 352

Part V. Quasicoherent sheaves 355

Chapter 13. Quasicoherent and coherent sheaves 357
13.1. Vector bundles and locally free sheaves 357
13.2. Quasicoherent sheaves 363
13.3. Characterizing quasicoherence using the distinguished affine base 365
13.4. Quasicoherent sheaves form an abelian category 369
13.5. Module-like constructions 371
13.6. Finite type and coherent sheaves 375
13.7. Pleasant properties of finite type and coherent sheaves 377
13.8. ⋆⋆ Coherent modules over non-Noetherian rings 381

Chapter 14. Line bundles: Invertible sheaves and divisors 385
14.1. Some line bundles on projective space 385
14.2. Line bundles and Weil divisors 387
14.3. ⋆ Effective Cartier divisors "=" invertible ideal sheaves 396

Chapter 15. Quasicoherent sheaves on projective A-schemes 399
15.1. The quasicoherent sheaf corresponding to a graded module 399
15.2. Invertible sheaves (line bundles) on projective A-schemes 400
15.3. Globally generated and base-point-free line bundles 401
15.4. Quasicoherent sheaves and graded modules 404

Chapter 16. Pushforwards and pullbacks of quasicoherent sheaves 409
16.1. Introduction 409
16.2. Pushforwards of quasicoherent sheaves 409
16.3. Pullbacks of quasicoherent sheaves 410
16.4. Line bundles and maps to projective schemes 416
16.5. The Curve-to-Projective Extension Theorem 423
16.6. Ample and very ample line bundles 424
16.7. ⋆ The Grassmannian as a moduli space 429

Chapter 17. Relative versions of Spec and Proj, and projective morphisms 435
17.1. Relative Spec of a (quasicoherent) sheaf of algebras 435
17.2. Relative Proj of a sheaf of graded algebras 438
17.3. Projective morphisms 441
17.4. Applications to curves 447

Chapter 18. Čech cohomology of quasicoherent sheaves 453
18.1. (Desired) properties of cohomology 453
18.2. Definitions and proofs of key properties 458
18.3. Cohomology of line bundles on projective space 463
18.4. Riemann-Roch, degrees of coherent sheaves, arithmetic genus, and Serre duality 465
18.5. A first glimpse of Serre duality 473
18.6. Hilbert functions, Hilbert polynomials, and genus 476
18.7. ⋆ Serre's cohomological characterization of ampleness 482
18.8. Higher direct image sheaves 485
18.9. ⋆ Chow's Lemma and Grothendieck's Coherence Theorem 489

Chapter 19. Application: Curves 493
19.1. A criterion for a morphism to be a closed embedding 493
19.2. A series of crucial tools 495
19.3. Curves of genus 0 498
19.4. Classical geometry arising from curves of positive genus 499
19.5. Hyperelliptic curves 501
19.6. Curves of genus 2 505
19.7. Curves of genus 3 506
19.8. Curves of genus 4 and 5 508
19.9. Curves of genus 1 511
19.10. Elliptic curves are group varieties 518
19.11. Counterexamples and pathologies using elliptic curves 523

Chapter 20. ⋆ Application: A glimpse of intersection theory 529
20.1. Intersecting n line bundles with an n-dimensional variety 529
20.2. Intersection theory on a surface 533
20.3. The Grothendieck group of coherent sheaves, and an algebraic version of homology 539
20.4. ⋆⋆ The Nakai-Moishezon and Kleiman criteria for ampleness 541

Chapter 21. Differentials 547
21.1. Motivation and game plan 547
21.2. Definitions and first properties 548
21.3. Smoothness of varieties revisited 561
21.4. Examples 564
21.5. Studying smooth varieties using their cotangent bundles 569
21.6. Unramified morphisms 574
21.7. The Riemann-Hurwitz Formula 575

Chapter 22. ⋆ Blowing up 583
22.1. Motivating example: blowing up the origin in the plane 583
22.2. Blowing up, by universal property 585

22.3. The blow-up exists, and is projective 589
22.4. Examples and computations 594

Part VI. More 603

Chapter 23. Derived functors 605
23.1. The Tor functors 605
23.2. Derived functors in general 609
23.3. Derived functors and spectral sequences 613
23.4. Derived functor cohomology of $\mathscr{O}$-modules 618
23.5. Čech cohomology and derived functor cohomology agree 621

Chapter 24. Flatness 627
24.1. Introduction 627
24.2. Easier facts 629
24.3. Flatness through Tor 634
24.4. Ideal-theoretic criteria for flatness 636
24.5. Topological aspects of flatness 643
24.6. Local criteria for flatness 647
24.7. Flatness implies constant Euler characteristic 651

Chapter 25. Smooth, étale, and unramified morphisms revisited 655
25.1. Some motivation 655
25.2. Different characterizations of smooth and étale morphisms 657
25.3. Generic smoothness and the Kleiman-Bertini Theorem 662

Chapter 26. Depth and Cohen-Macaulayness 667
26.1. Depth 667
26.2. Cohen-Macaulay rings and schemes 670
26.3. $\star\star$ Serre's R1 + S2 criterion for normality 673

Chapter 27. Twenty-seven lines 679
27.1. Introduction 679
27.2. Preliminary facts 680
27.3. Every smooth cubic surface (over $\overline{k}$) has 27 lines 682
27.4. Every smooth cubic surface (over $\overline{k}$) is a blown up plane 685

Chapter 28. Cohomology and base change theorems 689
28.1. Statements and applications 689
28.2. $\star\star$ Proofs of cohomology and base change theorems 695
28.3. Applying cohomology and base change to moduli problems 702

Chapter 29. Power series and the Theorem on Formal Functions 707
29.1. Introduction 707
29.2. Algebraic preliminaries 707
29.3. Defining types of singularities 711
29.4. The Theorem on Formal Functions 713
29.5. Zariski's Connectedness Lemma and Stein Factorization 715
29.6. Zariski's Main Theorem 717
29.7. Castelnuovo's criterion for contracting (-1)-curves 721

29.8. ⋆⋆ Proof of the Theorem on Formal Functions 29.4.2 724

Chapter 30. ⋆ Proof of Serre duality 729
30.1. Introduction 729
30.2. Ext groups and Ext sheaves for $\mathscr{O}$-modules 734
30.3. Serre duality for projective k-schemes 738
30.4. The adjunction formula for the dualizing sheaf, and $\omega_X = \mathscr{K}_X$ 742

Bibliography 747

Index 753

I can illustrate the ... approach with the ... image of a nut to be opened. The first analogy that came to my mind is of immersing the nut in some softening liquid, and why not simply water? From time to time you rub so the liquid penetrates better, and otherwise you let time pass. The shell becomes more flexible through weeks and months — when the time is ripe, hand pressure is enough, the shell opens like a perfectly ripened avocado!

A different image came to me a few weeks ago. The unknown thing to be known appeared to me as some stretch of earth or hard marl, resisting penetration ... the sea advances insensibly in silence, nothing seems to happen, nothing moves, the water is so far off you hardly hear it ... yet finally it surrounds the resistant substance.

— *A. Grothendieck* [**Gr6**, p. 552-3], *translation by C. McLarty* [**Mc**, p. 1]

Preface

This book is intended to give a serious and reasonably complete introduction to algebraic geometry, not just for (future) experts in the field. The exposition serves a narrow set of goals (see §0.4), and necessarily takes a particular point of view on the subject.

It has now been four decades since David Mumford wrote that algebraic geometry "seems to have acquired the reputation of being esoteric, exclusive, and very abstract, with adherents who are secretly plotting to take over all the rest of mathematics! In one respect this last point is accurate ..." ([**Mu4**, preface] and [**Mu7**, p. 227]). The revolution has now fully come to pass, and has fundamentally changed how we think about many fields of pure mathematics. A remarkable number of celebrated advances rely in some way on the insights and ideas forcefully articulated by Grothendieck, Serre, and others.

For a number of reasons, algebraic geometry has earned a reputation of being inaccessible. The power of the subject comes from rather abstract heavy machinery, and it is easy to lose sight of the intuitive nature of the objects and methods. Many in nearby fields have only vague ideas of the fundamental ideas of the subject. Algebraic geometry itself has fractured into many parts, and even within algebraic geometry, new researchers are often unaware of the basic ideas in subfields removed from their own.

But there is another more optimistic perspective to be taken. The ideas that allow algebraic geometry to connect several parts of mathematics are fundamental, and well-motivated. Many people in nearby fields would find it useful to develop a working knowledge of the foundations of the subject, and not just at a superficial level. Within algebraic geometry itself, there is a canon (at least for those approaching the subject from this particular direction), that everyone in the field can and should be familiar with. The rough edges of scheme theory have been sanded down over the past half century, although there remains an inescapable need to understand the subject on its own terms.

0.0.1. *The importance of exercises.* This book has a lot of exercises. I have found that unless I have some problems I can think through, ideas don't get fixed in my mind. Some exercises are trivial — some experts find this offensive, but I find this desirable. A very few necessary ones may be hard, but the reader should have been given the background to deal with them — they are not just an excuse to push hard material out of the text. The exercises are interspersed with the exposition, not left to the end. Most have been extensively field-tested. The point of view here is one I explored with Kedlaya and Poonen in [**KPV**], a book that was ostensibly about problems, but secretly a case for how one should learn and do and think about mathematics. Most people learn by doing, rather than just passively reading.

Judiciously chosen problems can be the best way of guiding the learner toward enlightenment.

0.0.2. Acknowledgments.
This one is going to be really hard, so I'll write this later. (Mike Stay is the author of Jokes 1.3.11 and 21.5.2.)

0.1 For the reader

This is your last chance. After this, there is no turning back. You take the blue pill, the story ends, you wake up in your bed and believe whatever you want to believe. You take the red pill, you stay in Wonderland and I show you how deep the rabbit-hole goes.
— *Morpheus*

The contents of this book are intended to be a collection of communal wisdom, necessarily distilled through an imperfect filter. I wish to say a few words on how you might use it, although it is not clear to me if you should or will follow this advice.

Before discussing details, I want to say clearly at the outset: the wonderful machine of modern algebraic geometry was created to understand basic and naive questions about geometry (broadly construed). The purpose of this book is to give you a thorough foundation in these powerful ideas. *Do not be seduced by the lotus-eaters into infatuation with untethered abstraction.* Hold tight to the your geometric motivation as you learn the formal structures which have proved to be so effective in studying fundamental questions. When introduced to a new idea, always ask why you should care. Do not expect an answer right away, but demand an answer eventually. Try at least to apply any new abstraction to some concrete example you can understand well.

Understanding algebraic geometry is often thought to be hard because it consists of large complicated pieces of machinery. In fact the opposite is true; to switch metaphors, rather than being narrow and deep, algebraic geometry is shallow but extremely broad. It is built out of a large number of very small parts, in keeping with Grothendieck's vision of mathematics. It is a challenge to hold the entire organic structure, with its messy interconnections, in your head.

A reasonable place to start is with the idea of "affine complex varieties": subsets of $\mathbb{C}^n$ cut out by some polynomial equations. Your geometric intuition can immediately come into play — you may already have some ideas or questions about dimension, or smoothness, or solutions over subfields such as $\mathbb{R}$ or $\mathbb{Q}$. Wiser heads would counsel spending time understanding complex varieties in some detail before learning about schemes. Instead, I encourage you to learn about schemes immediately, learning about affine complex varieties as the central (but not exclusive) example. This is not ideal, but can save time, and is surprisingly workable. An alternative is to learn about varieties elsewhere, and then come back later.

The intuition for schemes can be built on the intuition for affine complex varieties. Allen Knutson and Terry Tao have pointed out that this involves three different simultaneous generalizations, which can be interpreted as three large themes in mathematics. (i) We allow nilpotents in the ring of functions, which is basically analysis (looking at near-solutions of equations instead of exact solutions). (ii) We glue these affine schemes together, which is what we do in differential geometry (looking at manifolds instead of coordinate patches). (iii) Instead of working over $\mathbb{C}$ (or another algebraically closed field), we work more generally over a ring that isn't an algebraically closed field, or even a field at all, which is basically number theory (solving equations over number fields, rings of integers, etc.).

Because our goal is to be comprehensive, and to understand everything one should know after a first course, it will necessarily take longer to get to interesting sample applications. You may be misled into thinking that one has to work this hard to get to these applications — it is not true! You should deliberately keep an eye out for examples you would have cared about before. This will take some time and patience.

As you learn algebraic geometry, you should pay attention to crucial stepping stones. Of course, the steps get bigger the farther you go.

Chapter 1. Category theory is only language, but it is language with an embedded logic. Category theory is much easier once you realize that it is designed to formalize and abstract things you already know. The initial chapter on category theory prepares you to think cleanly. For example, when someone names something a "cokernel" or a "product", you should want to know why it deserves that name, and what the name really should mean. The conceptual advantages of thinking this way will gradually become apparent over time. Yoneda's Lemma — and more generally, the idea of understanding an object through the maps to it — will play an important role.

Chapter 2. The theory of sheaves again abstracts something you already understand well (see the motivating example of §2.1), and what is difficult is understanding how one best packages and works with the information of a sheaf (stalks, sheafification, sheaves on a base, etc.).

Chapters 1 and 2 are a risky gamble, and they attempt a delicate balance. Attempts to explain algebraic geometry often leave such background to the reader, refer to other sources the reader won't read, or punt it to a telegraphic appendix. Instead, this book attempts to explain everything necessary, but as little possible, and tries to get across how you should think about (and work with) these fundamental ideas, and why they are more grounded than you might fear.

Chapters 3–5. Armed with this background, you will be able to think cleanly about various sorts of "spaces" studied in different parts of geometry (including differentiable real manifolds, topological spaces, and complex manifolds), as ringed spaces that locally are of a certain form. A scheme is just another kind of "geometric space", and we are then ready to transport lots of intuition from "classical geometry" to this new setting. (This also will set you up to later think about other geometric kinds of spaces in algebraic geometry, such as complex analytic spaces, algebraic spaces, orbifolds, stacks, rigid analytic spaces, and formal schemes.) The ways in which schemes differ from your geometric intuition can be internalized, and your intuition can be expanded to accomodate them. There are many properties you will realize you will want, as well as other properties that will later prove important. These all deserve names. Take your time becoming familiar with them.

Chapters 6–10. Thinking categorically will lead you to ask about morphisms about schemes (and other spaces in geometry). One of Grothendieck's fundamental lessons is that the morphisms are central. Important geometric properties should really be understood as properties of morphisms. There are many classes of morphisms with special names, and in each case you should think through why that class deserves a name.

Chapters 11–12. You will then be in a good position to think about fundamental geometric properties of schemes: dimension and smoothness. You may be surprised that these are subtle ideas, but you should keep in mind that they are subtle everywhere in mathematics.

Chapters 13–21. Vector bundles are ubiquitous tools in geometry, and algebraic geometry is no exception. They lead us to the more general notion of quasicoherent sheaves, much as free modules over a ring lead us to modules more generally. We study their properties next, including cohomology. Chapter 19, applying these ideas ideas to study curves, may help make clear how useful they are.

Chapters 23–30. With this in hand, you are ready to learn more advanced tools widely used in the subject. Many examples of what you can do are given, and the classical story of the 27 lines on a smooth cubic surface (Chapter 27) is a good opportunity to see many ideas come together.

The rough logical dependencies among the chapters are shown in Figure 0.1. (Caution: this should be taken with a grain of salt. For example, you can avoid using much of Chapter 19 on curves in later chapters, but it is a crucial source of examples, and a great way to consolidate your understanding. And Chapter 29 on completions uses Chapters 19, 20 and 22 only in the discussion of Castelnuovo's Criterion 29.7.1.)

In general, I like having as few hypotheses as possible. Certainly a hypothesis that isn't necessary to the proof is a red herring. But if a reasonable hypothesis can make the proof cleaner and more memorable, I am willing to include it.

In particular, Noetherian hypotheses are handy when necessary, but are otherwise misleading. Even Noetherian-minded readers (normal human beings) are better off having the right hypotheses, as they will make clearer why things are true.

We often state results particular to varieties, especially when there are techniques unique to this situation that one should know. But restricting to algebraically closed fields is useful surprisingly rarely. Geometers needn't be afraid of arithmetic examples or of algebraic examples; a central insight of algebraic geometry is that the same formalism applies without change.

Pathological examples are useful to know. On mountain highways, there are tall sticks on the sides of the road designed for bad weather. In winter, you cannot see the road clearly, and the sticks serve as warning signs: if you cross this line, you will die! Pathologies and (counter)examples serve a similar goal. They also serve as a reality check, when confronting a new statement, theorem, or conjecture, whose veracity you may doubt.

When working through a book in algebraic geometry, it is particularly helpful to have other algebraic geometry books at hand, to see different approaches and to have alternate expositions when things become difficult. This book may serve

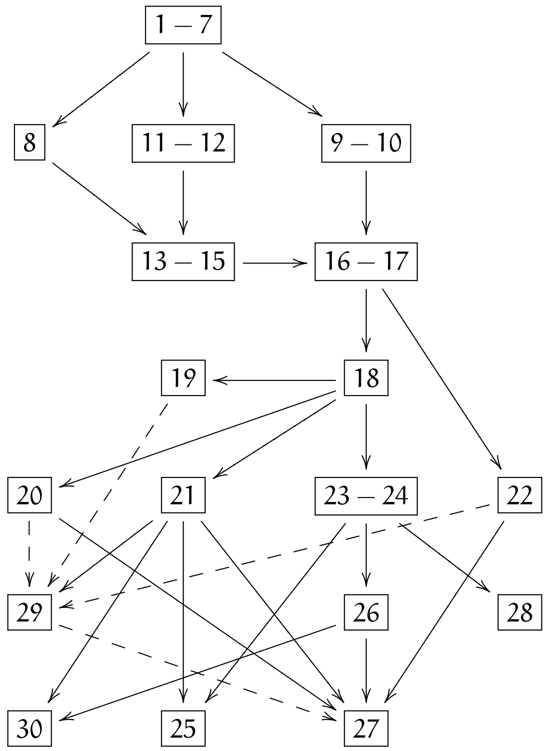


FIGURE 0.1. Important logical dependences among chapters (or more precisely, a directed graph showing which chapter should be read before which other chapter)

as a good secondary book. If it is your primary source, then two other excellent books with what I consider a similar philosophy are [**Liu**] and [**GW**]. De Jong's encyclopedic online reference [**Stacks**] is peerless. There are many other outstanding sources out there, perhaps one for each approach to the subject; you should browse around and find one you find sympathetic.

If you are looking for a correct or complete history of the subject, you have come to the wrong place. This book is not intended to be a complete guide to the literature, and many important sources are ignored or left out, due to my own ignorance and laziness.

Finally, if you attempt to read this without working through a significant number of exercises (see §0.0.1), I will come to your house and pummel you with [**Gr-EGA**] until you beg for mercy. It is important to not just have a vague sense of what is true, but to be able to actually get your hands dirty. As Mark Kisin has said, "You can wave your hands all you want, but it still won't make you fly."

0.2 For the expert

If you use this book for a course, you should of course adapt it to your own point of view and your own interests. In particular, you should think about an application or theorem you want to reach at the end of the course (which may well not be in this book), and then work toward it. You should feel no compulsion to sprint to the end; I advise instead taking more time, and ending at the right place for your students. (Figure 0.1, showing large-scale dependencies among the chapters, may help you map out a course.) I have found that the theory of curves (Chapter 19) and the 27 lines on the cubic surface (Chapter 27) have served this purpose well at the end of winter and spring quarters. This was true even if some of the needed background was not covered, and had to be taken by students as some sort of black box.

Faithfulness to the goals of §0.4 required a brutal triage, and I have made a number of decisions you may wish to reverse. I will briefly describe some choices made that may be controversial.

Decisions on how to describe things were made for the sake of the learners. If there were two approaches, and one was "correct" from an advanced point of view, and one was direct and natural from a naive point of view, I went with the latter.

On the other hand, the theory of varieties (over an algebraically closed field, say) was *not* done first and separately. This choice brought me close to tears, but in the end I am convinced that it can work well, if done in the right spirit.

Instead of spending the first part of the course on varieties, I spent the time in a different way. It is tempting to assume that students will either arrive with great comfort and experience with category theory and sheaf theory, or that they should pick up these ideas on their own time. I would love to live in that world. I encourage you to not skimp on these foundational issues. I have found that although these first lectures felt painfully slow to me, they were revelatory to a number of the students, and those with more experience were not bored and did not waste their time. This investment paid off in spades when I was able to rely on their ability to think cleanly and to use these tools in practice. Furthermore, if they left the course with nothing more than hands-on experience with these ideas, the world was still better off for it.

For the most part, we will state results in the maximal generality that the proof justifies, but we will not give a much harder proof where the generality of the stronger result will not be used. There are a few cases where we work harder to prove a somewhat more general result that many readers may not appreciate. For example, we prove a number of theorems for proper morphisms, not just projective morphisms. But in such cases, readers are invited or encouraged to ignore the subtleties required for the greater generality.

I consider line bundles (and maps to projective space) more fundamental than divisors. General Cartier divisors are not discussed (although *effective* Cartier divisors play an essential role).

Cohomology is done first using the Čech approach (as Serre first did), and derived functor cohomology is introduced only later. I am well aware that Grothendieck thinks of the fact that the agreement of Čech cohomology with derived functor cohomology "should be considered as an accidental phenomenon", and that "it is important for technical reasons not to take as *definition* of cohomology the Čech cohomology", [**Gr4**, p. 108]. But I am convinced that this is the right way for most people to see this kind of cohomology *for the first time*. (It is certainly true that many topics in algebraic geometry are best understood in the language of derived functors. But this is a view from the mountaintop, looking down, and not the best way to explore the forests. In order to appreciate derived functors appropriately, one must understand the homological algebra behind it, and not just take it as a black box.)

We restrict to the Noetherian case only when it is necessary, or (rarely) when it really saves effort. In this way, non-Noetherian people will clearly see where they should be careful, and Noetherian people will realize that non-Noetherian things are not so terrible. Moreover, even if you are interested primarily in Noetherian schemes, it helps to see "Noetherian" in the hypotheses of theorems only when necessary, as it will help you remember how and when this property gets used.

There are some cases where Noetherian readers will suffer a little more than they would otherwise. As an inflammatory example, instead of using Noetherian hypotheses, the notion of quasiseparated comes up early and often. The cost is that one extra word has to be remembered, on top of an overwhelming number of other words. But once that is done, it is not hard to remember that essentially every scheme anyone cares about is quasiseparated. Furthermore, whenever the hypotheses "quasicompact and quasiseparated" turn up, the reader will immediately guess a key idea of the proof. As another example, coherent sheaves and finite type (quasicoherent) sheaves are the same in the Noetherian situation, but are still worth distinguishing in statements of the theorems and exercises, for the same reason: to be clearer on what is used in the proof.

Many important topics are not discussed. Valuative criteria are not proved (see §12.7), and their statement is relegated to an optional section. Completely omitted: devissage, formal schemes, and cohomology with supports. Sorry!

0.3 Background and conventions

"Should you just be an algebraist or a geometer?" is like saying "Would you rather be deaf or blind?"

— *M. Atiyah*, [**At2**, p. 659]

All rings are assumed to be commutative unless explicitly stated otherwise. All rings are assumed to contain a unit, denoted 1. Maps of rings must send 1 to 1. We don't require that $0 \neq 1$; in other words, the "0-ring" (with one element) is a ring. (There is a ring map from any ring to the 0-ring; the 0-ring only maps to itself. The 0-ring is the final object in the category of rings.) The definition of "integral domain" includes $1 \neq 0$, so the 0-ring is not an integral domain. We accept the axiom of choice. In particular, any proper ideal in a ring is contained in a maximal ideal. (The axiom of choice also arises in the argument that the category of A-modules has enough injectives, see Exercise 23.2.G.)

The reader should be familiar with some basic notions in commutative ring theory, in particular the notion of ideals (including prime and maximal ideals) and localization. Tensor products and exact sequences of A-modules will be important. We will use the notation $(A, \mathfrak{m})$ or $(A, \mathfrak{m}, k)$ for local rings (rings with a unique maximal ideal) — A is the ring, $\mathfrak{m}$ its maximal ideal, and $k = A/\mathfrak{m}$ its residue field.

We will use the structure theorem for finitely generated modules over a principal ideal domain A: any such module can be written as the direct sum of principal modules $A/(a)$. Some experience with field theory will be helpful from time to time.

0.3.1. *Caution about foundational issues.* We will not concern ourselves with subtle foundational issues (set-theoretic issues, universes, etc.). It is true that some people should be careful about these issues. But is that really how you want to live your life? (If you are one of these rare people, a good start is [**KS2**, §1.1].)

0.3.2. *Further background.* It may be helpful to have books on other subjects at hand that you can dip into for specific facts, rather than reading them in advance. In commutative algebra, [**E**] is good for this. Other popular choices are [**AtM**] and [**Mat2**]. The book [**Al**] takes a point of view useful to algebraic geometry. For homological algebra, [**Wei**] is simultaneously detailed and readable.

Background from other parts of mathematics (topology, geometry, complex analysis, number theory, ...) will of course be helpful for intuition and grounding. Some previous exposure to topology is certainly essential.

0.3.3. *Nonmathematical conventions.* "Unimportant" means "unimportant for the current exposition", *not* necessarily unimportant in the larger scheme of things. Other words may be used idiosyncratically as well.

There are optional starred sections of topics worth knowing on a second or third (but not first) reading. They are marked with a star: $\star$. Starred sections are not necessarily harder, merely unimportant. You should not read double-starred sections ($\star\star$) unless you really really want to, but you should be aware of their existence. (It may be strange to have parts of a book that should *not* be read!)

Let's now find out if you are taking my advice about double-starred sections.

0.4 $\star\star$ The goals of this book

There are a number of possible introductions to the field of algebraic geometry: Riemann surfaces; complex geometry; the theory of varieties; a nonrigorous examples-based introduction; algebraic geometry for number theorists; an abstract functorial approach; and more. All have their place. Different approaches suit different students (and different advisors). This book takes only one route.

Our intent is to cover a canon completely and rigorously, with enough examples and calculations to help develop intuition for the machinery. This is often the content of a second course in algebraic geometry, and in an ideal world, people would learn this material over many years, after having background courses in commutative algebra, algebraic topology, differential geometry, complex analysis, homological algebra, number theory, and French literature. We do not live in an ideal world. For this reason, the book is written as a first introduction, but a challenging one.

This book seeks to do a very few things, but to try to do them well. Our goals and premises are as follows.

The core of the material should be digestible over a single year. After a year of blood, sweat, and tears, readers should have a broad familiarity with the foundations of the subject, and be ready to attend seminars, and learn more advanced material. They should not just have a vague intuitive understanding of the ideas of the subject; they should know interesting examples, know why they are interesting, and be able to work through their details. Readers in other fields of mathematics should know enough to understand the algebro-geometric ideas that arise in their area of interest.

This means that this book is not encyclopedic, and even beyond that, hard choices have to be made. (In particular, analytic aspects are essentially ignored, and are at best dealt with in passing without proof. This is a book about *algebraic* algebraic geometry.)

This book is usable (and has been used) for a course, but the course should (as always) take on the personality of the instructor. With a good course, people should be able to leave early and still get something useful from the experience. With this book, it is possible to leave without regret after learning about category theory, or about sheaves, or about geometric spaces, having become a better person.

The book is also usable (and has been used) for learning on your own. But most mortals cannot learn algebraic geometry fully on their own; ideally you should read in a group, and even if not, you should have someone you can ask questions to (both stupid and smart questions).

There is certainly more than a year's material here, but I have tried to make clear which topics are essential, and which are not. Those teaching a class will choose which "inessential" things are important for the point they wish to get across, and use them.

There is a canon (at least for this particular approach to algebraic geometry). I have been repeatedly surprised at how much people in different parts of algebraic geometry agree on what every civilized algebraic geometer should know after a first (serious) year. (There are of course different canons for different parts of the subject, e.g. complex algebraic geometry, combinatorial algebraic geometry, computational algebraic geometry, etc.) There are extra bells and whistles that different instructors might add on, to prepare students for their particular part of the field or their own point of view, but the core of the subject remains unified, despite the diversity and richness of the subject. There are some serious and painful compromises to be made to reconcile this goal with the previous one.

Algebraic geometry is for everyone (with the appropriate definition of "everyone"). Algebraic geometry courses tend to require a lot of background, which makes them inaccessible to all but those who know they will go deeply into the subject. Algebraic geometry is too important for that; it is essential that many of those in nearby fields develop some serious familiarity with the foundational ideas and tools of the subject, and not just at a superficial level. (Similarly, algebraic geometers uninterested in any nearby field are necessarily arid, narrow thinkers. Do not be such a person!)

For this reason, this book attempts to require as little background as possible. The background required will, in a technical sense, be surprisingly minimal — ideally just some commutative ring theory and point-set topology, and some comfort with things like prime ideals and localization. This is misleading of course — the more you know, the better. And the less background you have, the harder you will have to work — this is not a light read. On a related note...

The book is intended to be as self-contained as possible. I have tried to follow the motto: "if you use it, you must prove it". I have noticed that most students are human beings: if you tell them that some algebraic fact is in some late chapter of a book in commutative algebra, they will not immediately go and read it. Surprisingly often, what we need can be developed quickly from scratch, and even if people do not read it, they can see what is involved. The cost is that the book is much denser, and that significant sophistication and maturity is demanded of the reader. The benefit is that more people can follow it; they are less likely to reach a point where they get thrown. On the other hand, people who already have some familiarity with algebraic geometry, but want to understand the foundations more completely, should not be bored, and will focus on more subtle issues.

As just one example, Krull's Principal Ideal Theorem 11.3.3 is an important tool. I have included an essentially standard proof (§11.5). I do not want people to read it (unless they really really want to), and signal this by a double-star in the title: ⋆⋆. Instead, I want people to skim it and realize that they *could* read it, and that it is not seriously hard.

This is an important goal because it is important not just to know what is true, but to know why things are true, and what is hard, and what is not hard. Also, this helps the previous goal, by reducing the number of prerequisites.

The book is intended to build intuition for the formidable machinery of algebraic geometry. The exercises are central for this (see §0.0.1). Informal language can sometimes be helpful. Many examples are given. Learning how to think cleanly (and in particular categorically) is essential. The advantages of *appropriate* generality should be made clear by example, and not by intimidation. The motivation is more local than global. For example, there is no introductory chapter explaining why one might be interested in algebraic geometry, and instead there is an introductory chapter explaining why you should want to think categorically (and how to actually do this).

Balancing the above goals is already impossible. We must thus give up any hope of achieving any other desiderata. **There are no other goals.**

Part I

Preliminaries

CHAPTER 1

Some category theory

The introduction of the digit 0 *or the group concept was general nonsense too, and mathematics was more or less stagnating for thousands of years because nobody was around to take such childish steps...*
— *A. Grothendieck,* [**BP**, p. 4-5]

That which does not kill me, makes me stronger.
— *F. Nietzsche*

1.1 Motivation

Before we get to any interesting geometry, we need to develop a language to discuss things cleanly and effectively. This is best done in the language of categories. There is not much to know about categories to get started; it is just a very useful language. Like all mathematical languages, category theory comes with an embedded logic, which allows us to abstract intuitions in settings we know well to far more general situations.

Our motivation is as follows. We will be creating some new mathematical objects (such as schemes, and certain kinds of sheaves), and we expect them to act like objects we have seen before. We could try to nail down precisely what we mean by "act like", and what minimal set of things we have to check in order to verify that they act the way we expect. Fortunately, we don't have to — other people have done this before us, by defining key notions, such as *abelian categories*, which behave like modules over a ring.

Our general approach will be as follows. I will try to tell what you need to know, and no more. (This I promise: if I use the word "topoi", you can shoot me.) I will begin by telling you things you already know, and describing what is essential about the examples, in a way that we can abstract a more general definition. We will then see this definition in less familiar settings, and get comfortable with using it to solve problems and prove theorems.

For example, we will define the notion of *product* of schemes. We could just give a definition of product, but then you should want to know why this precise definition deserves the name of "product". As a motivation, we revisit the notion of product in a situation we know well: (the category of) sets. One way to define the product of sets U and V is as the set of ordered pairs $\{(u, v) \,:\, u \in U, v \in V\}$. But someone from a different mathematical culture might reasonably define it as the set of symbols $\{ {u \atop v} : \; u \in U, v \in V\}$. These notions are "obviously the same". Better: there is "an obvious bijection between the two".

This can be made precise by giving a better definition of product, in terms of a *universal property*. Given two sets M and N, a product is a set P, along with maps $\mu : P \to M$ and $\nu : P \to N$, such that for *any set* P' *with maps* $\mu' : P' \to M$ *and* $\nu' : P' \to N$, these maps must factor *uniquely* through P:

(1.1.0.1)
$$\begin{tikzcd}
P' \arrow[dr, dotted, "\exists!"] \arrow[drr, bend left, "\nu'"] \arrow[ddr, bend right, "\mu'"'] & & \\
& P \arrow[r, "\nu"] \arrow[d, "\mu"] & N \\
& M &
\end{tikzcd}$$

(The symbol $\exists$ means "there exists", and the symbol ! here means "unique".) Thus a **product** is a *diagram*

$$\begin{tikzcd}
P \arrow[r, "\nu"] \arrow[d, "\mu"'] & N \\
M &
\end{tikzcd}$$

and not just a set P, although the maps μ and ν are often left implicit.

This definition agrees with the traditional definition, with one twist: there isn't just a single product; but any two products come with a *unique* isomorphism between them. In other words, the product is unique up to unique isomorphism. Here is why: if you have a product

$$\begin{tikzcd}
P_1 \arrow[r, "\nu_1"] \arrow[d, "\mu_1"'] & N \\
M &
\end{tikzcd}$$

and I have a product

$$\begin{tikzcd}
P_2 \arrow[r, "\nu_2"] \arrow[d, "\mu_2"'] & N \\
M &
\end{tikzcd}$$

then by the universal property of my product (letting (P_2, μ_2, ν_2) play the role of (P, μ, ν), and (P_1, μ_1, ν_1) play the role of (P', μ', ν') in (1.1.0.1)), there is a unique map $f : P_1 \to P_2$ making the appropriate diagram commute (i.e., $\mu_1 = \mu_2 \circ f$ and $\nu_1 = \nu_2 \circ f$). Similarly by the universal property of your product, there is a unique map $g : P_2 \to P_1$ making the appropriate diagram commute. Now consider the universal property of my product, this time letting (P_2, μ_2, ν_2) play the role of both

(P, μ, ν) and (P', μ', ν') in (1.1.0.1). There is a unique map $h : P_2 \to P_2$ such that

$$\begin{tikzcd} P_2 \arrow[dr, "h"] \arrow[drr, bend left, "\nu_2"] \arrow[ddr, bend right, "\mu_2"'] & & \\ & P_2 \arrow[r, "\nu_2"'] \arrow[d, "\mu_2"] & N \\ & M & \end{tikzcd}$$

commutes. However, I can name two such maps: the identity map id_{P_2}, and $f \circ g$. Thus $f \circ g = \mathrm{id}_{P_2}$. Similarly, $g \circ f = \mathrm{id}_{P_1}$. Thus the maps f and g arising from the universal property are bijections. In short, there is a unique bijection between P_1 and P_2 preserving the "product structure" (the maps to M and N). This gives us the right to name any such product $M \times N$, since any two such products are uniquely identified.

This definition has the advantage that it works in many circumstances, and once we define categories, we will soon see that the above argument applies verbatim in any category to show that products, if they exist, are unique up to unique isomorphism. Even if you haven't seen the definition of category before, you can verify that this agrees with your notion of product in some category that you have seen before (such as the category of vector spaces, where the maps are taken to be linear maps; or the category of manifolds, where the maps are taken to be *submersions*, i.e., differentiable maps whose differential is everywhere surjective).

This is handy even in cases that you understand. For example, one way of defining the product of two manifolds M and N is to cut them both up into charts, then take products of charts, then glue them together. But if I cut up the manifolds in one way, and you cut them up in another, how do we know our resulting manifolds are the "same"? We could wave our hands, or make an annoying argument about refining covers, but instead, we should just show that they are "categorical products" and hence canonically the "same" (i.e., isomorphic). We will formalize this argument in §1.3.

Another set of notions we will abstract are categories that "behave like modules". We will want to define kernels and cokernels for new notions, and we should make sure that these notions behave the way we expect them to. This leads us to the definition of *abelian categories*, first defined by Grothendieck in his Tôhoku paper [**Gr1**].

In this chapter, we will give an informal introduction to these and related notions, in the hope of giving just enough familiarity to comfortably use them in practice.

1.2 Categories and functors

Before functoriality, people lived in caves.
— B. Conrad

We begin with an informal definition of categories and functors.

1.2.1. Categories.

A **category** consists of a collection of **objects**, and for each pair of objects, a set of **morphisms** (or **arrows**) between them. (For experts: technically, this is the definition of a *locally small category*. In the correct definition, the morphisms need only form a class, not necessarily a set, but see Caution 0.3.1.) Morphisms are often informally called **maps**. The collection of objects of a category $\mathscr{C}$ is often denoted obj($\mathscr{C}$), but we will usually denote the collection also by $\mathscr{C}$. If $A, B \in \mathscr{C}$, then the set of morphisms from A to B is denoted $\text{Mor}(A, B)$. A morphism is often written $f : A \to B$, and A is said to be the **source** of f, and B the **target** of f. (Of course, $\text{Mor}(A, B)$ is taken to be disjoint from $\text{Mor}(A', B')$ unless $A = A'$ and $B = B'$.)

Morphisms compose as expected: there is a composition $\text{Mor}(B, C) \times \text{Mor}(A, B) \to \text{Mor}(A, C)$, and if $f \in \text{Mor}(A, B)$ and $g \in \text{Mor}(B, C)$, then their composition is denoted $g \circ f$. Composition is associative: if $f \in \text{Mor}(A, B)$, $g \in \text{Mor}(B, C)$, and $h \in \text{Mor}(C, D)$, then $h \circ (g \circ f) = (h \circ g) \circ f$. For each object $A \in \mathscr{C}$, there is always an **identity morphism** $\text{id}_A : A \to A$, such that when you (left- or right-)compose a morphism with the identity, you get the same morphism. More precisely, for any morphisms $f : A \to B$ and $g : B \to C$, $\text{id}_B \circ f = f$ and $g \circ \text{id}_B = g$. (If you wish, you may check that "identity morphisms are unique": there is only one morphism deserving the name id_A.) This ends the definition of a category.

We have a notion of **isomorphism** between two objects of a category (a morphism $f : A \to B$ such that there exists some — necessarily unique — morphism $g : B \to A$, where $f \circ g$ and $g \circ f$ are the identity on B and A respectively), and a notion of **automorphism** of an object (an isomorphism of the object with itself).

1.2.2. *Example.* The prototypical example to keep in mind is the category of sets, denoted *Sets*. The objects are sets, and the morphisms are maps of sets. (Because Russell's paradox shows that there is no set of all sets, we did not say earlier that there is a set of all objects. But as stated in §0.3, we are deliberately omitting all set-theoretic issues.)

1.2.3. *Example.* Another good example is the category Vec_k of vector spaces over a given field k. The objects are k-vector spaces, and the morphisms are linear transformations. (What are the isomorphisms?)

1.2.A. UNIMPORTANT EXERCISE. A category in which each morphism is an isomorphism is called a **groupoid**. (This notion is not important in what we will discuss. The point of this exercise is to give you some practice with categories, by relating them to an object you know well.)
(a) A perverse definition of a group is: a groupoid with one object. Make sense of this.
(b) Describe a groupoid that is not a group.

1.2.B. EXERCISE. If A is an object in a category $\mathscr{C}$, show that the invertible elements of $\text{Mor}(A, A)$ form a group (called the **automorphism group of** A, denoted $\text{Aut}(A)$). What are the automorphism groups of the objects in Examples 1.2.2 and 1.2.3? Show that two isomorphic objects have isomorphic automorphism groups. (For readers with a topological background: if X is a topological space, then the fundamental groupoid is the category where the objects are points of X, and the morphisms $x \to y$ are paths from x to y, up to homotopy. Then the automorphism group of x_0 is the (pointed) fundamental group $\pi_1(X, x_0)$. In the case where X is connected, and $\pi_1(X)$ is not abelian, this illustrates the fact that for a connected groupoid — whose definition you can guess — the automorphism groups of the objects are all isomorphic, but not canonically isomorphic.)

1.2.4. *Example: abelian groups.* The abelian groups, along with group homomorphisms, form a category *Ab*.

1.2.5. *Important example: modules over a ring.* If A is a ring, then the A-modules form a category Mod_A. (This category has additional structure; it will be the prototypical example of an *abelian category*, see §1.6.) Taking $A = k$, we obtain Example 1.2.3; taking $A = \mathbb{Z}$, we obtain Example 1.2.4.

1.2.6. *Example: rings.* There is a category *Rings*, where the objects are rings, and the morphisms are maps of rings in the usual sense (maps of sets which respect addition and multiplication, and which send 1 to 1 by our conventions, §0.3).

1.2.7. *Example: topological spaces.* The topological spaces, along with continuous maps, form a category *Top*. The isomorphisms are homeomorphisms.

In all of the above examples, the objects of the categories were in obvious ways sets with additional structure (a **concrete category**, although we won't use this terminology). This needn't be the case, as the next example shows.

1.2.8. *Example: partially ordered sets.* A **partially ordered set**, or **poset**, is a set S along with a binary relation $\geq$ on S satisfying:

(i) $x \geq x$ (reflexivity),
(ii) $x \geq y$ and $y \geq z$ imply $x \geq z$ (transitivity), and
(iii) if $x \geq y$ and $y \geq x$ then $x = y$ (antisymmetry).

A partially ordered set $(S, \geq)$ can be interpreted as a category whose objects are the elements of S, and with a single morphism from x to y if and only if $x \geq y$ (and no morphism otherwise).

A trivial example is $(S, \geq)$ where $x \geq y$ if and only if $x = y$. Another example is

(1.2.8.1)
$$\begin{tikzcd} & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

Here there are three objects. The identity morphisms are omitted for convenience, and the two non-identity morphisms are depicted. A third example is

(1.2.8.2)
$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

Here the "obvious" morphisms are again omitted: the identity morphisms, and the morphism from the upper left to the lower right. Similarly,

$$\cdots \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet$$

depicts a partially ordered set, where again, only the "generating morphisms" are depicted.

1.2.9. *Example: the category of subsets of a set, and the category of open sets in a topological space.* If X is a set, then the subsets form a partially ordered set, where the order is given by inclusion. Informally, if $U \subset V$, then we have exactly one morphism $U \to V$ in the category (and otherwise none). Similarly, if X is a topological space, then the *open* sets form a partially ordered set, where the order is given by inclusion.

1.2.10. *Definition.* A **subcategory** $\mathscr{A}$ of a category $\mathscr{B}$ has as its objects some of the objects of $\mathscr{B}$, and some of the morphisms, such that the morphisms of $\mathscr{A}$ include the identity morphisms of the objects of $\mathscr{A}$, and are closed under composition. (For example, (1.2.8.1) is in an obvious way a subcategory of (1.2.8.2). Also, we have an obvious "inclusion functor" $i : \mathscr{A} \to \mathscr{B}$.)

1.2.11. Functors.

A **covariant functor** F from a category $\mathscr{A}$ to a category $\mathscr{B}$, denoted $F : \mathscr{A} \to \mathscr{B}$, is the following data. It is a map of objects $F : \mathrm{obj}(\mathscr{A}) \to \mathrm{obj}(\mathscr{B})$, and for each $A_1, A_2 \in \mathscr{A}$, and morphism $m : A_1 \to A_2$, a morphism $F(m) : F(A_1) \to F(A_2)$ in $\mathscr{B}$. We require that F preserves identity morphisms (for $A \in \mathscr{A}$, $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$), and that F preserves composition ($F(m_2 \circ m_1) = F(m_2) \circ F(m_1)$). (You may wish to verify that covariant functors send isomorphisms to isomorphisms.) A trivial example is the **identity functor** $\mathrm{id} : \mathscr{A} \to \mathscr{A}$, whose definition you can guess. Here are some less trivial examples.

1.2.12. *Example: a forgetful functor.* Consider the functor from the category of vector spaces (over a field k) Vec_k to *Sets*, that associates to each vector space its underlying set. The functor sends a linear transformation to its underlying map of sets. This is an example of a **forgetful functor**, where some additional structure is forgotten. Another example of a forgetful functor is $Mod_A \to Ab$ from A-modules to abelian groups, remembering only the abelian group structure of the A-module.

1.2.13. *Topological examples.* Examples of covariant functors include the fundamental group functor π_1, which sends a topological space X with choice of a point $x_0 \in X$ to a group $\pi_1(X, x_0)$ (what are the objects and morphisms of the source category?), and the ith homology functor $Top \to Ab$, which sends a topological space X to its ith homology group $H_i(X, \mathbb{Z})$. The covariance corresponds to the fact that a (continuous) morphism of pointed topological spaces $\phi : X \to Y$ with $\phi(x_0) = y_0$ induces a map of fundamental groups $\pi_1(X, x_0) \to \pi_1(Y, y_0)$, and similarly for homology groups.

1.2.14. *Example.* Suppose A is an object in a category $\mathscr{C}$. Then there is a functor $h^A : \mathscr{C} \to Sets$ sending $B \in \mathscr{C}$ to $\mathrm{Mor}(A, B)$, and sending $f : B_1 \to B_2$ to $\mathrm{Mor}(A, B_1) \to \mathrm{Mor}(A, B_2)$ described by

$$[g : A \to B_1] \mapsto [f \circ g : A \to B_1 \to B_2].$$

This seemingly silly functor ends up surprisingly being an important concept.

1.2.15. Definitions. If $F : \mathscr{A} \to \mathscr{B}$ and $G : \mathscr{B} \to \mathscr{C}$ are covariant functors, then we define a functor $G \circ F : \mathscr{A} \to \mathscr{C}$ (the **composition** of G and F) in the obvious way. Composition of functors is associative in an evident sense.

A covariant functor $F : \mathscr{A} \to \mathscr{B}$ is **faithful** if for all $A, A' \in \mathscr{A}$, the map $\mathrm{Mor}_{\mathscr{A}}(A, A') \to \mathrm{Mor}_{\mathscr{B}}(F(A), F(A'))$ is injective, and **full** if it is surjective. A functor that is full and faithful is **fully faithful**. A subcategory $i : \mathscr{A} \to \mathscr{B}$ is a **full subcategory** if i is full. Thus a subcategory $\mathscr{A}'$ of $\mathscr{A}$ is full if and only if for all $A, B \in \mathrm{obj}(\mathscr{A}')$, $\mathrm{Mor}_{\mathscr{A}'}(A, B) = \mathrm{Mor}_{\mathscr{A}}(A, B)$. For example, the forgetful functor $Vec_k \to Sets$ is faithful, but not full; and if A is a ring, the category of finitely generated A-modules is a full subcategory of the category Mod_A of A-modules.

1.2.16. Definition. A **contravariant functor** is defined in the same way as a covariant functor, except the arrows switch directions: in the above language, $F(A_1 \to A_2)$ is now an arrow from $F(A_2)$ to $F(A_1)$. (Thus $F(m_2 \circ m_1) = F(m_1) \circ F(m_2)$, not $F(m_2) \circ F(m_1)$.)

It is wise to state whether a functor is covariant or contravariant, unless the context makes it very clear. If it is not stated (and the context does not make it clear), the functor is often assumed to be covariant.

(Sometimes people describe a contravariant functor $\mathscr{C} \to \mathscr{D}$ as a covariant functor $\mathscr{C}^{\mathrm{opp}} \to \mathscr{D}$, where $\mathscr{C}^{\mathrm{opp}}$ is the same category as $\mathscr{C}$ except that the arrows go in the opposite direction. Here $\mathscr{C}^{\mathrm{opp}}$ is said to be the **opposite category** to $\mathscr{C}$.) One can define fullness, etc. for contravariant functors, and you should do so.

1.2.17. *Linear algebra example.* If Vec_k is the category of k-vector spaces (introduced in Example 1.2.3), then taking duals gives a contravariant functor $(\cdot)^\vee : Vec_k \to Vec_k$. Indeed, to each linear transformation $f : V \to W$, we have a dual transformation $f^\vee : W^\vee \to V^\vee$, and $(f \circ g)^\vee = g^\vee \circ f^\vee$.

1.2.18. *Topological example (cf. Example 1.2.13) for those who have seen cohomology.* The ith cohomology functor $H^i(\cdot, \mathbb{Z}) : Top \to Ab$ is a contravariant functor.

1.2.19. *Example.* There is a contravariant functor $Top \to Rings$ taking a topological space X to the ring of real-valued continuous functions on X. A morphism of topological spaces $X \to Y$ (a continuous map) induces the pullback map from functions on Y to functions on X.

1.2.20. *Example (the functor of points, cf. Example 1.2.14).* Suppose A is an object of a category $\mathscr{C}$. Then there is a contravariant functor $h_A : \mathscr{C} \to Sets$ sending $B \in \mathscr{C}$ to $\mathrm{Mor}(B, A)$, and sending the morphism $f : B_1 \to B_2$ to the morphism $\mathrm{Mor}(B_2, A) \to \mathrm{Mor}(B_1, A)$ via

$$[g : B_2 \to A] \mapsto [g \circ f : B_1 \to B_2 \to A].$$

This example initially looks weird and different, but Examples 1.2.17 and 1.2.19 may be interpreted as special cases; do you see how? What is A in each case? This functor might reasonably be called the *functor of maps* (to A), but is actually known as the **functor of points**. We will meet this functor again in §1.3.10 and (in the category of schemes) in Definition 6.3.7.

1.2.21. $\star$ **Natural transformations (and natural isomorphisms) of covariant functors, and equivalences of categories.**

(This notion won't come up in an essential way until at least Chapter 6, so you shouldn't read this section until then.) Suppose F and G are two covariant functors from $\mathscr{A}$ to $\mathscr{B}$. A **natural transformation of covariant functors** $F \to G$ is the data of a morphism $m_A : F(A) \to G(A)$ for each $A \in \mathscr{A}$ such that for each $f : A \to A'$ in $\mathscr{A}$, the diagram

$$\begin{tikzcd} F(A) \arrow[r, "F(f)"] \arrow[d, "m_A"'] & F(A') \arrow[d, "m_{A'}"] \\ G(A) \arrow[r, "G(f)"'] & G(A') \end{tikzcd}$$

commutes. A **natural isomorphism** of functors is a natural transformation such that each m_A is an isomorphism. (We make analogous definitions when F and G are both contravariant.)

The data of functors $F : \mathscr{A} \to \mathscr{B}$ and $F' : \mathscr{B} \to \mathscr{A}$ such that $F \circ F'$ is naturally isomorphic to the identity functor $\mathrm{id}_{\mathscr{B}}$ on $\mathscr{B}$ and $F' \circ F$ is naturally isomorphic to $\mathrm{id}_{\mathscr{A}}$ is said to be an **equivalence of categories**. "Equivalence of categories" is an equivalence relation on categories. The right notion of when two categories are "essentially the same" is not *isomorphism* (a functor giving bijections of objects and morphisms) but *equivalence*. Exercises 1.2.C and 1.2.D might give you some vague sense of this. Later exercises (for example, that "rings" and "affine schemes" are essentially the same, once arrows are reversed, Exercise 6.3.D) may help too.

Two examples might make this strange concept more comprehensible. The double dual of a finite-dimensional vector space V is *not* V, but we learn early to say that it is canonically isomorphic to V. We can make that precise as follows. Let $f.d.Vec_k$ be the category of finite-dimensional vector spaces over k. Note that this category contains oodles of vector spaces of each dimension.

1.2.C. EXERCISE. Let $(\cdot)^{\vee\vee} : f.d.Vec_k \to f.d.Vec_k$ be the double dual functor from the category of finite-dimensional vector spaces over k to itself. Show that $(\cdot)^{\vee\vee}$ is naturally isomorphic to the identity functor on $f.d.Vec_k$. (Without the finite-dimensionality hypothesis, we only get a natural transformation of functors from id to $(\cdot)^{\vee\vee}$.)

Let $\mathscr{V}$ be the category whose objects are the k-vector spaces k^n for each $n \geq 0$ (there is one vector space for each n), and whose morphisms are linear transformations. The objects of $\mathscr{V}$ can be thought of as vector spaces with bases, and the morphisms as matrices. There is an obvious functor $\mathscr{V} \to f.d.Vec_k$, as each k^n *is* a finite-dimensional vector space.

1.2.D. EXERCISE. Show that $\mathscr{V} \to f.d.Vec_k$ gives an equivalence of categories, by describing an "inverse" functor. (Recall that we are being cavalier about set-theoretic assumptions, see Caution 0.3.1, so feel free to simultaneously choose bases for each vector space in $f.d.Vec_k$. To make this precise, you will need to use Gödel-Bernays set theory or else replace $f.d.Vec_k$ with a very similar small category, but we won't worry about this.)

1.2.22. $\star\star$ *Aside for experts.* Your argument for Exercise 1.2.D will show that (modulo set-theoretic issues) this definition of equivalence of categories is the same as another one commonly given: a covariant functor $F : \mathscr{A} \to \mathscr{B}$ is an equivalence of categories if it is fully faithful and every object of $\mathscr{B}$ is isomorphic to an object of the form $F(A)$ for some $A \in \mathscr{A}$ (F is *essentially surjective*). Indeed, one can show that such a functor has a *quasiinverse*, i.e., a functor $G : \mathscr{B} \to \mathscr{A}$ (necessarily also an equivalence and unique up to unique isomorphism) for which $G \circ F \cong \mathrm{id}_{\mathscr{A}}$ and $F \circ G \cong \mathrm{id}_{\mathscr{B}}$, and conversely, any functor that has a quasiinverse is an equivalence.

1.3 Universal properties determine an object up to unique isomorphism

Given some category that we come up with, we often will have ways of producing new objects from old. In good circumstances, such a definition can be made using the notion of a *universal property*. Informally, we wish that there were an object with some property. We first show that if it exists, then it is essentially unique, or more precisely, is unique up to unique isomorphism. Then we go about constructing an example of such an object to show existence.

Explicit constructions are sometimes easier to work with than universal properties, but with a little practice, universal properties are useful in proving things quickly and slickly. Indeed, when learning the subject, people often find explicit constructions more appealing, and use them more often in proofs, but as they become more experienced, they find universal property arguments more elegant and insightful.

1.3.1. Products were defined by a universal property. We have seen one important example of a universal property argument already in §1.1: products. You should go back and verify that our discussion there gives a notion of product in any category, and shows that products, *if they exist*, are unique up to unique isomorphism.

1.3.2. Initial, final, and zero objects. Here are some simple but useful concepts that will give you practice with universal property arguments. An object of a category $\mathscr{C}$ is an **initial object** if it has precisely one map to every object. It is a **final object** if it has precisely one map from every object. It is a **zero object** if it is both an initial object and a final object.

1.3.A. EXERCISE. Show that any two initial objects are uniquely isomorphic. Show that any two final objects are uniquely isomorphic.

In other words, *if* an initial object exists, it is unique up to unique isomorphism, and similarly for final objects. This (partially) justifies the phrase "*the* initial object" rather than "*an* initial object", and similarly for "*the* final object" and "*the* zero object". (Convention: we often say "the", not "a", for anything defined up to unique isomorphism.)

1.3.B. EXERCISE. What are the initial and final objects in *Sets*, *Rings*, and *Top* (if they exist)? How about in the two examples of §1.2.9?

1.3.3. Localization of rings and modules. Another important example of a definition by universal property is the notion of *localization* of a ring. We first review a constructive definition, and then reinterpret the notion in terms of universal property. A multiplicative subset S of a ring A is a subset closed under multiplication containing 1. We define a ring $S^{-1}A$. The elements of $S^{-1}A$ are of the form a/s

where $a \in A$ and $s \in S$, and where $a_1/s_1 = a_2/s_2$ if (and only if) *for some* $s \in S$, $s(s_2a_1 - s_1a_2) = 0$. We define $(a_1/s_1) + (a_2/s_2) = (s_2a_1 + s_1a_2)/(s_1s_2)$, and $(a_1/s_1) \times (a_2/s_2) = (a_1a_2)/(s_1s_2)$. (If you wish, you may check that this equality of fractions really is an equivalence relation and the two binary operations on fractions are well-defined on equivalence classes and make $S^{-1}A$ into a ring.) We have a canonical ring map

$$A \to S^{-1}A \tag{1.3.3.1}$$

given by $a \mapsto a/1$. Note that if $0 \in S$, $S^{-1}A$ is the 0-ring.

There are two particularly important flavors of multiplicative subsets. The first is $\{1, f, f^2, \dots\}$, where $f \in A$. This localization is denoted A_f. The second is $A - \mathfrak{p}$, where $\mathfrak{p}$ is a prime ideal. This localization $S^{-1}A$ is denoted $A_{\mathfrak{p}}$. (Notational warning: If $\mathfrak{p}$ is a prime ideal, then $A_{\mathfrak{p}}$ means you're allowed to divide by elements not in $\mathfrak{p}$. However, if $f \in A$, A_f means you're allowed to divide by f. This can be confusing. For example, if (f) is a prime ideal, then $A_f \neq A_{(f)}$.)

Warning: sometimes localization is first introduced in the special case where A is an integral domain and $0 \notin S$. In that case, $A \hookrightarrow S^{-1}A$, but this isn't always true, as shown by the following exercise. (But we will see that noninjective localizations needn't be pathological, and we can sometimes understand them geometrically, see Exercise 3.2.L.)

1.3.C. EXERCISE. Show that $A \to S^{-1}A$ is injective if and only if S contains no zerodivisors. (A **zerodivisor** of a ring A is an element a such that there is a nonzero element b with $ab = 0$. The other elements of A are called **non-zerodivisors**. For example, an invertible element is never a zerodivisor. Counter-intuitively, 0 is a zerodivisor in every ring but the 0-ring.)

If A is an integral domain and $S = A - \{0\}$, then $S^{-1}A$ is called the **fraction field** of A, which we denote $K(A)$. The previous exercise shows that A is a subring of its fraction field $K(A)$. We now return to the case where A is a general (commutative) ring.

1.3.D. EXERCISE. Verify that $A \to S^{-1}A$ satisfies the following universal property: $S^{-1}A$ is initial among A-algebras B where every element of S is sent to an invertible element in B. (Recall: the data of "an A-algebra B" and "a ring map $A \to B$" are the same.) Translation: any map $A \to B$ where every element of S is sent to an invertible element must factor uniquely through $A \to S^{-1}A$. Another translation: a ring map out of $S^{-1}A$ is the same thing as a ring map from A that sends every element of S to an invertible element. Furthermore, an $S^{-1}A$-module is the same thing as an A-module for which $s \times \cdot : M \to M$ is an A-module isomorphism for all $s \in S$.

In fact, it is cleaner to *define* $A \to S^{-1}A$ by the universal property, and to show that it exists, and to use the universal property to check various properties $S^{-1}A$ has. Let's get some practice with this by *defining* localizations of modules by universal property. Suppose M is an A-module. We define the A-module map $\phi : M \to S^{-1}M$ as being initial among A-module maps $M \to N$ such that elements of S are invertible in N ($s \times \cdot : N \to N$ is an isomorphism for all $s \in S$). More precisely, any such map $\alpha : M \to N$ factors uniquely through ϕ:

$$\begin{tikzcd} M \arrow[r, "\phi"] \arrow[dr, "\alpha"'] & S^{-1}M \arrow[d, dotted, "\exists!"] \\ & N \end{tikzcd}$$

(Translation: $M \to S^{-1}M$ is universal (initial) among A-module maps from M to modules that are actually $S^{-1}A$-modules. Can you make this precise by defining clearly the objects and morphisms in this category?)

Notice: (i) this determines $\phi : M \to S^{-1}M$ up to unique isomorphism (you should think through what this means); (ii) we are defining not only $S^{-1}M$, but also the map ϕ at the same time; and (iii) essentially by definition the A-module structure on $S^{-1}M$ extends to an $S^{-1}A$-module structure.

1.3.E. EXERCISE. Show that $\phi : M \to S^{-1}M$ exists, by constructing something satisfying the universal property. Hint: define elements of $S^{-1}M$ to be of the form m/s where $m \in M$ and $s \in S$, and $m_1/s_1 = m_2/s_2$ if and only if for some $s \in S$, $s(s_2 m_1 - s_1 m_2) = 0$. Define the additive structure by $(m_1/s_1) + (m_2/s_2) = (s_2 m_1 + s_1 m_2)/(s_1 s_2)$, and the $S^{-1}A$-module structure (and hence the A-module structure) is given by $(a_1/s_1) \cdot (m_2/s_2) = (a_1 m_2)/(s_1 s_2)$.

1.3.F. EXERCISE.
(a) Show that localization commutes with finite products, or equivalently, with finite direct sums. In other words, if $M_1, \dots, M_n$ are A-modules, describe an isomorphism (of A-modules, and of $S^{-1}A$-modules) $S^{-1}(M_1 \times \cdots \times M_n) \to S^{-1}M_1 \times \cdots \times S^{-1}M_n$.
(b) Show that localization commutes with *arbitrary* direct sums.
(c) Show that "localization does not necessarily commute with infinite products": the obvious map $S^{-1}(\prod_i M_i) \to \prod_i S^{-1}M_i$ induced by the universal property of localization is not always an isomorphism. (Hint: $(1, 1/2, 1/3, 1/4, \dots) \in \mathbb{Q} \times \mathbb{Q} \times \cdots$.)

1.3.4. *Remark.* Localization does not always commute with Hom, see Example 1.6.8. But Exercise 1.6.G will show that in good situations (if the first argument of Hom is *finitely presented*), localization *does* commute with Hom.

1.3.5. Tensor products. Another important example of a universal property construction is the notion of a **tensor product** of A-modules

$$\otimes_A : \quad \operatorname{obj}(Mod_A) \times \operatorname{obj}(Mod_A) \longrightarrow \operatorname{obj}(Mod_A)$$

$$(M, N) \longmapsto M \otimes_A N$$

The subscript A is often suppressed when it is clear from context. The tensor product is often defined as follows. Suppose you have two A-modules M and N. Then elements of the tensor product $M \otimes_A N$ are finite A-linear combinations of symbols $m \otimes n$ ($m \in M$, $n \in N$), subject to relations $(m_1 + m_2) \otimes n = m_1 \otimes n + m_2 \otimes n$, $m \otimes (n_1 + n_2) = m \otimes n_1 + m \otimes n_2$, $a(m \otimes n) = (am) \otimes n = m \otimes (an)$ (where $a \in A$, $m_1, m_2 \in M$, $n_1, n_2 \in N$). More formally, $M \otimes_A N$ is the free A-module generated by $M \times N$, quotiented by the submodule generated by $(m_1 + m_2, n) - (m_1, n) - (m_2, n)$, $(m, n_1 + n_2) - (m, n_1) - (m, n_2)$, $a(m, n) - (am, n)$, and $a(m, n) - (m, an)$ for $a \in A$, $m, m_1, m_2 \in M$, $n, n_1, n_2 \in N$. The image of (m, n) in this quotient is $m \otimes n$.

If A is a field k, we recover the tensor product of vector spaces.

1.3.G. EXERCISE (IF YOU HAVEN'T SEEN TENSOR PRODUCTS BEFORE). Show that $\mathbb{Z}/(10) \otimes_{\mathbb{Z}} \mathbb{Z}/(12) \cong \mathbb{Z}/(2)$. (This exercise is intended to give some hands-on practice with tensor products.)

1.3.H. IMPORTANT EXERCISE: RIGHT-EXACTNESS OF $(\cdot) \otimes_A N$. Show that $(\cdot) \otimes_A N$ gives a covariant functor $Mod_A \to Mod_A$. Show that $(\cdot) \otimes_A N$ is a **right-exact functor**, i.e., if

$$M' \to M \to M'' \to 0$$

is an exact sequence of A-modules (which means $f : M \to M''$ is surjective, and M' surjects onto the kernel of f; see §1.6), then the induced sequence

$$M' \otimes_A N \to M \otimes_A N \to M'' \otimes_A N \to 0$$

is also exact. This exercise is repeated in Exercise 1.6.F, but you may get a lot out of doing it now. (You will be reminded of the definition of right-exactness in §1.6.5.)

In contrast, you can quickly check that tensor product is not left-exact: tensor the exact sequence of $\mathbb{Z}$-modules

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\times 2} \mathbb{Z} \longrightarrow \mathbb{Z}/(2) \longrightarrow 0$$

with $\mathbb{Z}/(2)$.

The constructive definition $\otimes$ is a weird definition, and really the "wrong" definition. To motivate a better one: notice that there is a natural A-bilinear map $M \times N \to M \otimes_A N$. (If $M, N, P \in Mod_A$, a map $f : M \times N \to P$ is A-**bilinear** if $f(m_1 + m_2, n) = f(m_1, n) + f(m_2, n)$, $f(m, n_1 + n_2) = f(m, n_1) + f(m, n_2)$, and $f(am, n) = f(m, an) = af(m, n)$.) *Any* A-bilinear map $M \times N \to P$ factors through the tensor product uniquely: $M \times N \to M \otimes_A N \to P$. (Think this through!)

We can take this as the *definition* of the tensor product as follows. It is an A-module T along with an A-bilinear map $t : M \times N \to T$, such that given any A-bilinear map $t' : M \times N \to T'$, there is a unique A-linear map $f : T \to T'$ such that $t' = f \circ t$.

$$\begin{tikzcd} M \times N \arrow[rr, "t"] \arrow[dr, "t'"'] & & T \arrow[dl, dotted, "\exists! f"] \\ & T' & \end{tikzcd}$$

1.3.I. EXERCISE. Show that $(T, t : M \times N \to T)$ is unique up to unique isomorphism. Hint: first figure out what "unique up to unique isomorphism" means for such pairs, using a category of pairs (T, t). Then follow the analogous argument for the product.

In short: given M and N, there is an A-bilinear map $t : M \times N \to M \otimes_A N$, unique up to unique isomorphism, defined by the following universal property:

for any A-bilinear map $t' : M \times N \to T'$ there is a unique A-linear map $f : M \otimes_A N \to T'$ such that $t' = f \circ t$.

As with all universal property arguments, this argument shows uniqueness *assuming existence*. To show existence, we need an explicit construction.

1.3.J. EXERCISE. Show that the construction of §1.3.5 satisfies the universal property of tensor product.

The three exercises below are useful facts about tensor products with which you should be familiar.

1.3.K. IMPORTANT EXERCISE.
(a) If M is an A-module and $A \to B$ is a morphism of rings, give $B \otimes_A M$ the structure of a B-module (this is part of the exercise). Show that this describes a functor $Mod_A \to Mod_B$.
(b) If further $A \to C$ is another morphism of rings, show that $B \otimes_A C$ has a natural structure of a ring. Hint: multiplication will be given by $(b_1 \otimes c_1)(b_2 \otimes c_2) = (b_1 b_2) \otimes (c_1 c_2)$. (Exercise 1.3.U will interpret this construction as a fibered coproduct.)

1.3.L. IMPORTANT EXERCISE. If S is a multiplicative subset of A and M is an A-module, describe a natural isomorphism $(S^{-1}A) \otimes_A M \cong S^{-1}M$ (as $S^{-1}A$-modules *and* as A-modules).

1.3.M. EXERCISE ($\otimes$ COMMUTES WITH $\oplus$). Show that tensor products commute with arbitrary direct sums: if M and $\{N_i\}_{i \in I}$ are all A-modules, describe an isomorphism

$$M \otimes (\oplus_{i \in I} N_i) \xrightarrow{\sim} \oplus_{i \in I} (M \otimes N_i).$$

1.3.6. Essential Example: Fibered products. Suppose we have morphisms $\alpha : X \to Z$ and $\beta : Y \to Z$ (in *any* category). Then the **fibered product** (or *fibred product*) is an object $X \times_Z Y$ along with morphisms $pr_X : X \times_Z Y \to X$ and $pr_Y : X \times_Z Y \to Y$, where the two compositions $\alpha \circ pr_X, \beta \circ pr_Y : X \times_Z Y \to Z$ agree, such that given any object W with maps to X and Y (whose compositions to Z agree), these maps factor through some unique $W \to X \times_Z Y$:

$$\begin{tikzcd}
W \arrow[dr, dotted, "\exists!"] \arrow[drr, bend left] \arrow[ddr, bend right] & & \\
& X \times_Z Y \arrow[r, "pr_Y"'] \arrow[d, "pr_X"] & Y \arrow[d, "\beta"] \\
& X \arrow[r, "\alpha"] & Z
\end{tikzcd}$$

(Warning: the definition of the fibered product depends on α and β, even though they are omitted from the notation $X \times_Z Y$.)

By the usual universal property argument, if it exists, it is unique up to unique isomorphism. (You should think this through until it is clear to you.) Thus the use of the phrase "the fibered product" (rather than "a fibered product") is reasonable, and we should reasonably be allowed to give it the name $X \times_Z Y$. We know what maps to it are: they are precisely maps to X and maps to Y that agree as maps to Z.

Depending on your religion, the diagram

$$\begin{tikzcd} X \times_Z Y \arrow[r, "pr_Y"'] \arrow[d, "pr_X"] & Y \arrow[d, "\beta"] \\ X \arrow[r, "\alpha"] & Z \end{tikzcd}$$

is called a **fibered/pullback/Cartesian diagram/square** (six possibilities — even more are possible if you prefer "fibred" to "fibered").

The right way to interpret the notion of fibered product is first to think about what it means in the category of sets.

1.3.N. EXERCISE. Show that in *Sets*,

$$X \times_Z Y = \{(x, y) \in X \times Y \;:\; \alpha(x) = \beta(y)\}.$$

More precisely, show that the right side, equipped with its evident maps to X and Y, satisfies the universal property of the fibered product. (This will help you build intuition for fibered products.)

1.3.O. EXERCISE. If X is a topological space, show that fibered products always exist in the category of open sets of X, by describing what a fibered product is. (Hint: it has a one-word description.)

1.3.P. EXERCISE. If Z is the final object in a category $\mathscr{C}$, and $X, Y \in \mathscr{C}$, show that "$X \times_Z Y = X \times Y$": "the" fibered product over Z is uniquely isomorphic to "the" product. Assume all relevant (fibered) products exist. (This is an exercise about unwinding the definition.)

1.3.Q. USEFUL EXERCISE: TOWERS OF FIBER DIAGRAMS ARE FIBER DIAGRAMS. If the two squares in the following commutative diagram are fiber diagrams, show that the "outside rectangle" (involving U, V, Y, and Z) is also a fiber diagram.

$$\begin{tikzcd} U \arrow[r] \arrow[d] & V \arrow[d] \\ W \arrow[r] \arrow[d] & X \arrow[d] \\ Y \arrow[r] & Z \end{tikzcd}$$

1.3.R. EXERCISE. Given morphisms $X_1 \to Y$, $X_2 \to Y$, and $Y \to Z$, show that there is a natural morphism $X_1 \times_Y X_2 \to X_1 \times_Z X_2$, assuming that both fibered products exist. (This is trivial once you figure out what it is saying. The point of this exercise is to see why it is trivial.)

1.3.S. USEFUL EXERCISE: THE MAGIC DIAGRAM. Suppose we are given morphisms $X_1, X_2 \to Y$ and $Y \to Z$. Show that the following diagram is a fibered square.

$$\begin{tikzcd} X_1 \times_Y X_2 \arrow[r] \arrow[d] & X_1 \times_Z X_2 \arrow[d] \\ Y \arrow[r] & Y \times_Z Y \end{tikzcd}$$

Assume all relevant (fibered) products exist. This diagram is surprisingly useful — so useful that we will call it the **magic diagram**.

1.3.7. *Coproducts.* Define **coproduct** in a category by reversing all the arrows in the definition of product. Define **fibered coproduct** in a category by reversing all the arrows in the definition of fibered product.

1.3.T. EXERCISE. Show that coproduct for *Sets* is disjoint union. This is why we use the notation $\coprod$ for disjoint union.

1.3.U. EXERCISE. Suppose $A \to B$ and $A \to C$ are two ring morphisms, so in particular B and C are A-modules. Recall (Exercise 1.3.K) that $B \otimes_A C$ has a ring structure. Show that there is a natural morphism $B \to B \otimes_A C$ given by $b \mapsto b \otimes 1$. (This is not necessarily an inclusion; see Exercise 1.3.G.) Similarly, there is a natural morphism $C \to B \otimes_A C$. Show that this gives a fibered coproduct on rings, i.e., that

$$\begin{tikzcd} B \otimes_A C & C \arrow[l] \\ B \arrow[u] & A \arrow[l] \arrow[u] \end{tikzcd}$$

satisfies the universal property of fibered coproduct.

1.3.8. Monomorphisms and epimorphisms.

1.3.9. Definition. A morphism $\pi : X \to Y$ is a **monomorphism** if any two morphisms $\mu_1 : Z \to X$ and $\mu_2 : Z \to X$ such that $\pi \circ \mu_1 = \pi \circ \mu_2$ must satisfy $\mu_1 = \mu_2$. In other words, there is at most one way of filling in the dotted arrow so that the diagram

$$\begin{tikzcd} Z \arrow[d, dotted, "\leq 1"'] \arrow[dr] & \\ X \arrow[r, "\pi"'] & Y \end{tikzcd}$$

commutes — for any object Z, the natural map $\mathrm{Mor}(Z, X) \to \mathrm{Mor}(Z, Y)$ is an injection. Intuitively, it is the categorical version of an injective map, and indeed this notion generalizes the familiar notion of injective maps of sets. (The reason we don't use the word "injective" is that in some contexts, "injective" will have an intuitive meaning which may not agree with "monomorphism". One example: in the category of divisible groups, the map $\mathbb{Q} \to \mathbb{Q}/\mathbb{Z}$ is a monomorphism but not injective. This is also the case with "epimorphism" vs. "surjective".)

1.3.V. EXERCISE. Show that the composition of two monomorphisms is a monomorphism.

1.3.W. EXERCISE. Prove that a morphism $\pi : X \to Y$ is a monomorphism if and only if the fibered product $X \times_Y X$ exists, and the induced morphism $X \to X \times_Y X$ is an isomorphism. We may then take this as the definition of monomorphism. (Monomorphisms aren't central to future discussions, although they will come up again. This exercise is just good practice.)

1.3.X. EASY EXERCISE. We use the notation of Exercise 1.3.R. Show that if $Y \to Z$ is a monomorphism, then the morphism $X_1 \times_Y X_2 \to X_1 \times_Z X_2$ you described in Exercise 1.3.R is an isomorphism. (Hint: for any object V, give a natural bijection between maps from V to the first and maps from V to the second. It is also possible to use the magic diagram, Exercise 1.3.S.)

The notion of an **epimorphism** is "dual" to the definition of monomorphism, where all the arrows are reversed. This concept will not be central for us, although it turns up in the definition of an abelian category. Intuitively, it is the categorical version of a surjective map. (But be careful when working with categories of objects that are sets with additional structure, as epimorphisms need not be surjective. Example: in the category *Rings*, $\mathbb{Z} \to \mathbb{Q}$ is an epimorphism, but obviously not surjective.)

1.3.10. Representable functors and Yoneda's lemma. Much of our discussion about universal properties can be cleanly expressed in terms of representable functors, under the rubric of "Yoneda's Lemma". Yoneda's lemma is an easy fact stated in a complicated way. Informally speaking, you can essentially recover an object in a category by knowing the maps into it. For example, we have seen that the data of maps to $X \times Y$ are naturally (canonically) the data of maps to X and to Y. Indeed, we have now taken this as the *definition* of $X \times Y$.

Recall Example 1.2.20. Suppose A is an object of category $\mathscr{C}$. For any object $C \in \mathscr{C}$, we have a set of morphisms $\mathrm{Mor}(C, A)$. If we have a morphism $f : B \to C$, we get a map of sets

$$\mathrm{Mor}(C, A) \to \mathrm{Mor}(B, A), \tag{1.3.10.1}$$

by composition: given a map from C to A, we get a map from B to A by precomposing with $f : B \to C$. Hence this gives a contravariant functor $h_A : \mathscr{C} \to \mathit{Sets}$. Yoneda's Lemma states that the functor h_A determines A up to unique isomorphism. More precisely:

1.3.Y. IMPORTANT EXERCISE THAT YOU SHOULD DO ONCE IN YOUR LIFE (YONEDA'S LEMMA).
(a) Suppose you have two objects A and A' in a category $\mathscr{C}$, and morphisms

$$i_C : \mathrm{Mor}(C, A) \to \mathrm{Mor}(C, A') \tag{1.3.10.2}$$

that commute with the maps (1.3.10.1). Show that the i_C (as C ranges over the objects of $\mathscr{C}$) are induced from a unique morphism $g : A \to A'$. More precisely, show that there is a unique morphism $g : A \to A'$ such that for all $C \in \mathscr{C}$, i_C is $u \mapsto g \circ u$.
(b) If furthermore the i_C are all bijections, show that the resulting g is an isomorphism. (Hint for both: This is much easier than it looks. This statement is so general that there are really only a couple of things that you could possibly try. For example, if you're hoping to find a morphism $A \to A'$, where will you find

it? Well, you are looking for an element $\mathrm{Mor}(A, A')$. So just plug in $C = A$ to (1.3.10.2), and see where the identity goes.)

There is an analogous statement with the arrows reversed, where instead of maps into A, you think of maps *from* A. The role of the contravariant functor h_A of Example 1.2.20 is played by the covariant functor h^A of Example 1.2.14. Because the proof is the same (with the arrows reversed), you needn't think it through.

The phrase "Yoneda's lemma" properly refers to a more general statement. Although it looks more complicated, it is no harder to prove.

1.3.Z. ⋆ EXERCISE.
(a) Suppose A and B are objects in a category $\mathscr{C}$. Give a bijection between the natural transformations $h^A \to h^B$ of covariant functors $\mathscr{C} \to$ *Sets* (see Example 1.2.14 for the definition) and the morphisms $B \to A$.
(b) State and prove the corresponding fact for contravariant functors h_A (see Example 1.2.20). Remark: A contravariant functor F from $\mathscr{C}$ to *Sets* is said to be **representable** if there is a natural isomorphism

$$\xi : F \xrightarrow{\sim} h_A .$$

Thus the representing object A is determined up to unique isomorphism by the pair (F, ξ). There is a similar definition for covariant functors. (We will revisit this in §6.6, and this problem will appear again as Exercise 6.6.C. The element $\xi^{-1}(\mathrm{id}_A) \in F(A)$ is often called the "universal object"; do you see why?)
(c) **Yoneda's lemma.** Suppose F is a covariant functor $\mathscr{C} \to$ *Sets*, and $A \in \mathscr{C}$. Give a bijection between the natural transformations $h^A \to F$ and $F(A)$. (The corresponding fact for contravariant functors is essentially Exercise 9.1.C.)

In fancy terms, Yoneda's lemma states the following. Given a category $\mathscr{C}$, we can produce a new category, called the *functor category* of $\mathscr{C}$, where the objects are contravariant functors $\mathscr{C} \to$ *Sets*, and the morphisms are natural transformations of such functors. We have a functor (which we can usefully call h) from $\mathscr{C}$ to its functor category, which sends A to h_A. Yoneda's Lemma states that this is a fully faithful functor, called the *Yoneda embedding*. (Fully faithful functors were defined in §1.2.15.)

1.3.11. *Joke.* The Yoda embedding, contravariant it is.

1.4 Limits and colimits

Limits and colimits are two important definitions determined by universal properties. They generalize a number of familiar constructions. I will give the definition first, and then show you why it is familiar. For example, fractions will be motivating examples of colimits (Exercise 1.4.B(a)), and the p-adic integers (Example 1.4.3) will be motivating examples of limits.

1.4.1. Limits. We say that a category is a **small category** if the objects and the morphisms are sets. (This is a technical condition intended only for experts.) Suppose $\mathscr{I}$ is any small category, and $\mathscr{C}$ is any category. Then a functor $F : \mathscr{I} \to \mathscr{C}$ (i.e., with an object $A_i \in \mathscr{C}$ for each element $i \in \mathscr{I}$, and appropriate commuting morphisms dictated by $\mathscr{I}$) is said to be a **diagram indexed by** $\mathscr{I}$. We call $\mathscr{I}$ an **index category**. Our index categories will usually be partially ordered sets (Example 1.2.8), in which in particular there is at most one morphism between any two objects. (But other examples are sometimes useful.) For example, if $\square$ is the category

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

and $\mathscr{A}$ is a category, then a functor $\square \to \mathscr{A}$ is precisely the data of a commuting square in $\mathscr{A}$.

Then the **limit of the diagram** is an object $\varprojlim_{\mathscr{I}} A_i$ of $\mathscr{C}$ along with morphisms $f_j : \varprojlim_{\mathscr{I}} A_i \to A_j$ for each $j \in \mathscr{I}$, such that if $m : j \to k$ is a morphism in $\mathscr{I}$, then

(1.4.1.1)
$$\begin{tikzcd} \varprojlim_{\mathscr{I}} A_i \arrow[d, "f_j"'] \arrow[dr, "f_k"] & \\ A_j \arrow[r, "F(m)"] & A_k \end{tikzcd}$$

commutes, and this object and maps to each A_i are universal (final) with respect to this property. More precisely, given any other object W along with maps $g_i : W \to A_i$ commuting with the $F(m)$ (if $m : j \to k$ is a morphism in $\mathscr{I}$, then $g_k = F(m) \circ g_j$), then there is a unique map $g : W \to \varprojlim_{\mathscr{I}} A_i$ so that $g_i = f_i \circ g$ for all i. (In some cases, the limit is sometimes called the **inverse limit** or **projective limit**. We won't use this language.) By the usual universal property argument, if the limit exists, it is unique up to unique isomorphism.

1.4.2. *Examples: products.* For example, if $\mathscr{I}$ is the partially ordered set

$$\begin{tikzcd} & \bullet \arrow[d] \\ \bullet \arrow[r] & \bullet \end{tikzcd}$$

we obtain the fibered product.

If $\mathscr{I}$ is

$$\bullet \qquad \bullet$$

we obtain the product.

If $\mathscr{I}$ is a set (i.e., the only morphisms are the identity maps), then the limit is called the **product** of the A_i, and is denoted $\prod_i A_i$. The special case where $\mathscr{I}$ has two elements is the example of the previous paragraph.

If $\mathscr{I}$ has an initial object e, then A_e is the limit, and in particular the limit always exists.

1.4.3. *Example: the* p*-adic integers.* For a prime number p, the p**-adic integers** (or more informally, p**-adics**), $\mathbb{Z}_p$, are often described informally (and somewhat unnaturally) as being of the form $\mathbb{Z}_p = a_0 + a_1 p + a_2 p^2 + \cdots$ (where $0 \leq a_i < p$).

They are an example of a limit in the category of rings:

$$
\begin{tikzcd}
 & & & \mathbb{Z}_p \arrow[dl] \arrow[dr] \arrow[drrr] & & & \\
\cdots \arrow[rr] & & \mathbb{Z}/p^3 \arrow[rr] & & \mathbb{Z}/p^2 \arrow[rr] & & \mathbb{Z}/p.
\end{tikzcd}
$$

(Warning: $\mathbb{Z}_p$ is sometimes is used to denote the integers modulo p, but $\mathbb{Z}/(p)$ or $\mathbb{Z}/p\mathbb{Z}$ is better to use for this, to avoid confusion. Worse: by §1.3.3, $\mathbb{Z}_p$ also denotes those rationals whose denominators are a power of p. Hopefully the meaning of $\mathbb{Z}_p$ will be clear from the context.) The p-adic integers are an example of a *completion*, the topic of Chapter 29.

Limits do not always exist for any index category $\mathscr{I}$. However, you can often easily check that limits exist if the objects of your category can be interpreted as sets with additional structure, and arbitrary products exist (respecting the set-like structure).

1.4.A. IMPORTANT EXERCISE. Show that in the category *Sets*,

$$\left\{ (a_i)_{i \in \mathscr{I}} \in \prod_i A_i : F(m)(a_j) = a_k \text{ for all } m \in \mathrm{Mor}_{\mathscr{I}}(j,k) \in \mathrm{Mor}(\mathscr{I}) \right\},$$

along with the obvious projection maps to each A_i, is the limit $\varprojlim_{\mathscr{I}} A_i$.

This clearly also works in the category Mod_A of A-modules (in particular Vec_k and *Ab*), as well as *Rings*.

From this point of view, $2 + 3p + 2p^2 + \cdots \in \mathbb{Z}_p$ can be understood as the sequence $(2, 2 + 3p, 2 + 3p + 2p^2, \dots)$.

1.4.4. Colimits. More immediately relevant for us will be the dual (arrow-reversed version) of the notion of limit (or inverse limit). We just flip the arrows f_i in (1.4.1.1), and get the notion of a **colimit**, which is denoted $\varinjlim_{\mathscr{I}} A_i$. (You should draw the corresponding diagram.) Again, if it exists, it is unique up to unique isomorphism. (In some cases, the colimit is sometimes called the **direct limit**, **inductive limit**, or **injective limit**. We won't use this language. I prefer using limit/colimit in analogy with kernel/cokernel and product/coproduct. This is more than analogy, as kernels and products may be interpreted as limits, and similarly with cokernels and coproducts. Also, I remember that kernels "map to", and cokernels are "mapped to", which reminds me that a limit maps *to* all the objects in the big commutative diagram indexed by $\mathscr{I}$; and a colimit has a map *from* all the objects.)

1.4.5. *Joke.* A comathematician is a device for turning cotheorems into ffee.

Even though we have just flipped the arrows, colimits behave quite differently from limits.

1.4.6. *Example.* The set $5^{-\infty}\mathbb{Z}$ of rational numbers whose denominators are powers of 5 is a colimit $\varinjlim 5^{-i}\mathbb{Z}$. More precisely, $5^{-\infty}\mathbb{Z}$ is the colimit of the diagram

$$\mathbb{Z} \longrightarrow 5^{-1}\mathbb{Z} \longrightarrow 5^{-2}\mathbb{Z} \longrightarrow \cdots$$

The colimit over an index *set* I is called the **coproduct**, denoted $\coprod_i A_i$, and is the dual (arrow-reversed) notion to the product.

1.4.B. EXERCISE.
(a) Interpret the statement "$\mathbb{Q} = \varinjlim \frac{1}{n}\mathbb{Z}$".
(b) Interpret the union of some subsets of a given set as a colimit. (Dually, the intersection can be interpreted as a limit.) The objects of the category in question are the subsets of the given set.

Colimits do not always exist, but there are two useful large classes of examples for which they do.

1.4.7. Definition. A nonempty partially ordered set $(S, \geq)$ is **filtered** (or is said to be a **filtered set**) if for each $x, y \in S$, there is a z such that $x \geq z$ and $y \geq z$. More generally, a nonempty category $\mathscr{I}$ is **filtered** if:

(i) for each $x, y \in \mathscr{I}$, there is a $z \in \mathscr{I}$ and arrows $x \to z$ and $y \to z$, and
(ii) for every two arrows $u, v : x \to y$, there is an arrow $w : y \to z$ such that $w \circ u = w \circ v$.

(Other terminologies are also commonly used, such as "directed partially ordered set" and "filtered index category", respectively.)

1.4.C. EXERCISE. Suppose $\mathscr{I}$ is filtered. (We will almost exclusively use the case where $\mathscr{I}$ is a filtered set.) Recall the symbol $\coprod$ for disjoint union of sets. Show that any diagram in *Sets* indexed by $\mathscr{I}$ has the following, with the obvious maps to it, as a colimit:

$$\left\{ (a_i, i) \in \coprod_{i \in \mathscr{I}} A_i \right\} \Big/ \left(\begin{array}{c} (a_i, i) \sim (a_j, j) \text{ if and only if there are } f\colon A_i \to A_k \text{ and} \\ g\colon A_j \to A_k \text{ in the diagram for which } f(a_i) = g(a_j) \text{ in } A_k \end{array} \right)$$

(You will see that the "$\mathscr{I}$ filtered" hypothesis is there is to ensure that $\sim$ is an equivalence relation.)

For example, in Example 1.4.6, each element of the colimit is an element of something upstairs, but you can't say in advance what it is an element of. For example, $17/125$ is an element of the $5^{-3}\mathbb{Z}$ (or $5^{-4}\mathbb{Z}$, or later ones), but not $5^{-2}\mathbb{Z}$.

This idea applies to many categories whose objects can be interpreted as sets with additional structure (such as abelian groups, A-modules, groups, etc.). For example, the colimit $\varinjlim M_i$ in the category of A-modules Mod_A can be described as follows. The set underlying $\varinjlim M_i$ is defined as in Exercise 1.4.C. To add the elements $m_i \in M_i$ and $m_j \in M_j$, choose an $\ell \in \mathscr{I}$ with arrows $u : i \to \ell$ and $v : j \to \ell$, and then define the sum of m_i and m_j to be $F(u)(m_i) + F(v)(m_j) \in M_\ell$. The element $m_i \in M_i$ is 0 if and only if there is some arrow $u : i \to k$ for which $F(u)(m_i) = 0$, i.e., if it becomes 0 "later in the diagram". Last, multiplication by an element of A is defined in the obvious way. (You can now reinterpret Example 1.4.6 as a colimit of groups, not just of sets.)

1.4.D. EXERCISE. Verify that the A-module described above is indeed the colimit. (Make sure you verify that addition is well-defined, i.e., is independent of the choice of representatives m_i and m_j, the choice of ℓ, and the choice of arrows u and v. Similarly, make sure that scalar multiplication is well-defined.)

1.4.E. USEFUL EXERCISE (LOCALIZATION AS A COLIMIT). Generalize Exercise 1.4.B(a) to interpret localization of an integral domain as a colimit over a filtered set: suppose S is a multiplicative set of A, and interpret $S^{-1}A = \varinjlim \frac{1}{s}A$ where the limit is over $s \in S$, and in the category of A-modules. (Aside: Can you make some version of this work even if A isn't an integral domain, e.g. $S^{-1}A = \varinjlim A_s$? This will work in the category of A-algebras.)

A variant of this construction works without the filtered condition, if you have another means of "connecting elements in different objects of your diagram". For example:

1.4.F. EXERCISE: COLIMITS OF A-MODULES WITHOUT THE FILTERED CONDITION. Suppose you are given a diagram of A-modules indexed by $\mathscr{I}$: $F : \mathscr{I} \to Mod_A$, where we let $M_i := F(i)$. Show that the colimit is $\oplus_{i \in \mathscr{I}} M_i$ modulo the relations $m_i - F(n)(m_i)$ for every $n : i \to j$ in $\mathscr{I}$ (i.e., for every arrow in the diagram). (Somewhat more precisely: "modulo" means "quotiented by the submodule generated by".)

1.4.8. Summary. One useful thing to informally keep in mind is the following. In a category where the objects are "set-like", an element of a limit can be thought of as a family of elements of each object in the diagram, that are "compatible" (Exercise 1.4.A). And an element of a colimit can be thought of ("has a representative that is") an element of a single object in the diagram (Exercise 1.4.C). Even though the definitions of limit and colimit are the same, just with arrows reversed, these interpretations are quite different.

1.4.9. *Small remark.* In fact, colimits exist in the category of sets for all reasonable ("small") index categories (see for example [E, Thm. A6.1]), but that won't matter to us.

1.5 Adjoints

We next come to a very useful notion closely related to universal properties. Just as a universal property "essentially" (up to unique isomorphism) determines an object in a category (assuming such an object exists), "adjoints" essentially determine a functor (again, assuming it exists). Two *covariant* functors $F : \mathscr{A} \to \mathscr{B}$ and $G : \mathscr{B} \to \mathscr{A}$ are **adjoint** if there is a natural bijection for all $A \in \mathscr{A}$ and $B \in \mathscr{B}$

$$\tau_{AB} : \mathrm{Mor}_{\mathscr{B}}(F(A), B) \to \mathrm{Mor}_{\mathscr{A}}(A, G(B)). \tag{1.5.0.1}$$

We say that (F, G) form an **adjoint pair**, and that F is **left-adjoint** to G (and G is **right-adjoint** to F). We say F is a **left adjoint** (and G is a **right adjoint**). By "natural" we mean the following. For all $f : A \to A'$ in $\mathscr{A}$, we require

$$\begin{tikzcd} \mathrm{Mor}_{\mathscr{B}}(F(A'), B) \arrow[r, "Ff^*"] \arrow[d, "\tau_{A'B}"] & \mathrm{Mor}_{\mathscr{B}}(F(A), B) \arrow[d, "\tau_{AB}"] \\ \mathrm{Mor}_{\mathscr{A}}(A', G(B)) \arrow[r, "f^*"] & \mathrm{Mor}_{\mathscr{A}}(A, G(B)) \end{tikzcd} \tag{1.5.0.2}$$

to commute, and for all $g : B \to B'$ in $\mathscr{B}$ we want a similar commutative diagram to commute. (Here f^* is the map induced by $f : A \to A'$, and Ff^* is the map induced by $Ff : F(A) \to F(A')$.)

1.5.A. EXERCISE. Write down what this diagram should be.

1.5.B. EXERCISE. Show that the map τ_{AB} (1.5.0.1) has the following properties. For each A there is a map $\eta_A : A \to GF(A)$ so that for any $g : F(A) \to B$, the corresponding $\tau_{AB}(g) : A \to G(B)$ is given by the composition

$$A \xrightarrow{\eta_A} GF(A) \xrightarrow{Gg} G(B).$$

Similarly, there is a map $\epsilon_B : FG(B) \to B$ for each B so that for any $f : A \to G(B)$, the corresponding map $\tau_{AB}^{-1}(f) : F(A) \to B$ is given by the composition

$$F(A) \xrightarrow{Ff} FG(B) \xrightarrow{\epsilon_B} B.$$

Here is a key example of an adjoint pair.

1.5.C. EXERCISE. Suppose M, N, and P are A-modules (where A is a ring). Describe a bijection $\mathrm{Hom}_A(M \otimes_A N, P) \leftrightarrow \mathrm{Hom}_A(M, \mathrm{Hom}_A(N, P))$. (Hint: try to use the universal property of $\otimes$.)

1.5.D. EXERCISE. Show that $(\cdot) \otimes_A N$ and $\mathrm{Hom}_A(N, \cdot)$ are adjoint functors.

1.5.E. EXERCISE. Suppose $B \to A$ is a morphism of rings. If M is an A-module, you can create a B-module M_B by considering it as a B-module. This gives a functor $\cdot_B : Mod_A \to Mod_B$. Show that this functor is right-adjoint to $\cdot \otimes_B A$. In other words, describe a bijection

$$\mathrm{Hom}_A(N \otimes_B A, M) \cong \mathrm{Hom}_B(N, M_B)$$

functorial in both arguments. (This adjoint pair is very important, and is the key player in Chapter 16.)

1.5.1. ⋆ *Fancier remarks we won't use.* You can check that the left adjoint determines the right adjoint up to natural isomorphism, and vice versa. The maps η_A and ϵ_B of Exercise 1.5.B are called the **unit** and **counit** of the adjunction. This leads to a different characterization of adjunction. Suppose functors $F : \mathscr{A} \to \mathscr{B}$ and $G : \mathscr{B} \to \mathscr{A}$ are given, along with natural transformations $\eta : \mathrm{id}_{\mathscr{A}} \to GF$ and $\epsilon : FG \to \mathrm{id}_{\mathscr{B}}$ with the property that $G\epsilon \circ \eta G = \mathrm{id}_G$ (for each $B \in \mathscr{B}$, the composition of $\eta_{G(B)} : G(B) \to GFG(B)$ and $G(\epsilon_B) : GFG(B) \to G(B)$ is the identity) and $\epsilon F \circ F\eta = \mathrm{id}_F$. Then you can check that F is left-adjoint to G. These facts aren't hard to check, so if you want to use them, you should verify everything for yourself.

1.5.2. *Examples from other fields.* For those familiar with representation theory: Frobenius reciprocity may be understood in terms of adjoints. Suppose V is a finite-dimensional representation of a finite group G, and W is a representation of a subgroup $H < G$. Then induction and restriction are an adjoint pair $(\mathrm{Ind}_H^G, \mathrm{Res}_H^G)$ between the category of G-modules and the category of H-modules.

Topologists' favorite adjoint pair may be the suspension functor and the loop space functor.

1.5.3. Example: groupification of abelian semigroups. Here is another motivating example: getting an abelian group from an abelian semigroup. (An **abelian semigroup** is just like an abelian group, except you don't require an identity or an inverse. Morphisms of abelian semigroups are maps of sets preserving the binary operation. One example is the non-negative integers $\mathbb{Z}^{\geq 0} = \{0, 1, 2, \dots\}$ under addition. Another is the positive integers $1, 2, \dots$ under multiplication. You may enjoy groupifying both.) From an abelian semigroup, you can create an abelian group. Here is a formalization of that notion. A **groupification** of a semigroup S is a map of abelian semigroups $\pi : S \to G$ such that G is an abelian group, and any map of abelian semigroups from S to an abelian *group* G' factors *uniquely* through G:

$$\begin{tikzcd} S \arrow[r, "\pi"] \arrow[dr] & G \arrow[d, dotted, "\exists!"] \\ & G' \end{tikzcd}$$

(Perhaps "abelian groupification" is better than "groupification".)

1.5.F. EXERCISE (A GROUP IS GROUPIFIED BY ITSELF). Show that if a semigroup is *already* a group then the identity morphism is the groupification. (More correct: the identity morphism is *a* groupification.) Note that you don't need to construct groupification (or even know that it exists in general) to solve this exercise.

1.5.G. EXERCISE. Construct groupification H from the category of nonempty abelian semigroups to the category of abelian groups. (One possible construction: given an abelian semigroup S, the elements of its groupification $H(S)$ are ordered pairs $(a, b) \in S \times S$, which you may think of as $a - b$, with the equivalence that $(a, b) \sim (c, d)$ if $a + d + e = b + c + e$ for some $e \in S$. Describe addition in this group, and show that it satisfies the properties of an abelian group. Describe the semigroup map $S \to H(S)$.) Let F be the forgetful functor from the category of abelian groups *Ab* to the category of abelian semigroups. Show that H is left-adjoint to F.

(Here is the general idea for experts: We have a full subcategory of a category. We want to "project" from the category to the subcategory. We have

$$\mathrm{Mor}_{\text{category}}(S, H) = \mathrm{Mor}_{\text{subcategory}}(G, H)$$

automatically; thus we are describing the left adjoint to the forgetful functor. How the argument worked: we constructed something which was in the smaller category, which automatically satisfies the universal property.)

1.5.H. EXERCISE (CF. EXERCISE 1.5.E). The purpose of this exercise is to give you more practice with "adjoints of forgetful functors", the means by which we get groups from semigroups, and sheaves from presheaves. Suppose A is a ring, and S is a multiplicative subset. Then $S^{-1}A$-modules are a fully faithful subcategory (§1.2.15) of the category of A-modules (via the obvious inclusion $Mod_{S^{-1}A} \hookrightarrow Mod_A$). Then $Mod_A \to Mod_{S^{-1}A}$ can be interpreted as an adjoint to the forgetful functor $Mod_{S^{-1}A} \to Mod_A$. State and prove the correct statements.

(Here is the larger story. Every $S^{-1}A$-module is an A-module, and this is an injective map, so we have a covariant forgetful functor $F : Mod_{S^{-1}A} \to Mod_A$. In fact this is a fully faithful functor: it is injective on objects, and the morphisms between any two $S^{-1}A$-modules *as* A*-modules* are just the same when they are considered as $S^{-1}A$-modules. Then there is a functor $G : Mod_A \to Mod_{S^{-1}A}$, which might reasonably be called "localization with respect to S", which is left-adjoint to the forgetful functor. Translation: If M is an A-module, and N is an $S^{-1}A$-module, then $\mathrm{Mor}(GM, N)$ (morphisms as $S^{-1}A$-modules, which are the same as morphisms as A-modules) are in natural bijection with $\mathrm{Mor}(M, FN)$ (morphisms as A-modules).)

Here is a table of adjoints that will come up for us.

situation	category $\mathscr{A}$	category $\mathscr{B}$	left adjoint $F : \mathscr{A} \to \mathscr{B}$	right adjoint $G : \mathscr{B} \to \mathscr{A}$
A-modules (Ex. 1.5.D)			$(\cdot) \otimes_A N$	$\mathrm{Hom}_A(N, \cdot)$
ring maps $B \to A$ (Ex. 1.5.E)	Mod_B	Mod_A	$(\cdot) \otimes_B A$ (extension of scalars)	$M \mapsto M_B$ (restriction of scalars)
(pre)sheaves on a topological space X (Ex. 2.4.L)	presheaves on X	sheaves on X	sheafification	forgetful
(semi)groups (§1.5.3)	semigroups	groups	groupification	forgetful
sheaves, $\pi : X \to Y$ (Ex. 2.6.B)	sheaves on Y	sheaves on X	π^{-1}	π_*
sheaves of abelian groups or $\mathscr{O}$-modules, open embeddings $\pi : U \hookrightarrow Y$ (Ex. 2.6.G)	sheaves on U	sheaves on Y	$\pi_!$	π^{-1}
quasicoherent sheaves, $\pi : X \to Y$ (Prop. 16.3.6)	$QCoh_Y$	$QCoh_X$	π^*	π_*
ring maps $B \to A$ (Ex. 30.3.A)	Mod_A	Mod_B	$M \mapsto M_B$ (restriction of scalars)	$N \mapsto \mathrm{Hom}_B(A, N)$
quasicoherent sheaves, affine $\pi : X \to Y$ (Ex. 30.3.B(b))	$QCoh_X$	$QCoh_Y$	π_*	$\pi^!_{sh}$

Other examples will also come up, such as the adjoint pair $(\sim, \Gamma_\bullet)$ between graded modules over a graded ring, and quasicoherent sheaves on the corresponding projective scheme (§15.4).

1.5.4. *Useful comment for experts.* One last comment only for people who have seen adjoints before: If (F, G) is an adjoint pair of functors, then F commutes with colimits, and G commutes with limits. Also, limits commute with limits and colimits commute with colimits. We will prove these facts (and a little more) in §1.6.12.

1.6 An introduction to abelian categories

Ton papier sur l'Algèbre homologique a été lu soigneusement, et a converti tout le monde (même Dieudonné, qui semble complètement fonctorisé!) à ton point de vue.

Your paper on homological algebra was read carefully and converted everyone (even Dieudonné, who seems to be completely functorised!) to your point of view.

— *J.-P. Serre, letter to A. Grothendieck* [**GrS**, p. 17-18]

Since learning linear algebra, you have been familiar with the notions and behaviors of kernels, cokernels, etc. Later in your life you saw them in the category of abelian groups, and later still in the category of A-modules. Each of these notions generalizes the previous one.

We will soon define some new categories (certain sheaves) that will have familiar-looking behavior, reminiscent of that of modules over a ring. The notions of kernels, cokernels, images, and more will make sense, and they will behave "the way we expect" from our experience with modules. This can be made precise through the notion of an *abelian category*. Abelian categories are the right general setting in which one can do "homological algebra", in which notions of kernel, cokernel, and so on are used, and one can work with complexes and exact sequences.

We will see enough to motivate the definitions that we will see in general: monomorphism (and subobject), epimorphism, kernel, cokernel, and image. But in this book we will avoid having to show that they behave "the way we expect" in a general abelian category because the examples we will see are directly interpretable in terms of modules over rings. In particular, it is not worth memorizing the definition of abelian category.

Two central examples of an abelian category are the category *Ab* of abelian groups, and the category Mod_A of A-modules. The first is a special case of the second (just take $A = \mathbb{Z}$). As we give the definitions, you should verify that Mod_A is an abelian category.

We first define the notion of *additive category*. We will use it only as a stepping stone to the notion of an abelian category. Two examples you can keep in mind while reading the definition: the category of free A-modules (where A is a ring), and real (or complex) Banach spaces.

1.6.1. Definition. A category $\mathscr{C}$ is said to be **additive** if it satisfies the following properties.

- Ad1. For each $A, B \in \mathscr{C}$, $\mathrm{Mor}(A, B)$ is an abelian group, such that composition of morphisms distributes over addition. (You should think about what this means — it translates to two distinct statements).
- Ad2. $\mathscr{C}$ has a zero object, denoted 0. (This is an object that is simultaneously an initial object and a final object, Definition 1.3.2.)
- Ad3. It has products of two objects (a product $A \times B$ for any pair of objects), and hence by induction, products of any finite number of objects.

In an additive category, the morphisms are often called **homomorphisms**, and Mor is denoted by Hom. In fact, this notation Hom is a good indication that you're working in an additive category. A functor between additive categories preserving the additive structure of Hom, is called an **additive functor**.

1.6.2. *Remarks.* It is a consequence of the definition of additive category that finite direct products are also finite direct sums (coproducts) — the details don't matter to us. The symbol $\oplus$ is used for this notion. Also, it is quick to show that additive functors send zero objects to zero objects (show that Z is a 0-object if and only if $\mathrm{id}_Z = 0_Z$; additive functors preserve both id and 0), and preserve products.

One motivation for the name 0-object is that the 0-morphism in the abelian group $\mathrm{Hom}(A, B)$ is the composition $A \to 0 \to B$. (We also remark that the notion of 0-morphism thus makes sense in any category with a 0-object.)

The category of A-modules Mod_A is clearly an additive category, but it has even more structure, which we now formalize as an example of an abelian category.

1.6.3. Definition. Let $\mathscr{C}$ be a category with a 0-object (and thus 0-morphisms). A **kernel** of a morphism $f : B \to C$ is a map $i : A \to B$ such that $f \circ i = 0$, and that is universal with respect to this property. Diagramatically:

$$\begin{tikzcd} Z \arrow[d, dotted, "\exists!"'] \arrow[dr] \arrow[drr, "0"] & & \\ A \arrow[r, "i"] \arrow[rr, bend right, "0"'] & B \arrow[r, "f"] & C \end{tikzcd}$$

(Note that the kernel is not just an object; it is a morphism of an object to B.) Hence it is unique up to unique isomorphism by universal property nonsense. The kernel is written $\ker f \to B$. A **cokernel** (denoted $\operatorname{coker} f$) is defined dually by reversing the arrows — do this yourself. The kernel of $f : B \to C$ is the limit (§1.4) of the diagram

(1.6.3.1)
$$\begin{tikzcd} & 0 \arrow[d] \\ B \arrow[r, "f"] & C \end{tikzcd}$$

and similarly the cokernel is a colimit (see (2.5.0.2)).

If $i : A \to B$ is a monomorphism, then we say that A is a **subobject** of B, where the map i is implicit. Dually, there is the notion of **quotient object**, defined dually to subobject.

An **abelian category** is an additive category satisfying three additional properties.

(1) Every map has a kernel and cokernel.
(2) Every monomorphism is the kernel of its cokernel.
(3) Every epimorphism is the cokernel of its kernel.

It is a nonobvious (and imprecisely stated) fact that every property you want to be true about kernels, cokernels, etc. follows from these three. (Warning: in part of the literature, additional hypotheses are imposed as part of the definition.)

The **image** of a morphism $f : A \to B$ is defined as $\operatorname{im}(f) = \ker(\operatorname{coker} f)$ whenever it exists (e.g. in every abelian category). The morphism $f : A \to B$ factors uniquely through $\operatorname{im} f \to B$ whenever $\operatorname{im} f$ exists, and $A \to \operatorname{im} f$ is an epimorphism and a cokernel of $\ker f \to A$ in every abelian category. The reader may want to verify this as a (hard!) exercise.

The cokernel of a monomorphism is called the **quotient**. The quotient of a monomorphism $A \to B$ is often denoted B/A (with the map from B implicit).

We will leave the foundations of abelian categories untouched. The key thing to remember is that if you understand kernels, cokernels, images and so on in the category of modules over a given ring, you can manipulate objects in any abelian category. This is made precise by Freyd-Mitchell Embedding Theorem (Remark 1.6.4).

However, the abelian categories we will come across will obviously be related to modules, and our intuition will clearly carry over, so we needn't invoke a theorem whose proof we haven't read. For example, we will show that sheaves of abelian groups on a topological space X form an abelian category (§2.5), and the interpretation in terms of "compatible germs" will connect notions of kernels, cokernels etc. of sheaves of abelian groups to the corresponding notions of abelian groups.

1.6.4. *Small remark on chasing diagrams.* It is useful to prove facts (and solve exercises) about abelian categories by chasing elements. This can be justified by the Freyd-Mitchell Embedding Theorem: If $\mathscr{C}$ is an abelian category such that $\mathrm{Hom}(X, Y)$ is a set for all $X, Y \in \mathscr{C}$, then there is a ring A and an exact, fully faithful functor from $\mathscr{C}$ into Mod_A, which embeds $\mathscr{C}$ as a full subcategory. (Unfortunately, the ring A need not be commutative.) A proof is sketched in [**Wei**, §1.6], and references to a complete proof are given there. A proof is also given in [**KS1**, §9.7]. The upshot is that to prove something about a diagram in some abelian category, we may assume that it is a diagram of modules over some ring, and we may then "diagram-chase" elements. Moreover, any fact about kernels, cokernels, and so on that holds in Mod_A holds in any abelian category.)

If invoking a theorem whose proof you haven't read bothers you, a short alternative is Mac Lane's "elementary rules for chasing diagrams", [**Mac**, Thm. 3, p. 200]; [**Mac**, Lem. 4, p. 201] gives a proof of the Five Lemma (Exercise 1.7.6) as an example.

But in any case, do what you need to do to put your mind at ease, so you can move forward. Do as little as your conscience will allow.

1.6.5. Complexes, exactness, and homology.

(In this entire discussion, we assume we are working in an abelian category.) We say a sequence

$$\cdots \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow \cdots \tag{1.6.5.1}$$

is a **complex at** B if $g \circ f = 0$, and is **exact at** B if $\ker g = \operatorname{im} f$. (More specifically, g has a kernel that is an image of f. Exactness at B implies being a complex at B — do you see why?) A sequence is a **complex** (resp. **exact**) if it is a complex (resp. exact) at each (internal) term. A **short exact sequence** is an exact sequence with five terms, the first and last of which are zeroes — in other words, an exact sequence of the form

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0.$$

For example, $0 \longrightarrow A \longrightarrow 0$ is exact if and only if $A = 0$;

$$0 \longrightarrow A \xrightarrow{f} B$$

is exact if and only if f is a monomorphism (with a similar statement for $A \xrightarrow{f} B \longrightarrow 0$);

$$0 \longrightarrow A \xrightarrow{f} B \longrightarrow 0$$

is exact if and only if f is an isomorphism; and

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C$$

is exact if and only if f is a kernel of g (with a similar statement for $A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$). To show some of these facts it may be helpful to prove that (1.6.5.1) is exact at B if and only if the cokernel of f is a cokernel of the kernel of g.

If you would like practice in playing with these notions before thinking about homology, you can prove the Snake Lemma (stated in Example 1.7.5, with a stronger version in Exercise 1.7.B), or the Five Lemma (stated in Example 1.7.6, with a stronger version in Exercise 1.7.C). (I would do this in the category of A-modules, but see [**KS1**, Lem. 12.1.1, Lem. 8.3.13] for proofs in general.)

If (1.6.5.1) is a complex at B, then its **homology at** B (often denoted by H) is $\ker g / \operatorname{im} f$. (More precisely, there is some monomorphism $\operatorname{im} f \hookrightarrow \ker g$, and that H is the cokernel of this monomorphism.) Therefore, (1.6.5.1) is exact at B if and only if its homology at B is 0. We say that elements of $\ker g$ (assuming the objects of the category are sets with some additional structure) are the **cycles**, and elements of $\operatorname{im} f$ are the **boundaries** (so homology is "cycles mod boundaries"). If the complex is indexed in decreasing order, the indices are often written as subscripts, and H_i is the homology at $A_{i+1} \to A_i \to A_{i-1}$. If the complex is indexed in increasing order, the indices are often written as superscripts, and the homology H^i at $A^{i-1} \to A^i \to A^{i+1}$ is often called **cohomology**.

An exact sequence

$$A^\bullet : \qquad \cdots \longrightarrow A^{i-1} \xrightarrow{f^{i-1}} A^i \xrightarrow{f^i} A^{i+1} \xrightarrow{f^{i+1}} \cdots \tag{1.6.5.2}$$

can be "factored" into short exact sequences

$$0 \longrightarrow \ker f^i \longrightarrow A^i \longrightarrow \ker f^{i+1} \longrightarrow 0$$

which is helpful in proving facts about long exact sequences by reducing them to facts about short exact sequences.

More generally, if (1.6.5.2) is assumed only to be a complex, then it can be "factored" into short exact sequences.

$$0 \longrightarrow \ker f^i \longrightarrow A^i \longrightarrow \operatorname{im} f^i \longrightarrow 0 \tag{1.6.5.3}$$

$$0 \longrightarrow \operatorname{im} f^{i-1} \longrightarrow \ker f^i \longrightarrow H^i(A^\bullet) \longrightarrow 0$$

1.6.A. EXERCISE. Describe exact sequences

$$0 \longrightarrow \operatorname{im} f^i \longrightarrow A^{i+1} \longrightarrow \operatorname{coker} f^i \longrightarrow 0 \tag{1.6.5.4}$$

$$0 \longrightarrow H^i(A^\bullet) \longrightarrow \operatorname{coker} f^{i-1} \longrightarrow \operatorname{im} f^i \longrightarrow 0$$

(These are somehow dual to (1.6.5.3). In fact in some mirror universe this might have been given as the standard definition of homology.) Assume the category is that of modules over a fixed ring for convenience, but be aware that the result is true for any abelian category.

1.6.B. Exercise and important definition. Suppose

$$0 \xrightarrow{d^0} A^1 \xrightarrow{d^1} \cdots \xrightarrow{d^{n-1}} A^n \xrightarrow{d^n} 0$$

is a complex of finite-dimensional k-vector spaces (often called $A^\bullet$ for short). Define $h^i(A^\bullet) := \dim H^i(A^\bullet)$. Show that $\sum(-1)^i \dim A^i = \sum(-1)^i h^i(A^\bullet)$. In particular, if $A^\bullet$ is exact, then $\sum(-1)^i \dim A^i = 0$. (If you haven't dealt much with cohomology, this will give you some practice.)

1.6.C. Important exercise. Suppose $\mathscr{C}$ is an abelian category. Define the category $Com_{\mathscr{C}}$ as follows. The objects are infinite complexes

$$A^\bullet : \qquad \cdots \longrightarrow A^{i-1} \xrightarrow{f^{i-1}} A^i \xrightarrow{f^i} A^{i+1} \xrightarrow{f^{i+1}} \cdots$$

in $\mathscr{C}$, and the morphisms $A^\bullet \to B^\bullet$ are commuting diagrams

$$\begin{tikzcd}
\cdots \arrow[r] & A^{i-1} \arrow[r, "f^{i-1}"] \arrow[d] & A^i \arrow[r, "f^i"] \arrow[d] & A^{i+1} \arrow[r, "f^{i+1}"] \arrow[d] & \cdots \\
\cdots \arrow[r] & B^{i-1} \arrow[r, "g^{i-1}"] & B^i \arrow[r, "g^i"] & B^{i+1} \arrow[r, "g^{i+1}"] & \cdots
\end{tikzcd} \tag{1.6.5.5}$$

Show that $Com_{\mathscr{C}}$ is an abelian category. Feel free to deal with the special case of modules over a fixed ring. (Remark for experts: Essentially the same argument shows that the functor category $\mathscr{C}^{\mathscr{I}}$ is an abelian category for any small category $\mathscr{I}$ and any abelian category $\mathscr{C}$. This immediately implies that the category of presheaves on a topological space X with values in an abelian category $\mathscr{C}$ is automatically an abelian category, cf. §2.3.4.)

1.6.D. Important Exercise. Show that (1.6.5.5) induces a map of homology $H^i(A^\bullet) \to H^i(B^\bullet)$. Show furthermore that H^i is a covariant functor $Com_{\mathscr{C}} \to \mathscr{C}$. (Again, feel free to deal with the special case Mod_A.)

We will later define when two maps of complexes are *homotopic* (§23.1), and show that homotopic maps induce isomorphisms on cohomology (Exercise 23.1.A), but we won't need that any time soon.

1.6.6. Theorem (Long exact sequence). — *A short exact sequence of complexes*

$$\begin{tikzcd}
0^\bullet \arrow[d] & : & \cdots \arrow[r] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & \cdots \\
A^\bullet \arrow[d] & : & \cdots \arrow[r] & A^{i-1} \arrow[r, "f^{i-1}"] \arrow[d] & A^i \arrow[r, "f^i"] \arrow[d] & A^{i+1} \arrow[r, "f^{i+1}"] \arrow[d] & \cdots \\
B^\bullet \arrow[d] & : & \cdots \arrow[r] & B^{i-1} \arrow[r, "g^{i-1}"] \arrow[d] & B^i \arrow[r, "g^i"] \arrow[d] & B^{i+1} \arrow[r, "g^{i+1}"] \arrow[d] & \cdots \\
C^\bullet \arrow[d] & : & \cdots \arrow[r] & C^{i-1} \arrow[r, "h^{i-1}"] \arrow[d] & C^i \arrow[r, "h^i"] \arrow[d] & C^{i+1} \arrow[r, "h^{i+1}"] \arrow[d] & \cdots \\
0^\bullet & : & \cdots \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & \cdots
\end{tikzcd}$$

induces a **long exact sequence in cohomology**

$$\begin{array}{l} \qquad\qquad\qquad \dots \longrightarrow H^{i-1}(C^\bullet) \longrightarrow \\ H^i(A^\bullet) \longrightarrow H^i(B^\bullet) \longrightarrow H^i(C^\bullet) \longrightarrow \\ H^{i+1}(A^\bullet) \longrightarrow \cdots \end{array}$$

(This requires a definition of the *connecting homomorphism* $H^{i-1}(C^\bullet) \to H^i(A^\bullet)$, which is natural in an appropriate sense.) In the category of modules over a ring, Theorem 1.6.6 will come out of our discussion of spectral sequences, see Exercise 1.7.F, but this is a somewhat perverse way of proving it. For a proof in general, see [**KS1**, Theorem 12.3.3]. You may want to prove it yourself, by first proving a weaker version of the Snake Lemma (Example 1.7.5), where in the hypotheses (1.7.5.1), the 0's in the bottom left and top right are removed, and in the conclusion (1.7.5.2), the first and last 0's are removed.

1.6.7. *Exactness of functors.* If $F : \mathscr{A} \to \mathscr{B}$ is a covariant additive functor from one abelian category to another, we say that F is **right-exact** if the exactness of

$$A' \longrightarrow A \longrightarrow A'' \longrightarrow 0,$$

in $\mathscr{A}$ implies that

$$F(A') \longrightarrow F(A) \longrightarrow F(A'') \longrightarrow 0$$

is also exact. Dually, we say that F is **left-exact** if the exactness of

$$0 \longrightarrow A' \longrightarrow A \longrightarrow A'' \qquad \text{implies}$$

$$0 \longrightarrow F(A') \longrightarrow F(A) \longrightarrow F(A'') \qquad \text{is exact.}$$

A contravariant functor is **left-exact** if the exactness of

$$A' \longrightarrow A \longrightarrow A'' \longrightarrow 0 \qquad \text{implies}$$

$$0 \longrightarrow F(A'') \longrightarrow F(A) \longrightarrow F(A') \qquad \text{is exact.}$$

The reader should be able to deduce what it means for a contravariant functor to be **right-exact**.

A covariant or contravariant functor is **exact** if it is both left-exact and right-exact.

1.6.E. Exercise. Suppose F is an exact functor. Show that applying F to an exact sequence preserves exactness. For example, if F is covariant, and $A' \to A \to A''$ is exact, then $FA' \to FA \to FA''$ is exact. (This will be generalized in Exercise 1.6.H(c).)

1.6.F. EXERCISE. Suppose A is a ring, $S \subset A$ is a multiplicative subset, and M is an A-module.
(a) Show that localization of A-modules $Mod_A \to Mod_{S^{-1}A}$ is an exact covariant functor.
(b) Show that $(\cdot) \otimes_A M$ is a right-exact covariant functor $Mod_A \to Mod_A$. (This is a repeat of Exercise 1.3.H.)
(c) Show that $\operatorname{Hom}(M, \cdot)$ is a left-exact covariant functor $Mod_A \to Mod_A$. If $\mathscr{C}$ is any abelian category, and $C \in \mathscr{C}$, show that $\operatorname{Hom}(C, \cdot)$ is a left-exact covariant functor $\mathscr{C} \to Ab$.
(d) Show that $\operatorname{Hom}(\cdot, M)$ is a left-exact contravariant functor $Mod_A \to Mod_A$. If $\mathscr{C}$ is any abelian category, and $C \in \mathscr{C}$, show that $\operatorname{Hom}(\cdot, C)$ is a left-exact contravariant functor $\mathscr{C} \to Ab$.

1.6.G. EXERCISE. Suppose M is a **finitely presented** A**-module**: M has a finite number of generators, and with these generators it has a finite number of relations; or usefully equivalently, fits in an exact sequence

$$A^{\oplus q} \to A^{\oplus p} \to M \to 0 \tag{1.6.7.1}$$

Use (1.6.7.1) and the left-exactness of Hom to describe an isomorphism

$$S^{-1} \operatorname{Hom}_A(M, N) \cong \operatorname{Hom}_{S^{-1}A}(S^{-1}M, S^{-1}N).$$

(You might be able to interpret this in light of a variant of Exercise 1.6.H below, for left-exact contravariant functors rather than right-exact covariant functors.)

1.6.8. *Example:* Hom *doesn't always commute with localization.* In the language of Exercise 1.6.G, take $A = N = \mathbb{Z}$, $M = \mathbb{Q}$, and $S = \mathbb{Z} \setminus \{0\}$.

1.6.9. $\star$ **Two useful facts in homological algebra.**

We now come to two (sets of) facts I wish I had learned as a child, as they would have saved me lots of grief. They encapsulate what is best and worst of abstract nonsense. The statements are so general as to be nonintuitive. The proofs are very short. They generalize some specific behavior it is easy to prove on an ad hoc basis. Once they are second nature to you, many subtle facts will become obvious to you as special cases. And you will see that they will get used (implicitly or explicitly) repeatedly.

1.6.10. $\star$ *Interaction of homology and (right/left-)exact functors.*

You might wait to prove this until you learn about cohomology in Chapter 18, when it will first be used in a serious way.

1.6.H. IMPORTANT EXERCISE (THE FHHF THEOREM). This result can take you far, and perhaps for that reason it has sometimes been called the Fernbahnhof (FernbaHnHoF) Theorem, notably in [**Vak**, Exer. 1.6.H]. ("From Here Hop Far"?) Suppose $F : \mathscr{A} \to \mathscr{B}$ is a covariant functor of abelian categories, and $C^\bullet$ is a complex in $\mathscr{A}$.

(a) (F *right-exact yields* $FH^\bullet \longrightarrow H^\bullet F$) If F is right-exact, describe a natural morphism $FH^\bullet \to H^\bullet F$. (More precisely, for each i, the left side is F applied to the cohomology at piece i of $C^\bullet$, while the right side is the cohomology at piece i of $FC^\bullet$.)

(b) *(F left-exact yields $FH^\bullet \longleftarrow H^\bullet F$)* If F is left-exact, describe a natural morphism $H^\bullet F \to FH^\bullet$.

(c) *(F exact yields $FH^\bullet \longleftrightarrow H^\bullet F$)* If F is exact, show that the morphisms of (a) and (b) are inverses and thus isomorphisms.

Hint for (a): use $C^i \xrightarrow{d^i} C^{i+1} \longrightarrow \operatorname{coker} d^i \longrightarrow 0$ to give an isomorphism $F \operatorname{coker} d^i \cong \operatorname{coker} Fd^i$. Then use the first line of (1.6.5.4) to give a epimorphism $F \operatorname{im} d^i \twoheadrightarrow \operatorname{im} Fd^i$. Then use the second line of (1.6.5.4) to give the desired map $FH^iC^\bullet \longrightarrow H^iFC^\bullet$. While you are at it, you may as well describe a map for the fourth member of the quartet $\{\operatorname{coker}, \operatorname{im}, H, \ker\}$: $F \ker d^i \longrightarrow \ker Fd^i$.

1.6.11. If this makes your head spin, you may prefer to think of it in the following specific case, where both $\mathscr{A}$ and $\mathscr{B}$ are the category of A-modules, and F is $(\cdot) \otimes N$ for some fixed N-module. Your argument in this case will translate without change to yield a solution to Exercise 1.6.H(a) and (c) in general. If $\otimes N$ is exact, then N is called a **flat** A-module. (The notion of flatness will turn out to be very important, and is discussed in detail in Chapter 24.)

For example, localization is exact (Exercise 1.6.F(a)), so $S^{-1}A$ is a *flat* A-algebra for all multiplicative sets S. Thus taking cohomology of a complex of A-modules commutes with localization — something you could verify directly.

1.6.12. *Interaction of adjoints, (co)limits, and (left- and right-) exactness.*

A surprising number of arguments boil down to the statement:

Limits commute with limits and right adjoints. In particular, in an abelian category, because kernels are limits, both right adjoints and limits are left-exact.

as well as its dual:

Colimits commute with colimits and left adjoints. In particular, because cokernels are colimits, both left adjoints and colimits are right-exact.

These statements were promised in §1.5.4, and will be proved below. The latter has a useful extension:

In Mod_A, colimits over filtered index categories are exact. "Filtered" was defined in §1.4.7. (Caution: it is not true that in abelian categories in general, colimits over filtered index categories are exact. ⋆⋆ Here is a counterexample. Because the axioms of abelian categories are self-dual, it suffices to give an example in which a filtered *limit* fails to be exact, and we do this. Fix a prime p. In the category *Ab* of abelian groups, for each positive integer n, we have an exact sequence $\mathbb{Z} \to \mathbb{Z}/(p^n) \to 0$. Taking the limit over all n in the obvious way, we obtain $\mathbb{Z} \to \mathbb{Z}_p \to 0$, which is certainly not exact.)

If you want to use these statements (for example, later in this book), you will have to prove them. Let's now make them precise.

1.6.I. Exercise (kernels commute with limits). Suppose $\mathscr{C}$ is an abelian category, and $a : \mathscr{I} \to \mathscr{C}$ and $b : \mathscr{I} \to \mathscr{C}$ are two diagrams in $\mathscr{C}$ indexed by $\mathscr{I}$. For convenience, let $A_i = a(i)$ and $B_i = b(i)$ be the objects in those two diagrams. Let $h_i : A_i \to B_i$ be maps commuting with the maps in the diagram. (Translation: h is a natural transformation of functors $a \to b$, see §1.2.21.) Then the $\ker h_i$ form another diagram in $\mathscr{C}$ indexed by $\mathscr{I}$. Describe a canonical isomorphism $\varprojlim \ker h_i \cong \ker(\varprojlim A_i \to \varprojlim B_i)$, assuming the limits exist.

Implicit in the previous exercise is the idea that limits should somehow be understood as functors.

1.6.J. EXERCISE. Make sense of the statement that "limits commute with limits" in a general category, and prove it. (Hint: recall that kernels are limits. The previous exercise should be a corollary of this one.)

1.6.13. Proposition (right adjoints commute with limits). — *Suppose* $(F : \mathscr{C} \to \mathscr{D}, G : \mathscr{D} \to \mathscr{C})$ *is a pair of adjoint functors. If* $A = \varprojlim A_i$ *is a limit in* $\mathscr{D}$ *of a diagram indexed by* $\mathscr{I}$*, then* $GA = \varprojlim GA_i$ *(with the corresponding maps* $GA \to GA_i$*) is a limit in* $\mathscr{C}$*.*

Proof. We must show that $GA \to GA_i$ satisfies the universal property of limits. Suppose we have maps $W \to GA_i$ commuting with the maps of $\mathscr{I}$. We wish to show that there exists a unique $W \to GA$ extending the $W \to GA_i$. By adjointness of F and G, we can restate this as: Suppose we have maps $FW \to A_i$ commuting with the maps of $\mathscr{I}$. We wish to show that there exists a unique $FW \to A$ extending the $FW \to A_i$. But this is precisely the universal property of the limit. □

Of course, the dual statements to Exercise 1.6.J and Proposition 1.6.13 hold by the dual arguments.

If F and G are additive functors between abelian categories, and (F, G) is an adjoint pair, then (as kernels are limits and cokernels are colimits) G is left-exact and F is right-exact.

1.6.K. EXERCISE. Show that in Mod_A, colimits over filtered index categories are exact. (Your argument will apply without change to any abelian category whose objects can be interpreted as "sets with additional structure".) Right-exactness follows from the above discussion, so the issue is left-exactness. (Possible hint: After you show that localization is exact, Exercise 1.6.F(a), or sheafification is exact, Exercise 2.5.D, in a hands-on way, you will be easily able to prove this. Conversely, if you do this exercise, those two will be easy.)

1.6.L. EXERCISE. Show that filtered colimits commute with homology in Mod_A. Hint: use the FHHF Theorem (Exercise 1.6.H), and the previous Exercise.

In light of Exercise 1.6.L, you may want to think about how limits (and colimits) commute with homology in general, and which way maps go. The statement of the FHHF Theorem should suggest the answer. (Are limits analogous to left-exact functors, or right-exact functors?) We won't directly use this insight.

Just as colimits are exact (not just right-exact) in especially good circumstances, limits are exact (not just left-exact) too. The following will be used twice in Chapter 29.

1.6.M. EXERCISE. Suppose

$$\begin{tikzcd}
& \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
0 \arrow[r] & A_{n+1} \arrow[r] \arrow[d] & B_{n+1} \arrow[r] \arrow[d] & C_{n+1} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A_n \arrow[r] \arrow[d] & B_n \arrow[r] \arrow[d] & C_n \arrow[r] \arrow[d] & 0 \\
& \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
0 \arrow[r] & A_0 \arrow[r] \arrow[d] & B_0 \arrow[r] \arrow[d] & C_0 \arrow[r] \arrow[d] & 0 \\
& 0 & 0 & 0 &
\end{tikzcd}$$

is an inverse system of exact sequences of modules over a ring, such that the maps $A_{n+1} \to A_n$ are surjective. (We say: "transition maps of the left term are surjective".) Show that the limit

$$0 \longrightarrow \varprojlim A_n \longrightarrow \varprojlim B_n \longrightarrow \varprojlim C_n \longrightarrow 0 \qquad (1.6.13.1)$$

is also exact. (You will need to define the maps in (1.6.13.1).)

1.6.14. *Unimportant remark.* Based on these ideas, you may suspect that right-exact functors always commute with colimits. The fact that tensor product commutes with infinite direct sums (Exercise 1.3.M) may reinforce this idea. Unfortunately, it is not true, see [**MO93716**].

1.6.15. ⋆ **Dreaming of derived functors.** When you see a left-exact functor, you should always dream that you are seeing the end of a long exact sequence. If

$$0 \to M' \to M \to M'' \to 0$$

is an exact sequence in abelian category $\mathscr{A}$, and $F : \mathscr{A} \to \mathscr{B}$ is a left-exact functor, then

$$0 \to FM' \to FM \to FM''$$

is exact, and you should always dream that it should continue in some natural way. For example, the next term should depend only on M', call it R^1FM', and if it is zero, then $FM \to FM''$ is an epimorphism. This remark holds true for left-exact and contravariant functors too. In good cases, such a continuation exists, and is incredibly useful. We will discuss this in Chapter 23.

1.7 ⋆ Spectral sequences

Spectral sequences are a powerful book-keeping tool for proving things involving complicated commutative diagrams. They were introduced by Leray in the 1940's at the same time as he introduced sheaves. They have a reputation for being abstruse and difficult. It has been suggested that the name 'spectral' was given because, like spectres, spectral sequences are terrifying, evil, and dangerous. I have heard no one disagree with this interpretation, which is perhaps not surprising since I just made it up.

Nonetheless, the goal of this section is to tell you enough that you can use spectral sequences without hesitation or fear, and why you shouldn't be frightened when they come up in a seminar. What is perhaps different in this presentation is that we will use spectral sequences to prove things that you may have already seen, and that you can prove easily in other ways. This will allow you to get some hands-on experience for how to use them. We will also see them only in the special case of double complexes (the version by far the most often used in algebraic geometry), and not in the general form usually presented (filtered complexes, exact couples, etc.). See [**Wei**, Ch. 5] for more detailed information if you wish.

You should *not* read this section when you are reading the rest of Chapter 1. Instead, you should read it just before you need it for the first time. When you finally *do* read this section, you *must* do the exercises.

For concreteness, we work in the category Mod_A of module over a ring A. However, everything we say will apply in any abelian category. (And if it helps you feel secure, work instead in the category Vec_k of vector spaces over a field k.)

1.7.1. Double complexes.

A **double complex** is a collection of A-modules $E^{p,q}$ $(p, q \in \mathbb{Z})$, and "rightward" morphisms $d_{\rightarrow}^{p,q} : E^{p,q} \rightarrow E^{p+1,q}$ and "upward" morphisms $d_{\uparrow}^{p,q} : E^{p,q} \rightarrow E^{p,q+1}$. In the superscript, the first entry denotes the column number (the "x-coordinate"), and the second entry denotes the row number (the "y-coordinate"). (Warning: this is opposite to the convention for matrices.) The subscript is meant to suggest the direction of the arrows. We will always write these as $d_{\rightarrow}$ and $d_{\uparrow}$ and ignore the superscripts. We require that $d_{\rightarrow}$ and $d_{\uparrow}$ satisfy (a) $d_{\rightarrow}^2 = 0$, (b) $d_{\uparrow}^2 = 0$, and one more condition: (c) either $d_{\rightarrow} d_{\uparrow} = d_{\uparrow} d_{\rightarrow}$ (all the squares commute) or $d_{\rightarrow} d_{\uparrow} + d_{\uparrow} d_{\rightarrow} = 0$ (they all anticommute). Both come up in nature, and you can switch from one to the other by replacing $d_{\uparrow}^{p,q}$ with $(-1)^p d_{\uparrow}^{p,q}$. So I will assume that all the squares anticommute, but that you know how to turn the commuting case into this one. (You will see that there is no difference in the recipe, basically because the image and kernel of a homomorphism f equal the image and kernel respectively of $-f$.)

$$\begin{tikzcd}
E^{p,q+1} \arrow[r, "d_{\rightarrow}^{p,q+1}"] \arrow[dr, phantom, "\text{anticommutes}"] & E^{p+1,q+1} \\
E^{p,q} \arrow[u, "d_{\uparrow}^{p,q}"] \arrow[r, "d_{\rightarrow}^{p,q}"] & E^{p+1,q} \arrow[u, "d_{\uparrow}^{p+1,q}"']
\end{tikzcd}$$

There are variations on this definition, where for example the vertical arrows go downwards, or some different subset of the $E^{p,q}$ are required to be zero, but I will leave these straightforward variations to you.

From the double complex we construct a corresponding (single) complex $E^\bullet$ with $E^k = \oplus_i E^{i,k-i}$, with $d = d_\rightarrow + d_\uparrow$. In other words, when there is a *single* superscript k, we mean a sum of the kth antidiagonal of the double complex. The single complex is sometimes called the **total complex**. Note that $d^2 = (d_\rightarrow + d_\uparrow)^2 = d_\rightarrow^2 + (d_\rightarrow d_\uparrow + d_\uparrow d_\rightarrow) + d_\uparrow^2 = 0$, so $E^\bullet$ is indeed a complex.

The **cohomology** of the single complex is sometimes called the **hypercohomology** of the double complex. We will instead use the phrase "cohomology of the double complex".

Our initial goal will be to find the cohomology of the double complex. You will see later that we secretly also have other goals.

A spectral sequence is a recipe for computing some information about the cohomology of the double complex. I won't yet give the full recipe. Surprisingly, this fragmentary bit of information is sufficent to prove lots of things.

1.7.2. Approximate Definition. A **spectral sequence** with **rightward orientation** is a sequence of tables or **pages** ${}_\rightarrow E_0^{p,q}, {}_\rightarrow E_1^{p,q}, {}_\rightarrow E_2^{p,q}, \dots$ ($p, q \in \mathbb{Z}$), where ${}_\rightarrow E_0^{p,q} = E^{p,q}$, along with a differential

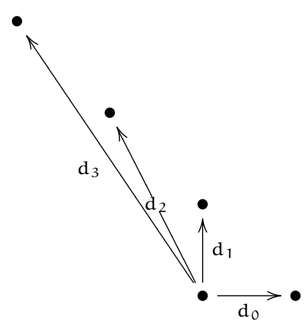
$${}_\rightarrow d_r^{p,q} : {}_\rightarrow E_r^{p,q} \rightarrow {}_\rightarrow E_r^{p-r+1,q+r}$$

with ${}_\rightarrow d_r^{p,q} \circ {}_\rightarrow d_r^{p+r-1,q-r} = 0$, and with an isomorphism of the cohomology of ${}_\rightarrow d_r$ at ${}_\rightarrow E_r^{p,q}$ (i.e., $\ker {}_\rightarrow d_r^{p,q} / \operatorname{im} {}_\rightarrow d_r^{p+r-1,q-r}$) with ${}_\rightarrow E_{r+1}^{p,q}$.

The orientation indicates that our 0th differential is the rightward one: $d_0 = d_\rightarrow$. The left subscript "$\rightarrow$" is usually omitted.

The order of the morphisms is best understood visually:

(1.7.2.1)



(the morphisms each apply to different pages). Notice that the map always is "degree 1" in terms of the grading of the single complex $E^\bullet$. (You should figure out what this informal statement really means.)

The actual definition describes what $E_r^{\bullet,\bullet}$ and $d_r^{\bullet,\bullet}$ really are, in terms of $E^{\bullet,\bullet}$. We will describe d_0, d_1, and d_2 below, and you should for now take on faith that this sequence continues in some natural way.

Note that $E_r^{p,q}$ is always a subquotient of the corresponding term on the ith page $E_i^{p,q}$ for all $i < r$. In particular, if $E^{p,q} = 0$, then $E_r^{p,q} = 0$ for all r.

Suppose now that $E^{\bullet,\bullet}$ is a **first quadrant double complex**, i.e., $E^{p,q} = 0$ for $p < 0$ or $q < 0$ (so $E_r^{p,q} = 0$ for all r unless $p, q \in \mathbb{Z}^{\geq 0}$). Then for any fixed p, q, once r is sufficiently large, $E_{r+1}^{p,q}$ is computed from $(E_r^{\bullet,\bullet}, d_r)$ using the complex

$$\begin{tikzcd} 0 & & \\ & E_r^{p,q} \arrow[ul, "d_r^{p,q}"] & \\ & & 0 \arrow[ul, "d_r^{p+r-1,q-r}"'] \end{tikzcd}$$

and thus we have canonical isomorphisms

$$E_r^{p,q} \cong E_{r+1}^{p,q} \cong E_{r+2}^{p,q} \cong \cdots .$$

We denote this module $E_\infty^{p,q}$. The same idea works in other circumstances, for example if the double complex is only nonzero in a finite number of rows — $E^{p,q} = 0$ unless $q_0 < q < q_1$. This will come up for example in the long exact sequence and mapping cone discussion (Exercises 1.7.F and 1.7.E below).

We now describe the first few pages of the spectral sequence explicitly. As stated above, the differential d_0 on $E_0^{\bullet,\bullet} = E^{\bullet,\bullet}$ is defined to be $d_\rightarrow$. The rows are complexes:

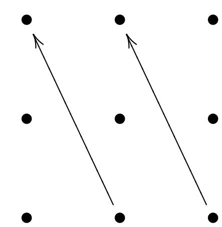
$$\text{The 0th page } E_0: \qquad \begin{tikzcd} \bullet \arrow[r] & \bullet \arrow[r] & \bullet \\ \bullet \arrow[r] & \bullet \arrow[r] & \bullet \\ \bullet \arrow[r] & \bullet \arrow[r] & \bullet \end{tikzcd}$$

and so E_1 is just the table of cohomologies of the rows. You should check that there are now vertical maps $d_1^{p,q} : E_1^{p,q} \to E_1^{p,q+1}$ of the row cohomology groups, induced by $d_\uparrow$, and that these make the columns into complexes. (This is essentially the fact that a map of complexes induces a map on homology.) We have "used up the horizontal morphisms", but "the vertical differentials live on".

$$\text{The 1st page } E_1: \qquad \begin{tikzcd} \bullet & \bullet & \bullet \\ \bullet \arrow[u] & \bullet \arrow[u] & \bullet \arrow[u] \\ \bullet \arrow[u] & \bullet \arrow[u] & \bullet \arrow[u] \end{tikzcd}$$

We take cohomology of d_1 on E_1, giving us a new table, $E_2^{p,q}$. It turns out that there are natural morphisms from each entry to the entry two above and one to the left, and that the composition of these two is 0. (It is a very worthwhile exercise to work out how this natural morphism d_2 should be defined. Your argument may be reminiscent of the connecting homomorphism in the Snake Lemma 1.7.5 or in the long exact sequence in cohomology arising from a short exact sequence of complexes, Exercise 1.6.C. This is no coincidence.)

The 2nd page E_2:



This is the beginning of a pattern.

Then it is a theorem that there is a filtration of $H^k(E^\bullet)$ by $E_\infty^{p,q}$ where $p+q=k$. (We can't yet state it as an official **Theorem** because we haven't precisely defined the pages and differentials in the spectral sequence.) More precisely, there is a filtration

$$E_\infty^{0,k} \xhookrightarrow{E_\infty^{1,k-1}} ? \xhookrightarrow{E_\infty^{2,k-2}} \cdots \xhookrightarrow{E_\infty^{k,0}} H^k(E^\bullet) \tag{1.7.2.2}$$

where the quotients are displayed above each inclusion. (Here is a tip for remember which way the quotients are supposed to go. The differentials on later and later pages point deeper and deeper into the filtration. Thus the entries in the direction of the later arrowheads are the subobjects, and the entries in the direction of the later "arrowtails" are quotients. This tip has the advantage of being independent of the details of the spectral sequence, e.g. the "quadrant" or the orientation.)

We say that the spectral sequence ${}_{\rightarrow}E_\bullet^{\bullet,\bullet}$ **converges** to $H^\bullet(E^\bullet)$. We often say that ${}_{\rightarrow}E_2^{\bullet,\bullet}$ (or any other page) **abuts** to $H^\bullet(E^\bullet)$.

Although the filtration gives only partial information about $H^\bullet(E^\bullet)$, sometimes one can find $H^\bullet(E^\bullet)$ precisely. One example is if all $E_\infty^{i,k-i}$ are zero, or if all but one of them are zero (e.g. if $E_r^{\bullet,\bullet}$ has precisely one nonzero row or column, in which case one says that the spectral sequence **collapses** at the rth step, although we will not use this term). Another example is in the category of vector spaces over a field, in which case we can find the dimension of $H^k(E^\bullet)$. Also, in lucky circumstances, E_2 (or some other small page) already equals E_∞.

1.7.A. EXERCISE: INFORMATION FROM THE SECOND PAGE. Show that $H^0(E^\bullet) = E_\infty^{0,0} = E_2^{0,0}$ and

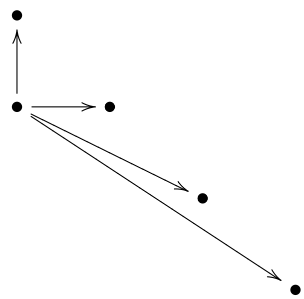
$$0 \longrightarrow E_2^{0,1} \longrightarrow H^1(E^\bullet) \longrightarrow E_2^{1,0} \xrightarrow{d_2^{1,0}} E_2^{0,2} \longrightarrow H^2(E^\bullet)$$

is exact.

1.7.3. *The other orientation.*

You may have observed that we could as well have done everything in the opposite direction, i.e., reversing the roles of horizontal and vertical morphisms. Then the sequences of arrows giving the spectral sequence would look like this (compare to (1.7.2.1)).

(1.7.3.1)



This spectral sequence is denoted ${}_{\uparrow}E_{\bullet}^{\bullet,\bullet}$ ("with the upward orientation"). Then we would again get pieces of a filtration of $H^{\bullet}(E^{\bullet})$ (where we have to be a bit careful with the order with which ${}_{\uparrow}E_{\infty}^{p,q}$ corresponds to the subquotients — it in the opposite order to that of (1.7.2.2) for ${}_{\rightarrow}E_{\infty}^{p,q}$). Warning: in general there is no isomorphism between ${}_{\rightarrow}E_{\infty}^{p,q}$ and ${}_{\uparrow}E_{\infty}^{p,q}$.

In fact, this observation that we can start with either the horizontal or vertical maps was our secret goal all along. Both algorithms compute information about the same thing ($H^{\bullet}(E^{\bullet})$), and usually we don't care about the final answer — we often care about the answer we get in one way, and we get at it by doing the spectral sequence in the *other* way.

1.7.4. Examples.

We are now ready to see how this is useful. The moral of these examples is the following. In the past, you may have proved various facts involving various sorts of diagrams, by chasing elements around. Now, you will just plug them into a spectral sequence, and let the spectral sequence machinery do your chasing for you.

1.7.5. *Example: Proving the Snake Lemma.* Consider the diagram

(1.7.5.1)
$$\begin{tikzcd} 0 \arrow[r] & D \arrow[r] & E \arrow[r] & F \arrow[r] & 0 \\ 0 \arrow[r] & A \arrow[r] \arrow[u, "\alpha"] & B \arrow[r] \arrow[u, "\beta"] & C \arrow[r] \arrow[u, "\gamma"] & 0 \end{tikzcd}$$

where the rows are exact in the middle (at A, B, C, D, E, F) and the squares commute. (Normally the Snake Lemma is described with the vertical arrows pointing downwards, but I want to fit this into my spectral sequence conventions.) We wish to show that there is an exact sequence

$$0 \to \ker\alpha \to \ker\beta \to \ker\gamma \to \operatorname{coker}\alpha \to \operatorname{coker}\beta \to \operatorname{coker}\gamma \to 0. \tag{1.7.5.2}$$

We plug this into our spectral sequence machinery. We first compute the cohomology using the rightward orientation, i.e., using the order (1.7.2.1). Then because the rows are exact, $E_1^{p,q} = 0$, so the spectral sequence has already converged: $E_{\infty}^{p,q} = 0$.

We next compute this "0" in another way, by computing the spectral sequence using the upward orientation. Then ${}_{\uparrow}E_1^{\bullet,\bullet}$ (with its differentials) is:

$$\begin{tikzcd} 0 \arrow[r] & \operatorname{coker} \alpha \arrow[r] & \operatorname{coker} \beta \arrow[r] & \operatorname{coker} \gamma \arrow[r] & 0 \end{tikzcd}$$

$$\begin{tikzcd} 0 \arrow[r] & \ker \alpha \arrow[r] & \ker \beta \arrow[r] & \ker \gamma \arrow[r] & 0. \end{tikzcd}$$

Then ${}_{\uparrow}E_2^{\bullet,\bullet}$ is of the form:

$$\begin{tikzcd}
0 \arrow[drr] & 0 \arrow[drr] & & & & & \\
& 0 \arrow[drr] & ?? \arrow[drr] & ? \arrow[drr] & ? & 0 & \\
& 0 & ? & ? \arrow[drr] & ?? \arrow[drr] & 0 & \\
& & & & & 0 & 0
\end{tikzcd}$$

We see that after ${}_{\uparrow}E_2$, all the terms will stabilize except for the double question marks — all maps to and from the single question marks are to and from 0-entries. And after ${}_{\uparrow}E_3$, even these two double-question-mark terms will stabilize. But in the end our complex must be the 0 complex. This means that in ${}_{\uparrow}E_2$, all the entries must be zero, except for the two double question marks, and these two must be isomorphic. This means that $0 \to \ker \alpha \to \ker \beta \to \ker \gamma$ and $\operatorname{coker} \alpha \to \operatorname{coker} \beta \to \operatorname{coker} \gamma \to 0$ are both exact (that comes from the vanishing of the single question marks), and

$$\operatorname{coker}(\ker \beta \to \ker \gamma) \cong \ker(\operatorname{coker} \alpha \to \operatorname{coker} \beta)$$

is an isomorphism (that comes from the equality of the double question marks). Taken together, we have proved the exactness of (1.7.5.2), and hence the Snake Lemma! (Notice: in the end we didn't really care about the double complex. We just used it as a prop to prove the snake lemma.)

Spectral sequences make it easy to see how to generalize results further. For example, if $A \to B$ is no longer assumed to be injective, how would the conclusion change?

1.7.B. UNIMPORTANT EXERCISE (GRAFTING EXACT SEQUENCES, A WEAKER VERSION OF THE SNAKE LEMMA). Extend the snake lemma as follows. Suppose we have a commuting diagram

$$\begin{tikzcd}
& 0 \arrow[r] & X' \arrow[r] & Y' \arrow[r] & Z' \arrow[r] & A' \arrow[r] & \cdots \\
\cdots \arrow[r] & W \arrow[r] \arrow[u] & X \arrow[r] \arrow[u, "a"] & Y \arrow[r] \arrow[u, "b"] & Z \arrow[r] \arrow[u, "c"] & 0. \arrow[u] &
\end{tikzcd}$$

where the top and bottom rows are exact. Show that the top and bottom rows can be "grafted together" to an exact sequence

$$\cdots \longrightarrow W \longrightarrow \ker a \longrightarrow \ker b \longrightarrow \ker c$$

$$\longrightarrow \operatorname{coker} a \longrightarrow \operatorname{coker} b \longrightarrow \operatorname{coker} c \longrightarrow A' \longrightarrow \cdots .$$

1.7.6. *Example: the Five Lemma.* Suppose

$$\begin{tikzcd} F \arrow[r] & G \arrow[r] & H \arrow[r] & I \arrow[r] & J \\ A \arrow[r] \arrow[u, "\alpha"] & B \arrow[r] \arrow[u, "\beta"] & C \arrow[r] \arrow[u, "\gamma"] & D \arrow[r] \arrow[u, "\delta"] & E \arrow[u, "\epsilon"] \end{tikzcd} \tag{1.7.6.1}$$

where the rows are exact and the squares commute.

Suppose α, β, δ, ϵ are isomorphisms. We will show that γ is an isomorphism.

We first compute the cohomology of the total complex using the rightward orientation (1.7.2.1). We choose this because we see that we will get lots of zeros. Then ${}_{\rightarrow}E_1^{\bullet,\bullet}$ looks like this:

$$\begin{tikzcd} ? & 0 & 0 & 0 & ? \\ ? \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & ? \arrow[u] \end{tikzcd}$$

Then ${}_{\rightarrow}E_2$ looks similar, and the sequence will converge by E_2, as we will never get any arrows between two nonzero entries in a table thereafter. We can't conclude that the cohomology of the total complex vanishes, but we can note that it vanishes in all but four degrees — and most important, it vanishes in the two degrees corresponding to the entries C and H (the source and target of γ).

We next compute this using the upward orientation (1.7.3.1). Then ${}_{\uparrow}E_1$ looks like this:

$$0 \longrightarrow 0 \longrightarrow ? \longrightarrow 0 \longrightarrow 0$$

$$0 \longrightarrow 0 \longrightarrow ? \longrightarrow 0 \longrightarrow 0$$

and the spectral sequence converges at this step. We wish to show that those two question marks are zero. But they are precisely the cohomology groups of the total complex that we just showed *were* zero — so we are done!

The best way to become comfortable with this sort of argument is to try it out yourself several times, and realize that it really is easy. So you should do the following exercises! Many can readily be done directly, but you should deliberately try to use this spectral sequence machinery in order to get practice and develop confidence.

1.7.C. EXERCISE: A SUBTLER FIVE LEMMA. By looking at the spectral sequence proof of the Five Lemma above, prove a subtler version of the Five Lemma, where one of the isomorphisms can instead just be required to be an injection, and another can instead just be required to be a surjection. (I am deliberately not telling you which ones, so you can see how the spectral sequence is telling you how to improve the result.)

1.7.D. EXERCISE: ANOTHER SUBTLE VERSION OF THE FIVE LEMMA. If β and δ (in (1.7.6.1)) are injective, and α is surjective, show that γ is injective. Give the dual statement (whose proof is of course essentially the same).

The next two exercises no longer involve first quadrant double complexes. You will have to think a little to realize why there is no reason for confusion or alarm.

1.7.E. EXERCISE (THE MAPPING CONE). Suppose $\mu : A^\bullet \to B^\bullet$ is a morphism of complexes. Suppose $C^\bullet$ is the single complex associated to the double complex $A^\bullet \to B^\bullet$. ($C^\bullet$ is called the *mapping cone* of μ.) Show that there is a long exact sequence of complexes:

$$\cdots \to H^{i-1}(C^\bullet) \to H^i(A^\bullet) \to H^i(B^\bullet) \to H^i(C^\bullet) \to H^{i+1}(A^\bullet) \to \cdots .$$

(There is a slight notational ambiguity here; depending on how you index your double complex, your long exact sequence might look slightly different.) In particular, we will use the fact that μ induces an isomorphism on cohomology if and only if the mapping cone is exact. (We won't use it until the proof of Theorem 18.2.4.)

1.7.F. EXERCISE. Use spectral sequences to show that a short exact sequence of complexes gives a long exact sequence in cohomology (Exercise 1.6.C). (This is a generalization of Exercise 1.7.E.)

The Grothendieck (or composition of functor) spectral sequence (Theorem 23.3.5) will be an important example of a spectral sequence that specializes in a number of useful ways.

You are now ready to go out into the world and use spectral sequences to your heart's content!

1.7.7. $\star\star$ **Complete definition of the spectral sequence, and proof.**

You should most definitely not read this section any time soon after reading the introduction to spectral sequences above. Instead, flip quickly through it to convince yourself that nothing fancy is involved.

1.7.8. *Remark: Spectral sequences are actually spectral functors.* It is useful to notice that the proof implies that spectral sequences are functorial in the 0th page: the spectral sequence formalism has good functorial properties in the double complex. Unfortunately the terminology "spectral functor" that Grothendieck used in [**Gr1**, §2.4] did not catch on.

1.7.9. *Goals.* We consider the rightward orientation. The upward orientation is of course a trivial variation of this. We wish to describe the pages and differentials of the spectral sequence explicitly, and prove that they behave the way we said they did. More precisely, we wish to:

(a) describe $E_r^{p,q}$ (and verify that $E_0^{p,q}$ is indeed $E^{p,q}$),
(b) verify that $H^k(E^\bullet)$ is filtered by $E_\infty^{p,k-p}$ as in (1.7.2.2),
(c) describe d_r and verify that $d_r^2 = 0$, and
(d) verify that $E_{r+1}^{p,q}$ is given by cohomology using d_r.

Before tackling these goals, you can impress your friends by giving this short description of the pages and differentials of the spectral sequence. We say that an element of $E^{\bullet,\bullet}$ is a (p,q)-*strip* if it is an element of $\oplus_{l\geq 0} E^{p-l,q+l}$ (see Fig. 1.1). Its nonzero entries lie on an "upper-leftwards" semi-infinite antidiagonal starting with position (p,q). We say that the (p,q)-entry (the projection to $E^{p,q}$) is the *leading term* of the (p,q)-strip. Let $\boxed{S^{p,q}} \subset E^{\bullet,\bullet}$ be the submodule of all the (p,q)-strips. Clearly $S^{p,q} \subset E^{p+q}$, and $S^{k,0} = E^k$.

$$
\begin{array}{ccccc}
\ddots & 0 & 0 & 0 & 0 \\
0 & *^{p-2,q+2} & 0 & 0 & 0 \\
0 & 0 & *^{p-1,q+1} & 0 & 0 \\
0 & 0 & 0 & *^{p,q} & 0 \\
0 & 0 & 0 & 0 & 0^{p+1,q-1}
\end{array}
$$

FIGURE 1.1. A (p,q)-strip (in $S^{p,q} \subset E^{p+q}$). Clearly $S^{k,0} = E^k$.

Note that the differential $d = d_\uparrow + d_\rightarrow$ sends a (p,q)-strip x to a $(p+1,q)$-strip dx. If dx is furthermore a $(p-r+1, q+r)$-strip ($r \in \mathbb{Z}^{\geq 0}$), we say that x is an r-*closed* (p,q)-*strip* — "the differential knocks x at least r terms deeper into the filtration". We denote the set of r-closed (p,q)-strips $\boxed{S_r^{p,q}}$ (so for example $S_0^{p,q} = S^{p,q}$, and $S_0^{k,0} = E^k$). An element of $S_r^{p,q}$ may be depicted as:

$$
\begin{tikzcd}
\ddots \arrow[r] & ? & & & \\
& *^{p-2,q+2} \arrow[u] \arrow[r] & 0 & & \\
& & *^{p-1,q+1} \arrow[u] \arrow[r] & 0 & \\
& & & *^{p,q} \arrow[u] \arrow[r] & 0
\end{tikzcd}
$$

1.7.10. *Preliminary definition of* $E_r^{p,q}$. We are now ready to give a first definition of $E_r^{p,q}$, which by construction should be a subquotient of $E^{p,q} = E_0^{p,q}$. We describe it as such by describing two submodules $Y_r^{p,q} \subset X_r^{p,q} \subset E^{p,q}$, and defining $E_r^{p,q} =$

$X_r^{p,q}/Y_r^{p,q}$. Let $X_r^{p,q}$ be those elements of $E^{p,q}$ that are the leading terms of r-closed (p,q)-strips. Note that by definition, d sends $(r-1)$-closed $(p+(r-1)-1, q-(r-1))$-strips to (p,q)-strips. Let $Y_r^{p,q}$ be the leading $((p,q))$-terms of the differential d of $(r-1)$-closed $(p+(r-1)-1, q-(r-1))$-strips (where the differential is considered as a (p,q)-strip).

1.7.G. EXERCISE (REALITY CHECK). Verify that $E_0^{p,q}$ is (canonically isomorphic to) $E^{p,q}$.

We next give the definition of the differential d_r of such an element $x \in X_r^{p,q}$. We take *any* r-closed (p,q)-strip with leading term x. Its differential d is a $(p-r+1, q+r)$-strip, and we take its leading term. The choice of the r-closed (p,q)-strip means that this is not a well-defined element of $E^{p,q}$. But it is well-defined modulo the differentials of the $(r-1)$-closed $(p-1, q+1)$-strips, and hence gives a map $E_r^{p,q} \to E_r^{p-r+1,q+r}$.

This definition is fairly short, but not much fun to work with, so we will forget it, and instead dive into a snakes' nest of subscripts and superscripts.

We begin with making some quick but important observations about (p,q)-strips.

1.7.H. EXERCISE (NOT HARD). Verify the following.

(a) $S^{p,q} = S^{p-1,q+1} \oplus E^{p,q}$.

(b) *(Any closed (p,q)-strip is r-closed for all r.)* Any element x of $S^{p,q} = S_0^{p,q}$ that is a cycle (i.e., $dx = 0$) is automatically in $S_r^{p,q}$ for all r. For example, this holds when x is a boundary (i.e., of the form dy).

(c) Show that for fixed p, q,
$$S_0^{p,q} \supset S_1^{p,q} \supset \cdots \supset S_r^{p,q} \supset \cdots$$
stabilizes for $r \gg 0$ (i.e., $S_r^{p,q} = S_{r+1}^{p,q} = \cdots$). Denote the stabilized module $S_\infty^{p,q}$. Show $S_\infty^{p,q}$ is the set of closed (p,q)-strips (those (p,q)-strips annihilated by d, i.e., the cycles). In particular, $S_\infty^{k,0}$ is the set of cycles in E^k.

1.7.11. *Defining* $E_r^{p,q}$.

Define $X_r^{p,q} := S_r^{p,q}/S_{r-1}^{p-1,q+1}$ and
$$Y_r^{p,q} := \left(dS_{r-1}^{p+(r-1)-1,q-(r-1)} + S_{r-1}^{p-1,q+1}\right)/S_{r-1}^{p-1,q+1}.$$

Then $Y_r^{p,q} \subset X_r^{p,q}$ by Exercise 1.7.H(b). We define

$$\boxed{E_r^{p,q} = \frac{X_r^{p,q}}{Y_r^{p,q}} = \frac{S_r^{p,q}}{dS_{r-1}^{p+(r-1)-1,q-(r-1)} + S_{r-1}^{p-1,q+1}}} \qquad (1.7.11.1)$$

We have completed Goal 1.7.9(a).

You are welcome to verify that these definitions of $X_r^{p,q}$ and $Y_r^{p,q}$ and hence $E_r^{p,q}$ agree with the earlier ones of §1.7.10 (and in particular $X_r^{p,q}$ and $Y_r^{p,q}$ are both submodules of $E^{p,q}$), but we won't need this fact.

1.7.I. EXERCISE: $E_\infty^{p,k-p}$ GIVES SUBQUOTIENTS OF $H^k(E^\bullet)$. By Exercise 1.7.H(c), $E_r^{p,q}$ stabilizes as $r \to \infty$. For $r \gg 0$, interpret $S_r^{p,q}/dS_{r-1}^{p+(r-1)-1,q-(r-1)}$ as the cycles in $S_\infty^{p,q} \subset E^{p+q}$ modulo those boundary elements of dE^{p+q-1} contained in $S_\infty^{p,q}$. Finally, show that $H^k(E^\bullet)$ is indeed filtered as described in (1.7.2.2).

We have completed Goal 1.7.9(b).

1.7.12. *Definition of* d_r.

We shall see that the map $d_r : E_r^{p,q} \to E_r^{p-r+1,q+r}$ is just induced by our differential d. Notice that d sends r-closed (p,q)-strips $S_r^{p,q}$ to $(p-r+1,q+r)$-strips $S^{p-r+1,q+r}$, by the definition "r-closed". By Exercise 1.7.H(b), the image lies in $S_r^{p-r+1,q+r}$.

1.7.J. EXERCISE. Verify that d sends

$$dS_{r-1}^{p+(r-1)-1,q-(r-1)} + S_{r-1}^{p-1,q+1} \to dS_{r-1}^{(p-r+1)+(r-1)-1,(q+r)-(r-1)} + S_{r-1}^{(p-r+1)-1,(q+r)+1}.$$

(The first term on the left goes to 0 from $d^2 = 0$, and the second term on the left goes to the first term on the right.)

Thus we may define

$$\boxed{d_r : E_r^{p,q} = \frac{S_r^{p,q}}{dS_{r-1}^{p+(r-1)-1,q-(r-1)} + S_{r-1}^{p-1,q+1}} \to}$$

$$\boxed{\frac{S_r^{p-r+1,q+r}}{dS_{r-1}^{p-1,q+1} + S_{r-1}^{p-r,q+r+1}} = E_r^{p-r+1,q+r}}$$

and clearly $d_r^2 = 0$ (as we may interpret it as taking an element of $S_r^{p,q}$ and applying d twice).

We have accomplished Goal 1.7.9(c).

1.7.13. *Verifying that the cohomology of* d_r *at* $E_r^{p,q}$ *is* $E_{r+1}^{p,q}$. We are left with the unpleasant job of verifying that the cohomology of

$$\frac{S_r^{p+r-1,q-r}}{dS_{r-1}^{p+2r-3,q-2r+1} + S_{r-1}^{p+r-2,q-r+1}} \xrightarrow{d_r} \frac{S_r^{p,q}}{dS_{r-1}^{p+r-2,q-r+1} + S_{r-1}^{p-1,q+1}} \xrightarrow{d_r} \frac{S_r^{p-r+1,q+r}}{dS_{r-1}^{p-1,q+1} + S_{r-1}^{p-r,q+r+1}} \tag{1.7.13.1}$$

is naturally identified with

$$\frac{S_{r+1}^{p,q}}{dS_r^{p+r-1,q-r} + S_r^{p-1,q+1}}$$

and this will conclude our final Goal 1.7.9(d).

We begin by understanding the kernel of the right map of (1.7.13.1). Suppose $a \in S_r^{p,q}$ is mapped to 0. This means that $da = db + c$, where $b \in S_{r-1}^{p-1,q+1}$. If $u = a - b$, then $u \in S^{p,q}$, while $du = c \in S_{r-1}^{p-r,q+r+1} \subset S^{p-r,q+r+1}$, from which u is $(r+1)$-closed, i.e., $u \in S_{r+1}^{p,q}$. Thus $a = b + u \in S_{r-1}^{p-1,q+1} + S_{r+1}^{p,q}$. Conversely, any $a \in S_{r-1}^{p-1,q+1} + S_{r+1}^{p,q}$ satisfies

$$da \in dS_{r-1}^{p-1,q+1} + dS_{r+1}^{p,q} \subset dS_{r-1}^{p-1,q+1} + S_{r-1}^{p-r,q+r+1}$$

(using $dS^{p,q}_{r+1} \subset S^{p-r,q+r+1}_0$ and Exercise 1.7.H(b)) so any such a is indeed in the kernel of

$$S^{p,q}_r \to \frac{S^{p-r+1,q+r}_r}{dS^{p-1,q+1}_{r-1} + S^{p-r,q+r+1}_{r-1}}.$$

Hence the kernel of the right map of (1.7.13.1) is

$$\ker = \frac{S^{p-1,q+1}_{r-1} + S^{p,q}_{r+1}}{dS^{p+r-2,q-r+1}_{r-1} + S^{p-1,q+1}_{r-1}}.$$

Next, the image of the left map of (1.7.13.1) is immediately

$$\operatorname{im} = \frac{dS^{p+r-1,q-r}_r + dS^{p+r-2,q-r+1}_{r-1} + S^{p-1,q+1}_{r-1}}{dS^{p+r-2,q-r+1}_{r-1} + S^{p-1,q+1}_{r-1}} = \frac{dS^{p+r-1,q-r}_r + S^{p-1,q+1}_{r-1}}{dS^{p+r-2,q-r+1}_{r-1} + S^{p-1,q+1}_{r-1}}$$

(as $S^{p+r-1,q-r}_r$ contains $S^{p+r-2,q-r+1}_{r-1}$).

Thus the cohomology of (1.7.13.1) is

$$\ker / \operatorname{im} = \frac{S^{p-1,q+1}_{r-1} + S^{p,q}_{r+1}}{dS^{p+r-1,q-r}_r + S^{p-1,q+1}_{r-1}} = \frac{S^{p,q}_{r+1}}{S^{p,q}_{r+1} \cap (dS^{p+r-1,q-r}_r + S^{p-1,q+1}_{r-1})}$$

where the equality on the right uses the fact that $dS^{p+r-1,q-r}_r \subset S^{p,q}_{r+1}$ and an isomorphism theorem. We thus must show

$$S^{p,q}_{r+1} \cap (dS^{p+r-1,q-r}_r + S^{p-1,q+1}_{r-1}) = dS^{p+r-1,q-r}_r + S^{p-1,q+1}_r.$$

However,

$$S^{p,q}_{r+1} \cap (dS^{p+r-1,q-r}_r + S^{p-1,q+1}_{r-1}) = dS^{p+r-1,q-r}_r + S^{p,q}_{r+1} \cap S^{p-1,q+1}_{r-1}$$

and $S^{p,q}_{r+1} \cap S^{p+1,q-1}_{r-1}$ consists of $(p-1,q+1)$-strips whose differential vanishes up to row $p+r$, from which $S^{p,q}_{r+1} \cap S^{p-1,q+1}_{r-1} = S^{p-1,q+1}_r$ as desired.

This completes the explanation of how spectral sequences work for a first-quadrant double complex. The argument applies without significant change to more general situations, including filtered complexes.

CHAPTER 2

Sheaves

It is perhaps surprising that geometric spaces are often best understood in terms of (nice) functions on them. For example, a differentiable manifold that is a subset of $\mathbb{R}^n$ can be studied in terms of its differentiable functions. Because "geometric spaces" can have few (everywhere-defined) functions, a more precise version of this insight is that the structure of the space can be well understood by considering all functions on all open subsets of the space. This information is encoded in something called a *sheaf*. Sheaves were introduced by Leray in the 1940's, and Serre introduced them to algebraic geometry. (The reason for the name will be somewhat explained in Remark 2.4.4.) We will define sheaves and describe useful facts about them. We will begin with a motivating example to convince you that the notion is not so foreign.

One reason sheaves are slippery to work with is that they keep track of a huge amount of information, and there are some subtle local-to-global issues. There are also three different ways of getting a hold of them:

- in terms of open sets (the definition §2.2) — intuitive but in some ways the least helpful;
- in terms of stalks (see §2.4.1); and
- in terms of a base of a topology (§2.7).

Knowing which to use requires experience, so it is essential to do a number of exercises on different aspects of sheaves in order to truly understand the concept. (Some people strongly prefer the espace étalé interpretation, §2.2.11, as well.)

2.1 Motivating example: The sheaf of differentiable functions.

Consider differentiable functions on the topological space $X = \mathbb{R}^n$ (or more generally on a manifold X). The sheaf of differentiable functions on X is the data of all differentiable functions on all open subsets on X. We will see how to manage these data, and observe some of their properties. On each open set $U \subset X$, we have a ring of differentiable functions. We denote this ring of functions $\mathscr{O}(U)$.

Given a differentiable function on an open set, you can restrict it to a smaller open set, obtaining a differentiable function there. In other words, if $U \subset V$ is an inclusion of open sets, we have a "restriction map" $\operatorname{res}_{V,U} : \mathscr{O}(V) \to \mathscr{O}(U)$.

Take a differentiable function on a big open set, and restrict it to a medium open set, and then restrict that to a small open set. The result is the same as if you restrict the differentiable function on the big open set directly to the small open set.

In other words, if $U \hookrightarrow V \hookrightarrow W$, then the following diagram commutes:

$$\begin{tikzcd} \mathscr{O}(W) \arrow[rr, "\mathrm{res}_{W,V}"] \arrow[dr, "\mathrm{res}_{W,U}"'] & & \mathscr{O}(V) \arrow[dl, "\mathrm{res}_{V,U}"] \\ & \mathscr{O}(U) & \end{tikzcd}$$

Next take two differentiable functions f_1 and f_2 on a big open set U, and an open cover of U by some collection of open subsets $\{U_i\}$. (We say $\{U_i\}$ **covers** U, or is an **open cover of** U, if $U = \cup U_i$.) Suppose that f_1 and f_2 agree on each of these U_i. Then they must have been the same function to begin with. In other words, if $\{U_i\}_{i \in I}$ is a cover of U, and $f_1, f_2 \in \mathscr{O}(U)$, and $\mathrm{res}_{U,U_i} f_1 = \mathrm{res}_{U,U_i} f_2$, then $f_1 = f_2$. Thus we can *identify* functions on an open set by looking at them on a covering by small open sets.

Finally, suppose you are given the same U and cover $\{U_i\}$, take a differentiable function on each of the U_i — a function f_1 on U_1, a function f_2 on U_2, and so on — and assume they agree on the pairwise overlaps. Then they can be "glued together" to make one differentiable function on all of U. In other words, given $f_i \in \mathscr{O}(U_i)$ for all i, such that $\mathrm{res}_{U_i, U_i \cap U_j} f_i = \mathrm{res}_{U_j, U_i \cap U_j} f_j$ for all i and j, then there is some $f \in \mathscr{O}(U)$ such that $\mathrm{res}_{U,U_i} f = f_i$ for all i.

The entire example above would have worked just as well with continuous functions, or smooth functions, or just plain functions. Thus all of these classes of "nice" functions share some common properties. We will soon formalize these properties in the notion of a sheaf.

2.1.1. The germ of a differentiable function. Before we do, we first give another definition, that of the germ of a differentiable function at a point $p \in X$. Intuitively, it is a "shred" of a differentiable function at p. Germs are objects of the form $\{(f, \text{open } U) \;:\; p \in U, f \in \mathscr{O}(U)\}$ modulo the relation that $(f, U) \sim (g, V)$ if there is some open set $W \subset U, V$ containing p where $f|_W = g|_W$ (i.e., $\mathrm{res}_{U,W} f = \mathrm{res}_{V,W}\, g$). In other words, two functions that are the same in a neighborhood of p (but may differ elsewhere) have the same germ. We call this set of germs the stalk at p, and denote it $\mathscr{O}_p$. Notice that the stalk is a ring: you can add two germs, and get another germ: if you have a function f defined on U, and a function g defined on V, then $f + g$ is defined on $U \cap V$. Moreover, $f + g$ is well-defined: if $\tilde{f}$ has the same germ as f, meaning that there is some open set W containing p on which they agree, and $\tilde{g}$ has the same germ as g, meaning they agree on some open W' containing p, then $\tilde{f} + \tilde{g}$ is the same function as $f + g$ on $U \cap V \cap W \cap W'$.

Notice also that if $p \in U$, you get a map $\mathscr{O}(U) \to \mathscr{O}_p$. Experts may already see that we are talking about germs as colimits.

We can see that $\mathscr{O}_p$ is a local ring as follows. Consider those germs vanishing at p, which we denote $\mathfrak{m}_p \subset \mathscr{O}_p$. They certainly form an ideal: $\mathfrak{m}_p$ is closed under addition, and when you multiply something vanishing at p by any function, the result also vanishes at p. We check that this ideal is maximal by showing that the quotient ring is a field:

$$0 \longrightarrow \mathfrak{m}_p := \text{ideal of germs vanishing at } p \longrightarrow \mathscr{O}_p \xrightarrow{f \mapsto f(p)} \mathbb{R} \longrightarrow 0 \tag{2.1.1.1}$$

2.1.A. EXERCISE. Show that this is the only maximal ideal of $\mathscr{O}_p$. (Hint: show that every element of $\mathscr{O}_p \setminus \mathfrak{m}_p$ is invertible.)

Note that we can interpret the value of a function at a point, or the value of a germ at a point, as an element of the local ring modulo the maximal ideal. (We will see that this doesn't work for more general sheaves, but *does* work for things behaving like sheaves of functions. This will be formalized in the notion of a *locally ringed space*, which we will see, briefly, in §6.3.)

2.1.2. *Aside.* Notice that $\mathfrak{m}_p/\mathfrak{m}_p^2$ is a module over $\mathscr{O}_p/\mathfrak{m} \cong \mathbb{R}$, i.e., it is a real vector space. It turns out to be naturally (whatever that means) the cotangent space to the manifold at p. This insight will prove handy later, when we define tangent and cotangent spaces of schemes.

2.1.B. $\star$ EXERCISE FOR THOSE WITH DIFFERENTIAL GEOMETRIC BACKGROUND. Prove this. (Rhetorical question for experts: what goes wrong if the sheaf of continuous functions is substituted for the sheaf of differentiable functions?)

2.2 Definition of sheaf and presheaf

We now formalize these notions, by defining presheaves and sheaves. Presheaves are simpler to define, and notions such as kernel and cokernel are straightforward. Sheaves are more complicated to define, and some notions such as cokernel require more thought. But sheaves are more useful because they are in some vague sense more geometric; you can get information about a sheaf locally.

2.2.1. Definition of sheaf and presheaf on a topological space X.

To be concrete, we will define sheaves of sets. However, in the definition the category *Sets* can be replaced by any category, and other important examples are abelian groups *Ab*, k-vector spaces Vec_k, rings *Rings*, modules over a ring Mod_A, and more. (You may have to think more when dealing with a category of objects that aren't "sets with additional structure", but there aren't any new complications. In any case, this won't be relevant for us, although people who want to do this should start by solving Exercise 2.2.C.) Sheaves (and presheaves) are often written in calligraphic font. The fact that $\mathscr{F}$ is a sheaf on a topological space X is often written as

$$\begin{array}{c} \mathscr{F} \\ \Big| \\ X \end{array}$$

2.2.2. Definition: Presheaf. A **presheaf** $\mathscr{F}$ on a topological space X is the following data.

• To each open set $U \subset X$, we have a set $\mathscr{F}(U)$ (e.g. the set of differentiable functions in our motivating example). (Notational warning: Several notations are in use, for various good reasons: $\mathscr{F}(U) = \Gamma(U, \mathscr{F}) = H^0(U, \mathscr{F})$. We will use them all.) The elements of $\mathscr{F}(U)$ are called **sections of $\mathscr{F}$ over U**. (§2.2.11 combined with Exercise 2.2.G gives a motivation for this terminology, although this isn't so important for us.)

• For each inclusion $U \hookrightarrow V$ of open sets, we have a **restriction map** $\mathrm{res}_{V,U} : \mathscr{F}(V) \to \mathscr{F}(U)$ (just as we did for differentiable functions).

The data is required to satisfy the following two conditions.

• The map $\mathrm{res}_{U,U}$ is the identity: $\mathrm{res}_{U,U} = \mathrm{id}_{\mathscr{F}(U)}$.

• If $U \hookrightarrow V \hookrightarrow W$ are inclusions of open sets, then the restriction maps commute, i.e.,

$$\begin{tikzcd} \mathscr{F}(W) \arrow[rr, "\mathrm{res}_{W,V}"] \arrow[dr, "\mathrm{res}_{W,U}"'] & & \mathscr{F}(V) \arrow[dl, "\mathrm{res}_{V,U}"] \\ & \mathscr{F}(U) & \end{tikzcd}$$

commutes.

2.2.A. Exercise for category-lovers: "A presheaf is the same as a contravariant functor". Given any topological space X, we have a "category of open sets" (Example 1.2.9), where the objects are the open sets and the morphisms are inclusions. Verify that the data of a presheaf is precisely the data of a contravariant functor from the category of open sets of X to the category of sets. (This interpretation is surprisingly useful.)

2.2.3. Definition: Stalks and germs. We define the stalk of a presheaf at a point in two equivalent ways. One will be hands-on, and the other will be as a colimit.

2.2.4. Define the **stalk** of a presheaf $\mathscr{F}$ at a point p to be the set of **germs** of $\mathscr{F}$ at p, denoted $\mathscr{F}_p$, as in the example of §2.1.1. Germs correspond to sections over some open set containing p, and two of these sections are considered the same if they agree on some smaller open set. More precisely: the stalk is

$$\{(f, \text{open } U) \ : \ p \in U, f \in \mathscr{F}(U)\}$$

modulo the relation that $(f, U) \sim (g, V)$ if there is some open set $W \subset U, V$ where $p \in W$ and $\mathrm{res}_{U,W} f = \mathrm{res}_{V,W} g$. (To explain the agricultural terminology: the French name "germe" is meant to suggest a tiny shoot sprouting from a seed, cf. "germinate".)

2.2.5. A useful equivalent definition of a stalk is as a colimit of all $\mathscr{F}(U)$ over all open sets U containing p:

$$\mathscr{F}_p = \varinjlim \mathscr{F}(U).$$

The index category is a filtered set (given any two such open sets, there is a third such set contained in both), so these two definitions are the same by Exercise 1.4.C. Hence we can define stalks for sheaves of sets, groups, rings, and other things for which colimits exist for directed sets. It is very helpful to simultaneously keep both definitions of stalk in mind at the same time.

If $p \in U$, and $f \in \mathscr{F}(U)$, then the image of f in $\mathscr{F}_p$ is called the **germ of** f **at** p. (Warning: unlike the example of §2.1.1, in general, the value of a section at a point doesn't make sense.)

2.2.6. Definition: Sheaf. A presheaf is a **sheaf** if it satisfies two more axioms. Notice that these axioms use the additional information of when some open sets cover another.

Identity axiom. If $\{U_i\}_{i \in I}$ is an open cover of U, and $f_1, f_2 \in \mathscr{F}(U)$, and $\mathrm{res}_{U,U_i} f_1 = \mathrm{res}_{U,U_i} f_2$ for all i, then $f_1 = f_2$.

(A presheaf satisfying the identity axiom is called a **separated presheaf**, but we will not use that notation in any essential way.)

Gluability axiom. If $\{U_i\}_{i \in I}$ is a open cover of U, then given $f_i \in \mathscr{F}(U_i)$ for all i, such that $\mathrm{res}_{U_i, U_i \cap U_j} f_i = \mathrm{res}_{U_j, U_i \cap U_j} f_j$ for all i, j, then there is some $f \in \mathscr{F}(U)$ such that $\mathrm{res}_{U,U_i} f = f_i$ for all i.

In mathematics, definitions often come paired: "at most one" and "at least one". In this case, identity means there is at most one way to glue, and gluability means that there is at least one way to glue.

(For experts and scholars of the empty set only: an additional axiom sometimes included is that $F(\varnothing)$ is a one-element set, and in general, for a sheaf with values in a category, $F(\varnothing)$ is required to be the final object in the category. This actually follows from the above definitions, assuming that the empty product is appropriately defined as the final object.)

Example. If U and V are disjoint, then $\mathscr{F}(U \cup V) = \mathscr{F}(U) \times \mathscr{F}(V)$. Here we use the fact that $F(\varnothing)$ is the final object.

The **stalk of a sheaf** at a point is just its stalk as a presheaf — the same definition applies — and similarly for the **germs** of a section of a sheaf.

2.2.B. UNIMPORTANT EXERCISE: PRESHEAVES THAT ARE NOT SHEAVES. Show that the following are presheaves on $\mathbb{C}$ (with the classical topology), but not sheaves: (a) bounded functions, (b) holomorphic functions admitting a holomorphic square root.

Both of the presheaves in the previous Exercise satisfy the identity axiom. A "natural" example failing even the identity axiom is implicit in Remark 2.7.5.

We now make a couple of points intended only for category-lovers.

2.2.7. *Interpretation in terms of the equalizer exact sequence.* The two axioms for a presheaf to be a sheaf can be interpreted as "exactness" of the "equalizer exact sequence": $\cdot \longrightarrow \mathscr{F}(U) \longrightarrow \prod \mathscr{F}(U_i) \rightrightarrows \prod \mathscr{F}(U_i \cap U_j)$. Identity is exactness at $\mathscr{F}(U)$, and gluability is exactness at $\prod \mathscr{F}(U_i)$. I won't make this precise, or even explain what the double right arrow means. (What is an exact sequence of sets?!) But you may be able to figure it out from the context.

2.2.C. EXERCISE. The identity and gluability axioms may be interpreted as saying that $\mathscr{F}(\cup_{i \in I} U_i)$ is a certain limit. What is that limit?

Here are a number of examples of sheaves.

2.2.D. EXERCISE.
(a) Verify that the examples of §2.1 are indeed sheaves (of differentiable functions, or continuous functions, or smooth functions, or functions on a manifold or $\mathbb{R}^n$).
(b) Show that real-valued continuous functions on (open sets of) a topological space X form a sheaf.

2.2.8. *Important Example: Restriction of a sheaf.* Suppose $\mathscr{F}$ is a sheaf on X, and U is an open subset of X. Define the **restriction of $\mathscr{F}$ to U**, denoted $\mathscr{F}|_U$, to be the collection $\mathscr{F}|_U(V) = \mathscr{F}(V)$ for all open subsets $V \subset U$. Clearly this is a sheaf on U. (Unimportant but fun fact: §2.6 will tell us how to restrict sheaves to arbitrary subsets.)

2.2.9. *Important Example: skyscraper sheaf.* Suppose X is a topological space, with $p \in X$, and S is a set. Let $i_p : p \to X$ be the inclusion. Then $i_{p,*}S$ defined by

$$i_{p,*}S(U) = \begin{cases} S & \text{if } p \in U, \text{and} \\ \{e\} & \text{if } p \notin U \end{cases}$$

forms a sheaf. Here $\{e\}$ is any one-element set. (Check this if it isn't clear to you — what are the restriction maps?) This is called a **skyscraper sheaf**, because the informal picture of it looks like a skyscraper at p. (Mild caution: this informal picture suggests that the only nontrivial stalk of a skyscraper sheaf is at p, which isn't the case. Exercise 13.2.A(b) gives an example, although it certainly isn't the simplest one.) There is an analogous definition for sheaves of abelian groups, except $i_{p,*}(S)(U) = \{0\}$ if $p \notin U$; and for sheaves with values in a category more generally, $i_{p,*}S(U)$ should be a final object.

(This notation is admittedly hideous, and the alternative $(i_p)_*S$ is equally bad. §2.2.12 explains this notation.)

2.2.10. *Constant presheaves and constant sheaves.* Let X be a topological space, and S a set. Define $\underline{S}^{\text{pre}}(U) = S$ for all open sets U. You will readily verify that $\underline{S}^{\text{pre}}$ forms a presheaf (with restriction maps the identity). This is called the **constant presheaf associated to** S. This isn't (in general) a sheaf. (It may be distracting to say why. Lovers of the empty set will insist that the sheaf axioms force the sections over the empty set to be the final object in the category, i.e., a one-element set. But even if we patch the definition by setting $\underline{S}^{\text{pre}}(\varnothing) = \{e\}$, if S has more than one element, and X is the two-point space with the **discrete topology**, i.e., where every subset is open, you can check that $\underline{S}^{\text{pre}}$ fails gluability.)

2.2.E. EXERCISE (CONSTANT SHEAVES). Now let $\mathscr{F}(U)$ be the maps to S that are *locally constant*, i.e., for any point p in U, there is a neighborhood of p where the function is constant. Show that this is a *sheaf*. (A better description is this: endow S with the discrete topology, and let $\mathscr{F}(U)$ be the continuous maps $U \to S$.) This is called the **constant sheaf** (associated to S); do not confuse it with the constant presheaf. We denote this sheaf $\underline{S}$.

2.2.F. EXERCISE ("MORPHISMS GLUE"). Suppose Y is a topological space. Show that "continuous maps to Y" form a sheaf of sets on X. More precisely, to each open set U of X, we associate the set of continuous maps of U to Y. Show that this forms a sheaf. (Exercise 2.2.D(b), with $Y = \mathbb{R}$, and Exercise 2.2.E, with $Y = S$ with the discrete topology, are both special cases.)

2.2.G. EXERCISE. This is a fancier version of the previous exercise.
(a) (sheaf of sections of a map) Suppose we are given a continuous map $\mu : Y \to X$. Show that "sections of μ" form a sheaf. More precisely, to each open set U of X, associate the set of continuous maps $s : U \to Y$ such that $\mu \circ s = \mathrm{id}|_U$. Show that this forms a sheaf. (For those who have heard of vector bundles, these are a good

example.) This is motivation for the phrase "section of a sheaf".
(b) (This exercise is for those who know what a topological group is. If you don't know what a topological group is, you might be able to guess.) Suppose that Y is a topological group. Show that continuous maps to Y form a sheaf of *groups*.

2.2.11. $\star$ *The space of sections (espace étalé) of a (pre)sheaf.* Depending on your background, you may prefer the following perspective on sheaves. Suppose $\mathscr{F}$ is a presheaf (e.g. a sheaf) on a topological space X. Construct a topological space F along with a continuous map $\pi : F \to X$ as follows: as a set, F is the disjoint union of all the stalks of $\mathscr{F}$. This naturally gives a map of sets $\pi : F \to X$. Topologize F as follows. Each s in $\mathscr{F}(U)$ determines a subset $\{(x, s_x) : x \in U\}$ of F. The topology on F is the weakest topology such that these subsets are open. (These subsets form a base of the topology. For each $y \in F$, there is a neighborhood V of y and a neighborhood U of $\pi(y)$ such that $\pi|_V$ is a homeomorphism from V to U. Do you see why these facts are true?) The topological space F could be thought of as the **space of sections** of $\mathscr{F}$ (and in french is called the **espace étalé** of $\mathscr{F}$). We will not discuss this construction at any length, but it can have some advantages: (a) It is always better to know as many ways as possible of thinking about a concept. (b) Pullback has a natural interpretation in this language (mentioned briefly in Exercise 2.6.C). (c) Sheafification has a natural interpretation in this language (see Remark 2.4.8).

2.2.H. IMPORTANT EXERCISE: THE PUSHFORWARD SHEAF OR DIRECT IMAGE SHEAF. Suppose $\pi : X \to Y$ is a continuous map, and $\mathscr{F}$ is a presheaf on X. Then define $\pi_*\mathscr{F}$ by $\pi_*\mathscr{F}(V) = \mathscr{F}(\pi^{-1}(V))$, where V is an open subset of Y. Show that $\pi_*\mathscr{F}$ is a presheaf on Y, and is a sheaf if $\mathscr{F}$ is. This is called the **direct image** or **pushforward** of $\mathscr{F}$. More precisely, $\pi_*\mathscr{F}$ is called the **pushforward of $\mathscr{F}$ by** π.

2.2.12. As the notation suggests, the skyscraper sheaf (Example 2.2.9) can be interpreted as the pushforward of the constant sheaf $\underline{S}$ on a one-point space p, under the inclusion morphism $i : \{p\} \to X$.

Once we realize that sheaves form a category, we will see that the pushforward is a functor from sheaves on X to sheaves on Y (Exercise 2.3.B).

2.2.I. EXERCISE (PUSHFORWARD INDUCES MAPS OF STALKS). Suppose $\pi : X \to Y$ is a continuous map, and $\mathscr{F}$ is a sheaf of sets (or rings or A-modules) on X. If $\pi(p) = q$, describe the natural morphism of stalks $(\pi_*\mathscr{F})_q \to \mathscr{F}_p$. (You can use the explicit definition of stalk using representatives, §2.2.4, or the universal property, §2.2.5. If you prefer one way, you should try the other.)

2.2.13. Important Example: Ringed spaces, and $\mathscr{O}_X$-modules. Suppose $\mathscr{O}_X$ is a sheaf of rings on a topological space X (i.e., a sheaf on X with values in the category of *Rings*). Then $(X, \mathscr{O}_X)$ is called a **ringed space**. The sheaf of rings is often denoted by $\mathscr{O}_X$, pronounced "oh-X". This sheaf is called the **structure sheaf** of the ringed space. Sections of the structure sheaf $\mathscr{O}_X$ over an open subset U are called **functions on** U. *(Caution: what we call "functions", others sometimes call "regular functions". Furthermore, we will later define "rational functions" on schemes in §5.5.6, which are not precisely functions in this sense; they are a particular type of "partially-defined function".)* The symbol $\mathscr{O}_X$ will always refer to the structure sheaf of a ringed space. (Note: the stalk of $\mathscr{O}_X$ at a point p is written "$\mathscr{O}_{X,p}$", because this looks less hideous than "$\mathscr{O}_{X_p}$".)

Just as we have modules over a ring, we have $\mathscr{O}_X$-modules over a sheaf of rings $\mathscr{O}_X$. There is only one possible definition that could go with the name **$\mathscr{O}_X$-module** — a sheaf of abelian groups $\mathscr{F}$ with the following additional structure. For each U, $\mathscr{F}(U)$ is an $\mathscr{O}_X(U)$-module. Furthermore, this structure should behave well with respect to restriction maps: if $U \subset V$, then

$$\begin{tikzcd} \mathscr{O}_X(V) \times \mathscr{F}(V) \arrow[r, "\text{action}"] \arrow[d, "\mathrm{res}_{V,U} \times \mathrm{res}_{V,U}"'] & \mathscr{F}(V) \arrow[d, "\mathrm{res}_{V,U}"] \\ \mathscr{O}_X(U) \times \mathscr{F}(U) \arrow[r, "\text{action}"] & \mathscr{F}(U) \end{tikzcd} \tag{2.2.13.1}$$

commutes. (You should convince yourself that I haven't forgotten anything.)

Recall that the notion of A-module generalizes the notion of abelian group, because an abelian group is the same thing as a $\mathbb{Z}$-module. Similarly, the notion of $\mathscr{O}_X$-module generalizes the notion of sheaf of abelian groups, because the latter is the same thing as a $\underline{\mathbb{Z}}$-module, where $\underline{\mathbb{Z}}$ is the constant sheaf associated to $\mathbb{Z}$. Hence when we are proving things about $\mathscr{O}_X$-modules, we are also proving things about sheaves of abelian groups.

2.2.J. EXERCISE. If $(X, \mathscr{O}_X)$ is a ringed space, and $\mathscr{F}$ is an $\mathscr{O}_X$-module, describe how for each $p \in X$, $\mathscr{F}_p$ is an $\mathscr{O}_{X,p}$-module.

2.2.14. *For those who know about vector bundles.* The motivating example of $\mathscr{O}_X$-modules is the sheaf of sections of a vector bundle. If $(X, \mathscr{O}_X)$ is a differentiable manifold (so $\mathscr{O}_X$ is the sheaf of differentiable functions), and $\pi : V \to X$ is a vector bundle over X, then the sheaf of differentiable sections $\sigma : X \to V$ is an $\mathscr{O}_X$-module. Indeed, given a section s of π over an open subset $U \subset X$, and a function f on U, we can multiply s by f to get a new section fs of π over U. Moreover, if V is a smaller subset, then we could multiply f by s and then restrict to V, or we could restrict both f and s to V and then multiply, and we would get the same answer. That is precisely the commutativity of (2.2.13.1).

2.3 Morphisms of presheaves and sheaves

2.3.1. Whenever one defines a new mathematical object, category theory teaches to try to understand maps between them. We now define morphisms of presheaves, and similarly for sheaves. In other words, we will describe the *category of presheaves* (of sets, abelian groups, etc.) and the *category of sheaves*.

A **morphism of presheaves** of sets (or indeed of presheaves with values in any category) on X, $\phi : \mathscr{F} \to \mathscr{G}$, is the data of maps $\phi(U) : \mathscr{F}(U) \to \mathscr{G}(U)$ for all U behaving well with respect to restriction: if $U \hookrightarrow V$ then

$$\begin{tikzcd} \mathscr{F}(V) \arrow[r, "\phi(V)"] \arrow[d, "\mathrm{res}_{V,U}"'] & \mathscr{G}(V) \arrow[d, "\mathrm{res}_{V,U}"] \\ \mathscr{F}(U) \arrow[r, "\phi(U)"] & \mathscr{G}(U) \end{tikzcd}$$

commutes. (Notice: the underlying space of both $\mathscr{F}$ and $\mathscr{G}$ is X.)

Morphisms of sheaves are defined identically: the morphisms from a sheaf $\mathscr{F}$ to a sheaf $\mathscr{G}$ are precisely the morphisms from $\mathscr{F}$ to $\mathscr{G}$ as presheaves. (Translation: The category of sheaves on X is a full subcategory of the category of presheaves on X.) If $(X, \mathscr{O}_X)$ is a ringed space, then morphisms of $\mathscr{O}_X$-modules have the obvious definition. (Can you write it down?)

An example of a morphism of sheaves is the map from the sheaf of differentiable functions on $\mathbb{R}$ to the sheaf of continuous functions. This is a "forgetful map": we are forgetting that these functions are differentiable, and remembering only that they are continuous.

We may as well set some notation: let $Sets_X$, Ab_X, etc. denote the category of sheaves of sets, abelian groups, etc. on a topological space X. Let $Mod_{\mathscr{O}_X}$ denote the category of $\mathscr{O}_X$-modules on a ringed space $(X, \mathscr{O}_X)$. Let $Sets_X^{\text{pre}}$, etc. denote the category of presheaves of sets, etc. on X.

2.3.2. *Aside for category-lovers.* If you interpret a presheaf on X as a contravariant functor (from the category of open sets), a morphism of presheaves on X is a natural transformation of functors (§1.2.21).

2.3.A. EXERCISE: MORPHISMS OF (PRE)SHEAVES INDUCE MORPHISMS OF STALKS. If $\phi : \mathscr{F} \to \mathscr{G}$ is a morphism of presheaves on X, and $p \in X$, describe an induced morphism of stalks $\phi_p : \mathscr{F}_p \to \mathscr{G}_p$. Translation: taking the stalk at p induces a functor $Sets_X \to Sets$. (Your proof will extend in obvious ways. For example, if ϕ is a morphism of $\mathscr{O}_X$-modules, then ϕ_p is a map of $\mathscr{O}_{X,p}$-modules.)

2.3.B. EXERCISE. Suppose $\pi : X \to Y$ is a continuous map of topological spaces (i.e., a morphism in the category of topological spaces). Show that pushforward gives a functor $\pi_* : Sets_X \to Sets_Y$. Here *Sets* can be replaced by other categories. (Watch out for some possible confusion: a presheaf is a functor, and presheaves form a category. It may be best to forget that presheaves are functors for now.)

2.3.C. IMPORTANT EXERCISE AND DEFINITION: "SHEAF Hom". Suppose $\mathscr{F}$ and $\mathscr{G}$ are two sheaves of sets on X. (In fact, it will suffice that $\mathscr{F}$ is a presheaf.) Let $\mathcal{H}om(\mathscr{F}, \mathscr{G})$ be the collection of data

$$\mathcal{H}om(\mathscr{F}, \mathscr{G})(U) := \operatorname{Mor}(\mathscr{F}|_U, \mathscr{G}|_U).$$

(Recall the notation $\mathscr{F}|_U$, the restriction of the sheaf to the open set U, Example 2.2.8.) Show that this is a sheaf of sets on X. This is called "sheaf $\mathcal{H}om$". (Strictly speaking, we should reserve Hom for when we are in an additive category, so this should possibly be called "sheaf Mor". But the terminology "sheaf Hom" is too established to uproot.) It will be clear from your construction that, like Hom, $\mathcal{H}om$ is a contravariant functor in its first argument and a covariant functor in its second argument.

Warning: $\mathcal{H}om$ does not commute with taking stalks. More precisely: it is not true that $\mathcal{H}om(\mathscr{F}, \mathscr{G})_p$ is isomorphic to $\operatorname{Hom}(\mathscr{F}_p, \mathscr{G}_p)$. (Can you think of a counterexample? There is at least a map from one of these to the other — in which direction?)

2.3.3. We will use many variants of the definition of $\mathcal{H}om$. For example, if $\mathscr{F}$ and $\mathscr{G}$ are sheaves of abelian groups on X, then $\mathcal{H}om_{Ab_X}(\mathscr{F}, \mathscr{G})$ is defined by taking

$\mathcal{H}om_{Ab_X}(\mathscr{F},\mathscr{G})(U)$ to be the maps *as sheaves of abelian groups* $\mathscr{F}|_U \to \mathscr{G}|_U$. (Note that $\mathcal{H}om_{Ab_X}(\mathscr{F},\mathscr{G})$ has the structure of a sheaf of abelian groups in a natural way.) Similarly, if $\mathscr{F}$ and $\mathscr{G}$ are $\mathscr{O}_X$-modules, we define $\mathcal{H}om_{Mod_{\mathscr{O}_X}}(\mathscr{F},\mathscr{G})$ in the analogous way (and it is an $\mathscr{O}_X$-module). Obnoxiously, the subscripts Ab_X and $Mod_{\mathscr{O}_X}$ are often dropped (here and in the literature), so be careful which category you are working in! We call $\mathcal{H}om_{Mod_{\mathscr{O}_X}}(\mathscr{F},\mathscr{O}_X)$ the *dual* of the $\mathscr{O}_X$-module $\mathscr{F}$, and denote it $\mathscr{F}^\vee$.

2.3.D. UNIMPORTANT EXERCISE (REALITY CHECK).

(a) If $\mathscr{F}$ is a sheaf of sets on X, then show that $\mathcal{H}om(\underline{\{p\}},\mathscr{F}) \cong \mathscr{F}$, where $\underline{\{p\}}$ is the constant sheaf associated to the one element set $\{p\}$.
(b) If $\mathscr{F}$ is a sheaf of abelian groups on X, then show that $\mathcal{H}om_{Ab_X}(\underline{\mathbb{Z}},\mathscr{F}) \cong \mathscr{F}$ (an isomorphism of sheaves of abelian groups).
(c) If $\mathscr{F}$ is an $\mathscr{O}_X$-module, then show that $\mathcal{H}om_{Mod_{\mathscr{O}_X}}(\mathscr{O}_X,\mathscr{F}) \cong \mathscr{F}$ (an isomorphism of $\mathscr{O}_X$-modules).

A key idea in (b) and (c) is that 1 "generates" (in some sense) $\underline{\mathbb{Z}}$ (in (b)) and $\mathscr{O}_X$ (in (c)).

2.3.4. Presheaves of abelian groups (and even "presheaf $\mathscr{O}_X$-modules") form an abelian category.

We can make module-like constructions using presheaves of abelian groups on a topological space X. (Throughout this section, all (pre)sheaves are of abelian groups.) For example, we can clearly add maps of presheaves and get another map of presheaves: if $\phi,\psi : \mathscr{F} \to \mathscr{G}$, then we define the map $f+g$ by $(\phi+\psi)(V) = \phi(V) + \psi(V)$. (There is something small to check here: that the result is indeed a map of presheaves.) In this way, presheaves of abelian groups form an additive category (Definition 1.6.1: the morphisms between any two presheaves of abelian groups form an abelian group; there is a 0-object; and one can take finite products). For exactly the same reasons, sheaves of abelian groups also form an additive category.

If $\phi : \mathscr{F} \to \mathscr{G}$ is a morphism of presheaves, define the **presheaf kernel** $\ker_{pre}\phi$ by $(\ker_{pre}\phi)(U) = \ker\phi(U)$.

2.3.E. EXERCISE. Show that $\ker_{pre}\phi$ is a presheaf. (Hint: if $U \hookrightarrow V$, define the restriction map by chasing the following diagram:

$$\begin{tikzcd} 0 \arrow[r] & \ker_{pre}\phi(V) \arrow[r] \arrow[d, dotted, "\exists!"] & \mathscr{F}(V) \arrow[r] \arrow[d, "res_{V,U}"] & \mathscr{G}(V) \arrow[d, "res_{V,U}"] \\ 0 \arrow[r] & \ker_{pre}\phi(U) \arrow[r] & \mathscr{F}(U) \arrow[r] & \mathscr{G}(U) \end{tikzcd}$$

You should check that the restriction maps compose as desired.)

Define the **presheaf cokernel** $\operatorname{coker}_{pre}\phi$ similarly. It is a presheaf by essentially the same (dual) argument.

2.3.F. EXERCISE: THE COKERNEL DESERVES ITS NAME. Show that the presheaf cokernel satisfies the universal property of cokernels (Definition 1.6.3) in the category of presheaves.

Similarly, $\ker_{\text{pre}} \phi \to \mathscr{F}$ satisfies the universal property for kernels in the category of presheaves.

It is not too tedious to verify that presheaves of abelian groups form an abelian category, and the reader is free to do so. The key idea is that all abelian-categorical notions may be defined and verified "open set by open set". We needn't worry about restriction maps — they "come along for the ride". Hence we can define terms such as **subpresheaf**, **image presheaf** (or **presheaf image**), and **quotient presheaf** (or **presheaf quotient**), and they behave as you would expect. You construct kernels, quotients, cokernels, and images open set by open set. Homological algebra (exact sequences and so forth) works, and also "works open set by open set". In particular:

2.3.G. EASY EXERCISE. Show (or observe) that for a topological space X with open set U, $\mathscr{F} \mapsto \mathscr{F}(U)$ gives a functor from presheaves of abelian groups on X, Ab_X^{pre}, to abelian groups, *Ab*. Then show that this functor is exact.

2.3.H. EXERCISE. Show that a sequence of presheaves $0 \to \mathscr{F}_1 \to \mathscr{F}_2 \to \cdots \to \mathscr{F}_n \to 0$ is exact if and only if $0 \to \mathscr{F}_1(U) \to \mathscr{F}_2(U) \to \cdots \to \mathscr{F}_n(U) \to 0$ is exact for all U.

The above discussion essentially carries over without change to presheaves with values in any abelian category. (Think this through if you wish.)

However, we are interested in more geometric objects, sheaves, where things can be understood in terms of their local behavior, thanks to the identity and gluing axioms. We will soon see that sheaves of abelian groups also form an abelian category, but a complication will arise that will force the notion of *sheafification* on us. Sheafification will be the answer to many of our prayers. We just haven't yet realized what we should be praying for.

To begin with, sheaves Ab_X form an additive category, as described in the first paragraph of §2.3.4.

Kernels work just as with presheaves:

2.3.I. IMPORTANT EXERCISE. Suppose $\phi : \mathscr{F} \to \mathscr{G}$ is a morphism of *sheaves*. Show that the presheaf kernel $\ker_{\text{pre}} \phi$ is in fact a sheaf. Show that it satisfies the universal property of kernels (Definition 1.6.3). (Hint: the second question follows immediately from the fact that $\ker_{\text{pre}} \phi$ satisfies the universal property in the category of *presheaves*.)

Thus if ϕ is a morphism of sheaves, we define

$$\ker \phi := \ker_{\text{pre}} \phi.$$

The problem arises with the cokernel.

2.3.J. IMPORTANT EXERCISE. Let X be $\mathbb{C}$ with the classical topology, let $\underline{\mathbb{Z}}$ be the constant sheaf on X associated to $\mathbb{Z}$, $\mathscr{O}_X$ the sheaf of holomorphic functions, and $\mathscr{F}$ the *presheaf* of functions admitting a holomorphic logarithm. Describe an exact sequence of presheaves on X:

$$0 \longrightarrow \underline{\mathbb{Z}} \longrightarrow \mathscr{O}_X \longrightarrow \mathscr{F} \longrightarrow 0$$

where $\underline{\mathbb{Z}} \to \mathscr{O}_X$ is the natural inclusion and $\mathscr{O}_X \to \mathscr{F}$ is given by $f \mapsto \exp(2\pi i f)$. (Be sure to verify exactness.) Show that $\mathscr{F}$ is *not* a sheaf. (Hint: $\mathscr{F}$ does not satisfy the gluability axiom. The problem is that there are functions that don't have a logarithm but locally have a logarithm.) This will come up again in Example 2.4.10.

We will have to put our hopes for understanding cokernels of sheaves on hold for a while. We will first learn to understand sheaves using stalks.

2.4 Properties determined at the level of stalks, and sheafification

2.4.1. Properties determined by stalks. We now come to the second way of understanding sheaves mentioned at the start of the chapter. In this section, we will see that lots of facts about sheaves can be checked "at the level of stalks". This isn't true for presheaves, and reflects the local nature of sheaves. We will see that sections and morphisms are determined "by their stalks", and the property of a morphism being an isomorphism may be checked at stalks. (The last one is the trickiest.)

2.4.A. IMPORTANT EASY EXERCISE **(sections are determined by germs)**. Prove that a section of a sheaf of sets is determined by its germs, i.e., the natural map

$$\mathscr{F}(U) \to \prod_{p \in U} \mathscr{F}_p \tag{2.4.1.1}$$

is injective. Hint 1: you won't use the gluability axiom, so this is true for separated presheaves. Hint 2: it is false for presheaves in general, see Exercise 2.4.F, so you *will* use the identity axiom. (Your proof will also apply to sheaves of groups, rings, etc. — to categories of "sets with additional structure". The same is true of many exercises in this section.)

2.4.2. *Definition: support of a section.* This motivates a concept we will find useful later. Suppose $\mathscr{F}$ is a sheaf (or indeed separated presheaf) of abelian groups on X, and s is a global section of $\mathscr{F}$. Then let the **support of** s, denoted $\mathrm{Supp}(s)$, be the points p of X where s has a nonzero germ:

$$\mathrm{Supp}\, s := \{p \in X \;:\; s_p \neq 0 \text{ in } \mathscr{F}_p\}.$$

We think of this as the subset of X where "the section s lives" — the complement is the locus where s is the 0-section. We could define this even if $\mathscr{F}$ is a presheaf, but without the inclusion of Exercise 2.4.A, we could have the strange situation where we have a nonzero section that "lives nowhere" (because it is 0 "near every point", i.e., is 0 in every stalk).

2.4.B. EXERCISE (THE SUPPORT OF A SECTION IS CLOSED). Show that $\mathrm{Supp}(s)$ is a closed subset of X.

Exercise 2.4.A suggests an important question: which elements of the right side of (2.4.1.1) are in the image of the left side?

2.4.3. Important definition. We say that an element $\prod_{p \in U} s_p$ of the right side $\prod_{p \in U} \mathscr{F}_p$ of (2.4.1.1) consists of **compatible germs** if for all $p \in U$, there is some representative $(U_p, s'_p \in \mathscr{F}(U_p))$ for s_p (where $p \in U_p \subset U$) such that the germ of

s'_p at all $y \in U_p$ is s_y. You will have to think about this a little. Clearly any section s of $\mathscr{F}$ over U gives a choice of compatible germs for U — take $(U_p, s'_p) = (U, s)$.

2.4.C. IMPORTANT EXERCISE. Prove that any choice of compatible germs for a sheaf of sets $\mathscr{F}$ over U is the image of a section of $\mathscr{F}$ over U. (Hint: you will use gluability.)

We have thus completely described the image of (2.4.1.1), in a way that we will find useful.

2.4.4. *Remark.* This perspective motivates the agricultural terminology "sheaf": it is (the data of) a bunch of stalks, bundled together appropriately.

Now we throw morphisms into the mix. Recall Exercise 2.3.A: morphisms of (pre)sheaves induce morphisms of stalks.

2.4.D. EXERCISE **(morphisms are determined by stalks)**. If ϕ_1 and ϕ_2 are morphisms from a presheaf of sets $\mathscr{F}$ to a sheaf of sets $\mathscr{G}$ that induce the same maps on each stalk, show that $\phi_1 = \phi_2$. Hint: consider the following diagram.

(2.4.4.1)
$$\begin{tikzcd} \mathscr{F}(U) \arrow[r] \arrow[d] & \mathscr{G}(U) \arrow[d, hook] \\ \prod_{p \in U} \mathscr{F}_p \arrow[r] & \prod_{p \in U} \mathscr{G}_p \end{tikzcd}$$

2.4.E. TRICKY EXERCISE **(isomorphisms are determined by stalks)**. Show that a morphism of sheaves of sets is an isomorphism if and only if it induces an isomorphism of all stalks. Hint: Use (2.4.4.1). Once you have injectivity, show surjectivity, perhaps using Exercise 2.4.C, or gluability in some other way; this is more subtle. Note: this question does *not* say that if two sheaves have isomorphic stalks, then they are isomorphic.

2.4.F. EXERCISE.
(a) Show that Exercise 2.4.A is false for general presheaves.
(b) Show that Exercise 2.4.D is false for general presheaves.
(c) Show that Exercise 2.4.E is false for general presheaves.
(General hint for finding counterexamples of this sort: consider a 2-point space with the discrete topology.)

2.4.5. Sheafification.

Every sheaf is a presheaf (and indeed by definition sheaves on X form a full subcategory of the category of presheaves on X). Just as groupification (§1.5.3) gives an abelian group that best approximates an abelian semigroup, sheafification gives the sheaf that best approximates a presheaf, with an analogous universal property. (One possible example to keep in mind is the sheafification of the presheaf of holomorphic functions admitting a square root on $\mathbb{C}$ with the classical topology. See also the exponential exact sequence, Exercise 2.4.10.)

2.4.6. *Definition.* If $\mathscr{F}$ is a presheaf on X, then a morphism of presheaves $\mathrm{sh} : \mathscr{F} \to \mathscr{F}^{\mathrm{sh}}$ on X is a **sheafification of** $\mathscr{F}$ if $\mathscr{F}^{\mathrm{sh}}$ is a sheaf, and for any sheaf $\mathscr{G}$, and any presheaf morphism $g : \mathscr{F} \to \mathscr{G}$, there *exists* a *unique* morphism of sheaves

$f : \mathscr{F}^{sh} \to \mathscr{G}$ making the diagram

$$\begin{tikzcd} \mathscr{F} \arrow[r, "sh"] \arrow[dr, "g"'] & \mathscr{F}^{sh} \arrow[d, "f"] \\ & \mathscr{G} \end{tikzcd}$$

commute.

We still have to show that it exists. The following two exercises require existence (which we will show shortly), but not the details of the construction.

2.4.G. EXERCISE. Show that sheafification is unique up to unique isomorphism, assuming it exists. Show that if $\mathscr{F}$ is a sheaf, then the sheafification is $\mathscr{F} \xrightarrow{\text{id}} \mathscr{F}$. (This should be second nature by now.)

2.4.H. EASY EXERCISE (SHEAFIFICATION IS A FUNCTOR). Assume for now that sheafification exists. Use the universal property to show that for any morphism of presheaves $\phi : \mathscr{F} \to \mathscr{G}$, we get a natural induced morphism of sheaves $\phi^{sh} : \mathscr{F}^{sh} \to \mathscr{G}^{sh}$. Show that sheafification is a functor from presheaves on X to sheaves on X.

2.4.7. *Construction.* We next show that any presheaf of sets (or groups, rings, etc.) has a sheafification. Suppose $\mathscr{F}$ is a *presheaf*. Define $\mathscr{F}^{sh}$ by defining $\mathscr{F}^{sh}(U)$ as the set of compatible germs of the presheaf $\mathscr{F}$ over U. Explicitly:

$$\begin{aligned} \mathscr{F}^{sh}(U) \quad := \quad & \{(f_x \in \mathscr{F}_x)_{x \in U} : \text{for all } x \in U, \text{ there exists } x \in V \subset U \text{ and } s \in \mathscr{F}(V) \\ & \text{with } s_y = f_y \text{ for all } y \in V\}. \end{aligned}$$

Here s_y means the image of s in the stalk $\mathscr{F}_y$. (Those who want to worry about the empty set are welcome to.)

2.4.I. EASY EXERCISE. Show that $\mathscr{F}^{sh}$ (using the tautological restriction maps) forms a sheaf.

2.4.J. EASY EXERCISE. Describe a natural map of presheaves $sh : \mathscr{F} \to \mathscr{F}^{sh}$.

2.4.K. EXERCISE. Show that the map sh satisfies the universal property of sheafification (Definition 2.4.6). (This is easier than you might fear.)

2.4.L. USEFUL EXERCISE, NOT JUST FOR CATEGORY-LOVERS. Show that the sheafification functor is left-adjoint to the forgetful functor from sheaves on X to presheaves on X. This is not difficult — it is largely a restatement of the universal property. But it lets you use results from §1.6.12, and can "explain" why you don't need to sheafify when taking kernel (why the presheaf kernel is already the sheaf kernel), and why you need to sheafify when taking cokernel and (soon, in Exercise 2.5.J) $\otimes$.

2.4.M. EXERCISE. Show $\mathscr{F} \to \mathscr{F}^{sh}$ induces an isomorphism of stalks. (Possible hint: Use the concrete description of the stalks. Another possibility once you read Remark 2.6.3: judicious use of adjoints.)

As a reality check, you may want to verify that "the sheafification of a constant presheaf is the corresponding constant sheaf" (see §2.2.10): if X is a topological space and S is a set, then $(\underline{S}^{\mathrm{pre}})^{\mathrm{sh}}$ may be naturally identified with $\underline{S}$.

2.4.8. ⋆ *Remark.* The "space of sections" (or "espace étalé") construction (§2.2.11) yields a different-sounding description of sheafification which may be preferred by some readers. The main idea is identical: if $\mathscr{F}$ is a presheaf, let F be the space of sections (or espace étalé) of $\mathscr{F}$. You may wish to show that if $\mathscr{F}$ is a presheaf, the sheaf of sections of $F \to X$ (defined in Exercise 2.2.G(a)) is the sheafification of $\mathscr{F}$. Exercise 2.2.E may be interpreted as an example of this construction. The "space of sections" construction of the sheafification is essentially the same as Construction 2.4.7.

2.4.9. Subsheaves and quotient sheaves.

We now discuss subsheaves and quotient sheaves from the perspective of stalks.

2.4.N. EXERCISE. Suppose $\phi : \mathscr{F} \to \mathscr{G}$ is a morphism of sheaves of sets on a topological space X. Show that the following are equivalent.

(a) ϕ is a monomorphism in the category of sheaves.
(b) ϕ is injective on the level of stalks: $\phi_p : \mathscr{F}_p \to \mathscr{G}_p$ is injective for all $p \in X$.
(c) ϕ is injective on the level of open sets: $\phi(U) : \mathscr{F}(U) \to \mathscr{G}(U)$ is injective for all open $U \subset X$.

(Possible hints: for (b) implies (a), recall that morphisms are determined by stalks, Exercise 2.4.D. For (a) implies (c), use the "indicator sheaf" with one section over every open set contained in U, and no section over any other open set.) If these conditions hold, we say that $\mathscr{F}$ is a **subsheaf** of $\mathscr{G}$ (where the "inclusion" ϕ is sometimes left implicit).

(You may later wish to extend your solution to Exercise 2.4.N to show that for any morphism of *presheaves*, if all maps of sections are injective, then all stalk maps are injective. And furthermore, if $\phi : \mathscr{F} \to \mathscr{G}$ is a morphism from a separated presheaf to an arbitrary presheaf, then injectivity on the level of stalks implies that ϕ is a monomorphism in the category of presheaves. This is useful in some approaches to Exercise 2.5.C.)

2.4.O. EXERCISE. Continuing the notation of the previous exercise, show that the following are equivalent.

(a) ϕ is an epimorphism in the category of sheaves.
(b) ϕ is surjective on the level of stalks: $\phi_p : \mathscr{F}_p \to \mathscr{G}_p$ is surjective for all $p \in X$.

(Possible hint: use a skyscraper sheaf.)

If these conditions hold, we say that $\mathscr{G}$ is a **quotient sheaf** of $\mathscr{F}$.

Thus *monomorphisms and epimorphisms — subsheafiness and quotient sheafiness — can be checked at the level of stalks.*

Both exercises generalize readily to sheaves with values in any reasonable category, where "injective" is replaced by "monomorphism" and "surjective" is replaced by "epimorphism".

Notice that there was no part (c) to Exercise 2.4.O, and Example 2.4.10 shows why. (But there is a version of (c) that *implies* (a) and (b): surjectivity on all open sets in the base of a topology implies that the corresponding map of sheaves is an epimorphism, Exercise 2.7.E.)

2.4.10. *Example (cf. Exercise 2.3.J).* Let $X = \mathbb{C}$ with the classical topology, and define $\mathscr{O}_X$ to be the sheaf of holomorphic functions, and $\mathscr{O}_X^*$ to be the sheaf of invertible (nowhere zero) holomorphic functions. This is a sheaf of abelian groups under multiplication. We have maps of sheaves

$$0 \longrightarrow \underline{\mathbb{Z}} \xrightarrow{\times 2\pi i} \mathscr{O}_X \xrightarrow{\exp} \mathscr{O}_X^* \longrightarrow 1 \tag{2.4.10.1}$$

where $\underline{\mathbb{Z}}$ is the constant sheaf associated to $\mathbb{Z}$. (You can figure out what the sheaves 0 and 1 mean; they are isomorphic, and are written in this way for reasons that may be clear.) We will soon interpret this as an exact sequence of sheaves of abelian groups (the *exponential exact sequence*, see Exercise 2.5.E), although we don't yet have the language to do so.

2.4.P. ENLIGHTENING EXERCISE. Show that $\mathscr{O}_X \xrightarrow{\exp} \mathscr{O}_X^*$ describes $\mathscr{O}_X^*$ as a quotient sheaf of $\mathscr{O}_X$. Find an open set on which this map is not surjective.

This is a great example to get a sense of what "surjectivity" means for sheaves: nowhere vanishing holomorphic functions have logarithms locally, but they need not globally.

2.5 Sheaves of abelian groups, and $\mathscr{O}_X$-modules, form abelian categories

We are now ready to see that sheaves of abelian groups, and their cousins, $\mathscr{O}_X$-modules, form abelian categories. In other words, we may treat them similarly to vector spaces, and modules over a ring. In the process of doing this, we will see that this is much stronger than an analogy; kernels, cokernels, exactness, and so forth can be understood at the level of germs (which are just abelian groups), and the compatibility of the germs will come for free.

The category of sheaves of abelian groups is clearly an additive category (Definition 1.6.1). In order to show that it is an abelian category, we must begin by showing that any morphism $\phi : \mathscr{F} \to \mathscr{G}$ has a kernel and a cokernel. We have already seen that ϕ has a kernel (Exercise 2.3.I): the presheaf kernel is a sheaf, and is a kernel in the category of sheaves.

2.5.A. EXERCISE. Show that the stalk of the kernel is the kernel of the stalks: there is a natural isomorphism

$$(\ker(\mathscr{F} \to \mathscr{G}))_x \cong \ker(\mathscr{F}_x \to \mathscr{G}_x).$$

We next address the issue of the cokernel. Now $\phi : \mathscr{F} \to \mathscr{G}$ has a cokernel in the category of presheaves; call it $\mathscr{H}^{\mathrm{pre}}$ (where the superscript is meant to remind us that this is a presheaf). Let $\mathscr{H}^{\mathrm{pre}} \xrightarrow{\mathrm{sh}} \mathscr{H}$ be its sheafification. Recall that the cokernel is defined using a universal property: it is the colimit of the diagram

$$\begin{tikzcd} \mathscr{F} \arrow[r, "\phi"] \arrow[d] & \mathscr{G} \\ 0 & \end{tikzcd} \tag{2.5.0.2}$$

in the category of presheaves (cf. (1.6.3.1) and the comment thereafter).

2.5.1. Proposition. — *The composition $\mathscr{G} \to \mathscr{H}$ is the cokernel of ϕ in the category of sheaves.*

Proof. We show that it satisfies the universal property. Given any sheaf $\mathscr{E}$ and a commutative diagram

$$\begin{tikzcd} \mathscr{F} \arrow[r, "\phi"] \arrow[d] & \mathscr{G} \arrow[d] \\ 0 \arrow[r] & \mathscr{E} \end{tikzcd}$$

We construct

$$\begin{tikzcd} \mathscr{F} \arrow[r, "\phi"] \arrow[d] & \mathscr{G} \arrow[d] \arrow[drr, bend left] & & \\ 0 \arrow[r] \arrow[drrr] & \mathscr{H}^{\mathrm{pre}} \arrow[r, "\mathrm{sh}"] & \mathscr{H} & \\ & & & \mathscr{E} \end{tikzcd}$$

We show that there is a unique morphism $\mathscr{H} \to \mathscr{E}$ making the diagram commute. As $\mathscr{H}^{\mathrm{pre}}$ is the cokernel in the category of presheaves, there is a unique morphism of presheaves $\mathscr{H}^{\mathrm{pre}} \to \mathscr{E}$ making the diagram commute. But then by the universal property of sheafification (Definition 2.4.6), there is a unique morphism of *sheaves* $\mathscr{H} \to \mathscr{E}$ making the diagram commute. □

2.5.B. EXERCISE. Show that the stalk of the cokernel is naturally isomorphic to the cokernel of the stalk.

We have now defined the notions of kernel and cokernel, and verified that they may be checked at the level of stalks. We have also verified that the properties of a morphism being a monomorphism or epimorphism are also determined at the level of stalks (Exercises 2.4.N and 2.4.O). Hence we have proved the following:

2.5.2. Theorem. — *Sheaves of abelian groups on a topological space X form an abelian category.*

That's all there is to it — what needs to be proved has been shifted to the stalks, where everything works because stalks are abelian groups!

And we see more: all structures coming from the abelian nature of this category may be checked at the level of stalks. For example:

2.5.C. EXERCISE. Suppose $\phi : \mathscr{F} \to \mathscr{G}$ is a morphism of sheaves of abelian groups. Show that the image sheaf $\operatorname{im} \phi$ is the sheafification of the image presheaf. (You must use the definition of image in an abelian category. In fact, this gives the accepted definition of image sheaf for a morphism of sheaves of sets.) Show that the stalk of the image is the image of the stalk.

As a consequence, **exactness of a sequence of sheaves may be checked at the level of stalks**. In particular:

2.5.D. IMPORTANT EXERCISE (CF. EXERCISE 2.3.A). Show that taking the stalk of a sheaf of abelian groups is an exact functor. More precisely, if X is a topological space and $p \in X$ is a point, show that taking the stalk at p defines an exact functor $Ab_X \to Ab$.

2.5.E. EXERCISE. Check that the exponential exact sequence (2.4.10.1) is exact.

2.5.F. EXERCISE: LEFT-EXACTNESS OF THE FUNCTOR OF "SECTIONS OVER U". Suppose $U \subset X$ is an open set, and $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H}$ is an exact sequence of sheaves of abelian groups. Show that

$$0 \to \mathscr{F}(U) \to \mathscr{G}(U) \to \mathscr{H}(U)$$

is exact. (You should do this "by hand", even if you realize there is a very fast proof using the left-exactness of the "forgetful" right adjoint to the sheafification functor.) Show that the section functor need not be exact: show that if $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is an exact sequence of sheaves of abelian groups, then

$$0 \to \mathscr{F}(U) \to \mathscr{G}(U) \to \mathscr{H}(U) \to 0$$

need not be exact. (Hint: the exponential exact sequence (2.4.10.1). But feel free to make up a different example.)

2.5.G. EXERCISE: LEFT-EXACTNESS OF PUSHFORWARD. Suppose $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H}$ is an exact sequence of sheaves of abelian groups on X. If $\pi : X \to Y$ is a continuous map, show that

$$0 \to \pi_*\mathscr{F} \to \pi_*\mathscr{G} \to \pi_*\mathscr{H}$$

is exact. (The previous exercise, dealing with the left-exactness of the global section functor can be interpreted as a special case of this, in the case where Y is a point.)

2.5.H. EXERCISE: LEFT-EXACTNESS OF $\mathcal{H}om$ (CF. EXERCISE 1.6.F(C) AND (D)). Suppose $\mathscr{F}$ is a sheaf of abelian groups on a topological space X. Show that $\mathcal{H}om(\mathscr{F}, \cdot)$ is a left-exact covariant functor $Ab_X \to Ab_X$. Show that $\mathcal{H}om(\cdot, \mathscr{F})$ is a left-exact contravariant functor $Ab_X \to Ab_X$.

2.5.3. $\mathscr{O}_X$-modules.

2.5.I. EXERCISE. Show that if $(X, \mathscr{O}_X)$ is a ringed space, then $\mathscr{O}_X$-modules form an abelian category. (There is a fair bit to check, but there aren't many new ideas.)

2.5.4. Many facts about sheaves of abelian groups carry over to $\mathscr{O}_X$-modules without change, because a sequence of $\mathscr{O}_X$-modules is exact if and only if the underlying sequence of sheaves of abelian groups is exact. You should be able to easily check that all of the statements of the earlier exercises in §2.5 also hold for $\mathscr{O}_X$-modules, when modified appropriately. For example (Exercise 2.5.H), $\mathcal{H}om_{\mathscr{O}_X}$ is a left-exact contravariant functor in its first argument and a left-exact covariant functor in its second argument.

We end with a useful construction using some of the ideas in this section.

2.5.J. IMPORTANT EXERCISE: TENSOR PRODUCTS OF $\mathscr{O}_X$-MODULES.
(a) Suppose $\mathscr{O}_X$ is a sheaf of rings on X. Define (categorically) what we should mean by **tensor product of two $\mathscr{O}_X$-modules**. Give an explicit construction, and show that it satisfies your categorical definition. *Hint:* take the "presheaf tensor product" — which needs to be defined — and sheafify. Note: $\otimes_{\mathscr{O}_X}$ is often written $\otimes$ when the subscript is clear from the context. (An example showing sheafification is necessary will arise in Example 14.1.1.)
(b) Show that the tensor product of stalks is the stalk of tensor product. (If you can show this, you may be able to make sense of the phrase "colimits commute with tensor products".)

2.5.5. Conclusion. Just as presheaves are abelian categories because all abelian-categorical notions make sense open set by open set, sheaves are abelian categories because all abelian-categorical notions make sense stalk by stalk.

2.6 The inverse image sheaf

We next describe a notion that is fundamental, but rather intricate. We will not need it for some time, so this may be best left for a second reading. Suppose we have a continuous map $\pi : X \to Y$. If $\mathscr{F}$ is a sheaf on X, we have defined the pushforward or direct image sheaf $\pi_*\mathscr{F}$, which is a sheaf on Y. There is also a notion of inverse image sheaf. (We will not call it the pullback sheaf, reserving that name for a later construction for quasicoherent sheaves, §16.3.) This is a covariant functor π^{-1} from sheaves on Y to sheaves on X. If the sheaves on Y have some additional structure (e.g. group or ring), then this structure is respected by π^{-1}.

2.6.1. *Definition by adjoint: elegant but abstract.* We define π^{-1} as the left adjoint to π_*.

This isn't really a definition; we need a construction to show that the adjoint exists. Note that we then get canonical maps $\pi^{-1}\pi_*\mathscr{F} \to \mathscr{F}$ (associated to the identity in $\mathrm{Mor}_Y(\pi_*\mathscr{F}, \pi_*\mathscr{F})$) and $\mathscr{G} \to \pi_*\pi^{-1}\mathscr{G}$ (associated to the identity in $\mathrm{Mor}_X(\pi^{-1}\mathscr{G}, \pi^{-1}\mathscr{G})$).

$$\begin{tikzcd}
 & \pi^{-1}\mathscr{G} \arrow[r] \arrow[dl, dash] & \mathscr{F} \arrow[dll, dash] \\
X \arrow[d, "\pi"'] & \mathscr{G} \arrow[r] \arrow[dl, dash] & \pi_*\mathscr{F} \arrow[dll, dash] \\
Y & &
\end{tikzcd}$$

2.6.2. *Construction: concrete but ugly.* Define the temporary notation

$$\pi^{-1}\mathscr{G}^{\mathrm{pre}}(U) = \varinjlim_{V \supset \pi(U)} \mathscr{G}(V).$$

(Recall the explicit description of colimit: sections are sections on open sets containing $\pi(U)$, with an equivalence relation. Note that $\pi(U)$ won't be an open set in general.)

2.6.A. EXERCISE. Show that this defines a presheaf on X. Show that it needn't form a sheaf. (Hint: map 2 points to 1 point.)

Now define the **inverse image of** $\mathscr{G}$ by $\pi^{-1}\mathscr{G} := (\pi^{-1}\mathscr{G}^{\text{pre}})^{\text{sh}}$. Note that π^{-1} is a functor from sheaves on Y to sheaves on X. The next exercise shows that π^{-1} is indeed left-adjoint to π_*. But you may wish to try the later exercises first, and come back to Exercise 2.6.B later. (For the later exercises, try to give two proofs, one using the universal property, and the other using the explicit description.)

2.6.B. IMPORTANT TRICKY EXERCISE. If $\pi : X \to Y$ is a continuous map, and $\mathscr{F}$ is a sheaf on X and $\mathscr{G}$ is a sheaf on Y, describe a bijection

$$\mathrm{Mor}_X(\pi^{-1}\mathscr{G}, \mathscr{F}) \leftrightarrow \mathrm{Mor}_Y(\mathscr{G}, \pi_*\mathscr{F}).$$

Observe that your bijection is "natural" in the sense of the definition of adjoints (i.e., functorial in both $\mathscr{F}$ and $\mathscr{G}$). Thus Construction 2.6.2 satisfies the universal property of Definition 2.6.1. Possible hint: Show that both sides agree with the following third construction, which we denote $\mathrm{Mor}_{YX}(\mathscr{G}, \mathscr{F})$. A collection of maps $\phi_{VU} : \mathscr{G}(V) \to \mathscr{F}(U)$ (as U runs through all open sets of X, and V runs through all open sets of Y containing $\pi(U)$) is said to be *compatible* if for all open $U' \subset U \subset X$ and all open $V' \subset V \subset Y$ with $\pi(U) \subset V$, $\pi(U') \subset V'$, the diagram

(2.6.2.1)
$$\begin{tikzcd} \mathscr{G}(V) \arrow[r, "\phi_{VU}"] \arrow[d, "\mathrm{res}_{V,V'}"'] & \mathscr{F}(U) \arrow[d, "\mathrm{res}_{U,U'}"] \\ \mathscr{G}(V') \arrow[r, "\phi_{V'U'}"] & \mathscr{F}(U') \end{tikzcd}$$

commutes. Define $\mathrm{Mor}_{YX}(\mathscr{G}, \mathscr{F})$ to be the set of all compatible collections $\phi = \{\phi_{VU}\}$.

2.6.3. *Remark ("stalk and skyscraper are an adjoint pair").* As a special case, if X is a point $p \in Y$, we see that $\pi^{-1}\mathscr{G}$ is the stalk $\mathscr{G}_p$ of $\mathscr{G}$, and maps from the stalk $\mathscr{G}_p$ to a set S are the same as maps of sheaves on Y from $\mathscr{G}$ to the skyscraper sheaf with set S supported at p. You may prefer to prove this special case by hand directly before solving Exercise 2.6.B, as it is enlightening. (It can also be useful — can you use it to solve Exercises 2.4.M and 2.4.O?)

2.6.C. EXERCISE. Show that the stalks of $\pi^{-1}\mathscr{G}$ are the same as the stalks of $\mathscr{G}$. More precisely, if $\pi(p) = q$, describe a natural isomorphism $\mathscr{G}_q \cong (\pi^{-1}\mathscr{G})_p$. (Possible hint: use the concrete description of the stalk, as a colimit. Recall that stalks are preserved by sheafification, Exercise 2.4.M. Alternatively, use adjointness.) This, along with the notion of compatible stalks, may give you a simple way of thinking about (and perhaps visualizing) inverse image sheaves. (Those preferring the "espace étalé" or "space of sections" perspective, §2.2.11, can check that the pullback of the "space of sections" is the "space of sections" of the pullback.)

2.6.D. EXERCISE (EASY BUT USEFUL). If U is an open subset of Y, $i : U \to Y$ is the inclusion, and $\mathscr{G}$ is a sheaf on Y, show that $i^{-1}\mathscr{G}$ is naturally isomorphic to $\mathscr{G}|_U$.

2.6.4. *Definition.* If $\mathscr{G}$ is a sheaf on Y, and U is an open subset of Y, then $\mathscr{G}|_U$ is called the **restriction of** $\mathscr{G}$ **to** U. The restriction of $\mathscr{O}_Y$ to U is denoted $\mathscr{O}_U$. (We will later call $(U, \mathscr{O}_U) \to (Y, \mathscr{O}_Y)$ an *open embedding* of ringed spaces, see Definition 6.2.1.)

2.6.E. EXERCISE. Show that π^{-1} is an exact functor from sheaves of abelian groups on Y to sheaves of abelian groups on X (cf. Exercise 2.5.D). (Hint: exactness can be checked on stalks, and by Exercise 2.6.C, the stalks are the same.) Essentially the same argument will show that π^{-1} is an exact functor from $\mathscr{O}_Y$-modules (on Y) to $(\pi^{-1}\mathscr{O}_Y)$-modules (on X), but don't bother writing that down. (Remark for experts: π^{-1} is a left adjoint, hence right-exact by abstract nonsense, as discussed in §1.6.12. Left-exactness holds because colimits over filtered index sets are exact.)

2.6.F. EXERCISE.
(a) Suppose $Z \subset Y$ is a closed subset, and $i : Z \hookrightarrow Y$ is the inclusion. If $\mathscr{F}$ is a sheaf of sets on Z, then show that the stalk $(i_*\mathscr{F})_q$ is a one element-set if $q \notin Z$, and $\mathscr{F}_q$ if $q \in Z$.
(b) *Definition:* Define the **support** of a sheaf $\mathscr{G}$ of sets, denoted $\operatorname{Supp} \mathscr{G}$, as the locus where the stalks are not the one-element set:

$$\operatorname{Supp} \mathscr{G} := \{p \in X \ : \ |\mathscr{G}_p| \neq 1\}.$$

(More generally, if the sheaf has value in some category, the support consists of points where the stalk is not the final object. For a sheaf $\mathscr{G}$ of abelian groups, the support consists of points with nonzero stalks — $\operatorname{Supp} \mathscr{G} = \{p \in X \ : \ \mathscr{G}_p \neq 0\}$ — or equivalently is the union of supports of sections over all open sets, see Definition 2.4.2.) Suppose $\operatorname{Supp} \mathscr{G} \subset Z$ where Z is closed. Show that the natural map $\mathscr{G} \to i_* i^{-1} \mathscr{G}$ is an isomorphism. Thus a sheaf supported on a closed subset can be considered a sheaf on that closed subset. ("Support of a sheaf" is a useful notion, and will arise again in §13.7.D.)

The final exercise of this section may be best left for later, when you will realize why you care about it. (We will start to use it in Chapter 23 — more precisely, in Exercise 23.4.F.)

2.6.G. EXERCISE (EXTENSION BY ZERO $\pi_!$: AN OCCASIONAL *left adjoint* TO π^{-1}). In addition to always being a left adjoint, π^{-1} can sometimes be a right adjoint. Suppose $i : U \hookrightarrow Y$ is an inclusion of an open set into Y. Define the **extension of** i **by zero** $i_! : \mathit{Mod}_{\mathscr{O}_U} \to \mathit{Mod}_{\mathscr{O}_Y}$ as follows. Suppose $\mathscr{F}$ is an $\mathscr{O}_U$-module. For open $W \subset Y$, define $(i_!^{\text{pre}} \mathscr{F})(W) = \mathscr{F}(W)$ if $W \subset U$, and 0 otherwise (with the obvious restriction maps). This is clearly a presheaf $\mathscr{O}_Y$-module. Define $i_!$ as $(i_!^{\text{pre}})^{\text{sh}}$. Note that $i_! \mathscr{F}$ is an $\mathscr{O}_Y$-module, and that this defines a functor. (The symbol "!" is read as "shriek". I have no idea why, but I suspect it is because people often shriek when they see it. Thus "$i_!$" is read as "i-lower-shriek".)
(a) Show that $i_!^{\text{pre}} \mathscr{F}$ need not be a sheaf. (We won't need this, but it may give some insight into why this is called "extension by zero". Possible source for an example: continuous functions on $\mathbb{R}$.)
(b) For $q \in Y$, show that $(i_! \mathscr{F})_q = \mathscr{F}_q$ if $q \in U$, and 0 otherwise.
(c) Show that $i_!$ is an exact functor.
(d) If $\mathscr{G}$ is an $\mathscr{O}_Y$-module, describe an inclusion $i_! i^{-1} \mathscr{G} \hookrightarrow \mathscr{G}$. (Interesting remark we won't need: Let Z be the complement of U, and $j : Z \to Y$ the natural inclusion.

Then there is a short exact sequence

$$0 \to i_! i^{-1}\mathscr{G} \to \mathscr{G} \to j_* j^{-1}\mathscr{G} \to 0.$$

This is best checked by describing the maps, then checking exactness at stalks.)
(e) Show that $(i_!, i^{-1})$ is an adjoint pair, so there is a natural bijection $\mathrm{Hom}_{\mathscr{O}_Y}(i_!\mathscr{F}, \mathscr{G}) \leftrightarrow \mathrm{Hom}_{\mathscr{O}_U}(\mathscr{F}, \mathscr{G}|_U)$ for any $\mathscr{O}_U$-module $\mathscr{F}$ and $\mathscr{O}_Y$-module $\mathscr{G}$. (In particular, the sections of $\mathscr{G}$ over U can be identified with $\mathrm{Hom}_{\mathscr{O}_Y}(i_!\mathscr{O}_U, \mathscr{G})$.)

2.7 Recovering sheaves from a "sheaf on a base"

Sheaves are natural things to want to think about, but hard to get our hands on. We like the identity and gluability axioms, but they make proving things trickier than for presheaves. We have discussed how we can understand sheaves using stalks (using "compatible germs"). We now introduce a second way of getting a hold of sheaves, by introducing the notion of a *sheaf on a base*. Warning: this way of understanding an entire sheaf from limited information is confusing. It may help to keep sight of the central insight that this partial information is enough to understand germs, and the notion of when they are compatible (with nearby germs).

First, we define the notion of a **base of a topology**. Suppose we have a topological space X, i.e., we know which subsets U_i of X are open. Then a base of a topology is a subcollection of the open sets $\{B_j\} \subset \{U_i\}$, such that each U_i is a union of the B_j. Here is one example that you have seen early in your mathematical life. Suppose $X = \mathbb{R}^n$. Then the way the classical topology is often first defined is by defining *open balls* $B_r(x) = \{y \in \mathbb{R}^n \ : \ |y - x| < r\}$, and declaring that any union of open balls is open. So the balls form a base of the classical topology — we say they *generate* the classical topology. As an application of how we use them, to check continuity of some map $\pi : X \to \mathbb{R}^n$, you need only think about the pullback of balls on $\mathbb{R}^n$ — part of the traditional δ-ϵ definition of continuity.

Now suppose we have a sheaf $\mathscr{F}$ on a topological space X, and a base $\{B_i\}$ of open sets on X. Then consider the information

$$(\{\mathscr{F}(B_i)\}, \{\mathrm{res}_{B_i,B_j} : \mathscr{F}(B_i) \to \mathscr{F}(B_j)\}),$$

which is a subset of the information contained in the sheaf — we are only paying attention to the information involving elements of the base, not all open sets.

We can recover the entire sheaf from this information. This is because we can determine the stalks from this information, and we can determine when germs are compatible.

2.7.A. IMPORTANT EXERCISE. Make this precise. How can you recover a sheaf $\mathscr{F}$ from this partial information?

This suggests a notion, called a **sheaf on a base**. A sheaf of sets (or abelian groups, rings, ...) on a base $\{B_i\}$ is the following. For each B_i in the base, we have a set $F(B_i)$. If $B_i \subset B_j$, we have maps $\mathrm{res}_{B_j,B_i} : F(B_j) \to F(B_i)$, with $\mathrm{res}_{B_i,B_i} = \mathrm{id}_{F(B_i)}$. (Things called "B" are always assumed to be in the base.) If $B_i \subset B_j \subset B_k$, then $\mathrm{res}_{B_k,B_i} = \mathrm{res}_{B_j,B_i} \circ \mathrm{res}_{B_k,B_j}$. So far we have defined a **presheaf on a base** $\{B_i\}$.

We also require the **base identity** axiom: If $B = \cup B_i$, then if $f, g \in F(B)$ are such that $\mathrm{res}_{B,B_i} f = \mathrm{res}_{B,B_i} g$ for all i, then $f = g$.

We require the **base gluability** axiom too: If $B = \cup B_i$, and we have $f_i \in F(B_i)$ such that f_i agrees with f_j on any basic open set contained in $B_i \cap B_j$ (i.e., $\mathrm{res}_{B_i,B_k} f_i = \mathrm{res}_{B_j,B_k} f_j$ for all $B_k \subset B_i \cap B_j$) then there exists $f \in F(B)$ such that $\mathrm{res}_{B,B_i} f = f_i$ for all i.

2.7.1. Theorem. — *Suppose $\{B_i\}$ is a base on X, and F is a sheaf of sets on this base. Then there is a sheaf $\mathscr{F}$ extending F (with isomorphisms $\mathscr{F}(B_i) \cong F(B_i)$ agreeing with the restriction maps). This sheaf $\mathscr{F}$ is unique up to unique isomorphism.*

Proof. We will define $\mathscr{F}$ as the sheaf of compatible germs of F.

Define the **stalk** of a base presheaf F at $p \in X$ by

$$F_p = \varinjlim F(B_i)$$

where the colimit is over all B_i (in the base) containing p.

We will say a family of germs in an open set U is compatible near p if there is a section s of F over some B_i containing p such that the germs over B_i are precisely the germs of s. More formally, define

$$\begin{aligned}\mathscr{F}(U) \quad := \quad \{(f_p \in F_p)_{p\in U} : & \text{ for all } p \in U, \text{ there exists } B \text{ with } p \in B \subset U, s \in F(B), \\ & \text{with } s_q = f_q \text{ for all } q \in B\}\end{aligned}$$

where each B is in our base.

This is a sheaf (for the same reasons that the sheaf of compatible germs was, cf. Exercise 2.4.I).

I next claim that if B is in our base, the natural map $F(B) \to \mathscr{F}(B)$ is an isomorphism.

2.7.B. EXERCISE. Verify that $F(B) \to \mathscr{F}(B)$ is an isomorphism, likely by showing that it is injective and surjective (or else by describing the inverse map and verifying that it is indeed inverse). Possible hint: elements of $\mathscr{F}(B)$ are determined by stalks, as are elements of $F(B)$.

It will be clear from your solution to Exercise 2.7.B that the restriction maps for F are the same as the restriction maps of $\mathscr{F}$ (for elements of the base).

Finally, you should verify to your satisfaction that $\mathscr{F}$ is indeed unique up to unique isomorphism. (First be sure that you understand what this means!) $\square$

Theorem 2.7.1 shows that sheaves on X can be recovered from their "restriction to a base". It is clear from the argument (and in particular the solution to the Exercise 2.7.B) that if $\mathscr{F}$ is a sheaf and F is the corresponding sheaf on the base B, then for any $p \in X$, $\mathscr{F}_p$ is naturally isomorphic to F_p.

Theorem 2.7.1 is a statement about *objects* in a category, so we should hope for a similar statement about *morphisms*.

2.7.C. IMPORTANT EXERCISE: MORPHISMS OF SHEAVES CORRESPOND TO MORPHISMS OF SHEAVES ON A BASE. Suppose $\{B_i\}$ is a base for the topology of X. A morphism $F \to G$ of sheaves on the base is a collection of maps $F(B_k) \to G(B_k)$

such that the diagram

$$\begin{tikzcd} F(B_i) \arrow[r] \arrow[d, "\mathrm{res}_{B_i,B_j}"'] & G(B_i) \arrow[d, "\mathrm{res}_{B_i,B_j}"] \\ F(B_j) \arrow[r] & G(B_j) \end{tikzcd}$$

commutes for all $B_j \hookrightarrow B_i$.
(a) Verify that a morphism of sheaves is determined by the induced morphism of sheaves on the base.
(b) Show that a morphism of sheaves on the base gives a morphism of the induced sheaves. (Possible hint: compatible stalks.)

2.7.2. *Remark.* The above constructions and arguments describe an equivalence of categories (§1.2.21) between sheaves on X and sheaves on a given base of X. There is no new content to this statement, but you may wish to think through what it means. What are the functors in each direction? Why aren't their compositions the identity?

2.7.3. *Remark.* It will be useful to extend these notions to $\mathscr{O}_X$-modules (see for example Exercise 13.3.C). You will readily be able to verify that there is a correspondence (really, equivalence of categories) between $\mathscr{O}_X$-modules and "$\mathscr{O}_X$-modules on a base". Rather than working out the details, you should just informally think through the main points: what is an "$\mathscr{O}_X$-module on a base"? Given an $\mathscr{O}_X$-module on a base, why is the corresponding sheaf naturally an $\mathscr{O}_X$-module? Later, if you are forced at gunpoint to fill in details, you will be able to.

2.7.D. IMPORTANT EXERCISE. Suppose $X = \cup U_i$ is an open cover of X, and we have sheaves $\mathscr{F}_i$ on U_i along with isomorphisms $\phi_{ij} : \mathscr{F}_i|_{U_i \cap U_j} \to \mathscr{F}_j|_{U_i \cap U_j}$ (with ϕ_{ii} the identity) that agree on triple overlaps, i.e., $\phi_{jk} \circ \phi_{ij} = \phi_{ik}$ on $U_i \cap U_j \cap U_k$ (this is called the **cocycle condition**, for reasons we ignore). Show that these sheaves can be glued together into a sheaf $\mathscr{F}$ on X (unique up to unique isomorphism), such that $\mathscr{F}_i \cong \mathscr{F}|_{U_i}$, and the isomorphisms over $U_i \cap U_j$ are the obvious ones. (Thus we can "glue sheaves together", using limited patching information.) Warning: we are not assuming this is a finite cover, so you cannot use induction. For this reason this exercise can be perplexing. (You can use the ideas of this section to solve this problem, but you don't necessarily need to. Hint: As the base, take those open sets contained in *some* U_i. Small observation: the hypothesis on ϕ_{ii} is extraneous, as it follows from the cocycle condition.)

2.7.4. *Important remark.* We will repeatedly see the theme of constructing some object by gluing, in many different contexts. Keep an eye out for it! In each case, we carefully consider what information we need in order to glue.

2.7.5. *Remark for experts.* Exercise 2.7.D almost says that the "set" of sheaves forms a sheaf itself, but not quite. Making this precise leads one to the notion of a *stack*.

2.7.E. UNIMPORTANT EXERCISE. Suppose a morphism of sheaves $F \to G$ on a base B_i is surjective for all B_i (i.e., $F(B_i) \to G(B_i)$ is surjective for all i). Show that the corresponding morphism of sheaves (*not* on the base) is surjective (or more precisely: an epimorphism). The converse is not true, unlike the case for injectivity.

This gives a useful sufficient criterion for "surjectivity": a morphism of sheaves is an epimorphism ("surjective") if it is surjective for sections on a base. You may enjoy trying this out with Example 2.4.10 (dealing with holomorphic functions in the classical topology on $X = \mathbb{C}$), showing that the exponential map $\exp : \mathscr{O}_X \to \mathscr{O}_X^*$ is surjective, using the base of contractible open sets.

Part II

Schemes

The very idea of scheme is of infantile simplicity — so simple, so humble, that no one before me thought of stooping so low. So childish, in short, that for years, despite all the evidence, for many of my erudite colleagues, it was really "not serious"!

— *A. Grothendieck* [**Gr6**, p. P32], *translated by C. McLarty* [**Mc**, p. 313]

CHAPTER 3

Toward affine schemes: the underlying set, and topological space

There is no serious historical question of how Grothendieck found his definition of schemes. It was in the air. Serre has well said that no one invented schemes... . The question is, what made Grothendieck believe he should use this definition to simplify an 80 page paper by Serre into some 1000 pages of Élements de Géométrie Algébrique?
— *C. McLarty* [**Mc**, p. 313]

3.1 Toward schemes

We are now ready to consider the notion of a *scheme*, which is the type of geometric space central to algebraic geometry. We should first think through what we mean by "geometric space". You have likely seen the notion of a manifold, and we wish to abstract this notion so that it can be generalized to other settings, notably so that we can deal with nonsmooth and arithmetic objects.

The key insight behind this generalization is the following: we can understand a geometric space (such as a manifold) well by understanding the functions on this space. More precisely, we will understand it through the sheaf of functions on the space. If we are interested in differentiable manifolds, we will consider differentiable functions; if we are interested in manifolds, we will consider smooth functions; and so on.

Thus we will define a scheme to be the following data

- *The set:* the points of the scheme
- *The topology:* the open sets of the scheme
- *The structure sheaf:* the sheaf of "algebraic functions" (a sheaf of rings) on the scheme.

Recall that a topological space with a sheaf of rings is called a *ringed space* (§2.2.13).

We will try to draw pictures throughout. Pictures can help develop geometric intuition, which can guide the algebraic development (and, eventually, vice versa). Some people find pictures very helpful, while others are repulsed or nonplussed or confused.

We will try to make all three notions as intuitive as possible. For the set, in the key example of complex (affine) varieties (roughly, things cut out in $\mathbb{C}^n$ by polynomials), we will see that the points are the "traditional points" (n-tuples of complex numbers), plus some extra points that will be handy to have around. For the topology, we will require that "the subset where an algebraic function vanishes must be closed", and require nothing else. For the sheaf of algebraic functions (the structure sheaf), we will expect that in the complex plane, $(3x^2+y^2)/(2x+4xy+1)$ should be an algebraic function on the open set consisting of points where the denominator doesn't vanish, and this will largely motivate our definition.

3.1.1. Example: Differentiable manifolds. As motivation, we return to our example of differentiable manifolds, reinterpreting them in this light. We will be quite informal in this discussion. Suppose X is a manifold. It is a topological space, and has a *sheaf of differentiable functions* $\mathscr{O}_X$ (see §2.1). This gives X the structure of a ringed space. We have observed that evaluation at a point $p \in X$ gives a surjective map from the stalk to $\mathbb{R}$

$$\mathscr{O}_{X,p} \twoheadrightarrow \mathbb{R},$$

so the kernel, the (germs of) functions vanishing at p, is a maximal ideal $\mathfrak{m}_{X,p}$ (see §2.1.1).

We could *define* a differentiable real manifold as a topological space X with a sheaf of rings. We would require that there is a cover of X by open sets such that on each open set the ringed space is isomorphic to a ball around the origin in $\mathbb{R}^n$ (with the sheaf of differentiable functions on that ball). With this definition, the ball is the basic patch, and a general manifold is obtained by gluing these patches together. (Admittedly, a great deal of geometry comes from how one chooses to patch the balls together!) In the algebraic setting, the basic patch is the notion of an *affine scheme*, which we will discuss soon. (In the definition of manifold, there is an additional requirement that the topological space be Hausdorff, to avoid pathologies. Schemes are often required to be "separated" to avoid essentially the same pathologies. Separatedness will be discussed in Chapter 10.)

Functions are determined by their values at points. This is an obvious statement, but won't be true for schemes in general. We will see an example in Exercise 3.2.A(a), and discuss this behavior further in §3.2.11.

Morphisms of manifolds. How can we describe differentiable maps of manifolds $\pi : X \to Y$? They are certainly continuous maps — but which ones? We can pull back functions along continuous maps. Differentiable functions pull back to differentiable functions. More formally, we have a map $\pi^{-1}\mathscr{O}_Y \to \mathscr{O}_X$. (The inverse image sheaf π^{-1} was defined in §2.6.) Inverse image is left-adjoint to pushforward, so we also get a map $\pi^\sharp : \mathscr{O}_Y \to \pi_*\mathscr{O}_X$.

Certainly given a differentiable map of manifolds, differentiable functions pull back to differentiable functions. It is less obvious that *this is a sufficient condition for a continuous function to be differentiable*.

3.1.A. IMPORTANT EXERCISE FOR THOSE WITH A LITTLE EXPERIENCE WITH MANIFOLDS. Suppose that $\pi : X \to Y$ is a continuous map of differentiable manifolds (as topological spaces — not a priori differentiable). Show that π is differentiable if differentiable functions pull back to differentiable functions, i.e., if pullback by π gives a map $\mathscr{O}_Y \to \pi_*\mathscr{O}_X$. (Hint: check this on small patches. Once you figure out what you are trying to show, you will realize that the result is immediate.)

3.1.B. EXERCISE. Show that a morphism of differentiable manifolds $\pi : X \to Y$ with $\pi(p) = q$ induces a morphism of stalks $\pi^\sharp : \mathscr{O}_{Y,q} \to \mathscr{O}_{X,p}$. Show that $\pi^\sharp(\mathfrak{m}_{Y,q}) \subset \mathfrak{m}_{X,p}$. In other words, if you pull back a function that vanishes at q, you get a function that vanishes at p — not a huge surprise. (In §6.3, we formalize this by saying that maps of differentiable manifolds are maps of locally ringed spaces.)

3.1.2. *Aside.* Here is a little more for experts: Notice that π induces a map on tangent spaces (see Aside 2.1.2)

$$(\mathfrak{m}_{X,p}/\mathfrak{m}_{X,p}^2)^\vee \to (\mathfrak{m}_{Y,q}/\mathfrak{m}_{Y,q}^2)^\vee.$$

This is the tangent map you would geometrically expect. Again, it is interesting that the cotangent map $\mathfrak{m}_{Y,q}/\mathfrak{m}_{Y,q}^2 \to \mathfrak{m}_{X,p}/\mathfrak{m}_{X,p}^2$ is algebraically more natural than the tangent map (there are no "duals").

Experts are now free to try to interpret other differential-geometric information using only the map of topological spaces and map of sheaves. For example: how can one check if f is a submersion of manifolds? How can one check if f is an immersion? (We will see that the algebro-geometric version of these notions are *smooth morphism* and *unramified morphism*; see Chapter 25, although they will be defined earlier.)

3.1.3. *Side Remark.* Manifolds are covered by disks that are all isomorphic. This isn't true for schemes (even for "smooth complex varieties"). There are examples of two "smooth complex curves" (the algebraic version of Riemann surfaces) X and Y so that no nonempty open subset of X is isomorphic to a nonempty open subset of Y (see Exercise 6.5.K). And there is a Riemann surface X such that no two open subsets of X are isomorphic (see Exercise 19.7.D). Informally, this is because in the Zariski topology on schemes, all nonempty open sets are "huge" and have more "structure".

3.1.4. Other examples. If you are interested in differential geometry, you will be interested in differentiable manifolds, on which the functions under consideration are differentiable functions. Similarly, if you are interested in topology, you will be interested in topological spaces, on which you will consider the continuous function. If you are interested in complex geometry, you will be interested in complex manifolds (or possibly "complex analytic varieties"), on which the functions are holomorphic functions. In each of these cases of interesting "geometric spaces", the topological space and sheaf of functions is clear. The notion of scheme fits naturally into this family.

3.2 The underlying set of affine schemes

For any ring A, we are going to define something called $\operatorname{Spec} A$, the **spectrum of** A. In this section, we will define it as a set, but we will soon endow it with a topology, and later we will define a sheaf of rings on it (the structure sheaf). Such an object is called an *affine scheme*. Later $\operatorname{Spec} A$ will denote the set along with the topology, and a sheaf of functions. But for now, as there is no possibility of confusion, $\operatorname{Spec} A$ will just be the set.

3.2.1. The set $\operatorname{Spec} A$ is the set of prime ideals of A. The prime ideal $\mathfrak{p}$ of A when considered as an element of $\operatorname{Spec} A$ will be denoted $[\mathfrak{p}]$, to avoid confusion. Elements $a \in A$ will be called **functions on** $\operatorname{Spec} A$, and their **value** at the point $[\mathfrak{p}]$

will be $a \pmod{\mathfrak{p}}$. *This is weird: a function can take values in different rings at different points — the function* 5 *on* $\operatorname{Spec} \mathbb{Z}$ *takes the value* $1 \pmod 2$ *at* $[(2)]$ *and* $2 \pmod 3$ *at* $[(3)]$. "An element a of the ring lying in a prime ideal $\mathfrak{p}$" translates to "a function a that is 0 at the point $[\mathfrak{p}]$" or "a function a vanishing at the point $[\mathfrak{p}]$", and we will use these phrases interchangeably. Notice that if you add or multiply two functions, you add or multiply their values at all points; this is a translation of the fact that $A \to A/\mathfrak{p}$ is a ring morphism. These translations are important — make sure you are very comfortable with them! They should become second nature.

If A is generated over a base field (or base ring) by elements $x_1, \dots, x_r$, the elements $x_1, \dots, x_r$ are often called **coordinates**, because we will later be able to reinterpret them as restrictions of "coordinates on r-space", via the idea of §3.2.9, made precise in Exercise 6.2.D.

3.2.2. *Some glimpses of the future.* In §4.1.3 we will interpret functions on $\operatorname{Spec} A$ as global sections of the "structure sheaf", i.e. as a function on a ringed space, in the sense of §2.2.13. We repeat a caution from §2.2.13: what we will call "functions", others may call "regular functions". And we will later define "rational functions" (§5.5.6), which are not precisely functions in this sense; they are a particular type of "partially-defined function".

The notion of "value of a function" will be later interpreted as a value of a function on a particular locally ringed space, see Definition 4.3.6.

3.2.3. We now give some examples.

Example 1 (the complex affine line): $\mathbb{A}^1_{\mathbb{C}} := \operatorname{Spec} \mathbb{C}[x]$. Let's find the prime ideals of $\mathbb{C}[x]$. As $\mathbb{C}[x]$ is an integral domain, 0 is prime. Also, $(x-a)$ is prime, for any $a \in \mathbb{C}$: it is even a maximal ideal, as the quotient by this ideal is a field:

$$0 \longrightarrow (x-a) \longrightarrow \mathbb{C}[x] \xrightarrow{f \mapsto f(a)} \mathbb{C} \longrightarrow 0$$

(This exact sequence may remind you of (2.1.1.1) in our motivating example of manifolds.)

We now show that there are no other prime ideals. We use the fact that $\mathbb{C}[x]$ has a division algorithm, and is a unique factorization domain. Suppose $\mathfrak{p}$ is a prime ideal. If $\mathfrak{p} \neq (0)$, then suppose $f(x) \in \mathfrak{p}$ is a nonzero element of smallest degree. It is not constant, as prime ideals can't contain 1. If $f(x)$ is not linear, then factor $f(x) = g(x)h(x)$, where $g(x)$ and $h(x)$ have positive degree. (Here we use that $\mathbb{C}$ is algebraically closed.) Then $g(x) \in \mathfrak{p}$ or $h(x) \in \mathfrak{p}$, contradicting the minimality of the degree of f. Hence there is a linear element $x - a$ of $\mathfrak{p}$. Then I claim that $\mathfrak{p} = (x-a)$. Suppose $f(x) \in \mathfrak{p}$. Then the division algorithm would give $f(x) = g(x)(x-a) + m$ where $m \in \mathbb{C}$. Then $m = f(x) - g(x)(x-a) \in \mathfrak{p}$. If $m \neq 0$, then $1 \in \mathfrak{p}$, giving a contradiction.

Thus we can and should (and must!) have a picture of $\mathbb{A}^1_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x]$ (see Figure 3.1). This is just the first illustration of a point of view of Sophie Germain [Ge]: "L'algèbre n'est qu'une géométrie écrite; la géométrie n'est qu'une algèbre figurée." (Algebra is but written geometry; geometry is but drawn algebra.)

There is one "traditional" point for each complex number, plus one extra ("bonus") point $[(0)]$. We can mostly picture $\mathbb{A}^1_{\mathbb{C}}$ as $\mathbb{C}$: the point $[(x-a)]$ we will reasonably associate to $a \in \mathbb{C}$. Where should we picture the point $[(0)]$? The best way of thinking about it is somewhat zen. It is somewhere on the complex line, but nowhere in particular. Because (0) is contained in all of these primes, we will somehow associate it with this line passing through all the other points. $[(0)]$ is called the "generic point" of the line; it is "generically on the line" but you can't pin it down any further than that. (We will formally define "generic point" in §3.6.) We will place it far to the right for lack of anywhere better to put it. You will notice that we sketch $\mathbb{A}^1_{\mathbb{C}}$ as one-(real-)dimensional (even though we picture it as an enhanced version of $\mathbb{C}$); this is to later remind ourselves that this will be a one-dimensional space, where dimensions are defined in an algebraic (or complex-geometric) sense. (Dimension will be defined in Chapter 11.)

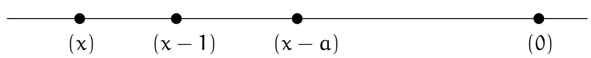


FIGURE 3.1. A picture of $\mathbb{A}^1_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x]$

To give you some feeling for this space, we make some statements that are currently undefined, but suggestive. The functions on $\mathbb{A}^1_{\mathbb{C}}$ are the polynomials. So $f(x) = x^2 - 3x + 1$ is a function. What is its value at $[(x-1)]$, which we think of as the point $1 \in \mathbb{C}$? Answer: $f(1)$! Or equivalently, we can evaluate $f(x)$ modulo $x - 1$ — this is the same thing by the division algorithm. (What is its value at (0)? It is $f(x) \pmod 0$, which is just $f(x)$.)

Here is a more complicated example: $g(x) = (x-3)^3/(x-2)$ is a "rational function". It is defined everywhere but $x = 2$. (When we know what the structure sheaf is, we will be able to say that it is an element of the structure sheaf on the open set $\mathbb{A}^1_{\mathbb{C}} \setminus \{2\}$.) We want to say that $g(x)$ has a triple zero at 3, and a single pole at 2, and we will be able to after §12.5.

Example 2 (the affine line over $k = \overline{k}$): $\mathbb{A}^1_k := \operatorname{Spec} k[x]$ where k is an algebraically closed field. This is called the affine line over k. All of our discussion in the previous example carries over without change. We will use the same picture, which is after all intended to just be a metaphor.

Example 3: $\operatorname{Spec} \mathbb{Z}$. An amazing fact is that from our perspective, this will look a lot like the affine line $\mathbb{A}^1_{\overline{k}}$. The integers, like $\overline{k}[x]$, form a unique factorization domain, with a division algorithm. The prime ideals are: (0), and (p) where p is prime. Thus everything from Example 1 carries over without change, even the picture. Our picture of $\operatorname{Spec} \mathbb{Z}$ is shown in Figure 3.2.

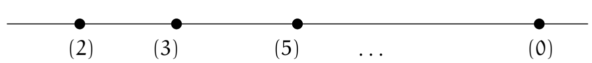


FIGURE 3.2. A "picture" of $\operatorname{Spec} \mathbb{Z}$, which looks suspiciously like Figure 3.1

Let's blithely carry over our discussion of functions to this space. 100 is a function on $\operatorname{Spec} \mathbb{Z}$. Its value at (3) is "$1 \pmod 3$". Its value at (2) is "$0 \pmod 2$", and in fact it has a double zero. $27/4$ is a "rational function" on $\operatorname{Spec}\mathbb{Z}$, defined away from (2). We want to say that it has a double pole at (2), and a triple zero at (3). Its value at (5) is

$$27 \times 4^{-1} \equiv 2 \times (-1) \equiv 3 \pmod 5.$$

(We will gradually make this discussion precise over time.)

Example 4: silly but important examples, and the German word for bacon. The set $\operatorname{Spec} k$ where k is any field is boring: one point. $\operatorname{Spec} 0$, where 0 is the zero-ring, is the empty set, as 0 has no prime ideals.

3.2.A. A SMALL EXERCISE ABOUT SMALL SCHEMES.
(a) Describe the set $\operatorname{Spec} k[\epsilon]/(\epsilon^2)$. The ring $k[\epsilon]/(\epsilon^2)$ is called the ring of **dual numbers**, and will turn out to be quite useful. You should think of ϵ as a very small number, so small that its square is 0 (although it itself is not 0). It is a nonzero function whose value at all points is zero, thus giving our first example of functions not being determined by their values at points. We will discuss this phenomenon further in §3.2.11.
(b) Describe the set $\operatorname{Spec} k[x]_{(x)}$ (see §1.3.3 for a discussion of localization). We will see this scheme again repeatedly, starting with §3.2.8 and Exercise 3.4.K. You might later think of it as a shred of a particularly nice "smooth curve".

In Example 2, we restricted to the case of algebraically closed fields for a reason: things are more subtle if the field is not algebraically closed.

Example 5 (the affine line over $\mathbb{R}$): $\mathbb{R}[x]$. Using the fact that $\mathbb{R}[x]$ is a unique factorization domain, similar arguments to those of Examples 1–3 show that the primes are (0), $(x-a)$ where $a \in \mathbb{R}$, and (x^2+ax+b) where x^2+ax+b is an irreducible quadratic. The latter two are maximal ideals, i.e., their quotients are fields. For example: $\mathbb{R}[x]/(x-3) \cong \mathbb{R}$, $\mathbb{R}[x]/(x^2+1) \cong \mathbb{C}$.

3.2.B. UNIMPORTANT EXERCISE. Show that for the last type of prime, of the form (x^2+ax+b), the quotient is *always* isomorphic to $\mathbb{C}$.

So we have the points that we would normally expect to see on the real line, corresponding to real numbers; the generic point 0; and new points which we may interpret as *conjugate pairs* of complex numbers (the roots of the quadratic). This last type of point should be seen as more akin to the real numbers than to the generic point. You can picture $\mathbb{A}^1_{\mathbb{R}}$ as the complex plane, folded along the real axis. But the key point is that Galois-conjugate points (such as i and $-i$) are considered glued.

Let's explore functions on this space. Consider the function $f(x) = x^3 - 1$. Its value at the point $[(x-2)]$ is $f(x) = 7$, or perhaps better, $7 \pmod{x-2}$. How about at (x^2+1)? We get

$$x^3 - 1 \equiv -x - 1 \pmod{x^2+1},$$

which may be profitably interpreted as $-i-1$.

One moral of this example is that we can work over a non-algebraically closed field if we wish. It is more complicated, but we can recover much of the information we care about.

3.2.C. IMPORTANT EXERCISE. Describe the set $\mathbb{A}^1_{\mathbb{Q}}$. (This is harder to picture in a way analogous to $\mathbb{A}^1_{\mathbb{R}}$. But the rough cartoon of points on a line, as in Figure 3.1, remains a reasonable sketch.)

Example 6 (the affine line over $\mathbb{F}_p$): $\mathbb{A}^1_{\mathbb{F}_p} = \operatorname{Spec} \mathbb{F}_p[x]$. As in the previous examples, $\mathbb{F}_p[x]$ is a Euclidean domain, so the prime ideals are of the form (0) or $(f(x))$ where $f(x) \in \mathbb{F}_p[x]$ is an irreducible polynomial, which can be of any degree. Irreducible polynomials correspond to sets of Galois conjugates in $\overline{\mathbb{F}}_p$.

Note that $\operatorname{Spec} \mathbb{F}_p[x]$ has p points corresponding to the elements of $\mathbb{F}_p$, but also many more (infinitely more, see Exercise 3.2.D). This makes this space much richer than simply p points. For example, a polynomial $f(x)$ is not determined by its values at the p elements of $\mathbb{F}_p$, but it *is* determined by its values at the points of $\operatorname{Spec} \mathbb{F}_p[x]$. (As we have mentioned before, this is not true for all schemes.)

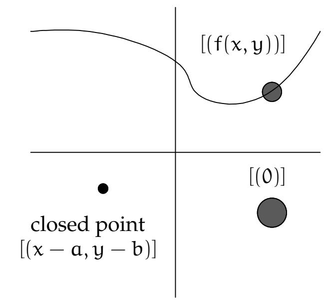
You should think about this, even if you are a geometric person — this intuition will later turn up in geometric situations. Even if you think you are interested only in working over an algebraically closed field (such as $\mathbb{C}$), you will have non-algebraically closed fields (such as $\mathbb{C}(x)$) forced upon you.

3.2.D. EXERCISE. If k is a field, show that $\operatorname{Spec} k[x]$ has infinitely many points. (Hint: Euclid's proof of the infinitude of primes of $\mathbb{Z}$.)

Example 7 (the complex affine plane): $\mathbb{A}^2_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x, y]$. (As with Examples 1 and 2, our discussion will apply with $\mathbb{C}$ replaced by *any* algebraically closed field.) Sadly, $\mathbb{C}[x, y]$ is not a principal ideal domain: (x, y) is not a principal ideal. We can quickly name *some* prime ideals. One is (0), which has the same flavor as the (0) ideals in the previous examples. $(x-2, y-3)$ is prime, and indeed maximal, because $\mathbb{C}[x, y]/(x-2, y-3) \cong \mathbb{C}$, where this isomorphism is via $f(x, y) \mapsto f(2, 3)$. More generally, $(x-a, y-b)$ is prime for any $(a, b) \in \mathbb{C}^2$. Also, if $f(x, y)$ is an irreducible polynomial (e.g. $y - x^2$ or $y^2 - x^3$) then $(f(x, y))$ is prime.

3.2.E. EXERCISE. Show that we have identified all the prime ideals of $\mathbb{C}[x, y]$. Hint: Suppose $\mathfrak{p}$ is a prime ideal that is not principal. Show you can find $f(x, y), g(x, y) \in \mathfrak{p}$ with no common factor. By considering the Euclidean algorithm in the Euclidean domain $\mathbb{C}(x)[y]$, show that you can find a nonzero $h(x) \in (f(x, y), g(x, y)) \subset \mathfrak{p}$. Using primality, show that one of the linear factors of $h(x)$, say $(x - a)$, is in $\mathfrak{p}$. Similarly show there is some $(y - b) \in \mathfrak{p}$.

We now attempt to draw a picture of $\mathbb{A}^2_{\mathbb{C}}$ (see Figure 3.3). The maximal primes of $\mathbb{C}[x, y]$ correspond to the traditional points in $\mathbb{C}^2$: $[(x-a, y-b)]$ corresponds to $(a, b) \in \mathbb{C}^2$. We now have to visualize the "bonus points". $[(0)]$ somehow lives behind all of the traditional points; it is somewhere on the plane, but nowhere in particular. So for example, it does not lie on the parabola $y = x^2$. The point $[(y - x^2)]$ lies on the parabola $y = x^2$, but nowhere in particular on it. (Figure 3.3 is a bit misleading. For example, the point $[(0)]$ isn't in the fourth quadrant; it is somehow near every other point, which is why it is depicted as a somewhat diffuse large dot.) You can see from this picture that we already are implicitly thinking about "dimension". The primes $(x-a, y-b)$ are somehow of dimension 0, the primes $(f(x, y))$ are of dimension 1, and (0) is of dimension 2. (All of our dimensions here are *complex* or *algebraic* dimensions. The complex plane $\mathbb{C}^2$ has real dimension 4, but complex dimension 2. Complex dimensions are in general



[change "closed point" to [begin]"closed point", see Definition 3.6.8[end])

FIGURE 3.3. Picturing $\mathbb{A}^2_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x, y]$

half of real dimensions.) We won't define dimension precisely until Chapter 11, but you should feel free to keep it in mind before then.

Note too that maximal ideals correspond to the "smallest" points. Smaller ideals correspond to "bigger" points. "One prime ideal contains another" means that the points "have the opposite containment." All of this will be made precise once we have a topology. This order-reversal is a little confusing, and will remain so even once we have made the notions precise.

We now come to the obvious generalization of Example 7.

Example 8 (complex affine n-space — important!): Let $\mathbb{A}^n_{\mathbb{C}} := \operatorname{Spec} \mathbb{C}[x_1, \dots, x_n]$. (More generally, $\mathbb{A}^n_A$ is defined to be $\operatorname{Spec} A[x_1, \dots, x_n]$, where A is an arbitrary ring. When the base ring is clear from context, the subscript A is often omitted. For pedants: the notation $\mathbb{A}^n_A$ implicitly includes the data of n "coordinate functions".) For concreteness, let's consider $n = 3$. We now have an interesting question in what at first appears to be pure algebra: What are the prime ideals of $\mathbb{C}[x, y, z]$?

Analogously to before, $(x - a, y - b, z - c)$ is a prime ideal. This is a maximal ideal, because its residue ring is a field ($\mathbb{C}$); we think of these as "0-dimensional points". We will often write (a, b, c) for $[(x-a, y-b, z-c)]$ because of our geometric interpretation of these ideals. There are no more maximal ideals, by Hilbert's Weak Nullstellensatz.

3.2.4. Hilbert's Weak Nullstellensatz. — *If k is an algebraically closed field, then the maximal ideals of $k[x_1, \dots, x_n]$ are precisely those ideals of the form $(x_1 - a_1, \dots, x_n - a_n)$, where $a_i \in k$.*

We may as well state a slightly stronger version now.

3.2.5. Hilbert's Nullstellensatz. — *If k is any field, every maximal ideal of $k[x_1, \dots, x_n]$ has residue field a finite extension of k. Translation: any field extension of k that is finitely generated as a ring is necessarily also finitely generated as a module (i.e., is a finite field extension).*

3.2.F. EXERCISE. Show that the Nullstellensatz 3.2.5 implies the Weak Nullstellensatz 3.2.4.

We will prove the Nullstellensatz in §7.4.3, and again in Exercise 11.2.B.

The following fact is a useful accompaniment to the Nullstellensatz.

3.2.G. EXERCISE (NOT REQUIRING THE NULLSTELLENSATZ). Any integral domain A which is a finite k-algebra (i.e., a finite-dimensional vector space) must be a field. Hint: for any $x \in A$, show $\times x : A \to A$ is an isomorphism. (Thus, in combination the Nullstellensatz 3.2.5, we see that *prime ideals of* $k[x_1, \dots, x_n]$ *with finite residue ring are the same as maximal ideals of* $k[x_1, \dots, x_n]$. This is worth remembering.)

There are other prime ideals of $\mathbb{C}[x, y, z]$ too. We have (0), which corresponds to a "3-dimensional point". We have $(f(x, y, z))$, where f is irreducible. To this we associate the "hypersurface" $f = 0$, so this is "2-dimensional" in nature. But we have not found them all! One clue: we have prime ideals of "dimension" 0, 2, and 3 — we are missing "dimension 1". Here is one such prime ideal: (x, y). We picture this as the locus where $x = y = 0$, which is the z-axis. This is a prime ideal, as the corresponding quotient $\mathbb{C}[x, y, z]/(x, y) \cong \mathbb{C}[z]$ is an integral domain (and should be interpreted as the functions on the z-axis). There are lots of "1-dimensional primes", and it is not possible to classify them in a reasonable way. It will turn out that they correspond to things that we think of as irreducible curves. Thus remarkably the answer to the purely algebraic question ("what are the primes of $\mathbb{C}[x, y, z]$") is fundamentally geometric!

The fact that the points of $\mathbb{A}^1_{\mathbb{Q}}$ corresponding to maximal ideals of the ring $\mathbb{Q}[x]$ (what we will soon call "closed points", see Definition 3.6.8) can be interpreted as points of $\overline{\mathbb{Q}}$ where Galois-conjugates are glued together (Exercise 3.2.C) extends to $\mathbb{A}^n_{\mathbb{Q}}$. For example, in $\mathbb{A}^2_{\mathbb{Q}}$, $(\sqrt{2}, \sqrt{2})$ is glued to $(-\sqrt{2}, -\sqrt{2})$ but not to $(\sqrt{2}, -\sqrt{2})$. The following exercise will give you some idea of how this works.

3.2.H. EXERCISE. Describe the maximal ideal of $\mathbb{Q}[x, y]$ corresponding to $(\sqrt{2}, \sqrt{2})$ and $(-\sqrt{2}, -\sqrt{2})$. Describe the maximal ideal of $\mathbb{Q}[x, y]$ corresponding to $(\sqrt{2}, -\sqrt{2})$ and $(-\sqrt{2}, \sqrt{2})$. What are the residue fields in both cases?

The description of "closed points" of $\mathbb{A}^2_{\mathbb{Q}}$ (those points corresponding to maximal ideals of the ring $\mathbb{Q}[x, y]$) as Galois-orbits can even be extended to other "non-closed" points, as follows.

3.2.I. UNIMPORTANT AND TRICKY BUT FUN EXERCISE. Consider the map of sets $\phi : \mathbb{C}^2 \to \mathbb{A}^2_{\mathbb{Q}}$ defined as follows. (z_1, z_2) is sent to the prime ideal of $\mathbb{Q}[x, y]$ consisting of polynomials vanishing at (z_1, z_2).
(a) What is the image of (π, π^2)?
$\star$(b) Show that ϕ is surjective. (Warning: You will need some ideas we haven't discussed in order to solve this. Once we define the Zariski topology on $\mathbb{A}^2_{\mathbb{Q}}$, you will be able to check that ϕ is continuous, where we give $\mathbb{C}^2$ the classical topology. This example generalizes.)

3.2.6. Quotients and localizations. Two natural ways of getting new rings from old — quotients and localizations — have interpretations in terms of spectra.

3.2.7. *Quotients:* $\operatorname{Spec} A/I$ *as a subset of* $\operatorname{Spec} A$. It is an important fact that the primes of A/I are in bijection with the primes of A containing I.

3.2.J. ESSENTIAL ALGEBRA EXERCISE (MANDATORY IF YOU HAVEN'T SEEN IT BEFORE). Suppose A is a ring, and I an ideal of A. Let $\phi : A \to A/I$. Show that ϕ^{-1} gives an inclusion-preserving bijection between primes of A/I and primes of A containing I. Thus we can picture $\operatorname{Spec} A/I$ as a subset of $\operatorname{Spec} A$.

As an important motivational special case, you now have a picture of *affine complex varieties*. Suppose A is a finitely generated $\mathbb{C}$-algebra, generated by $x_1, \dots, x_n$, with relations $f_1(x_1, \dots, x_n) = \cdots = f_r(x_1, \dots, x_n) = 0$. Then this description in terms of generators and relations naturally gives us an interpretation of $\operatorname{Spec} A$ as a subset of $\mathbb{A}^n_{\mathbb{C}}$, which we think of as "traditional points" (n-tuples of complex numbers) along with some "bonus" points we haven't yet fully described. To see which of the traditional points are in $\operatorname{Spec} A$, we simply solve the equations $f_1 = \cdots = f_r = 0$. For example, $\operatorname{Spec} \mathbb{C}[x, y, z]/(x^2+y^2-z^2)$ may be pictured as shown in Figure 3.4. (Admittedly this is just a "sketch of the $\mathbb{R}$-points", but we will still find it helpful later.) This entire picture carries over (along with the Nullstellensatz) with $\mathbb{C}$ replaced by any algebraically closed field. Indeed, the picture of Figure 3.4 can be said to depict $k[x, y, z]/(x^2 + y^2 - z^2)$ for most algebraically closed fields k (although it is misleading in characteristic 2, because of the coincidence $x^2 + y^2 - z^2 = (x + y + z)^2$).

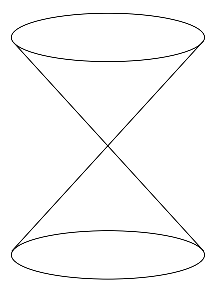


FIGURE 3.4. A "picture" of $\operatorname{Spec} \mathbb{C}[x, y, z]/(x^2 + y^2 - z^2)$

3.2.8. *Localizations:* $\operatorname{Spec} S^{-1}A$ *as a subset of* $\operatorname{Spec} A$. The following exercise shows how prime ideals behave under localization.

3.2.K. ESSENTIAL ALGEBRA EXERCISE (MANDATORY IF YOU HAVEN'T SEEN IT BEFORE). Suppose S is a multiplicative subset of A. Describe an order-preserving bijection of the primes of $S^{-1}A$ with the primes of A that *don't meet* the multiplicative set S.

Recall from §1.3.3 that there are two important flavors of localization. The first is $A_f = \{1, f, f^2, \dots\}^{-1}A$ where $f \in A$. A motivating example is $A = \mathbb{C}[x, y]$, $f = y - x^2$. The second is $A_{\mathfrak{p}} = (A - \mathfrak{p})^{-1}A$, where $\mathfrak{p}$ is a prime ideal. A motivating example is $A = \mathbb{C}[x, y]$, $S = A - (x, y)$.

If $S = \{1, f, f^2, \dots\}$, the primes of $S^{-1}A$ are just those primes not containing f — the points where "f doesn't vanish". (In §3.5, we will call this a *distinguished open set*, once we know what open sets are.) So to picture $\operatorname{Spec} \mathbb{C}[x, y]_{y-x^2}$, we picture the affine plane, and throw out those points on the parabola $y - x^2$ — the points (a, a^2) for $a \in \mathbb{C}$ (by which we mean $[(x - a, y - a^2)]$), as well as the "new kind of point" $[(y - x^2)]$.

It can be initially confusing to think about localization in the case where zerodivisors are inverted, because localization $A \to S^{-1}A$ is not injective (Exercise 1.3.C). Geometric intuition can help. Consider the case $A = \mathbb{C}[x, y]/(xy)$ and $f = x$. What is the localization A_f? The space $\operatorname{Spec} \mathbb{C}[x, y]/(xy)$ "is" the union of the two axes in the plane. Localizing means throwing out the locus where x vanishes. So we are left with the x-axis, minus the origin, so we expect $\operatorname{Spec} \mathbb{C}[x]_x$. So there should be some natural isomorphism $(\mathbb{C}[x, y]/(xy))_x \cong \mathbb{C}[x]_x$.

3.2.L. EXERCISE. Show that these two rings are isomorphic. (You will see that y on the left goes to 0 on the right.)

If $S = A - \mathfrak{p}$, the primes of $S^{-1}A$ are just the primes of A contained in $\mathfrak{p}$. In our example $A = \mathbb{C}[x, y]$, $\mathfrak{p} = (x, y)$, we keep all those points corresponding to "things through the origin", i.e., the 0-dimensional point (x, y), the 2-dimensional point (0), and those 1-dimensional points $(f(x, y))$ where $f(0, 0) = 0$, i.e., those "irreducible curves through the origin". You can think of this being a shred of the plane near the origin; anything not actually "visible" at the origin is discarded (see Figure 3.5).

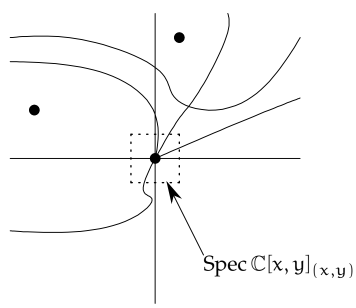


FIGURE 3.5. Picturing $\operatorname{Spec} \mathbb{C}[x, y]_{(x,y)}$ as a "shred of $\mathbb{A}^2_{\mathbb{C}}$". Only those points near the origin remain.

Another example is when $A = k[x]$, and $\mathfrak{p} = (x)$ (or more generally when $\mathfrak{p}$ is any maximal ideal). Then $A_{\mathfrak{p}}$ has only two prime ideals (Exercise 3.2.A(b)). You should see this as the germ of a "smooth curve", where one point is the "classical point", and the other is the "generic point of the curve". This is an example of a discrete valuation ring, and indeed all discrete valuation rings should be visualized in such a way. We will discuss discrete valuation rings in §12.5. By then we will have justified the use of the words "smooth" and "curve". (Reality check: try to picture Spec of $\mathbb{Z}$ localized at (2) and at (0). How do the two pictures differ?)

3.2.9. Important fact: Maps of rings induce maps of spectra (as sets). We now make an observation that will later grow up to be the notion of morphisms of schemes.

3.2.M. IMPORTANT EASY EXERCISE. If $\phi : B \to A$ is a map of rings, and $\mathfrak{p}$ is a prime ideal of A, show that $\phi^{-1}(\mathfrak{p})$ is a prime ideal of B.

Hence a map of rings $\phi : B \to A$ induces a map of sets $\operatorname{Spec} A \to \operatorname{Spec} B$ "in the opposite direction". This gives a contravariant functor from the category of rings to the category of sets: the composition of two maps of rings induces the composition of the corresponding maps of spectra.

3.2.N. EASY EXERCISE (REALITY CHECK). Let B be a ring.
(a) Suppose $I \subset B$ is an ideal. Show that the map $\operatorname{Spec} B/I \to \operatorname{Spec} B$ is the inclusion of §3.2.7.
(b) Suppose $S \subset B$ is a multiplicative set. Show that the map $\operatorname{Spec} S^{-1}B \to \operatorname{Spec} B$ is the inclusion of §3.2.8.

3.2.10. *An explicit example.* In the case of "affine complex varieties" (or indeed affine varieties over any algebraically closed field), the translation between maps given by explicit formulas and maps of rings is quite direct. For example, consider a map from the parabola in $\mathbb{C}^2$ (with coordinates a and b) given by $b = a^2$, to the "curve" in $\mathbb{C}^3$ (with coordinates x, y, and z) cut out by the equations $y = x^2$ and $z = y^2$. Suppose the map sends the point $(a, b) \in \mathbb{C}^2$ to the point $(a, b, b^2) \in \mathbb{C}^3$. In our new language, we have a map

$$\operatorname{Spec} \mathbb{C}[a,b]/(b - a^2) \longrightarrow \operatorname{Spec} \mathbb{C}[x,y,z]/(y - x^2, z - y^2)$$

given by

$$\mathbb{C}[a,b]/(b - a^2) \longleftarrow \mathbb{C}[x,y,z]/(y - x^2, z - y^2)$$

$$(a, b, b^2) \longleftarrow\!\shortmid (x, y, z),$$

i.e., $x \mapsto a$, $y \mapsto b$, and $z \mapsto b^2$. If the idea is not yet clear, the following two exercises are very much worth doing — they can be very confusing the first time you see them, and very enlightening (and finally, trivial) when you finally figure them out.

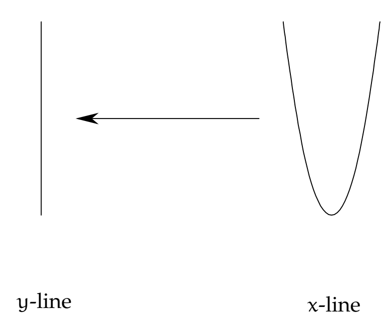


FIGURE 3.6. The map $\mathbb{C} \to \mathbb{C}$ given by $x \mapsto y = x^2$

3.2.O. IMPORTANT EXERCISE (SPECIAL CASE). Consider the map of complex manifolds sending $\mathbb{C} \to \mathbb{C}$ via $x \mapsto y = x^2$. We interpret the "source" $\mathbb{C}$ as the "x-line", and the "target" $\mathbb{C}$ the "y-line". You can picture it as the projection of the parabola $y = x^2$ in the xy-plane to the y-axis (see Figure 3.6). Interpret the corresponding map of rings as given by $\mathbb{C}[y] \to \mathbb{C}[x]$ by $y \mapsto x^2$. Verify that the preimage (the fiber) above the point $a \in \mathbb{C}$ is the point(s) $\pm\sqrt{a} \in \mathbb{C}$, using the definition given above. (A more sophisticated version of this example appears in Example 9.3.3. Warning: the roles of x and y are swapped there, in order to picture double covers in a certain way.)

3.2.P. IMPORTANT EXERCISE (GENERALIZING EXAMPLE 3.2.10). Suppose k is a field, and $f_1, \dots, f_n \in k[x_1, \dots, x_m]$ are given. Let $\phi : k[y_1, \dots, y_n] \to k[x_1, \dots, x_m]$ be the ring morphism defined by $y_i \mapsto f_i$.
(a) Show that ϕ induces a map of sets $\operatorname{Spec} k[x_1, \dots, x_m]/I \to \operatorname{Spec} k[y_1, \dots, y_n]/J$ for any ideals $I \subset k[x_1, \dots, x_m]$ and $J \subset k[y_1, \dots, y_n]$ such that $\phi(J) \subset I$. (You may wish to consider the case $I = 0$ and $J = 0$ first. In fact, part (a) has nothing to do with k-algebras; you may wish to prove the statement when the rings $k[x_1, \dots, x_m]$ and $k[y_1, \dots, y_n]$ are replaced by general rings A and B.)
(b) Show that the map of part (a) sends the point $(a_1, \dots, a_m) \in k^m$ (or more precisely, $[(x_1 - a_1, \dots, x_m - a_m)] \in \operatorname{Spec} k[x_1, \dots, x_m]$) to

$$(f_1(a_1, \dots, a_m), \dots, f_n(a_1, \dots, a_m)) \in k^n.$$

3.2.Q. EXERCISE: PICTURING $\mathbb{A}^n_{\mathbb{Z}}$. Consider the map of sets $\pi : \mathbb{A}^n_{\mathbb{Z}} \to \operatorname{Spec} \mathbb{Z}$, given by the ring map $\mathbb{Z} \to \mathbb{Z}[x_1, \dots, x_n]$. If p is prime, describe a bijection between the fiber $\pi^{-1}([(p)])$ and $\mathbb{A}^n_{\mathbb{F}_p}$. (You won't need to describe either set! Which is good because you can't.) This exercise may give you a sense of how to picture maps (see Figure 3.7), and in particular why you can think of $\mathbb{A}^n_{\mathbb{Z}}$ as an "$\mathbb{A}^n$-bundle" over $\operatorname{Spec} \mathbb{Z}$. (Can you interpret the fiber over $[(0)]$ as $\mathbb{A}^n_k$ for some field k?)

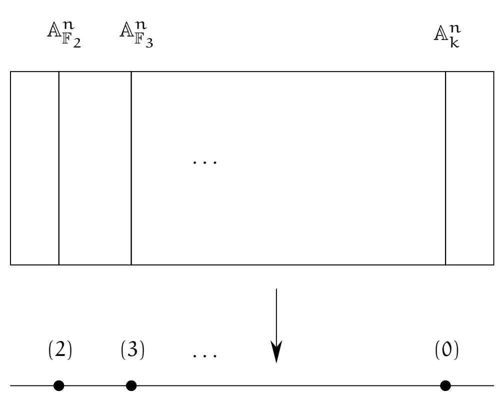


FIGURE 3.7. A picture of $\mathbb{A}^n_{\mathbb{Z}} \to \operatorname{Spec} \mathbb{Z}$ as a "family of $\mathbb{A}^n$'s", or an "$\mathbb{A}^n$-bundle over $\operatorname{Spec} \mathbb{Z}$". What is k?

3.2.11. Functions are not determined by their values at points: the fault of nilpotents. We conclude this section by describing some strange behavior. We are developing machinery that will let us bring our geometric intuition to algebra. There is one serious serious point where your intuition will be false, so you should know now, and adjust your intuition appropriately. As noted by Mumford ([**Mu2**, p. 12]), "it is this aspect of schemes which was most scandalous when Grothendieck defined them."

Suppose we have a function (ring element) vanishing at all points. Then it is not necessarily the zero function! The translation of this question is: is the intersection of all prime ideals necessarily just 0? The answer is no, as is shown by the example of the ring of dual numbers $k[\epsilon]/(\epsilon^2)$: $\epsilon \neq 0$, but $\epsilon^2 = 0$. (We saw this ring in Exercise 3.2.A(a).) Any function whose power is zero certainly lies in the intersection of all prime ideals.

3.2.R. EXERCISE. Ring elements that have a power that is 0 are called **nilpotents**.
(a) Show that if I is an ideal of nilpotents, then the inclusion $\operatorname{Spec} B/I \to \operatorname{Spec} B$ of Exercise 3.2.J is a bijection. Thus nilpotents don't affect the underlying set. (We will soon see in §3.4.5 that they won't affect the topology either — the difference will be in the structure sheaf.)
(b) Show that the nilpotents of a ring B form an ideal. This ideal is called the **nilradical**, and is denoted $\mathfrak{N} = \mathfrak{N}(B)$.

Thus the nilradical is contained in the intersection of all the prime ideals. The converse is also true:

3.2.12. Theorem. — *The nilradical $\mathfrak{N}(A)$ is the intersection of all the primes of A. Geometrically: a function on $\operatorname{Spec} A$ vanishes everywhere if and only if it is nilpotent.*

3.2.S. EXERCISE. If you don't know this theorem, then look it up, or better yet, prove it yourself. (Hint: Use the fact that any proper ideal of A is contained in a maximal ideal, which requires Zorn's lemma. Possible further hint: Suppose $x \notin \mathfrak{N}(A)$. We wish to show that there is a prime ideal not containing x. Show that A_x is not the 0-ring, by showing that $1 \neq 0$.)

3.2.13. In particular, although it is upsetting that functions are not determined by their values at points, we have precisely specified what the failure of this intuition is: two functions have the same values at points if and only if they differ by a nilpotent. You should think of this geometrically: a function vanishes at every point of the spectrum of a ring if and only if it has a power that is zero. And if there are no nonzero nilpotents — if $\mathfrak{N} = (0)$ — then functions *are* determined by their values at points. If a ring has no nonzero nilpotents, we say that it is **reduced**.

3.2.T. FUN UNIMPORTANT EXERCISE: DERIVATIVES WITHOUT DELTAS AND EPSILONS (OR AT LEAST WITHOUT DELTAS). Suppose we have a polynomial $f(x) \in k[x]$. Instead, we work in $k[x, \epsilon]/(\epsilon^2)$. What then is $f(x + \epsilon)$? (Do a couple of examples, then prove the pattern you observe.) This is a hint that nilpotents will be important in defining differential information (Chapter 21).

3.3 Visualizing schemes I: generic points

A heavy warning used to be given that pictures are not rigorous; this has never had its bluff called and has permanently frightened its victims into playing for safety. Some pictures, of course, are not rigorous, but I should say most are (and I use them whenever possible myself).

— *J. E. Littlewood,* [**Lit**, p. 54]

For years, you have been able to picture $x^2 + y^2 = 1$ in the plane, and you now have an idea of how to picture $\operatorname{Spec} \mathbb{Z}$. If we are claiming to understand rings as geometric objects (through the Spec functor), then we should wish to develop geometric insight into them. To develop geometric intuition about schemes, it is helpful to have pictures in your mind, extending your intuition about geometric spaces you are already familiar with. As we go along, we will empirically develop some idea of what schemes should look like. This section summarizes what we have gleaned so far.

Some mathematicians prefer to think completely algebraically, and never think in terms of pictures. Others will be disturbed by the fact that this is an art, not a science. And finally, this hand-waving will necessarily never be used in the rigorous development of the theory. For these reasons, you may wish to skip these sections. However, having the right picture in your mind can greatly help understanding what facts should be true, and how to prove them.

Our starting point is the example of "affine complex varieties" (things cut out by equations involving a finite number variables over $\mathbb{C}$), and more generally similar examples over arbitrary algebraically closed fields. We begin with notions that are intuitive ("traditional" points behaving the way you expect them to), and then add in the two features which are new and disturbing, generic points and nonreduced behavior. You can then extend this notion to seemingly different spaces, such as $\operatorname{Spec} \mathbb{Z}$.

Hilbert's Weak Nullstellensatz 3.2.4 shows that the "traditional points" are present as points of the scheme, and this carries over to any algebraically closed field. If the field is not algebraically closed, the traditional points are glued together into clumps by Galois conjugation, as in Examples 5 (the real affine line) and 6 (the affine line over $\mathbb{F}_p$) in §3.2. This is a geometric interpretation of Hilbert's Nullstellensatz 3.2.5.

But we have some additional points to add to the picture. You should remember that they "correspond" to "irreducible" "closed" (algebraic) subsets. As motivation, consider the case of the complex affine plane (Example 7): we had one for each irreducible polynomial, plus one corresponding to the entire plane. We will make "closed" precise when we define the Zariski topology (in the next section). You may already have an idea of what "irreducible" should mean; we make that precise at the start of §3.6. By "correspond" we mean that each closed irreducible subset has a corresponding point sitting on it, called its *generic point* (defined in §3.6). It is a new point, distinct from all the other points in the subset. (The correspondence is described in Exercise 3.7.E for $\operatorname{Spec} A$, and in Exercise 5.1.B for schemes in general.) We don't know precisely where to draw the generic point, so we may stick it arbitrarily anywhere, but you should think of it as being "almost everywhere", and in particular, near every other point in the subset.

In §3.2.7, we saw how the points of $\operatorname{Spec} A/I$ should be interpreted as a subset of $\operatorname{Spec} A$. So for example, when you see $\operatorname{Spec} \mathbb{C}[x,y]/(x+y)$, you should picture this not just as a line, but as a line in the xy-plane; the choice of generators x and y of the algebra $\mathbb{C}[x,y]$ implies an inclusion into affine space.

In §3.2.8, we saw how the points of $\operatorname{Spec} S^{-1}A$ should be interpreted as subsets of $\operatorname{Spec} A$. The two most important cases were discussed. The points of $\operatorname{Spec} A_f$ correspond to the points of $\operatorname{Spec} A$ where f doesn't vanish; we will later (§3.5) interpret this as a distinguished open set.

If $\mathfrak{p}$ is a prime ideal, then $\operatorname{Spec} A_{\mathfrak{p}}$ should be seen as a "shred of the space $\operatorname{Spec} A$ near the subset corresponding to $\mathfrak{p}$". The simplest nontrivial case of this is $\mathfrak{p} = (x) \subset \operatorname{Spec} k[x] = A$ (see Exercise 3.2.A, which we discuss again in Exercise 3.4.K).

"If any one of them can explain it," said Alice, (she had grown so large in the last few minutes that she wasn't a bit afraid of interrupting him), "I'll give him sixpence. I don't believe there's an atom of meaning in it." ...

"If there's no meaning in it," said the King, "that saves a world of trouble, you know, as we needn't try to find any."

— *Lewis Carroll* [**Carr**, Ch. XII]

3.4 The underlying topological space of an affine scheme

We next introduce the *Zariski topology* on the spectrum of a ring. When you first hear the definition, it seems odd, but with a little experience it becomes reasonable. As motivation, consider $\mathbb{A}^2_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x,y]$, the complex plane (with a few extra points). In algebraic geometry, we will only be allowed to consider algebraic functions, i.e., polynomials in x and y. The locus where a polynomial vanishes should reasonably be a closed set, and the Zariski topology is defined by saying that the only sets we should consider closed should be these sets, and other sets forced to be closed by these. In other words, it is the coarsest topology where these sets are closed.

In particular, although topologies are often described using open subsets, it will be more convenient for us to define this topology in terms of closed subsets. If S is a subset of a ring A, define the V**anishing set** of S by

$$V(S) := \{[\mathfrak{p}] \in \operatorname{Spec} A \ : \ S \subset \mathfrak{p}\}.$$

It is the set of points on which all elements of S are zero. (It should now be second nature to equate "vanishing at a point" with "contained in a prime".) We declare that these — and no other — are the closed subsets.

For example, consider $V(xy, yz) \subset \mathbb{A}^3_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x,y,z]$. Which points are contained in this locus? We think of this as solving $xy = yz = 0$. Of the "traditional" points (interpreted as ordered triples of complex numbers, thanks to the Hilbert's Nullstellensatz 3.2.4), we have the points where $y = 0$ or $x = z = 0$: the xz-plane and the y-axis respectively. Of the "new" points, we have the generic point of the xz-plane (also known as the point $[(y)]$), and the generic point of the y-axis (also known as the point $[(x,z)]$). You might imagine that we also have a number of "one-dimensional" points contained in the xz-plane.

3.4.A. EASY EXERCISE. Check that the x-axis is contained in $V(xy, yz)$. (The **x-axis** is defined by $y = z = 0$, and the y-axis and z-axis are defined analogously.)

Let's return to the general situation. The following exercise lets us restrict attention to vanishing sets of *ideals*.

3.4.B. EASY EXERCISE. Show that if (S) is the ideal generated by S, then $V(S) = V((S))$.

We define the **Zariski topology** by declaring that $V(S)$ is closed for all S. Let's check that this is a topology:

3.4.C. EXERCISE.
(a) Show that $\varnothing$ and $\operatorname{Spec} A$ are both open.
(b) If I_i is a collection of ideals (as i runs over some index set), show that $\cap_i V(I_i) = V(\sum_i I_i)$. Hence the union of any collection of open sets is open.
(c) Show that $V(I_1) \cup V(I_2) = V(I_1 I_2)$. (The **product of two ideals** I_1 and I_2 of A are finite A-linear combinations of products of elements of I_1 and I_2, i.e., elements of the form $\sum_{j=1}^n i_{1,j} i_{2,j}$, where $i_{k,j} \in I_k$. Equivalently, it is the ideal generated by products of elements of I_1 and I_2. You should quickly check that this is an ideal, and that products are associative, i.e., $(I_1 I_2) I_3 = I_1(I_2 I_3)$.) Hence the intersection of any finite number of open sets is open.

3.4.1. Properties of the "vanishing set" function $V(\cdot)$. The function $V(\cdot)$ is obviously inclusion-reversing: If $S_1 \subset S_2$, then $V(S_2) \subset V(S_1)$. Warning: We could have equality in the second inclusion without equality in the first, as the next exercise shows.

3.4.D. EXERCISE/DEFINITION. If $I \subset A$ is an ideal, then define its **radical** by

$$\sqrt{I} := \{r \in A \ : \ r^n \in I \text{ for some } n \in \mathbb{Z}^{>0}\}.$$

For example, the nilradical $\mathfrak{N}$ (§3.2.R) is $\sqrt{(0)}$. Show that $\sqrt{I}$ is an ideal (cf. Exercise 3.2.R(b)). Show that $V(\sqrt{I}) = V(I)$. We say *an ideal is* **radical** if it equals its own radical. Show that $\sqrt{\sqrt{I}} = \sqrt{I}$, and that prime ideals are radical.

Here are two useful consequences. As $(I \cap J)^2 \subset IJ \subset I \cap J$ (products of ideals were defined in Exercise 3.4.C), we have that $V(IJ) = V(I \cap J)$ $(= V(I) \cup V(J)$ by Exercise 3.4.C(c)). Also, combining this with Exercise 3.4.B, we see $V(S) = V((S)) = V(\sqrt{(S)})$.

3.4.E. EXERCISE (RADICALS COMMUTE WITH FINITE INTERSECTIONS). If $I_1, \dots, I_n$ are ideals of a ring A, show that $\sqrt{\cap_{i=1}^n I_i} = \cap_{i=1}^n \sqrt{I_i}$. We will use this property repeatedly without referring back to this exercise.

3.4.F. EXERCISE FOR LATER USE. Show that $\sqrt{I}$ is the intersection of all the prime ideals containing I. (Hint: Use Theorem 3.2.12 on an appropriate ring.)

3.4.2. Examples. Let's see how this meshes with our examples from the previous section.

Recall that $\mathbb{A}^1_{\mathbb{C}}$, as a set, was just the "traditional" points (corresponding to maximal ideals, in bijection with $a \in \mathbb{C}$), and one "new" point (0). The Zariski topology on $\mathbb{A}^1_{\mathbb{C}}$ is not that exciting: the open sets are the empty set, and $\mathbb{A}^1_{\mathbb{C}}$ minus a finite number of maximal ideals. (It "almost" has the cofinite topology. Notice that the open sets are determined by their intersections with the "traditional points". The "new" point (0) comes along for the ride, which is a good sign that it is harmless. Ignoring the "new" point, observe that the topology on $\mathbb{A}^1_{\mathbb{C}}$ is a coarser topology than the classical topology on $\mathbb{C}$.)

3.4.G. EXERCISE. Describe the topological space $\mathbb{A}^1_k$ (cf. Exercise 3.2.D).

The case $\operatorname{Spec}\mathbb{Z}$ is similar. The topology is "almost" the cofinite topology in the same way. The open sets are the empty set, and $\operatorname{Spec}\mathbb{Z}$ minus a finite number of "ordinary" ((p) where p is prime) primes.

3.4.3. *Closed subsets of* $\mathbb{A}^2_{\mathbb{C}}$. The case $\mathbb{A}^2_{\mathbb{C}}$ is more interesting. You should think through where the "one-dimensional primes" fit into the picture. In Exercise 3.2.E, we identified all the prime ideals of $\mathbb{C}[x, y]$ (i.e., the points of $\mathbb{A}^2_{\mathbb{C}}$) as the maximal ideals $[(x - a, y - b)]$ (where $a, b \in \mathbb{C}$ — "zero-dimensional points"), the "one-dimensional points" $[(f(x, y))]$ (where $f(x, y)$ is irreducible), and the "two-dimensional point" $[(0)]$.

Then the closed subsets are of the following form:

(a) the entire space (the closure of the "two-dimensional point" $[(0)]$), and

(b) a finite number (possibly none) of "curves" (each the closure of a "one-dimensional point" — the "one-dimensional point" along with the "zero-dimensional points" "lying on it") and a finite number (possibly none) of "zero-dimensional" points (what we will soon call "closed points", see Definition 3.6.8).

We will soon know enough to verify this using general theory, but you can prove it yourself now, using ideas in Exercise 3.2.E. (The key idea: if $f(x, y)$ and $g(x, y)$ are irreducible polynomials that are not multiples of each other, why do their zero sets intersect in a finite number of points?)

3.4.4. Important fact: Maps of rings induce continuous maps of topological spaces. We saw in §3.2.9 that a map of rings $\phi : B \to A$ induces a map of sets $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$.

3.4.H. IMPORTANT EASY EXERCISE. By showing that closed sets pull back to closed sets, show that π is a *continuous* map. Interpret Spec as a contravariant functor *Rings* $\to$ *Top*.

Not all continuous maps arise in this way. Consider for example the continuous map on $\mathbb{A}^1_{\mathbb{C}}$ that is the identity except 0 and 1 (i.e., $[(x)]$ and $[(x - 1)]$) are swapped; no polynomial can manage this marvellous feat.

In §3.2.9, we saw that $\operatorname{Spec} B/I$ and $\operatorname{Spec} S^{-1}B$ are naturally *subsets* of $\operatorname{Spec} B$. It is natural to ask if the Zariski topology behaves well with respect to these inclusions, and indeed it does.

3.4.I. IMPORTANT EXERCISE (CF. EXERCISE 3.2.N). Suppose that $I, S \subset B$ are an ideal and multiplicative subset respectively.
(a) Show that $\operatorname{Spec} B/I$ is naturally a *closed* subset of $\operatorname{Spec} B$. If $S = \{1, f, f^2, \dots\}$ ($f \in B$), show that $\operatorname{Spec} S^{-1}B$ is naturally an *open* subset of $\operatorname{Spec} B$. Show that for arbitrary S, $\operatorname{Spec} S^{-1}B$ need not be open or closed. (Hint: $\operatorname{Spec}\mathbb{Q} \subset \operatorname{Spec}\mathbb{Z}$, or possibly Figure 3.5.)

(b) Show that the Zariski topology on $\operatorname{Spec} B/I$ (resp. $\operatorname{Spec} S^{-1}B$) is the subspace topology induced by inclusion in $\operatorname{Spec} B$. (Hint: compare closed subsets.)

3.4.5. In particular, if $I \subset \mathfrak{N}$ is an ideal of nilpotents, the bijection $\operatorname{Spec} B/I \to \operatorname{Spec} B$ (Exercise 3.2.R) is a homeomorphism. Thus nilpotents don't affect the topological space. (The difference will be in the structure sheaf.)

3.4.J. USEFUL EXERCISE FOR LATER. Suppose $I \subset B$ is an ideal. Show that f vanishes on $V(I)$ if and only if $f \in \sqrt{I}$ (i.e., $f^n \in I$ for some $n \geq 1$). (Hint: Exercise 3.4.F. If you are stuck, you will get another hint when you see Exercise 3.5.E.)

3.4.K. EASY EXERCISE (CF. EXERCISE 3.2.A). Describe the topological space $\operatorname{Spec} k[x]_{(x)}$.

3.5 A base of the Zariski topology on Spec A: Distinguished open sets

If $f \in A$, define the **distinguished open set** $D(f) = \{[\mathfrak{p}] \in \operatorname{Spec} A : f \notin \mathfrak{p}\}$. It is the locus where f doesn't vanish. (I often privately write this as $D(f \neq 0)$ to remind myself of this. I also privately call this a "Doesn't-vanish set" in analogy with $V(f)$ being the Vanishing set.) We have already seen this set when discussing $\operatorname{Spec} A_f$ as a subset of $\operatorname{Spec} A$. For example, we have observed that the Zariski-topology on the distinguished open set $D(f) \subset \operatorname{Spec} A$ coincides with the Zariski topology on $\operatorname{Spec} A_f$ (Exercise 3.4.I).

The reason these sets are important is that they form a particularly nice base for the (Zariski) topology:

3.5.A. EASY EXERCISE. Show that the distinguished open sets form a base for the (Zariski) topology. (Hint: Given a subset $S \subset A$, show that the complement of $V(S)$ is $\cup_{f \in S} D(f)$.)

Here are some important but not difficult exercises to give you a feel for this concept.

3.5.B. EXERCISE. Suppose $f_i \in A$ as i runs over some index set J. Show that $\cup_{i \in J} D(f_i) = \operatorname{Spec} A$ if and only if $(f_i) = A$, or equivalently and very usefully, there are a_i ($i \in J$), all but finitely many 0, such that $\sum_{i \in J} a_i f_i = 1$. (One of the directions will use the fact that any proper ideal of A is contained in some maximal ideal.)

3.5.C. EXERCISE. Show that if $\operatorname{Spec} A$ is an infinite union of distinguished open sets $\cup_{j \in J} D(f_j)$, then in fact it is a union of a finite number of these, i.e., there is a finite subset J' so that $\operatorname{Spec} A = \cup_{j \in J'} D(f_j)$. (Hint: exercise 3.5.B.)

3.5.D. EASY EXERCISE. Show that $D(f) \cap D(g) = D(fg)$.

3.5.E. IMPORTANT EXERCISE (CF. EXERCISE 3.4.J). Show that $D(f) \subset D(g)$ if and only if $f^n \in (g)$ for some $n \geq 1$, if and only if g is an invertible element of A_f.

We will use Exercise 3.5.E often. You can solve it thinking purely algebraically, but the following geometric interpretation may be helpful. (You should try to draw your own picture to go with this discussion.) Inside $\operatorname{Spec} A$, we have the closed subset $V(g) = \operatorname{Spec} A/(g)$, where g vanishes, and its complement $D(g)$, where g doesn't vanish. Then f is a function on this closed subset $V(g)$ (or more precisely, on $\operatorname{Spec} A/(g)$), and by assumption it vanishes at all points of the closed subset. Now any function vanishing at every point of the spectrum of a ring must be nilpotent (Theorem 3.2.12). In other words, there is some n such that $f^n = 0$ in $A/(g)$, i.e., $f^n \equiv 0 \pmod g$ in A, i.e., $f^n \in (g)$.

3.5.F. EASY EXERCISE. Show that $D(f) = \varnothing$ if and only if $f \in \mathfrak{N}$.

3.6 Topological (and Noetherian) properties

Many topological notions are useful when applied to the topological space $\operatorname{Spec} A$, and later, to schemes.

3.6.1. Possible topological attributes of $\operatorname{Spec} A$: connectedness, irreducibility, quasicompactness.

3.6.2. *Connectedness.*
A topological space X is **connected** if it cannot be written as the disjoint union of two nonempty open sets. Exercise 3.6.A below gives an example of a nonconnected $\operatorname{Spec} A$, and the subsequent remark explains that all examples are of this form.

3.6.A. EXERCISE. If $A = A_1 \times A_2 \times \cdots \times A_n$, describe a homeomorphism $\operatorname{Spec} A_1 \coprod \operatorname{Spec} A_2 \coprod \cdots \coprod \operatorname{Spec} A_n \to \operatorname{Spec} A$ for which each $\operatorname{Spec} A_i$ is mapped onto a distinguished open subset $D(f_i)$ of $\operatorname{Spec} A$. Thus $\operatorname{Spec} \prod_{i=1}^n A_i = \coprod_{i=1}^n \operatorname{Spec} A_i$ as topological spaces. (Hint: reduce to $n = 2$ for convenience. Let $f_1 = (1, 0)$ and $f_2 = (0, 1)$.)

3.6.3. *Remark.* An extension of Exercise 3.6.A (that you can prove if you wish) is that $\operatorname{Spec} A$ is not connected if and only if A is isomorphic to the product of nonzero rings A_1 and A_2. The key idea is to show that both conditions are equivalent to there existing nonzero $a_1, a_2 \in A$ for which $a_1^2 = a_1$, $a_2^2 = a_2$, $a_1 + a_2 = 1$, and hence $a_1 a_2 = 0$. An element $a \in A$ satisfying $a^2 = a$ is called an *idempotent*. This will appear as Exercise 9.5.J.

3.6.4. *Irreducibility.*
A topological space is said to be **irreducible** if it is nonempty, and it is not the union of two proper closed subsets. In other words, a nonempty topological space X is irreducible if whenever $X = Y \cup Z$ with Y and Z closed in X, we have $Y = X$ or $Z = X$. This is a less useful notion in classical geometry — $\mathbb{C}^2$ is **reducible** (i.e., not irreducible), but we will see that $\mathbb{A}^2_{\mathbb{C}}$ is irreducible (Exercise 3.6.C).

3.6.B. EASY EXERCISE.
(a) Show that in an irreducible topological space, any nonempty open set is dense. (For this reason, you will see that unlike in the classical topology, in the Zariski

topology, nonempty open sets are all "huge".)
(b) If X is a topological space, and Z (with the subspace topology) is an irreducible subset, then the closure $\overline{Z}$ in X is irreducible as well.

3.6.C. EASY EXERCISE. If A is an integral domain, show that $\operatorname{Spec} A$ is irreducible. (Hint: pay attention to the generic point $[(0)]$.) We will generalize this in Exercise 3.7.F.

3.6.D. EXERCISE. Show that an irreducible topological space is connected.

3.6.E. EXERCISE. Give (with proof!) an example of a ring A where $\operatorname{Spec} A$ is connected but reducible. (Possible hint: a picture may help. The symbol "$\times$" has two "pieces" yet is connected.)

3.6.F. TRICKY EXERCISE.
(a) Suppose $I = (wz-xy, wy-x^2, xz-y^2) \subset k[w,x,y,z]$. Show that $\operatorname{Spec} k[w,x,y,z]/I$ is irreducible, by showing that $k[w,x,y,z]/I$ is an integral domain. (This is hard, so here is one of several possible hints: Show that $k[w,x,y,z]/I$ is isomorphic to the subring of $k[a,b]$ generated by monomials of degree divisible by 3. There are other approaches as well, some of which we will see later. This is an example of a hard question: how do you tell if an ideal is prime?) We will later see this as the cone over the *twisted cubic curve* (the twisted cubic curve is defined in Exercise 8.2.A, and is a special case of a Veronese embedding, §8.2.6).
$\star$ (b) Note that the generators of the ideal of part (a) may be rewritten as the equations ensuring that

$$\operatorname{rank}\begin{pmatrix} w & x & y \\ x & y & z \end{pmatrix} \leq 1,$$

i.e., as the determinants of the 2×2 submatrices. Generalize part (a) to the ideal of rank one $2 \times n$ matrices. This notion will correspond to the cone (§8.2.12) over the *degree n rational normal curve* (Exercise 8.2.J).

3.6.5. *Quasicompactness.*
A topological space X is **quasicompact** if given any cover $X = \cup_{i\in I} U_i$ by open sets, there is a finite subset S of the index set I such that $X = \cup_{i \in S} U_i$. Informally: every open cover has a finite subcover. We will like this condition, because we are afraid of infinity. Depending on your definition of "compactness", this is the definition of compactness, minus possibly a Hausdorff condition. However, this isn't really the algebro-geometric analog of "compact" (we certainly wouldn't want $\mathbb{A}^1_{\mathbb{C}}$ to be compact) — the right analog is "properness" (§10.3).

3.6.G. EXERCISE.
(a) Show that $\operatorname{Spec} A$ is quasicompact. (Hint: Exercise 3.5.C.)
$\star$(b) (less important) Show that in general $\operatorname{Spec} A$ can have nonquasicompact open sets. Possible hint: let $A = k[x_1, x_2, x_3, \dots]$ and $\mathfrak{m} = (x_1, x_2, \dots) \subset A$, and consider the complement of $V(\mathfrak{m})$. This example will be useful to construct other "counterexamples" later, e.g. Exercises 7.1.C and 5.1.J. In Exercise 3.6.T, we will see that such weird behavior doesn't happen for "suitably nice" (Noetherian) rings.

3.6.H. EXERCISE.
(a) If X is a topological space that is a finite union of quasicompact spaces, show

that X is quasicompact.
(b) Show that every closed subset of a quasicompact topological space is quasicompact.

3.6.6. $\star\star$ *Fun but irrelevant remark.* Exercise 3.6.A shows that $\coprod_{i=1}^n \operatorname{Spec} A_i \cong \operatorname{Spec} \prod_{i=1}^n A_i$, but this *never* holds if "n is infinite" and all A_i are nonzero, as Spec of any ring is quasicompact (Exercise 3.6.G(a)). This leads to an interesting phenomenon. We show that $\operatorname{Spec} \prod_{i=1}^\infty A_i$ is "strictly bigger" than $\coprod_{i=1}^\infty \operatorname{Spec} A_i$ where each A_i is isomorphic to the field k. First, we have an inclusion of sets $\coprod_{i=1}^\infty \operatorname{Spec} A_i \hookrightarrow \operatorname{Spec} \prod_{i=1}^\infty A_i$, as there is a maximal ideal of $\prod A_i$ corresponding to each i (precisely, those elements 0 in the ith component.) But there are other maximal ideals of $\prod A_i$. Hint: describe a proper ideal not contained in any of these maximal ideals. (One idea: consider elements $\prod a_i$ that are "eventually zero", i.e., $a_i = 0$ for $i \gg 0$.) This leads to the notion of *ultrafilters*, which are very useful, but irrelevant to our current discussion.

3.6.7. Possible topological properties of points of $\operatorname{Spec} A$**.**

3.6.8. *Definition.* A point of a topological space $p \in X$ is said to be a **closed point** if $\{p\}$ is a closed subset. In the classical topology on $\mathbb{C}^n$, all points are closed. In $\operatorname{Spec} \mathbb{Z}$ and $\operatorname{Spec} k[t]$, all the points are closed except for $[(0)]$.

3.6.I. EXERCISE. Show that the closed points of $\operatorname{Spec} A$ correspond to the maximal ideals.

3.6.9. *Connection to the classical theory of varieties.* Hilbert's Nullstellensatz lets us interpret the closed points of $\mathbb{A}^n_{\mathbb{C}}$ as the n-tuples of complex numbers. More generally, the closed points of $\operatorname{Spec} \overline{k}[x_1, \dots, x_n]/(f_1, \dots, f_r)$ are naturally interpreted as those points in $\overline{k}^n$ satisfying the equations $f_1 = \cdots = f_r = 0$ (see Exercises 3.2.J and 3.2.N(a) for example). Hence from now on we will say "closed point" instead of "traditional point" and "non-closed point" instead of "bonus" point when discussing subsets of $\mathbb{A}^n_k$.

3.6.J. EXERCISE.
(a) Suppose that k is a field, and A is a finitely generated k-algebra. Show that closed points of $\operatorname{Spec} A$ are dense, by showing that if $f \in A$, and $D(f)$ is a nonempty (distinguished) open subset of $\operatorname{Spec} A$, then $D(f)$ contains a closed point of $\operatorname{Spec} A$. Hint: note that A_f is *also* a finitely generated k-algebra. Use the Nullstellensatz 3.2.5 to recognize closed points of Spec of a finitely generated k-algebra B as those for which the residue field is a finite extension of k. Apply this to both $B = A$ and $B = A_f$.
(b) Show that if A is a k-algebra that is not finitely generated the closed points need not be dense. (Hint: Exercise 3.4.K.)

3.6.K. EXERCISE. Suppose k is an algebraically closed field, and $A = k[x_1, \dots, x_n]/I$ is a finitely generated k-algebra with $\mathfrak{N}(A) = \{0\}$ (so the discussion of §3.2.13 applies). Consider the set $X = \operatorname{Spec} A$ as a subset of $\mathbb{A}^n_k$. The space $\mathbb{A}^n_k$ contains the "classical" points k^n. Show that functions on X are determined by their values on the closed points (by the weak Nullstellensatz 3.2.4, the "classical" points

$k^n \cap \operatorname{Spec} A$ of $\operatorname{Spec} A$). Hint: if f and g are different functions on X, then $f - g$ is nowhere zero on an open subset of X. Use Exercise 3.6.J(a).

Once we know what a variety is (Definition 10.1.7), this will immediately imply that *a function on a variety over an algebraically closed field is determined by its values on the "classical points".* (Before the advent of scheme theory, functions on varieties — over algebraically closed fields — were thought of as functions on "classical" points, and Exercise 3.6.K basically shows that there is no harm in thinking of "traditional" varieties as a particular flavor of schemes.)

3.6.10. *Specialization and generization.* Given two points x, y of a topological space X, we say that x is a **specialization** of y, and y is a **generization** of x, if $x \in \overline{\{y\}}$. This (and Exercise 3.6.L) now makes precise our hand-waving about "one point containing another". It is of course nonsense for a point to contain another. But it is not nonsense to say that the closure of a point contains another. For example, in $\mathbb{A}^2_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x, y]$, $[(y - x^2)]$ is a generization of $[(x - 2, y - 4)] = (2, 4) \in \mathbb{C}^2$, and $(2, 4)$ is a specialization of $[(y - x^2)]$ (see Figure 3.8.

[to be made]

FIGURE 3.8. $(2, 4) = [(x - 2y - 4)]$ is a specialization of $[(y - x^2)]$. $[(y - x^2)]$ is a generization of $(2, 4)$.

3.6.L. EXERCISE. If $X = \operatorname{Spec} A$, show that $[\mathfrak{q}]$ is a specialization of $[\mathfrak{p}]$ if and only if $\mathfrak{p} \subset \mathfrak{q}$. Hence show that $V(\mathfrak{p}) = \overline{\{[\mathfrak{p}]\}}$.

3.6.11. *Definition.* We say that a point $p \in X$ is a **generic point** for a closed subset K if $\overline{\{p\}} = K$.

This important notion predates Grothendieck. The early twentieth-century Italian algebraic geometers had a notion of "generic points" of a variety, by which they meant points with no special properties, so that anything proved of "a generic point" was true of "almost all" the points on that variety. The modern "generic point" has the same intuitive meaning. If something is "generically" or "mostly" true for the points of an irreducible subset, in the sense of being true for a dense open subset (for "almost all points"), then it is true for the generic point, and vice versa. (This is a statement of principle, not of fact. An interesting case is "reducedness", for which this principle does not hold in general, but *does* hold for "reasonable" schemes such as varieties, see Remark 5.2.2.) For example, a function has value zero at the generic point of an integral scheme if and only if it has the value zero at all points. (See Exercise 5.5.C and the paragraph following it, although you will be able to prove this yourself well before then.) You should keep an eye out for other examples of this.

The phrase **general point** does not have the same meaning. The phrase "the general point of K satisfies such-and-such a property" means "every point of some dense open subset of X satisfies such-and-such a property". Be careful not to confuse "general" and "generic". But be warned that accepted terminology does not always follow this convention; witness "generic freeness", "generic flatness", and "generic smoothness".

3.6.M. EXERCISE. Verify that $[(y-x^2)] \in \mathbb{A}^2$ is a generic point for $V(y-x^2)$.

As more motivation for the terminology "generic": we think of $[(y-x^2)]$ as being some nonspecific point on the parabola (with the closed points $(a, a^2) \in \mathbb{C}^2$, i.e., $(x-a, y-a)^2$ for $a \in \mathbb{C}$, being "specific points"); it is "generic" in the conventional sense of the word. We might "specialize it" to a specific point of the parabola; hence for example $(2, 4)$ is a specialization of $[(y-x^2)]$. (Again, see Figure 3.8.)

We will soon see (Exercise 3.7.E) that there is a natural bijection between points of Spec A and irreducible closed subsets of Spec A, sending each point to its closure, and each irreducible closed subset to its (unique) generic point. You can prove this now, but we will wait until we have developed some convenient terminology.

3.6.12. Irreducible and connected components, and Noetherian conditions.

An **irreducible component** of a topological space is a maximal irreducible subset (an irreducible subset not contained in any larger irreducible subset). Irreducible components are closed (as the closure of irreducible subsets are irreducible, Exercise 3.6.B(b)), and it can be helpful to think of irreducible components of a topological space X as maximal among the irreducible *closed* subsets of X. We think of these as the "pieces of X" (see Figure 3.9).

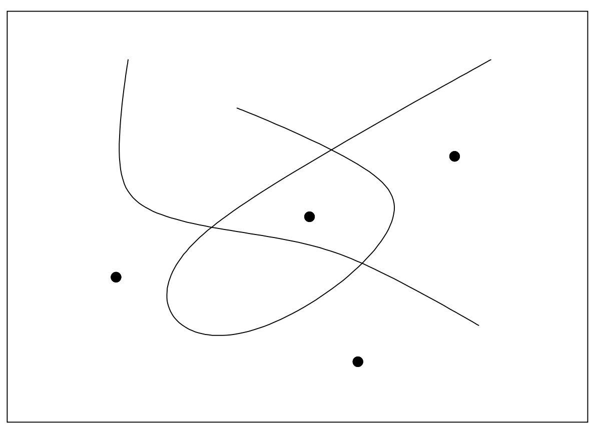


FIGURE 3.9. This closed subset of $\mathbb{A}^2_{\mathbb{C}}$ has six irreducible components

Similarly, a subset Y of a topological space X is a **connected component** if it is a maximal connected subset (a connected subset not contained in any larger connected subset).

3.6.N. EXERCISE (EVERY TOPOLOGICAL SPACE IS THE UNION OF IRREDUCIBLE COMPONENTS). Show that every point x of a topological space X is contained in an irreducible component of X. Hint: Zorn's Lemma. More precisely, consider the partially ordered set $\mathscr{S}$ of irreducible closed subsets of X containing x. Show that there exists a maximal totally ordered subset $\{Z_\alpha\}$ of $\mathscr{S}$. Show that $\cup Z_\alpha$ is irreducible.

3.6.13. *Remark.* Every point is contained in a connected component, and connected components are always closed. You can prove this now, but we deliberately postpone asking this as an exercise until we need it, in an optional starred section (Exercise 9.5.H). On the other hand, connected components need not be open, see [**Stacks**, tag 004T]. An example of an affine scheme with connected components that are not open is $\operatorname{Spec}(\prod_1^\infty \mathbb{F}_2)$.

3.6.14. In the examples we have considered, the spaces have naturally broken up into a finite number of irreducible components. For example, the locus $xy = 0$ in $\mathbb{A}^2_{\mathbb{C}}$ we think of as having two "pieces" — the two axes. The reason for this is that their underlying topological spaces (as we shall soon establish) are *Noetherian*. A topological space X is called **Noetherian** if it satisfies the **descending chain condition** for closed subsets: any sequence $Z_1 \supset Z_2 \supset \cdots \supset Z_n \supset \cdots$ of closed subsets eventually stabilizes: there is an r such that $Z_r = Z_{r+1} = \cdots$. Here is a first example (which you should work out explicitly, not using Noetherian rings).

3.6.O. EXERCISE. Show that $\mathbb{A}^2_{\mathbb{C}}$ is a Noetherian topological space: any decreasing sequence of closed subsets of $\mathbb{A}^2_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x, y]$ must eventually stabilize. Note that it can take arbitrarily long to stabilize. (The closed subsets of $\mathbb{A}^2_{\mathbb{C}}$ were described in §3.4.3.) Show that $\mathbb{C}^2$ with the classical topology is *not* a Noetherian topological space.

3.6.15. Proposition. — *Suppose X is a Noetherian topological space. Then every nonempty closed subset Z can be expressed uniquely as a finite union $Z = Z_1 \cup \cdots \cup Z_n$ of irreducible closed subsets, none contained in any other.*

Translation: any closed subset Z has a finite number of "pieces".

Proof. The following technique is called **Noetherian induction**, for reasons that will be clear. We will use it again, many times.

Consider the collection of closed subsets of X that *cannot* be expressed as a finite union of irreducible closed subsets. We will show that it is empty. Otherwise, let Y_1 be one such. If Y_1 properly contains another such, then choose one, and call it Y_2. If Y_2 properly contains another such, then choose one, and call it Y_3, and so on. By the descending chain condition, this must eventually stop, and we must have some Y_r that cannot be written as a finite union of irreducible closed subsets, but every closed subset properly contained in it can be so written. But then Y_r is not itself irreducible, so we can write $Y_r = Y' \cup Y''$ where Y' and Y'' are both proper closed subsets. Both of these by hypothesis can be written as the union of a finite number of irreducible subsets, and hence so can Y_r, yielding a contradiction. Thus each closed subset can be written as a finite union of irreducible closed subsets. We can assume that none of these irreducible closed subsets contain any others, by discarding some of them.

We now show uniqueness. Suppose

$$Z = Z_1 \cup Z_2 \cup \cdots \cup Z_r = Z'_1 \cup Z'_2 \cup \cdots \cup Z'_s$$

are two such representations. Then $Z_1' \subset Z_1 \cup Z_2 \cup \cdots \cup Z_r$, so $Z_1' = (Z_1 \cap Z_1') \cup \cdots \cup (Z_r \cap Z_1')$. Now Z_1' is irreducible, so one of these is Z_1' itself, say (without loss of generality) $Z_1 \cap Z_1'$. Thus $Z_1' \subset Z_1$. Similarly, $Z_1 \subset Z_a'$ for some a; but because $Z_1' \subset Z_1 \subset Z_a'$, and Z_1' is contained in no other Z_i', we must have $a = 1$, and $Z_1' = Z_1$. Thus each element of the list of Z's is in the list of Z''s, and vice versa, so they must be the same list. □

3.6.P. EXERCISE. Show that every connected component of a topological space X is the union of irreducible components. Show that any subset of X that is simultaneously open and closed must be the union of some of the connected components of X. If X is a *Noetherian* topological space, show that the union of any subset of the connected components of X is always open and closed in X. (In particular, connected components of Noetherian topological spaces are always open, which is not true for more general topological spaces, see Remark 3.6.13.)

3.6.16. Noetherian rings. It turns out that all of the spectra we have considered (except in starred Exercise 3.6.G(b)) are Noetherian topological spaces, but that isn't true of the spectra of all rings. The key characteristic all of our examples have had in common is that the rings were *Noetherian*. A ring is **Noetherian** if every ascending sequence $I_1 \subset I_2 \subset \cdots$ of ideals eventually stabilizes: there is an r such that $I_r = I_{r+1} = \cdots$. (This is called the **ascending chain condition** on ideals.)

Here are some quick facts about Noetherian rings. You should be able to prove them all.

- Fields are Noetherian. $\mathbb{Z}$ is Noetherian.
- If A is Noetherian, and $\phi : A \to B$ is any ring homomorphism, then $\phi(A)$ is Noetherian. Equivalently, quotients of Noetherian rings are Noetherian.
- If A is Noetherian, and S is any multiplicative set, then $S^{-1}A$ is Noetherian.

An important related notion is that of a Noetherian *module*. Although we won't use this notion for some time (§9.7.3), we will develop their most important properties in §3.6.18, while Noetherian ideals are still fresh in your mind.

3.6.Q. IMPORTANT EXERCISE. Show that a ring A is Noetherian if and only if every ideal of A is finitely generated.

The next fact is nontrivial.

3.6.17. The Hilbert Basis Theorem. — *If A is Noetherian, then so is $A[x]$.*

Hilbert proved this in the epochal paper [**Hil**] where he also proved the Hilbert Syzygy Theorem (§15.3.2), and defined Hilbert functions and showed that they are eventually polynomial (§18.6).

By the results described above, any polynomial ring over any field, or over the integers, is Noetherian — and also any quotient or localization thereof. Hence for example any finitely generated algebra over k or $\mathbb{Z}$, or any localization thereof, is Noetherian. Most "nice" rings are Noetherian, but not all rings are Noetherian: $k[x_1, x_2, \ldots]$ is not, because $(x_1) \subset (x_1, x_2) \subset (x_1, x_2, x_3) \subset \cdots$ is a strictly ascending chain of ideals (cf. Exercise 3.6.G(b)).

Proof of the Hilbert Basis Theorem 3.6.17. We show that any ideal $I \subset A[x]$ is finitely generated. We inductively produce a set of generators $f_1, \dots$ as follows. For $n > 0$, if $I \neq (f_1, \dots, f_{n-1})$, let f_n be any nonzero element of $I - (f_1, \dots, f_{n-1})$ of lowest degree. Thus f_1 is any element of I of lowest degree, assuming $I \neq (0)$. If this procedure terminates, we are done. Otherwise, let $a_n \in A$ be the initial coefficient of f_n for $n > 0$. Then as A is Noetherian, $(a_1, a_2, \dots) = (a_1, \dots, a_N)$ for some N. Say $a_{N+1} = \sum_{i=1}^N b_i a_i$. Then

$$f_{N+1} - \sum_{i=1}^{N} b_i f_i x^{\deg f_{N+1} - \deg f_i}$$

is an element of I that is nonzero (as $f_{N+1} \notin (f_1, \dots, f_N)$), and of lower degree than f_{N+1}, yielding a contradiction. $\square$

3.6.R. $\star\star$ UNIMPORTANT EXERCISE. Show that if A is Noetherian, then so is $A[[x]] := \varprojlim A[x]/x^n$, the ring of power series in x. (Possible hint: Suppose $I \subset A[[x]]$ is an ideal. Let $I_n \subset A$ be the coefficients of x^n that appear in the elements of I. Show that I_n is an ideal. Show that $I_n \subset I_{n+1}$, and that I is determined by $(I_0, I_1, I_2, \dots)$.)

We now connect Noetherian rings and Noetherian topological spaces.

3.6.S. EXERCISE. If A is Noetherian, show that $\operatorname{Spec} A$ is a Noetherian topological space. Describe a ring A such that $\operatorname{Spec} A$ is not a Noetherian topological space. (Aside: if $\operatorname{Spec} A$ is a Noetherian topological space, A need not be Noetherian. One example is $A = k[x_1, x_2, x_3, \dots]/(x_1, x_2^2, x_3^3, \dots)$. Then $\operatorname{Spec} A$ has one point, so is Noetherian. But A is not Noetherian as $(x_1) \subsetneq (x_1, x_2) \subsetneq (x_1, x_2, x_3) \subsetneq \cdots$ in A.)

3.6.T. EXERCISE (PROMISED IN EXERCISE 3.6.G(B)). Show that every open subset of a Noetherian topological space is quasicompact. Hence if A is Noetherian, every open subset of $\operatorname{Spec} A$ is quasicompact.

3.6.18. For future use: Noetherian conditions for modules. If A is any ring, not necessarily Noetherian, we say **an A-module is Noetherian** if it satisfies the ascending chain condition for submodules. Thus for example a ring A is Noetherian if and only if it is a Noetherian A-module.

3.6.U. EXERCISE. Show that if M is a Noetherian A-module, then any submodule of M is a finitely generated A-module.

3.6.V. EXERCISE. If $0 \to M' \to M \to M'' \to 0$ is exact, show that M' and M'' are Noetherian if and only if M is Noetherian. (Hint: Given an ascending chain in M, we get two simultaneous ascending chains in M' and M''. Possible further hint: prove that if $M' \longrightarrow M \xrightarrow{\phi} M''$ is exact, and $N \subset N' \subset M$, and $N \cap M' = N' \cap M'$ and $\phi(N) = \phi(N')$, then $N = N'$.)

3.6.W. EXERCISE. Show that if A is a Noetherian ring, then $A^{\oplus n}$ is a Noetherian A-module.

3.6.X. EXERCISE. Show that if A is a Noetherian ring and M is a finitely generated A-module, then M is a Noetherian module. Hence by Exercise 3.6.U, any submodule of a finitely generated module over a Noetherian ring is finitely generated.

3.6.19. Why you should not worry about Noetherian hypotheses. Should you work hard to eliminate Noetherian hypotheses? Should you worry about Noetherian hypotheses? Should you stay up at night thinking about non-Noetherian rings? For the most part, the answer to all of these questions is "no". Most people will never need to worry about non-Noetherian rings, but there are reasons to be open to them. First, they can actually come up. For example, fibered products of Noetherian schemes over Noetherian schemes (and even fibered products of Noetherian points over Noetherian points!) can be non-Noetherian (Warning 9.1.4), and the normalization of Noetherian rings can be non-Noetherian (Warning 9.7.4). You can either work hard to show that the rings or schemes you care about don't have this pathology, or you can just relax and not worry about it. Second, there is often no harm in working with schemes in general. Knowing when Noetherian conditions are needed will help you remember why results are true, because you will have some sense of where Noetherian conditions enter into arguments. Finally, for some people, non-Noetherian rings naturally come up. For example, adeles are not Noetherian. And many valuation rings that naturally arise in arithmetic and tropical geometry are not Noetherian.

3.7 The function $I(\cdot)$, taking subsets of $\operatorname{Spec} A$ to ideals of A

We now introduce a notion that is in some sense "inverse" to the vanishing set function $V(\cdot)$. Given a subset $S \subset \operatorname{Spec} A$, $I(S)$ is the set of functions vanishing on S. In other words, $I(S) = \bigcap_{[\mathfrak{p}] \in S} \mathfrak{p} \subset A$ (at least when S is nonempty).

We make three quick observations. (Do you see why they are true?)

- $I(S)$ is clearly an ideal of A.
- $I(\cdot)$ is inclusion-reversing: if $S_1 \subset S_2$, then $I(S_2) \subset I(S_1)$.
- $I(\overline{S}) = I(S)$.

3.7.A. EXERCISE. Let $A = k[x, y]$. If $S = \{[(x)], [(x-1, y)]\}$ (see Figure 3.10), then $I(S)$ consists of those polynomials vanishing on the y-axis, and at the point $(1, 0)$. Give generators for this ideal.

3.7.B. EXERCISE. Suppose $S \subset \mathbb{A}^3_{\mathbb{C}}$ is the union of the three axes. Give generators for the ideal $I(S)$. Be sure to prove it! We will see in Exercise 12.1.F that this ideal is not generated by less than three elements.

3.7.C. EXERCISE. Show that $V(I(S)) = \overline{S}$. Hence $V(I(S)) = S$ for a closed set S. (Compare this to Exercise 3.7.D.)

Note that $I(S)$ is always a radical ideal — if $f \in \sqrt{I(S)}$, then f^n vanishes on S for some $n > 0$, so then f vanishes on S, so $f \in I(S)$.

3.7.D. EASY EXERCISE. Prove that if $J \subset A$ is an ideal, then $I(V(J)) = \sqrt{J}$. (Huge hint: Exercise 3.4.J.)

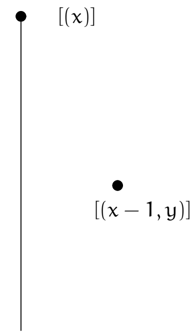


FIGURE 3.10. The set S of Exercise/example 3.7.A, pictured as a subset of $\mathbb{A}^2$

Exercises 3.7.C and 3.7.D show that V and I are "almost" inverse. More precisely:

3.7.1. Theorem. — *$V(\cdot)$ and $I(\cdot)$ give an inclusion-reversing bijection between closed subsets of* $\operatorname{Spec} A$ *and radical ideals of* A *(where a closed subset gives a radical ideal by $I(\cdot)$, and a radical ideal gives a closed subset by $V(\cdot)$).*

Theorem 3.7.1 is sometimes called Hilbert's Nullstellensatz, but we reserve that name for Theorem 3.2.5.

3.7.E. IMPORTANT EXERCISE (CF. EXERCISE 3.7.F). Show that $V(\cdot)$ and $I(\cdot)$ give a bijection between *irreducible closed subsets* of $\operatorname{Spec} A$ and *prime* ideals of A. From this conclude that in $\operatorname{Spec} A$ there is a bijection between points of $\operatorname{Spec} A$ and irreducible closed subsets of $\operatorname{Spec} A$ (where a point determines an irreducible closed subset by taking the closure). Hence *each irreducible closed subset of* $\operatorname{Spec} A$ *has precisely one generic point* — any irreducible closed subset Z can be written uniquely as $\overline{\{z\}}$.

3.7.F. EXERCISE/DEFINITION. A prime of a ring A is a **minimal prime** if it is minimal with respect to inclusion. (For example, the only minimal prime of $k[x, y]$ is (0).) If A is any ring, show that the irreducible components of $\operatorname{Spec} A$ are in bijection with the minimal primes of A. In particular, $\operatorname{Spec} A$ is irreducible if and only if A has only one minimal prime ideal; this generalizes Exercise 3.6.C.

Proposition 3.6.15, Exercise 3.6.S, and Exercise 3.7.F imply that every Noetherian ring has a finite number of minimal primes: an algebraic fact is now revealed to be really a "geometric" fact!

3.7.G. EXERCISE. What are the minimal primes of $k[x, y]/(xy)$ (where k is a field)?

CHAPTER 4

The structure sheaf, and the definition of schemes in general

4.1 The structure sheaf of an affine scheme

The final ingredient in the definition of an affine scheme is the *structure sheaf* $\mathscr{O}_{\operatorname{Spec} A}$, which we think of as the "sheaf of algebraic functions". You should keep in your mind the example of "algebraic functions" on $\mathbb{C}^n$, which you understand well. For example, in $\mathbb{A}^2$, we expect that on the open set $D(xy)$ (away from the two axes), $(3x^4 + y + 4)/x^7y^3$ should be an algebraic function.

These functions will have values at points, but won't be determined by their values at points. But like all sections of sheaves, they will be determined by their germs (see §4.3.5).

It suffices to describe the structure sheaf as a sheaf (of rings) on the base of distinguished open sets (Theorem 2.7.1 and Exercise 3.5.A).

4.1.1. *Definition.* Define $\mathscr{O}_{\operatorname{Spec} A}(D(f))$ to be the localization of A at the multiplicative set of all functions that do not vanish outside of $V(f)$ (i.e., those $g \in A$ such that $V(g) \subset V(f)$, or equivalently $D(f) \subset D(g)$, cf. Exercise 3.5.E). This depends only on $D(f)$, and not on f itself.

4.1.A. Great Exercise. Show that the natural map $A_f \to \mathscr{O}_{\operatorname{Spec} A}(D(f))$ is an isomorphism. (Possible hint: Exercise 3.5.E.)

If $D(f') \subset D(f)$, define the restriction map $\operatorname{res}_{D(f),D(f')} : \mathscr{O}_{\operatorname{Spec} A}(D(f)) \to \mathscr{O}_{\operatorname{Spec} A}(D(f'))$ in the obvious way: the latter ring is a further localization of the former ring. The restriction maps obviously commute: this is a "presheaf on the distinguished base".

4.1.2. Theorem. — *The data just described give a sheaf on the distinguished base, and hence determine a sheaf on the topological space* $\operatorname{Spec} A$.

4.1.3. This sheaf is called the **structure sheaf**, and will be denoted $\mathscr{O}_{\operatorname{Spec} A}$, or sometimes $\mathscr{O}$ if the subscript is clear from the context. Such a topological space, with sheaf, will be called an **affine scheme** (Definition 4.3.1). The notation $\operatorname{Spec} A$ will hereafter denote the data of a topological space with a structure sheaf. An important lesson of Theorem 4.1.2 is not just that $\mathscr{O}_{\operatorname{Spec} A}$ is a sheaf, but also that the distinguished base provides a good way of working with $\mathscr{O}_{\operatorname{Spec} A}$. Notice also that we have justified interpreting elements of A as functions on $\operatorname{Spec} A$.

Proof. We must show the base identity and base gluability axioms hold (§2.7). We show that they both hold for the open set that is the entire space $\operatorname{Spec} A$, and leave to you the trick which extends them to arbitrary distinguished open sets (Exercises 4.1.B and 4.1.C). Suppose $\operatorname{Spec} A = \cup_{i \in I} D(f_i)$, or equivalently (Exercise 3.5.B) the ideal generated by the f_i is the entire ring A.

(Aside: experts familiar with the equalizer exact sequence of §2.2.7 will realize that we are showing exactness of

$$0 \to A \to \prod_{i \in I} A_{f_i} \to \prod_{i \neq j \in I} A_{f_i f_j} \tag{4.1.3.1}$$

where $\{f_i\}_{i \in I}$ is a set of functions with $(f_i)_{i \in I} = A$. Signs are involved in the right-hand map: the map $A_{f_i} \to A_{f_i f_j}$ is the "obvious one" if $i < j$, and negative of the "obvious one" if $i > j$. Base identity corresponds to injectivity at A, and gluability corresponds to exactness at $\prod_i A_{f_i}$.)

We check identity on the base. Suppose that $\operatorname{Spec} A = \cup_{i \in I} D(f_i)$ where i runs over some index set I. Then there is some finite subset of I, which we name $\{1, \dots, n\}$, such that $\operatorname{Spec} A = \cup_{i=1}^n D(f_i)$, i.e., $(f_1, \dots, f_n) = A$ (quasicompactness of $\operatorname{Spec} A$, Exercise 3.5.C). Suppose we are given $s \in A$ such that $\operatorname{res}_{\operatorname{Spec} A, D(f_i)} s = 0$ in A_{f_i} for all i. We wish to show that $s = 0$. The fact that $\operatorname{res}_{\operatorname{Spec} A, D(f_i)} s = 0$ in A_{f_i} implies that there is some m such that for each $i \in \{1, \dots, n\}$, $f_i^m s = 0$. Now $(f_1^m, \dots, f_n^m) = A$ (for example, from $\operatorname{Spec} A = \cup D(f_i) = \cup D(f_i^m)$), so there are $r_i \in A$ with $\sum_{i=1}^n r_i f_i^m = 1$ in A, from which

$$s = \left(\sum r_i f_i^m\right) s = \sum r_i (f_i^m s) = 0.$$

Thus we have checked the "base identity" axiom for $\operatorname{Spec} A$. (Serre has described this as a "partition of unity" argument, and if you look at it in the right way, his insight is very enlightening.)

4.1.B. EXERCISE. Make tiny changes to the above argument to show base identity for any distinguished open $D(f)$. (Hint: judiciously replace A by A_f in the above argument.)

We next show base gluability. Suppose again $\cup_{i \in I} D(f_i) = \operatorname{Spec} A$, where I is a index set (possibly horribly infinite). Suppose we are given elements in each A_{f_i} that agree on the overlaps $A_{f_i f_j}$. Note that intersections of distinguished open sets are also distinguished open sets.

Assume first that I is finite, say $I = \{1, \dots, n\}$. We have elements $a_i / f_i^{l_i} \in A_{f_i}$ agreeing on overlaps $A_{f_i f_j}$ (see Figure 4.1(a)). Letting $g_i = f_i^{l_i}$, using $D(f_i) = D(g_i)$, we can simplify notation by considering our elements as of the form $a_i / g_i \in A_{g_i}$ (Figure 4.1(b)).

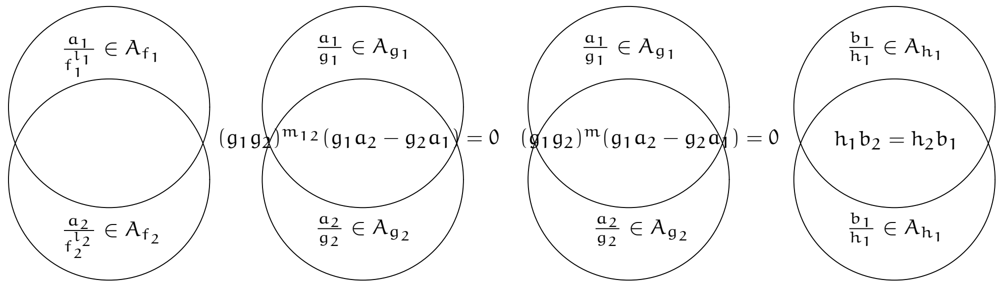
The fact that a_i / g_i and a_j / g_j "agree on the overlap" (i.e., in $A_{g_i g_j}$) means that for some m_{ij},

$$(g_i g_j)^{m_{ij}} (g_j a_i - g_i a_j) = 0$$

in A. By taking $m = \max m_{ij}$ (here we use the finiteness of I), we can simplify notation:

$$(g_i g_j)^m (g_j a_i - g_i a_j) = 0$$

for all i, j (Figure 4.1(c)). Let $b_i = a_i g_i^m$ for all i, and $h_i = g_i^{m+1}$ (so $D(h_i) = D(g_i)$). Then we can simplify notation even more (Figure 4.1(d)): on each $D(h_i)$, we have



(a) (b) (c) (d)

FIGURE 4.1. Base gluability of the structure sheaf

a function b_i/h_i, and the overlap condition is

$$h_jb_i = h_ib_j. \tag{4.1.3.2}$$

Now $\cup_i D(h_i) = \operatorname{Spec} A$, implying that $1 = \sum_{i=1}^n r_ih_i$ for some $r_i \in A$. Define

$$r = \sum r_ib_i. \tag{4.1.3.3}$$

This will be the element of A that restricts to each b_j/h_j. Indeed, from the overlap condition (4.1.3.2),

$$rh_j = \sum_i r_ib_ih_j = \sum_i r_ih_ib_j = b_j.$$

We next deal with the case where I is infinite. Choose a finite subset $\{1, \dots, n\} \subset I$ with $(f_1, \dots, f_n) = A$ (or equivalently, use quasicompactness of $\operatorname{Spec} A$ to choose a finite subcover by $D(f_i)$). Construct r as above, using (4.1.3.3). We will show that for any $\alpha \in I - \{1, \dots, n\}$, r restricts to the desired element a_α of A_{f_α}. Repeat the entire process above with $\{1, \dots, n, \alpha\}$ in place of $\{1, \dots, n\}$, to obtain $r' \in A$ which restricts to α_α for $i \in \{1, \dots, n, \alpha\}$. Then by base identity, $r' = r$. (Note that we use base identity to *prove* base gluability. This is an example of how the identity axiom is "prior" to the gluability axiom.) Hence r restricts to $a_\alpha/f_\alpha^{l_\alpha}$ as desired.

4.1.C. EXERCISE. Alter this argument appropriately to show base gluability for any distinguished open $D(f)$.

We have now completed the proof of Theorem 4.1.2. $\square$

The following generalization of Theorem 4.1.2 will be essential in the definition of a *quasicoherent sheaf* in Chapter 13.

4.1.D. IMPORTANT EXERCISE/DEFINITION. Suppose M is an A-module. Show that the following construction describes a sheaf $\widetilde{M}$ on the distinguished base. Define $\widetilde{M}(D(f))$ to be the localization of M at the multiplicative set of all functions that do not vanish outside of $V(f)$. Define restriction maps $\operatorname{res}_{D(f),D(g)}$ in the analogous way to $\mathscr{O}_{\operatorname{Spec} A}$. Show that this defines a sheaf on the distinguished base, and hence a sheaf on $\operatorname{Spec} A$. Then show that this is an $\mathscr{O}_{\operatorname{Spec} A}$-module.

4.1.4. *Remark.* In the course of answering the previous exercise, you will show that if $(f_i)_{i \in I} = A$,

$$0 \to M \to \prod_{i \in I} M_{f_i} \to \prod_{i \neq j \in I} M_{f_i f_j}$$

(cf. (4.1.3.1)) is exact. In particular, M can be identified with a specific submodule of $M_{f_1} \times \cdots \times M_{f_r}$. Even though $M \to M_{f_i}$ may not be an inclusion for any f_i, $M \to M_{f_1} \times \cdots \times M_{f_r}$ *is* an inclusion. This will be useful later: we will want to show that if M has some nice property, then M_f does too, which will be easy. We will also want to show that if $(f_1, \dots, f_n) = A$, and the M_{f_i} have this property, then M does too. This idea will be made precise in the Affine Communication Lemma 5.3.2.

4.1.5. $\star$ *Remark.* Definition 4.1.1 and Theorem 4.1.2 suggests a potentially slick way of describing sections of $\mathscr{O}_{\operatorname{Spec} A}$ over *any* open subset: perhaps $\mathscr{O}_{\operatorname{Spec} A}(U)$ is the localization of A at the multiplicative set of all functions that do not vanish at any point of U. This is not true. A counterexample (that you will later be able to make precise): let $\operatorname{Spec} A$ be two copies of $\mathbb{A}^2_k$ glued together at their origins (see (24.4.7.1) for explicit equations) and let U be the complement of the origin(s). Then the function which is 1 on the first copy of $\mathbb{A}^2_k \setminus \{(0,0)\}$ and 0 on the second copy of $\mathbb{A}^2_k \setminus \{(0,0)\}$ is not of this form.

4.1.6. (Counter)examples. The example of two planes meeting at a point will appear many times as an example or a counterexample. This is as good a time as any to collect a list of those examples that will repeatedly come up (often in counterexamples). Many of these you will not have the language to understand yet, but you can come back to this list later.

4.2 Visualizing schemes II: nilpotents

The price of metaphor is eternal vigilance.
— A. Rosenbluth and N. Wiener (attribution by [**Leo**, p. 4]*)*

In §3.3, we discussed how to visualize the underlying set of schemes, adding in generic points to our previous intuition of "classical" (or closed) points. Our later discussion of the Zariski topology fit well with that picture. In our definition of the "affine scheme" $(\operatorname{Spec} A, \mathscr{O}_{\operatorname{Spec} A})$, we have the additional information of nilpotents, which are invisible on the level of points (§3.2.11), so now we figure out to picture them. We will then readily be able to glue them together to picture schemes in general, once we have made the appropriate definitions. As we are building intuition, we cannot be rigorous or precise.

As motivation, note that we have incidence-reversing bijections

radical ideals of $A \longleftrightarrow$ closed subsets of $\operatorname{Spec} A$ (Theorem 3.7.1)

prime ideals of $A \longleftrightarrow$ irreducible closed subsets of $\operatorname{Spec} A$ (Exercise 3.7.E)

If we take the things on the right as "pictures", our goal is to figure out how to picture ideals that are not radical:

$$\text{ideals of } A \longleftrightarrow ???$$

(We will later fill this in rigorously in a different way with the notion of a *closed subscheme*, the scheme-theoretic version of closed subsets, §8.1. But our goal now is to create a picture.)

As motivation, when we see the expression, $\operatorname{Spec} \mathbb{C}[x]/(x(x-1)(x-2))$, we immediately interpret it as a closed subset of $\mathbb{A}^1_{\mathbb{C}}$, namely $\{0, 1, 2\}$. In particular, that the map $\mathbb{C}[x] \to \mathbb{C}[x]/(x(x-1)(x-2))$ can be interpreted (via the Chinese remainder theorem) as: take a function on $\mathbb{A}^1$, and restrict it to the three points 0, 1, and 2.

This will guide us in how to visualize a nonradical ideal. The simplest example to consider is $\operatorname{Spec} \mathbb{C}[x]/(x^2)$ (Exercise 3.2.A(a)). As a subset of $\mathbb{A}^1$, it is just the origin $0 = [(x)]$, which we are used to thinking of as $\operatorname{Spec} \mathbb{C}[x]/(x)$ (i.e., corresponding to the ideal (x), not (x^2)). We want to enrich this picture in some way. We should picture $\mathbb{C}[x]/(x^2)$ in terms of the information the quotient remembers. The image of a polynomial $f(x)$ is the information of its value at 0, and its derivative (cf. Exercise 3.2.T). We thus picture this as being the point, plus a little bit more — a little bit of infinitesimal "fuzz" on the point (see Figure 4.2). The sequence of restrictions $\mathbb{C}[x] \to \mathbb{C}[x]/(x^2) \to \mathbb{C}[x]/(x)$ should be interpreted as nested pictures.

$$\begin{tikzcd}
\mathbb{C}[x] \arrow[r, two heads] & \mathbb{C}[x]/(x^2) \arrow[r, two heads] & \mathbb{C}[x]/(x) \\
f(x) \arrow[rr, mapsto] & & f(0),
\end{tikzcd}$$

Similarly, $\mathbb{C}[x]/(x^3)$ remembers even more information — the second derivative as well. Thus we picture this as the point 0 with even more fuzz.

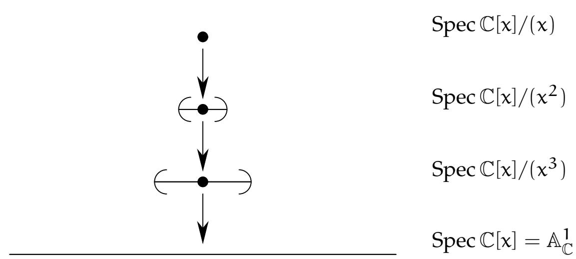


FIGURE 4.2. Picturing quotients of $\mathbb{C}[x]$

More subtleties arise in two dimensions (see Figure 4.3). Consider $\operatorname{Spec}\mathbb{C}[x,y]/(x,y)^2$, which is sandwiched between two rings we know well:

$$\begin{tikzcd} \mathbb{C}[x,y] \arrow[r, two heads] & \mathbb{C}[x,y]/(x,y)^2 \arrow[r, two heads] & \mathbb{C}[x,y]/(x,y) \\ f(x,y) \arrow[rr, mapsto] & & f(0). \end{tikzcd}$$

Again, taking the quotient by $(x,y)^2$ remembers the first derivative, "in all directions". We picture this as fuzz around the point, in the shape of a circle (no direction is privileged). Similarly, $(x,y)^3$ remembers the second derivative "in all directions" — bigger circular fuzz.

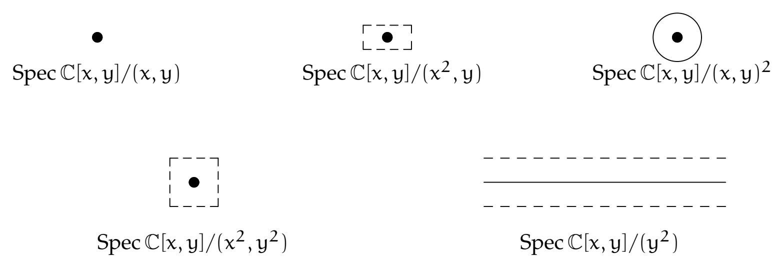


FIGURE 4.3. Picturing quotients of $\mathbb{C}[x,y]$

Consider instead the ideal (x^2,y). What it remembers is the derivative only in the x direction — given a polynomial, we remember its value at 0, and the coefficient of x. We remember this by picturing the fuzz only in the x direction.

This gives us some handle on picturing more things of this sort, but now it becomes more an art than a science. For example, $\operatorname{Spec}\mathbb{C}[x,y]/(x^2,y^2)$ we might picture as a fuzzy square around the origin. (Could you believe that this square is circumscribed by the circular fuzz $\operatorname{Spec}\mathbb{C}[x,y]/(x,y)^3$, and inscribed by the circular fuzz $\operatorname{Spec}\mathbb{C}[x,y]/(x,y)^2$?) One feature of this example is that given two ideals I and J of a ring A (such as $\mathbb{C}[x,y]$), your fuzzy picture of $\operatorname{Spec} A/(I,J)$ should be the "intersection" of your picture of $\operatorname{Spec} A/I$ and $\operatorname{Spec} A/J$ in $\operatorname{Spec} A$. (You will make this precise in Exercise 8.1.J(a).) For example, $\operatorname{Spec}\mathbb{C}[x,y]/(x^2,y^2)$ should be the intersection of two thickened lines. (How would you picture $\operatorname{Spec}\mathbb{C}[x,y]/(x^5,y^3)$? $\operatorname{Spec}\mathbb{C}[x,y,z]/(x^3,y^4,z^5,(x+y+z)^2)$? $\operatorname{Spec}\mathbb{C}[x,y]/((x,y)^5,y^3)$?)

One final example that will motivate us in §5.5 is $\operatorname{Spec}\mathbb{C}[x,y]/(y^2,xy)$. Knowing what a polynomial in $\mathbb{C}[x,y]$ is modulo (y^2,xy) is the same as knowing its value on the x-axis, as well as first-order differential information around the origin. This is worth thinking through carefully: do you see how this information is captured (however imperfectly) in Figure 4.4?

Our pictures capture useful information that you already have some intuition for. For example, consider the intersection of the parabola $y=x^2$ and the x-axis (in the xy-plane), see Figure 4.5. You already have a sense that the intersection has multiplicity two. In terms of this visualization, we interpret this as intersecting (in

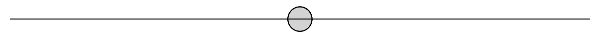


FIGURE 4.4. A picture of the scheme $\operatorname{Spec} k[x,y]/(y^2, xy)$. The fuzz at the origin indicates where "the nonreducedness lives".

$\operatorname{Spec} \mathbb{C}[x,y]$):

$$\begin{aligned} \operatorname{Spec} \mathbb{C}[x,y]/(y-x^2) \cap \operatorname{Spec} \mathbb{C}[x,y]/(y) &= \operatorname{Spec} \mathbb{C}[x,y]/(y-x^2, y) \\ &= \operatorname{Spec} \mathbb{C}[x,y]/(y, x^2) \end{aligned}$$

which we interpret as the fact that the parabola and line not just meet with multiplicity two, but that the "multiplicity 2" part is in the direction of the x-axis. You will make this example precise in Exercise 8.1.J(b).

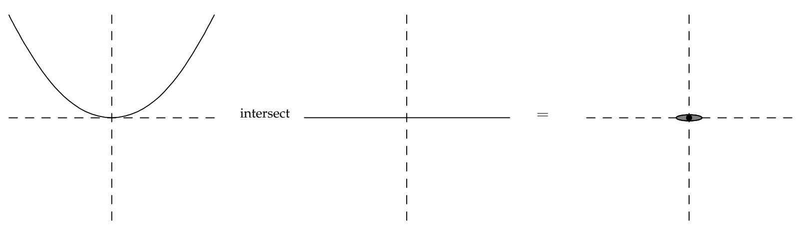


FIGURE 4.5. The "scheme-theoretic" intersection of the parabola $y = x^2$ and the x-axis is a nonreduced scheme (with fuzz in the x-direction)

4.2.1. We will later make the location of the fuzz somewhat more precise when we discuss associated points (§5.5). We will see that in reasonable circumstances, the fuzz is concentrated on closed subsets (Remark 13.7.2).

On a bien souvent répété que la Géométrie est l'art de bien raisonner sur des figures mal faites; encore ces figures, pour ne pas nous tromper, doivent-elles satisfaire à certaines conditions; les proportions peuvent être grossièrement altérées, mais les positions relatives des diverses parties ne doivent pas être bouleversées.

It is often said that geometry is the art of reasoning well from badly made figures; however, these figures, if they are not to deceive us, must satisfy certain conditions; the proportions may be grossly altered, but the relative positions of the different parts must not be upset.

— *H. Poincaré,* [**Po1**, p. 2] *(see Stillwell's translation* [**Po2**, p. ix]*)*

4.3 Definition of schemes

4.3.1. *Definitions.* We can now define *scheme* in general. First, define an **isomorphism of ringed spaces** $(X, \mathscr{O}_X)$ and $(Y, \mathscr{O}_Y)$ as (i) a homeomorphism $\pi : X \to Y$, and (ii) an isomorphism of sheaves $\mathscr{O}_X$ and $\mathscr{O}_Y$, considered to be on the same space via π. (Part (ii), more precisely, is an isomorphism $\mathscr{O}_Y \to \pi_* \mathscr{O}_X$ of sheaves on Y, or equivalently by adjointness, $\pi^{-1}\mathscr{O}_Y \to \mathscr{O}_X$ of sheaves on X.) In other words, we have a "correspondence" of sets, topologies, and structure sheaves. An **affine scheme** is a ringed space that is isomorphic to $(\operatorname{Spec} A, \mathscr{O}_{\operatorname{Spec} A})$ for some A. A **scheme** $(X, \mathscr{O}_X)$ is a ringed space such that any point of X has a neighborhood U such that $(U, \mathscr{O}_X|_U)$ is an affine scheme. The topology on a scheme is called the **Zariski topology**. The scheme can be denoted $(X, \mathscr{O}_X)$, although it is often denoted X, with the structure sheaf implicit.

An **isomorphism of two schemes** $(X, \mathscr{O}_X)$ and $(Y, \mathscr{O}_Y)$ is an isomorphism as ringed spaces. Recall the definition of $\Gamma(\cdot, \cdot)$ in §2.2.2. If $U \subset X$ is an open subset, then the elements of $\Gamma(U, \mathscr{O}_X)$ are said to be the **functions on** U; this generalizes in an obvious way the definition of functions on an affine scheme, §3.2.1.

4.3.2. *Remark.* From the definition of the structure sheaf on an affine scheme, several things are clear. First of all, if we are told that $(X, \mathscr{O}_X)$ is an affine scheme, we may recover its ring (i.e., find the ring A such that $\operatorname{Spec} A = X$) by taking the ring of global sections, as $X = D(1)$, so:

$$\begin{aligned} \Gamma(X, \mathscr{O}_X) &= \Gamma(D(1), \mathscr{O}_{\operatorname{Spec} A}) \quad \text{as } D(1) = \operatorname{Spec} A \\ &= A. \end{aligned}$$

(You can verify that we get more, and can "recognize X as the scheme $\operatorname{Spec} A$": we get an isomorphism $\pi : (\operatorname{Spec} \Gamma(X, \mathscr{O}_X), \mathscr{O}_{\operatorname{Spec} \Gamma(X, \mathscr{O}_X)}) \to (X, \mathscr{O}_X)$. For example, if $\mathfrak{m}$ is a maximal ideal of $\Gamma(X, \mathscr{O}_X)$, $\{\pi([\mathfrak{m}])\} = V(\mathfrak{m})$.) The following exercise will give you a chance to make these ideas rigorous — they are subtler than they appear.

4.3.A. Enlightening Exercise (which can be strangely confusing). Describe a bijection between the isomorphisms $\operatorname{Spec} A \to \operatorname{Spec} A'$ and the ring isomorphisms $A' \to A$. Hint: the hardest part is to show that if an isomorphism $\pi : \operatorname{Spec} A \to \operatorname{Spec} A'$ induces an isomorphism $\pi^\sharp : A' \to A$, which in turn induces an isomorphism $\rho : \operatorname{Spec} A \to \operatorname{Spec} A'$, then $\pi = \rho$. First show this on the level of points; this is (surprisingly) the trickiest part. Then show $\pi = \rho$ as maps of topological spaces. Finally, to show $\pi = \rho$ on the level of structure sheaves, use the distinguished base. Feel free to use insights from later in this section, but be careful to avoid circular arguments. Even struggling with this exercise and failing (until reading later sections) will be helpful.

More generally, given $f \in A$, $\Gamma(D(f), \mathscr{O}_{\operatorname{Spec} A}) \cong A_f$. Thus under the natural inclusion of sets $\operatorname{Spec} A_f \hookrightarrow \operatorname{Spec} A$, the Zariski topology on $\operatorname{Spec} A$ restricts to give the Zariski topology on $\operatorname{Spec} A_f$ (Exercise 3.4.I), and the structure sheaf of $\operatorname{Spec} A$ restricts to the structure sheaf of $\operatorname{Spec} A_f$, as the next exercise shows.

4.3.B. Important but easy exercise. Suppose $f \in A$. Show that under the identification of $D(f)$ in $\operatorname{Spec} A$ with $\operatorname{Spec} A_f$ (§3.5), there is a natural isomorphism of ringed spaces $(D(f), \mathscr{O}_{\operatorname{Spec} A}|_{D(f)}) \cong (\operatorname{Spec} A_f, \mathscr{O}_{\operatorname{Spec} A_f})$. Hint: notice that distinguished open sets of $\operatorname{Spec} A_f$ are already distinguished open sets in $\operatorname{Spec} A$.

4.3.C. EASY EXERCISE. If X is a scheme, and U is *any* open subset, prove that $(U, \mathscr{O}_X|_U)$ is also a scheme.

4.3.3. *Definitions.* We say $(U, \mathscr{O}_X|_U)$ is an **open subscheme of** X. If U is also an affine scheme, we often say U is an **affine open subset**, or an **affine open subscheme**, or sometimes informally just an **affine open**. For example, $D(f)$ is an affine open subscheme of $\operatorname{Spec} A$. (Unimportant remark: it is not true that every affine open subscheme of $\operatorname{Spec} A$ is of the form $D(f)$; see §19.11.10 for an example.)

4.3.D. EASY EXERCISE. Show that if X is a scheme, then the affine open sets form a base for the Zariski topology.

4.3.E. EASY EXERCISE. The **disjoint union of schemes** is defined as you would expect: it is the disjoint union of sets, with the expected topology (thus it is the disjoint union of topological spaces), with the expected sheaf. Once we know what morphisms are, it will be immediate (Exercise 9.1.A) that (just as for sets and topological spaces) disjoint union is the coproduct in the category of schemes.
(a) Show that the disjoint union of a *finite* number of affine schemes is also an affine scheme. (Hint: Exercise 3.6.A.)
(b) *(a first example of a non-affine scheme)* Show that an infinite disjoint union of (nonempty) affine schemes is not an affine scheme. (Hint: affine schemes are quasicompact, Exercise 3.6.G(a). This is basically answered in Remark 3.6.6.)

4.3.4. *Remark: A first glimpse of closed subschemes.* Open subsets of a scheme come with a natural scheme structure (Definition 4.3.3). For comparison, closed subsets can have many "natural" scheme structures. We will discuss this later (in §8.1), but for now, it suffices for you to know that a closed subscheme of X is, informally, a particular kind of scheme structure on a closed subset of X. As an example: if $I \subset A$ is an ideal, then $\operatorname{Spec} A/I$ endows the closed subset $V(I) \subset \operatorname{Spec} A$ with a scheme structure; but note that there can be different ideals with the same vanishing set (for example (x) and (x^2) in $k[x]$).

4.3.5. Stalks of the structure sheaf: germs, values at a point, and the residue field of a point. Like every sheaf, the structure sheaf has stalks, and we shouldn't be surprised if they are interesting from an algebraic point of view. In fact, we have seen them before.

4.3.F. IMPORTANT EASY EXERCISE. Show that the stalk of $\mathscr{O}_{\operatorname{Spec} A}$ at the point $[\mathfrak{p}]$ is the local ring $A_{\mathfrak{p}}$.

Essentially the same argument will show that the stalk of the sheaf $\widetilde{M}$ (defined in Exercise 4.1.D) at $[\mathfrak{p}]$ is $M_{\mathfrak{p}}$. Here is an interesting consequence, or if you prefer, a geometric interpretation of an algebraic fact. A section is determined by its germs (Exercise 2.4.A), meaning that $M \to \prod_{\mathfrak{p}} M_{\mathfrak{p}}$ is an inclusion. So for example an A-module is zero if and only if all its localizations at primes are zero.

4.3.6. *Definition.* We say a ringed space is a **locally ringed space** if its stalks are local rings. Thus Exercise 4.3.F shows that schemes are locally ringed spaces. Manifolds are another example of locally ringed spaces, see §2.1.1. In both cases, taking quotient by the maximal ideal may be interpreted as evaluating at the point.

The maximal ideal of the local ring $\mathscr{O}_{X,p}$ is denoted $\mathfrak{m}_{X,p}$ or $\mathfrak{m}_p$, and the **residue field** $\mathscr{O}_{X,p}/\mathfrak{m}_p$ is denoted $\kappa(p)$. Functions on an open subset U of a locally ringed space have **values** at each point of U. The value at p of such a function lies in $\kappa(p)$. As usual, we say that a function **vanishes** at a point p if its value at p is 0. (This generalizes our notion of the value of a function on $\operatorname{Spec} A$, defined in §3.2.1.)

4.3.G. USEFUL EXERCISE.
(a) If f is a function on a locally ringed space X, show that the subset of X where f doesn't vanish is open. (Hint: show that if f is a function on a *ringed space* X, show that the subset of X where the germ of f is invertible is open.)
(b) Show that if f is a function on a locally ringed space that vanishes nowhere, then f is invertible.

Consider a point $[\mathfrak{p}]$ of an affine scheme $\operatorname{Spec} A$. (Of course, any point of a scheme can be interpreted in this way, as each point has an affine neighborhood.) The residue field at $[\mathfrak{p}]$ is $A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p}$, which is isomorphic to $K(A/\mathfrak{p})$, the fraction field of the quotient. It is useful to note that localization at $\mathfrak{p}$ and taking quotient by $\mathfrak{p}$ "commute", i.e., the following diagram commutes.

(4.3.6.1)
$$\begin{tikzcd} & A_\mathfrak{p} \arrow[dr, "\text{quotient}"] & \\ A \arrow[ur, "\text{localize}"] \arrow[dr, "\text{quotient}"'] & & A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p} = K(A/\mathfrak{p}) \\ & A/\mathfrak{p} \arrow[ur, "\text{localize, i.e., } K(\cdot)"'] & \end{tikzcd}$$

For example, consider the scheme $\mathbb{A}^2_k = \operatorname{Spec} k[x,y]$, where k is a field of characteristic not 2. Then $(x^2+y^2)/x(y^2-x^5)$ is a function away from the y-axis and the curve y^2-x^5. Its value at $(2,4)$ (by which we mean $[(x-2,y-4)]$) is $(2^2+4^2)/(2(4^2-2^5))$, as
$$\frac{x^2+y^2}{x(y^2-x^5)} \equiv \frac{2^2+4^2}{2(4^2-2^5)}$$
in the residue field — check this if it seems mysterious. And its value at $[(y)]$, the generic point of the x-axis, is $\frac{x^2}{-x^6} = -1/x^4$, which we see by setting y to 0. This is indeed an element of the fraction field of $k[x,y]/(y)$, i.e., $k(x)$. (If you think you care only about algebraically closed fields, let this example be a first warning: $A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p}$ won't be algebraically closed in general, even if A is a finitely generated $\mathbb{C}$-algebra!)

If anything makes you nervous, you should make up an example to make you feel better. Here is one: $27/4$ is a function on $\operatorname{Spec}\mathbb{Z} - \{[(2)],[(7)]\}$ or indeed on an even bigger open set. What is its value at $[(5)]$? Answer: $2/(-1) \equiv -2 \pmod 5$. What is its value at the generic point $[(0)]$? Answer: $27/4$. Where does it vanish? At $[(3)]$.

4.3.7. *Stray definition: the fiber of an $\mathscr{O}_X$-module at a point.* If $\mathscr{F}$ is an $\mathscr{O}_X$-module on a scheme X (or more generally, a locally ringed space), define the **fiber** (or *fibre*)**of $\mathscr{F}$ at a point** $p \in X$ by
$$\mathscr{F}|_p := \mathscr{F}_p \otimes_{\mathscr{O}_{X,p}} \kappa(p).$$

For example, $\mathscr{O}_X|_p$ is $\kappa(p)$. (This notion will start to come into play in §13.7.)

4.4 Three examples

We now give three extended examples. Our short-term goal is to see that we can really work with the structure sheaf, and can compute the ring of sections of interesting open sets that aren't just distinguished open sets of affine schemes. Our long-term goal is to meet interesting examples that will come up repeatedly in the future.

4.4.1. First example: The plane minus the origin. This example will show you that the distinguished base is something that you can work with. Let $A = k[x, y]$, so $\operatorname{Spec} A = \mathbb{A}^2_k$. Let's work out the space of functions on the open set $U = \mathbb{A}^2 - \{(0, 0)\} = \mathbb{A}^2 - \{[(x, y)]\}$.

It is not immediately obvious whether this is a distinguished open set. (In fact it is not — you may be able to figure out why within a few paragraphs, if you can't right now. It is not enough to show that (x, y) is not a principal ideal.) But in any case, we can describe it as the union of two things which *are* distinguished open sets: $U = D(x) \cup D(y)$. We will find the functions on U by gluing together functions on $D(x)$ and $D(y)$.

The functions on $D(x)$ are, by Definition 4.1.1, $A_x = k[x, y, 1/x]$. The functions on $D(y)$ are $A_y = k[x, y, 1/y]$. Note that A injects into its localizations (if 0 is not inverted), as it is an integral domain (Exercise 1.3.C), so A injects into both A_x and A_y, and both inject into A_{xy} (and indeed $k(x, y) = K(A)$). So we are looking for functions on $D(x)$ and $D(y)$ that agree on $D(x) \cap D(y) = D(xy)$, i.e., we are interpreting $A_x \cap A_y$ in A_{xy} (or in $k(x, y)$). Clearly those rational functions with only powers of x in the denominator, and also with only powers of y in the denominator, are the polynomials. Translation: $A_x \cap A_y = A$. Thus we conclude:

$$\Gamma(U, \mathscr{O}_{\mathbb{A}^2}) \equiv k[x, y]. \tag{4.4.1.1}$$

In other words, we get no extra functions by removing the origin. Notice how easy that was to calculate!

4.4.2. *Aside.* Notice that any function on $\mathbb{A}^2 - \{(0, 0)\}$ extends over all of $\mathbb{A}^2$. This is an analog of *Hartogs's Lemma* in complex geometry: you can extend a holomorphic function defined on the complement of a set of codimension at least two on a complex manifold over the missing set. This will work more generally in the algebraic setting: you can extend over points in codimension at least 2 not only if they are "smooth", but also if they are mildly singular — what we will call *normal*. We will make this precise in §11.3.10. This fact will be very useful for us.

4.4.3. We now show an interesting fact: $(U, \mathscr{O}_{\mathbb{A}^2}|_U)$ is a scheme, but it is not an affine scheme. (This is confusing, so you will have to pay attention.) Here's why: otherwise, if $(U, \mathscr{O}_{\mathbb{A}^2}|_U) = (\operatorname{Spec} A, \mathscr{O}_{\operatorname{Spec} A})$, then we can recover A by taking global sections:

$$A = \Gamma(U, \mathscr{O}_{\mathbb{A}^2}|_U),$$

which we have already identified in (4.4.1.1) as $k[x, y]$. So if U is affine, then $U \cong \mathbb{A}^2_k$. But this bijection between primes in a ring and points of the spectrum is more constructive than that: *given the prime ideal* I, *you can recover the point as the generic point of the closed subset cut out by* I, *i.e.,* $V(I)$, *and given the point* $\mathfrak{p}$, *you can recover the ideal as those functions vanishing at* $\mathfrak{p}$, *i.e.,* $I(\mathfrak{p})$. In particular, the prime ideal (x, y) of A should cut out a point of $\operatorname{Spec} A$. But on U, $V(x) \cap V(y) = \varnothing$. Conclusion: U is *not* an affine scheme. (If you are ever looking for a counterexample to something, and you are expecting one involving a non-affine scheme, keep this example in mind!)

4.4.4. Gluing two copies of $\mathbb{A}^1$ together in two different ways. We have now seen two examples of non-affine schemes: an infinite disjoint union of nonempty schemes (Exercise 4.3.E) and $\mathbb{A}^2 - \{(0,0)\}$. I want to give you two more examples. They are important because they are the first examples of fundamental behavior, the first pathological, and the second central.

First, I need to tell you how to glue two schemes together. Before that, you should review how to glue topological spaces together along isomorphic open sets. Given two topological spaces X and Y, and open subsets $U \subset X$ and $V \subset Y$ along with a homeomorphism $U \cong V$, we can create a new topological space W, that we think of as gluing X and Y together along $U \cong V$. It is the quotient of the disjoint union $X \coprod Y$ by the equivalence relation $U \cong V$, where the quotient is given the quotient topology. Then X and Y are naturally (identified with) open subsets of W, and indeed cover W. Can you restate this cleanly with an arbitrary (not necessarily finite) number of topological spaces?

Now that we have discussed gluing topological spaces, let's glue schemes together. (This applies without change more generally to ringed spaces.) Suppose you have two schemes $(X, \mathscr{O}_X)$ and $(Y, \mathscr{O}_Y)$, and open subsets $U \subset X$ and $V \subset Y$, along with a homeomorphism $f : U \xrightarrow{\sim} V$, and an isomorphism of structure sheaves $\mathscr{O}_V \xrightarrow{\sim} f_*\mathscr{O}_U$ (i.e., an isomorphism *of schemes* $(U, \mathscr{O}_X|_U) \cong (V, \mathscr{O}_Y|_V)$). Then we can glue these together to get a single scheme. Reason: let W be X and Y glued together using the isomorphism $U \cong V$. Then Exercise 2.7.D shows that the structure sheaves can be glued together to get a sheaf of rings. Note that this is indeed a scheme: any point has a neighborhood that is an affine scheme. (Do you see why?)

4.4.A. Essential Exercise (cf. Exercise 2.7.D). Show that you can glue an arbitrary collection of schemes together. Suppose we are given:

- schemes X_i (as i runs over some index set I, not necessarily finite),
- open subschemes $X_{ij} \subset X_i$ with $X_{ii} = X_i$,
- isomorphisms $f_{ij} : X_{ij} \to X_{ji}$ with f_{ii} the identity

such that

- (the cocycle condition) the isomorphisms "agree on triple intersections", i.e., $f_{ik}|_{X_{ij} \cap X_{ik}} = f_{jk}|_{X_{ji} \cap X_{jk}} \circ f_{ij}|_{X_{ij} \cap X_{ik}}$ (so implicitly, to make sense of the right side, $f_{ij}(X_{ik} \cap X_{ij}) \subset X_{jk}$).

(The cocycle condition ensures that f_{ij} and f_{ji} are inverses. In fact, the hypothesis that f_{ii} is the identity also follows from the cocycle condition.) Show that there is a unique scheme X (up to unique isomorphism) along with open subsets isomorphic to the X_i respecting this gluing data in the obvious sense. (Hint: what is X as a set? What is the topology on this set? In terms of your description of the open sets of X, what are the sections of this sheaf over each open set?)

I will now give you two non-affine schemes. Both are handy to know. In both cases, I will glue together two copies of the affine line $\mathbb{A}^1_k$. Let $X = \operatorname{Spec} k[t]$, and $Y = \operatorname{Spec} k[u]$. Let $U = D(t) = \operatorname{Spec} k[t, 1/t] \subset X$ and $V = D(u) = \operatorname{Spec} k[u, 1/u] \subset Y$. We will get both examples by gluing X and Y together along U and V. The difference will be in how we glue.

4.4.5. Second example: the affine line with the doubled origin. Consider the isomorphism $U \cong V$ via the isomorphism $k[t, 1/t] \cong k[u, 1/u]$ given by $t \leftrightarrow u$ (cf. Exercise 4.3.A). The resulting scheme is called the **affine line with doubled origin**. Figure 4.6 is a picture of it.

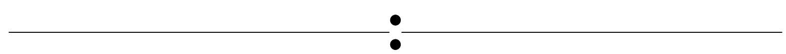


FIGURE 4.6. The affine line with doubled origin

As the picture suggests, intuitively this is an analog of a failure of Hausdorffness. Now $\mathbb{A}^1$ itself is not Hausdorff, so we can't say that it is a failure of Hausdorffness. We see this as weird and bad, so we will want to make a definition that will prevent this from happening. This will be the notion of *separatedness* (to be discussed in Chapter 10). This will answer other of our prayers as well. For example, on a separated scheme, the "affine base of the Zariski topology" is nice — the intersection of two affine open sets will be affine (Proposition 10.1.8).

4.4.B. EXERCISE. Show that the affine line with doubled origin is not affine. Hint: calculate the ring of global sections, and look back at the argument for $\mathbb{A}^2 - \{(0,0)\}$.

4.4.C. EASY EXERCISE. Do the same construction with $\mathbb{A}^1$ replaced by $\mathbb{A}^2$. You will have defined the **affine plane with doubled origin**. Describe two affine open subsets of this scheme whose intersection is not an affine open subset. (An "infinite-dimensional" version comes up in Exercise 5.1.J.)

4.4.6. Third example: the projective line. Consider the isomorphism $U \cong V$ via the isomorphism $k[t, 1/t] \cong k[u, 1/u]$ given by $t \leftrightarrow 1/u$. Figure 4.7 is a suggestive picture of this gluing. The resulting scheme is called the **projective line over the field** k, and is denoted $\mathbb{P}^1_k$.

Notice how the points glue. Let me assume that k is algebraically closed for convenience. (You can think about how this changes otherwise.) On the first affine line, we have the closed ("traditional") points $[(t-a)]$, which we think of as "a on the t-line", and we have the generic point $[(0)]$. On the second affine line, we have closed points that are "b on the u-line", and the generic point. Then a on the t-line is glued to $1/a$ on the u-line (if $a \neq 0$ of course), and the generic point is glued to the generic point (the ideal (0) of $k[t]$ becomes the ideal (0) of $k[t, 1/t]$ upon localization, and the ideal (0) of $k[u]$ becomes the ideal (0) of $k[u, 1/u]$. And (0) in $k[t, 1/t]$ is (0) in $k[u, 1/u]$ under the isomorphism $t \leftrightarrow 1/u$).

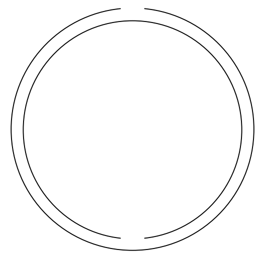


FIGURE 4.7. Gluing two affine lines together to get $\mathbb{P}^1$

4.4.7. If k is algebraically closed, we can interpret the closed points of $\mathbb{P}^1_k$ in the following way, which may make this sound closer to the way you have seen projective space defined earlier. The points are of the form $[a, b]$, where a and b are not both zero, and $[a, b]$ is identified with $[ac, bc]$ where $c \in k^\times$. Then if $b \neq 0$, this is identified with a/b on the t-line, and if $a \neq 0$, this is identified with b/a on the u-line.

4.4.8. Proposition. — $\mathbb{P}^1_k$ *is not affine.*

Proof. We do this by calculating the ring of global sections. The global sections correspond to sections over X and sections over Y that agree on the overlap. A section on X is a polynomial $f(t)$. A section on Y is a polynomial $g(u)$. If we restrict $f(t)$ to the overlap, we get something we can still call $f(t)$; and similarly for $g(u)$. Now we want them to be equal: $f(t) = g(1/t)$. But the only polynomials in t that are at the same time polynomials in $1/t$ are the constants k. Thus $\Gamma(\mathbb{P}^1, \mathscr{O}_{\mathbb{P}^1}) = k$. If $\mathbb{P}^1$ were affine, then it would be $\operatorname{Spec}\Gamma(\mathbb{P}^1, \mathscr{O}_{\mathbb{P}^1}) = \operatorname{Spec} k$, i.e., one point. But it isn't — it has lots of points. $\square$

We have proved an analog of an important theorem: the only holomorphic functions on $\mathbb{CP}^1$ are the constants! (See §10.3.7 for a serious yet easy generalization.)

4.4.9. Important example: Projective space. We now make a preliminary definition of **projective n-space over a field** k, denoted $\mathbb{P}^n_k$, by gluing together $n+1$ open sets each isomorphic to $\mathbb{A}^n_k$. Judicious choice of notation for these open sets will make our life easier. Our motivation is as follows. In the construction of $\mathbb{P}^1$ above, we thought of points of projective space as $[x_0, x_1]$, where (x_0, x_1) are only determined up to scalars, i.e., (x_0, x_1) is considered the same as $(\lambda x_0, \lambda x_1)$. Then the first patch can be interpreted by taking the locus where $x_0 \neq 0$, and then we consider the points $[1, t]$, and we think of t as x_1/x_0; even though x_0 and x_1 are not well-defined, x_1/x_0 *is*. The second corresponds to where $x_1 \neq 0$, and we consider the points $[u, 1]$, and we think of u as x_0/x_1. It will be useful to instead use the notation $x_{1/0}$ for t and $x_{0/1}$ for u.

For $\mathbb{P}^n$, we glue together $n+1$ open sets, one for each of $i = 0, \dots, n$. The ith open set U_i will have coordinates $x_{0/i}, \dots, x_{(i-1)/i}, x_{(i+1)/i}, \dots, x_{n/i}$. It will be convenient to write this as

$$\operatorname{Spec} k[x_{0/i}, x_{1/i}, \dots, x_{n/i}]/(x_{i/i} - 1) \tag{4.4.9.1}$$

(so we have introduced a "dummy variable" $x_{i/i}$ which we immediately set to 1). We glue the distinguished open set $D(x_{j/i})$ of U_i to the distinguished open set $D(x_{i/j})$ of U_j, by identifying these two schemes by describing the identification of rings

$$\operatorname{Spec} k[x_{0/i}, x_{1/i}, \dots, x_{n/i}, 1/x_{j/i}]/(x_{i/i} - 1) \cong$$
$$\operatorname{Spec} k[x_{0/j}, x_{1/j}, \dots, x_{n/j}, 1/x_{i/j}]/(x_{j/j} - 1)$$

via $x_{k/i} = x_{k/j}/x_{i/j}$ and $x_{k/j} = x_{k/i}/x_{j/i}$ (which implies $x_{i/j}x_{j/i} = 1$). We need to check that this gluing information agrees over triple overlaps.

4.4.D. EXERCISE. Check this, as painlessly as possible. (Possible hint: the triple intersection is affine; describe the corresponding ring.)

4.4.10. *Definition.* Note that our definition does not use the fact that k is a field. Hence we may as well define $\mathbb{P}^n_A$ for any *ring* A. This will be useful later.

4.4.E. EXERCISE. Show that the only functions on $\mathbb{P}^n_k$ are constants ($\Gamma(\mathbb{P}^n_k, \mathscr{O}) \cong k$), and hence that $\mathbb{P}^n_k$ is not affine if $n > 0$. Hint: you might fear that you will need some delicate interplay among all of your affine open sets, but you will only need two of your open sets to see this. There is even some geometric intuition behind this: the complement of the union of two open sets has codimension 2. But "Algebraic Hartogs's Lemma" (discussed informally in §4.4.2, and to be stated rigorously in Theorem 11.3.10) says that any function defined on this union extends to be a function on all of projective space. Because we are expecting to see only constants as functions on all of projective space, we should already see this for this union of our two affine open sets.

4.4.F. EXERCISE (GENERALIZING §4.4.7). Show that if k is algebraically closed, the closed points of $\mathbb{P}^n_k$ may be interpreted in the traditional way: the points are of the form $[a_0, \dots, a_n]$, where the a_i are not all zero, and $[a_0, \dots, a_n]$ is identified with $[\lambda a_0, \dots, \lambda a_n]$ where $\lambda \in k^\times$.

We will later give other definitions of projective space (Definition 4.5.8, §16.4.2). Our first definition here will often be handy for computing things. But there is something unnatural about it — projective space is highly symmetric, and that isn't clear from our current definition. Furthermore, as noted by Herman Weyl [**Wey**, p. 90], "The introduction of numbers as coordinates is an act of violence."

4.4.11. Fun aside: The Chinese Remainder Theorem is a *geometric* fact. The Chinese Remainder Theorem is embedded in what we have done. We will see this in a single example, but you should then figure out the general statement. The Chinese Remainder Theorem says that knowing an integer modulo 60 is the same as knowing an integer modulo 3, 4, and 5. Here is how to see this in the language of schemes. What is $\operatorname{Spec}\mathbb{Z}/(60)$? What are the primes of this ring? Answer: those prime ideals containing (60), i.e., those primes dividing 60, i.e., (2), (3), and (5). Figure 4.8 is a sketch of $\operatorname{Spec}\mathbb{Z}/(60)$. They are all closed points, as these are all maximal ideals, so the topology is the discrete topology. What are the stalks? You can check that they are $\mathbb{Z}/4$, $\mathbb{Z}/3$, and $\mathbb{Z}/5$. The nilpotents "at (2)" are indicated by the "fuzz" on that point. (We discussed visualizing nilpotents with "infinitesimal fuzz" in §4.2.) So what are global sections on this scheme? They are sections on this open set (2), this other open set (3), and this third open set (5). In other words, we have a natural isomorphism of rings

$$\mathbb{Z}/60 \to \mathbb{Z}/4 \times \mathbb{Z}/3 \times \mathbb{Z}/5.$$

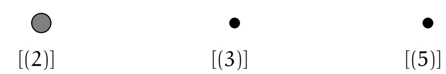


FIGURE 4.8. A picture of the scheme $\operatorname{Spec} \mathbb{Z}/(60)$

4.4.12. ⋆ *Example.* Here is an example of a function on an open subset of a scheme with some surprising behavior. On $X = \operatorname{Spec} k[w, x, y, z]/(wz - xy)$, consider the open subset $D(y) \cup D(w)$. Clearly the function z/y on $D(y)$ agrees with x/w on $D(w)$ on the overlap $D(y) \cap D(w)$. Hence they glue together to give a section. You may have seen this before when thinking about analytic continuation in complex geometry — we have a "holomorphic" function which has the description z/y on an open set, and this description breaks down elsewhere, but you can still "analytically continue" it by giving the function a different definition on different parts of the space.

Follow-up for curious experts. This function has no "single description" as a well-defined expression in terms of w, x, y, z! There is a lot of interesting geometry here, and this scheme will be a constant source of (counter)examples for us (look in the index under "cone over smooth quadric surface"). Here is a glimpse, in terms of words we have not yet defined. The space $\operatorname{Spec} k[w, x, y, z]$ is $\mathbb{A}^4$, and is, not surprisingly, 4-dimensional. We are working with the subset X, which is a hypersurface, and is 3-dimensional. It is a cone over a smooth quadric surface in $\mathbb{P}^3$ (flip to Figure 8.2). The open subset $D(y) \subset X$ is X minus some hypersurface, so we are throwing away a codimension 1 locus. $D(w)$ involves throwing away another codimension 1 locus. You might think that the intersection of these two discarded loci is then codimension 2, and that failure of extending this weird function to a global polynomial comes because of a failure of our Hartogs's Lemma-type theorem. But that's not true — $V(y) \cap V(w)$ is in fact codimension 1. Here is what is actually going on. $V(y)$ involves throwing away the (cone over the) union of two lines ℓ and m_1, one in each "ruling" of the surface, and $V(w)$ also involves throwing away the (cone over the) union of two lines ℓ and m_2. The intersection is the (cone over the) line ℓ, which is a codimension 1 set. Remarkably, despite being "pure codimension 1" the cone over ℓ is not cut out even set-theoretically by a single equation (see Exercise 14.2.R). This means that any expression $f(w, x, y, z)/g(w, x, y, z)$ for our function cannot correctly describe our function on $D(y) \cup D(w)$ — at some point of $D(y) \cup D(w)$ it must be $0/0$. Here's why. Our function can't be defined on $V(y) \cap V(w)$, so g must vanish here. But g can't vanish just on the cone over ℓ — it must vanish elsewhere too.

4.5 Projective schemes, and the Proj construction

Projective schemes are important for a number of reasons. Here are a few. Schemes that were of "classical interest" in geometry — and those that you would have cared about before knowing about schemes — are all projective or quasiprojective. Moreover, schemes of "current interest" tend to be projective or quasiprojective. In fact, it is very hard to even give an example of a scheme satisfying basic properties — for example, finite type and "Hausdorff" ("separated") over a field — that is provably not quasiprojective. For complex geometers: it is hard to find a compact complex variety that is provably not projective (see Remark 10.3.6), and it is quite hard to come up with a complex variety that is provably not an open subset of a projective variety. So projective schemes are really ubiquitous. Also a projective k-scheme is a good approximation of the algebro-geometric version of compactness ("properness", see §10.3).

Finally, although projective schemes may be obtained by gluing together affine schemes, and we know that keeping track of gluing can be annoying, there is a simple means of dealing with them without worrying about gluing. Just as there is a rough dictionary between rings and affine schemes, we will have a (slightly looser) analogous dictionary between graded rings and projective schemes. Just as one can work with affine schemes by instead working with rings, one can work with projective schemes by instead working with graded rings.

4.5.1. Motivation from classical geometry.
For geometric intuition, we recall how one thinks of projective space "classically" (in the classical topology, over the real numbers). $\mathbb{P}^n$ can be interpreted as the lines through the origin in $\mathbb{R}^{n+1}$. Thus subsets of $\mathbb{P}^n$ correspond to unions of lines through the origin of $\mathbb{R}^{n+1}$, and closed subsets correspond to such unions which are closed. (The same is not true with "closed" replaced by "open"!)

One often pictures $\mathbb{P}^n$ as being the "points at infinite distance" in $\mathbb{R}^{n+1}$, where the points infinitely far in one direction are associated with the points infinitely far in the opposite direction. We can make this more precise using the decomposition

$$\mathbb{P}^{n+1} = \mathbb{R}^{n+1} \coprod \mathbb{P}^n$$

by which we mean that there is an open subset in $\mathbb{P}^{n+1}$ identified with $\mathbb{R}^{n+1}$ (the points with last "projective coordinate" nonzero), and the complementary closed subset identified with $\mathbb{P}^n$ (the points with last "projective coordinate" zero). (The phrase "projective coordinate" will be formally defined in §4.5.8, but we will use it even before then, in Exercises 4.5.A and 4.5.B.)

Then for example any equation cutting out some set V of points in $\mathbb{P}^n$ will also cut out some set of points in $\mathbb{R}^{n+1}$ that will be a closed union of lines. We call this the *affine cone* of V. These equations will cut out some union of $\mathbb{P}^1$'s in $\mathbb{P}^{n+1}$, and we call this the *projective cone* of V. The projective cone is the disjoint union of the affine cone and V. For example, the affine cone over $x^2 + y^2 = z^2$ in $\mathbb{P}^2$ is just the "classical" picture of a cone in $\mathbb{R}^3$, see Figure 4.9. We will make this analogy precise in our algebraic setting in §8.2.12.

4.5.2. Projective schemes, a first description.
We now describe a construction of projective schemes, which will help motivate the Proj construction. We begin by giving an algebraic interpretation of the

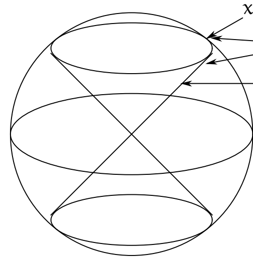


FIGURE 4.9. The affine and projective cone of $x^2 + y^2 = z^2$ in classical geometry

cone just described. We switch coordinates from x, y, z to x_0, x_1, x_2 in order to use the notation of §4.4.9.

4.5.A. EXERCISE (WORTH DOING BEFORE READING THE REST OF THIS SECTION). Consider $\mathbb{P}^2_k$, with projective coordinates x_0, x_1, and x_2. (The terminology "projective coordinate" will not be formally defined until §4.5.8, but you should be able to solve this problem anyway.) Think through how to define a scheme that should be interpreted as $x_0^2 + x_1^2 - x_2^2 = 0$ "in $\mathbb{P}^2_k$". Hint: in the affine open subset corresponding to $x_2 \neq 0$, it should (in the language of 4.4.9) be cut out by $x_{0/2}^2 + x_{1/2}^2 - 1 = 0$, i.e., it should "be" the scheme $\operatorname{Spec} k[x_{0/2}, x_{1/2}]/(x_{0/2}^2 + x_{1/2}^2 - 1)$. You can similarly guess what it should be on the other two standard open sets, and show that the three schemes glue together.

4.5.3. *Remark: Degree* d *hypersurfaces in* $\mathbb{P}^n$. We informally observe that degree d homogeneous polynomials in $n+1$ variables over a field form a vector space of dimension $\binom{n+d}{d}$. (This is essentially the content of Exercise 8.2.K and Exercise 14.1.C.) Two polynomials cut out the same subset of $\mathbb{P}^n_k$ if one is a nonzero multiple of the other. You will later be able to check that two polynomials cut out the same *closed subscheme* (whatever that means) if *and only if* one is a nonzero multiple of the other. The zero polynomial doesn't really cut out a hypersurface in any reasonable sense of the word. Thus we informally imagine that "degree d hypersurfaces in $\mathbb{P}^n$ are parametrized by $\mathbb{P}^{\binom{n+d}{d}-1}$". This intuition will come up repeatedly (in special cases), and we will give it some precise meaning in §28.3.5. (We will properly define *hypersurfaces* in §8.2.2, once we have the language of closed subschemes. At that time we will also define line, hyperplane, quadric hypersurfaces, conic curves, and other wondrous notions.)

4.5.B. EXERCISE. More generally, consider $\mathbb{P}^n_A$, with projective coordinates x_0, $\dots$, x_n. Given a collection of homogeneous polynomials $f_i \in A[x_0, \dots, x_n]$, make sense of the scheme "cut out in $\mathbb{P}^n_A$ by the f_i." (This will later be made precise as an example of a "vanishing scheme", see Exercise 4.5.P.) Hint: you will be able to piggyback on Exercise 4.4.D to make this quite straightforward.

This can be taken as the definition of a *projective* A*-scheme*, but we will wait until §4.5.9 to state it a little better.

4.5.4. Preliminaries on graded rings.

The Proj construction produces a scheme out of a graded ring. We now give some background on graded rings.

4.5.5. *$\mathbb{Z}$-graded rings.* A **$\mathbb{Z}$-graded ring** is a ring $S_\bullet = \oplus_{n\in\mathbb{Z}} S_n$ (the subscript is called the **grading**), where multiplication respects the grading, i.e., sends $S_m \times S_n$ to S_{m+n}. Clearly S_0 is a subring, each S_n is an S_0-module, and $S_\bullet$ is a S_0-algebra. Suppose for the remainder of §4.5.5 that $S_\bullet$ is a $\mathbb{Z}$-graded ring. Those elements of some S_n are called **homogeneous elements** of $S_\bullet$; nonzero homogeneous elements have an obvious **degree**. An ideal I of $S_\bullet$ is a **homogeneous ideal** if it is generated by homogeneous elements.

4.5.C. EXERCISE.
(a) Show that an ideal I is homogeneous if and only if it contains the degree n piece of each of its elements for each n. (Hence I can be decomposed into homogeneous pieces, $I = \oplus I_n$, and S/I has a natural $\mathbb{Z}$-graded structure.)
(b) Show that homogeneous ideals are closed under sum, product, intersection, and radical.
(c) Show that a homogeneous ideal $I \subset S_\bullet$ is prime if $I \neq S_\bullet$, and if for any *homogeneous* $a, b \in S$, if $ab \in I$, then $a \in I$ or $b \in I$.

If T is a multiplicative subset of $S_\bullet$ containing only homogeneous elements, then $T^{-1}S_\bullet$ has a natural structure as a $\mathbb{Z}$-graded ring.

(Everything in §4.5.5 can be generalized: $\mathbb{Z}$ can be replaced by an arbitrary abelian group.)

4.5.6. *$\mathbb{Z}^{\geq 0}$-graded rings, graded ring over* A*, and finitely generated graded rings.* A **$\mathbb{Z}^{\geq 0}$-graded ring** is a $\mathbb{Z}$-graded ring with no elements of negative degree.

For the remainder of the book, *graded ring* will refer to a $\mathbb{Z}^{\geq 0}$-graded ring. Warning: this convention is nonstandard (for good reason).

From now on, unless otherwise stated, $S_\bullet$ is assumed to be a graded ring. Fix a ring A, which we call the **base ring**. If $S_0 = A$, we say that $S_\bullet$ is a **graded ring over** A. A key example is $A[x_0, \dots, x_n]$, or more generally $A[x_0, \dots, x_n]/I$ where I is a homogeneous ideal (cf. Exercise 4.5.B). Here we take the conventional grading on $A[x_0, \dots, x_n]$, where each x_i has weight 1.

The subset $S_+ := \oplus_{i>0} S_i \subset S_\bullet$ is an ideal, called the **irrelevant ideal**. The reason for the name "irrelevant" will be clearer in a few paragraphs. If the irrelevant ideal S_+ is a finitely generated ideal, we say that $S_\bullet$ is a **finitely generated graded ring over** A. If $S_\bullet$ is generated by S_1 as an A-algebra, we say that $S_\bullet$ is **generated in degree** 1. (We will later find it useful to interpret "$S_\bullet$ is generated in degree 1" as "the natural map $\mathrm{Sym}^\bullet S_1 \to S_\bullet$ is a surjection". The *symmetric algebra* construction will be briefly discussed in §13.5.3.)

4.5.D. EXERCISE.
(a) Show that a graded ring $S_\bullet$ over A is a finitely generated graded ring (over A) if and only if $S_\bullet$ is a finitely generated graded A-algebra, i.e., generated over $A = S_0$ by a finite number of homogeneous elements of positive degree. (Hint

for the forward implication: show that the generators of S_+ as an ideal are also generators of $S_\bullet$ as an algebra.)
(b) Show that a graded ring $S_\bullet$ over A is Noetherian if and only if $A = S_0$ is Noetherian and $S_\bullet$ is a finitely generated graded ring.

4.5.7. The Proj **construction.**
We now define a scheme $\operatorname{Proj} S_\bullet$, where $S_\bullet$ is a ($\mathbb{Z}^{\geq 0}$-)graded ring. Here are two examples, to provide a light at the end of the tunnel. If $S_\bullet = A[x_0, \dots, x_n]$, we will recover $\mathbb{P}^n_A$; and if $S_\bullet = A[x_0, \dots, x_n]/(f(x_0, \dots, x_n))$ where f is homogeneous, we will construct something "cut out in $\mathbb{P}^n_A$ by the equation $f = 0$" (cf. Exercise 4.5.B).

As we did with Spec of a ring, we will build $\operatorname{Proj} S_\bullet$ first as a set, then as a topological space, and finally as a ringed space. In our preliminary definition of $\mathbb{P}^n_A$, we glued together $n+1$ well-chosen affine pieces, but we don't want to make any choices, so we do this by simultaneously considering "all possible" affine open sets. Our affine building blocks will be as follows. For each homogeneous $f \in S_+$, note that the localization $(S_\bullet)_f$ is naturally a $\mathbb{Z}$-graded ring, where $\deg(1/f) = -\deg f$. Consider

$$\operatorname{Spec}((S_\bullet)_f)_0. \tag{4.5.7.1}$$

where $((S_\bullet)_f)_0$ means the 0-graded piece of the graded ring $(S_\bullet)_f$. (These will be our affine building blocks, as f varies over the homogeneous elements of S_+.) The notation $((S_\bullet)_f)_0$ is admittedly horrible — the first and third subscripts refer to the grading, and the second refers to localization. As motivation for considering this construction: applying this to $S_\bullet = k[x_0, \dots, x_n]$, with $f = x_i$, we obtain the ring appearing in (4.4.9.1):

$$k[x_{0/i}, x_{1/i}, \dots, x_{n/i}]/(x_{i/i} - 1).$$

(Before we begin the construction: another possible way of defining $\operatorname{Proj} S_\bullet$ is by gluing together affines of this form, by jumping straight to Exercises 4.5.K and 4.5.L. If you prefer that, by all means do so.)

The *points* of $\operatorname{Proj} S_\bullet$ are the set of homogeneous prime ideals of $S_\bullet$ not containing the irrelevant ideal S_+ (the "relevant prime ideals").

4.5.E. IMPORTANT AND TRICKY EXERCISE. Suppose $f \in S_+$ is homogeneous.
(a) Give a bijection between the primes of $((S_\bullet)_f)_0$ and the homogeneous prime ideals of $(S_\bullet)_f$. Hint: Avoid notational confusion by proving instead that if A is a $\mathbb{Z}$-graded ring with a homogeneous invertible element f in positive degree, then there is a bijection between prime ideals of A_0 and homogeneous prime ideals of A. Using the ring map $A_0 \to A$, from each homogeneous prime ideal of A we find a prime ideal of A_0. The reverse direction is the harder one. Given a prime ideal $P_0 \subset A_0$, define $P \subset A$ (a priori only a subset) as $\oplus Q_i$, where $Q_i \subset A_i$, and $a \in Q_i$ if and only if $a^{\deg f}/f^i \in P_0$. Note that $Q_0 = P_0$. Show that $a \in Q_i$ if and only if $a^2 \in Q_{2i}$; show that if $a_1, a_2 \in Q_i$ then $a_1^2 + 2a_1a_2 + a_2^2 \in Q_{2i}$ and hence $a_1 + a_2 \in Q_i$; then show that P is a homogeneous ideal of A; then show that P is prime.
(b) Interpret the set of prime ideals of $((S_\bullet)_f)_0$ as a subset of $\operatorname{Proj} S_\bullet$.

The correspondence of the points of $\operatorname{Proj} S_\bullet$ with homogeneous prime ideals helps us picture $\operatorname{Proj} S_\bullet$. For example, if $S_\bullet = k[x, y, z]$ with the usual grading, then we picture the homogeneous prime ideal $(z^2 - x^2 - y^2)$ first as a subset of $\operatorname{Spec} S_\bullet$;

it is a cone (see Figure 4.9). As in §4.5.1, we picture $\mathbb{P}^2_k$ as the "plane at infinity". Thus we picture this equation as cutting out a conic "at infinity" (in $\operatorname{Proj} S_\bullet$). We will make this intuition somewhat more precise in §8.2.12.

Motivated by the affine case, if T is a set of homogeneous elements of $S_\bullet$ *of positive degree*, define the (projective) **vanishing set of** T, $V(T) \subset \operatorname{Proj} S_\bullet$, to be those homogeneous prime ideals containing T but not S_+. Define $V(f)$ if f is a homogeneous element of positive degree, and $V(I)$ if I is a homogeneous ideal contained in S_+, in the obvious way. Let $D(f) = \operatorname{Proj} S_\bullet \setminus V(f)$ (the **projective distinguished open set**) be the complement of $V(f)$. Once we define a scheme structure on $\operatorname{Proj} S_\bullet$, we will (without comment) use $D(f)$ to refer to the open *subscheme*, not just the open subset. (These definitions can certainly be extended to remove the positive degree hypotheses. For example, the definition of $V(T)$ makes sense for any subset T of $S_\bullet$, and the definition of $D(f)$ makes sense even if f has degree 0. In what follows, we deliberately stick to these narrower definitions. For example, we will want the $D(f)$ to form an affine cover, and if f has degree 0, then $D(f)$ needn't be affine.)

4.5.F. EXERCISE. Show that $D(f)$ "is" (or more precisely, "corresponds to") the subset $\operatorname{Spec}((S_\bullet)_f)_0$ you described in Exercise 4.5.E(b). For example, in §4.4.9, the $D(x_i)$ are the standard open sets covering projective space.

As in the affine case, the $V(I)$'s satisfy the axioms of the closed set of a topology, and we call this the **Zariski topology** on $\operatorname{Proj} S_\bullet$. (Other definitions given in the literature may look superficially different, but can be easily shown to be the same.) Many statements about the Zariski topology on Spec of a ring carry over to this situation with little extra work. Clearly $D(f) \cap D(g) = D(fg)$, by the same immediate argument as in the affine case (Exercise 3.5.D).

4.5.G. EASY EXERCISE. Verify that the projective distinguished open sets $D(f)$ (as f runs through the homogeneous elements of S_+) form a base of the Zariski topology.

4.5.H. EXERCISE. Fix a graded ring $S_\bullet$.
(a) Suppose I is any homogeneous ideal of $S_\bullet$ contained in S_+, and f is a homogeneous element of positive degree. Show that f vanishes on $V(I)$ (i.e., $V(I) \subset V(f)$) if and only if $f^n \in I$ for some n. (Hint: Mimic the affine case; see Exercise 3.4.J.) In particular, as in the affine case (Exercise 3.5.E), if $D(f) \subset D(g)$, then $f^n \in (g)$ for some n, and vice versa. (Here g is also homogeneous of positive degree.)
(b) If $Z \subset \operatorname{Proj} S_\bullet$, define $I(Z) \subset S_+$. Show that it is a homogeneous ideal of $S_\bullet$. For any two subsets, show that $I(Z_1 \cup Z_2) = I(Z_1) \cap I(Z_2)$.
(c) For any subset $Z \subset \operatorname{Proj} S_\bullet$, show that $V(I(Z)) = \overline{Z}$.

4.5.I. EXERCISE (CF. EXERCISE 3.5.B). Fix a graded ring $S_\bullet$, and a homogeneous ideal I. Show that the following are equivalent.

(a) $V(I) = \varnothing$.
(b) For any f_i (as i runs through some index set) generating I, $\cup D(f_i) = \operatorname{Proj} S_\bullet$.
(c) $\sqrt{I} \supset S_+$.

This is more motivation for the ideal S_+ being "irrelevant": any ideal whose radical contains it is "geometrically irrelevant".

We now construct $\operatorname{Proj} S_\bullet$ as a *scheme*.

4.5.J. EXERCISE. Suppose some homogeneous $f \in S_+$ is given. Via the inclusion

$$D(f) = \operatorname{Spec}((S_\bullet)_f)_0 \hookrightarrow \operatorname{Proj} S_\bullet$$

of Exercise 4.5.F, show that the Zariski topology on $\operatorname{Proj} S_\bullet$ restricts to the Zariski topology on $\operatorname{Spec}((S_\bullet)_f)_0$.

Now that we have defined $\operatorname{Proj} S_\bullet$ as a topological space, we are ready to define the structure sheaf. On $D(f)$, we wish it to be the structure sheaf of $\operatorname{Spec}((S_\bullet)_f)_0$. We will glue these sheaves together using Exercise 2.7.D on gluing sheaves.

4.5.K. EXERCISE. If $f, g \in S_+$ are homogeneous and nonzero, describe an isomorphism between $\operatorname{Spec}((S_\bullet)_{fg})_0$ and the distinguished open subset $D(g^{\deg f}/f^{\deg g})$ of $\operatorname{Spec}((S_\bullet)_f)_0$.

Similarly, $\operatorname{Spec}((S_\bullet)_{fg})_0$ is identified with a distinguished open subset of $\operatorname{Spec}((S_\bullet)_g)_0$. We then glue the various $\operatorname{Spec}((S_\bullet)_f)_0$ (as f varies) altogether, using these pairwise gluings.

4.5.L. EXERCISE. By checking that these gluings behave well on triple overlaps (see Exercise 2.7.D), finish the definition of the scheme $\operatorname{Proj} S_\bullet$.

4.5.M. EXERCISE. (Some will find this essential, others will prefer to ignore it.) (Re)interpret the structure sheaf of $\operatorname{Proj} S_\bullet$ in terms of compatible stalks.

4.5.8. *Definition.* We (re)define **projective space** (over a ring A) by $\mathbb{P}^n_A := \operatorname{Proj} A[x_0, \dots, x_n]$. This definition involves no messy gluing, or special choice of patches. Note that the variables $x_0, \dots, x_n$, which we call the **projective coordinates** on $\mathbb{P}^n_A$, are part of the definition. (They may have other names than x's, depending on the context.)

4.5.N. EXERCISE. Check that this agrees with our earlier construction of $\mathbb{P}^n_A$ (§4.4.9). (How do you know that the $D(x_i)$ cover $\operatorname{Proj} A[x_0, \dots, x_n]$?)

Notice that with our old definition of projective space, it would have been a nontrivial exercise to show that $D(x^2 + y^2 - z^2) \subset \mathbb{P}^2_k$ (the complement of a plane conic) is affine; with our new perspective, it is immediate — it is $\operatorname{Spec}(k[x, y, z]_{(x^2+y^2-z^2)})_0$.

4.5.O. EXERCISE. Suppose that k is an algebraically closed field. We know from Exercise 4.4.F that the closed points of $\mathbb{P}^n_k$, as defined in §4.4.9, are in bijection with the points of "classical" projective space. By Exercise 4.5.N, the scheme $\mathbb{P}^n_k$ as defined in §4.4.9 is isomorphic to $\operatorname{Proj} k[x_0, \dots, x_n]$. Therefore, each point $[a_0, \dots, a_n]$ of classical projective space corresponds to a homogeneous prime ideal of $k[x_0, \dots, x_n]$. Which homogeneous prime ideal is it?

We now figure out the "right definition" of the vanishing scheme, in analogy with the vanishing *set* $V(\cdot)$ defined at the start of §3.4. You will be defining a *closed subscheme* (mentioned in Remark 4.3.4, and to be properly defined in §8.1).

4.5.P. EXERCISE. If $S_\bullet$ is generated in degree 1, and $f \in S_+$ is homogeneous, explain how to define $V(f)$ "in" $\operatorname{Proj} S_\bullet$, the **vanishing scheme of** f. (Warning: f in general isn't a function on $\operatorname{Proj} S_\bullet$. We will later interpret it as something close: a section of a line bundle, see for example §14.1.2.) Hence define $V(I)$ for any homogeneous ideal I of S_+. (Another solution in more general circumstances will be given in Exercise 13.1.I.)

4.5.9. Projective and quasiprojective schemes.

We call a scheme of the form (i.e., isomorphic to) $\operatorname{Proj} S_\bullet$, where $S_\bullet$ is a *finitely generated* graded ring over A, a **projective scheme over** A, or a **projective** A**-scheme**. A **quasiprojective** A**-scheme** is a quasicompact open subscheme of a projective A-scheme. The "A" is omitted if it is clear from the context; often A is a field.

4.5.10. *Unimportant remarks.* (i) Note that $\operatorname{Proj} S_\bullet$ makes sense even when $S_\bullet$ is not finitely generated. This can be useful. For example, you will later be able to do Exercise 6.4.D without worrying about Exercise 6.4.H.

(ii) The quasicompact requirement in the definition of quasiprojectivity is of course redundant in the Noetherian case (cf. Exercise 3.6.T), which is all that matters to most.

4.5.11. *Silly example.* Note that $\mathbb{P}^0_A = \operatorname{Proj} A[T] \cong \operatorname{Spec} A$. Thus "$\operatorname{Spec} A$ is a projective A-scheme".

4.5.12. Example: projectivization of a vector space $\mathbb{P}V$. We can make this definition of projective space even more choice-free as follows. Let V be an $(n+1)$-dimensional vector space over k. (Here k can be replaced by any ring A as usual.) Define

$$\operatorname{Sym}^\bullet V^\vee = k \oplus V^\vee \oplus \operatorname{Sym}^2 V^\vee \oplus \cdots .$$

(The reason for the dual is explained by the next exercise. For a reminder of the definition of Sym, flip to §13.5.3.) If for example V is the dual of the vector space with basis associated to $x_0, \dots, x_n$, we would have $\operatorname{Sym}^\bullet V^\vee = k[x_0, \dots, x_n]$. Then we can define $\mathbb{P}V := \operatorname{Proj}(\operatorname{Sym}^\bullet V^\vee)$. In this language, we have an interpretation for $x_0, \dots, x_n$: they are linear functionals on the underlying vector space V. (Warning: some authors use the definition $\mathbb{P}V = \operatorname{Proj}(\operatorname{Sym}^\bullet V)$, so be cautious.)

4.5.Q. UNIMPORTANT EXERCISE. Suppose k is algebraically closed. Describe a natural bijection between one-dimensional subspaces of V and the closed points of $\mathbb{P}V$. Thus this construction canonically (in a basis-free manner) describes the one-dimensional subspaces of the vector space V.

Unimportant remark: you may be surprised at the appearance of the dual in the definition of $\mathbb{P}V$. This is partially explained by the previous exercise. Most normal (traditional) people define the projectivization of a vector space V to be the space of one-dimensional subspaces of V. Grothendieck considered the projectivization to be the space of one-dimensional *quotients*. One motivation for this is that it gets rid of the annoying dual in the definition above. There are better reasons, that we won't go into here. In a nutshell, quotients tend to be better-behaved than subobjects for coherent sheaves, which generalize the notion of vector bundle. (Coherent sheaves are discussed in Chapter 13.)

On another note related to Exercise 4.5.Q: you can also describe a natural bijection between points of V and the closed points of $\operatorname{Spec}(\operatorname{Sym}^\bullet V^\vee)$. This construction respects the affine/projective cone picture of §8.2.12.

4.5.13. *The Grassmannian.* At this point, we could describe the fundamental geometric object known as the *Grassmannian*, and give the "wrong" (but correct) definition of it. We will instead wait until §6.7 to give the wrong definition, when we will know enough to sense that something is amiss. The right definition will be given in §16.7.

CHAPTER 5

Some properties of schemes

5.1 Topological properties

We will now define some useful properties of schemes. As you see each example, you should try these out in specific examples of your choice, such as particular schemes of the form $\operatorname{Spec} \mathbb{C}[x_1 \ldots, x_n]/(f_1, \ldots, f_r)$.

The definitions of *connected*, *connected component*, *(ir)reducible*, *quasicompact*, *closed point*, *specialization*, *generization*, *generic point*, and *irreducible component* were given in §3.6. You should have pictures in your mind of each of these notions.

Exercise 3.6.C shows that $\mathbb{A}^n$ is irreducible (it was easy). This argument "behaves well under gluing", yielding:

5.1.A. EASY EXERCISE. Show that $\mathbb{P}^n_k$ is irreducible.

5.1.B. EXERCISE. Exercise 3.7.E showed that there is a bijection between irreducible closed subsets and points for affine schemes. Show that this is true of schemes as well.

5.1.C. EASY EXERCISE. Prove that if X is a scheme that has a finite cover $X = \cup_{i=1}^n \operatorname{Spec} A_i$ where A_i is Noetherian, then X is a Noetherian topological space (§3.6.14). (We will soon call a scheme with such a cover a *Noetherian scheme*, §5.3.4.) Hint: show that a topological space that is a finite union of Noetherian subspaces is itself Noetherian.

Thus $\mathbb{P}^n_k$ and $\mathbb{P}^n_{\mathbb{Z}}$ are Noetherian topological spaces: we built them by gluing together a finite number of spectra of Noetherian rings.

5.1.D. EASY EXERCISE. Show that a scheme X is quasicompact if and only if it can be written as a finite union of affine open subschemes. (Hence $\mathbb{P}^n_A$ is quasicompact for any ring A.)

5.1.E. IMPORTANT EXERCISE: QUASICOMPACT SCHEMES HAVE CLOSED POINTS. Show that if X is a quasicompact scheme, then every point has a closed point in its closure. Show that every nonempty closed subset of X contains a closed point of X. In particular, every nonempty quasicompact scheme has a closed point. (Warning: there exist nonempty schemes with no closed points — see for example [**Liu**, Exer. 3.27] or [**Schw**] — so your argument had better use the quasicompactness hypothesis!)

This exercise will often be used in the following way. If there is some property P of points of a scheme that is "open" (if a point p has P, then there is some neighborhood U of p such that all the points in U have P), then to check if *all* points of a quasicompact scheme have P, it suffices to check only the closed points. (A first example of this philosophy is Exercise 5.2.D.) This provides a connection between schemes and the classical theory of varieties — the points of traditional varieties are the *closed* points of the corresponding schemes (essentially by the Nullstellensatz, see §3.6.9 and Exercise 5.3.E). In many good situations, the closed points are dense (such as for varieties, see §3.6.9 and Exercise 5.3.E again), but this is not true in some fundamental cases (see Exercise 3.6.J(b)).

5.1.1. Quasiseparated schemes. Quasiseparatedness is a weird notion that comes in handy for certain people. (Warning: we will later realize that this is really a property of *morphisms*, not of schemes §7.3.1.) Most people, however, can ignore this notion, as the schemes they will encounter in real life will all have this property. A topological space is **quasiseparated** if the intersection of any two quasicompact open sets is quasicompact. (The motivation for the "separatedness" in the name is that it is a weakened version of *separated*, for which the intersection of any two *affine* open sets is affine, see Proposition 10.1.8.)

5.1.F. SHORT EXERCISE. Show that a scheme is quasiseparated if and only if the intersection of any two affine open subsets is a finite union of affine open subsets.

We will see later that this will be a useful hypothesis in theorems (in conjunction with quasicompactness), and that various interesting kinds of schemes (affine, locally Noetherian, separated, see Exercises 5.1.G, 5.3.A, and 10.1.H respectively) are quasiseparated, and this will allow us to state theorems more succinctly (e.g. "if X is quasicompact and quasiseparated" rather than "if X is quasicompact, and either this or that or the other thing hold").

5.1.G. EXERCISE. Show that affine schemes are quasiseparated.

"Quasicompact and quasiseparated" means something concrete:

5.1.H. EXERCISE. Show that a scheme X is quasicompact and quasiseparated if and only if X can be covered by a finite number of affine open subsets, any two of which have intersection also covered by a finite number of affine open subsets.

So when you see "quasicompact and quasiseparated" as hypotheses in a theorem, you should take this as a clue that you will use this interpretation, and that finiteness will be used in an essential way.

5.1.I. EASY EXERCISE. Show that all projective A-schemes are quasicompact and quasiseparated. (Hint: use the fact that the graded ring in the definition is finitely generated — those finite number of generators will lead you to a covering set.)

5.1.J. EXERCISE (A NONQUASISEPARATED SCHEME). Let $X = \operatorname{Spec} k[x_1, x_2, \dots]$, and let U be $X - [\mathfrak{m}]$ where $\mathfrak{m}$ is the maximal ideal $(x_1, x_2, \dots)$. Take two copies of X, glued along U ("affine ∞-space with a doubled origin", see Example 4.4.5 and Exercise 4.4.C for "finite-dimensional" versions). Show that the result is not quasiseparated. Hint: This open embedding $U \subset X$ came up earlier in Exercise 3.6.G(b) as an example of a nonquasicompact open subset of an affine scheme.

5.1.2. Dimension. One very important topological notion is *dimension*. (It is amazing that this is a *topological* idea.) But despite being intuitively fundamental, it is more difficult, so we postpone it until Chapter 11.

5.2 Reducedness and integrality

Recall that one of the alarming things about schemes is that functions are not determined by their values at points, and that was because of the presence of nilpotents (§3.2.11).

5.2.1. *Definition.* A ring is said to be *reduced* if it has no nonzero nilpotents (§3.2.13). A scheme X is **reduced** if $\mathscr{O}_X(U)$ is reduced for every open set U of X.

5.2.A. EXERCISE (REDUCEDNESS IS A **stalk-local** PROPERTY, I.E., CAN BE CHECKED AT STALKS). Show that a scheme is reduced if and only if none of the stalks have nonzero nilpotents. Hence show that if f and g are two functions (global sections of $\mathscr{O}_X$) on a reduced scheme that agree at all points, then $f = g$. (Two hints: $\mathscr{O}_X(U) \hookrightarrow \prod_{p \in U} \mathscr{O}_{X,p}$ from Exercise 2.4.A, and the nilradical is intersection of all prime ideals from Theorem 3.2.12.)

5.2.B. EXERCISE. If A is a reduced ring, show that $\operatorname{Spec} A$ is reduced. Show that $\mathbb{A}^n_k$ and $\mathbb{P}^n_k$ are reduced.

The scheme $\operatorname{Spec} k[x,y]/(y^2, xy)$ is nonreduced. When we sketched it in Figure 4.4, we indicated that the fuzz represented nonreducedness at the origin. The following exercise is a first stab at making this precise.

5.2.C. EXERCISE. Show that $\left(k[x,y]/(y^2,xy)\right)_x$ has no nonzero nilpotent elements. (Possible hint: show that it is isomorphic to another ring, by considering the geometric picture. Exercise 3.2.L may give another hint.) Show that the only point of $\operatorname{Spec} k[x,y]/(y^2,xy)$ with a nonreduced stalk is the origin.

5.2.D. EXERCISE. If X is a quasicompact scheme, show that it suffices to check reducedness at closed points. Hint: Do not try to show that reducedness is an open condition (see Remark 5.2.2). Instead show that any nonreduced point has a nonreduced closed point in its closure, using Exercise 5.1.E. (This result is interesting, but we won't use it.)

5.2.2. $\star$ *Remark.* Reducedness is not in general an open condition. You may be able to identify the underlying topological space of

$$X = \operatorname{Spec} \mathbb{C}[x, y_1, y_2, \dots]/(y_1^2, y_2^2, y_3^2, \dots, (x-1)y_1, (x-2)y_2, (x-3)y_3, \cdots)$$

with that of $\operatorname{Spec} \mathbb{C}[x]$, and then to show that the nonreduced points of X are precisely the closed points corresponding to positive integers. The complement of this set is not Zariski open. (This ring will come up again in §5.5, in the paragraph after the statement of property **(B)**.) However, we will see in Remark 5.5.5 (and again in Remark 13.7.2) that if X is a locally Noetherian scheme, then the reduced locus is indeed open.

5.2.3. $\star$ *Another warning for experts.* If a scheme X is reduced, then from the definition of reducedness, its ring of global sections is reduced. However, the converse is not true; the example of the scheme X cut out by $x^2 = 0$ in $\mathbb{P}^2_k$ will come up in §18.1.6, and you already know enough to verify that $\Gamma(X, \mathscr{O}_X) \cong k$, and that X is nonreduced.

5.2.E. EXERCISE. Suppose X is quasicompact, and f is a function that vanishes at all points of X. Show that there is some n such that $f^n = 0$. Show that this may fail if X is not quasicompact. (This exercise is less important, but shows why we like quasicompactness, and gives a standard pathology when quasicompactness doesn't hold.) Hint: take an infinite disjoint union of $\operatorname{Spec} A_n$ with $A_n := k[\epsilon]/(\epsilon^n)$. This scheme arises again in §8.3.2 (see Figure 8.4 for a picture) and in Caution/Example 8.3.11.

5.2.4. *Definition.* A scheme X is **integral** if it is nonempty, and $\mathscr{O}_X(U)$ is an integral domain for every nonempty open set U of X.

5.2.F. IMPORTANT EXERCISE. Show that a scheme X is integral if and only if it is irreducible and reduced. (Thus we picture integral schemes as: "one piece, no fuzz". Possible hint: Exercise 4.3.G.)

5.2.G. EXERCISE. Show that an affine scheme $\operatorname{Spec} A$ is integral if and only if A is an integral domain.

5.2.H. EXERCISE. Suppose X is an integral scheme. Then X (being irreducible) has a generic point η. Suppose $\operatorname{Spec} A$ is any nonempty affine open subset of X. Show that the stalk at η, $\mathscr{O}_{X,\eta}$ is naturally identified with $K(A)$, the fraction field of A. This is called the **function field** $K(X)$ of X. It can be computed on any nonempty open set of X, as any such open set contains the generic point. The reason for the name: we will soon think of this as the field of *rational functions on* X (Definition 5.5.6 and Exercise 5.5.Q).

5.2.I. EXERCISE. Suppose X is an integral scheme. Show that the restriction maps $\operatorname{res}_{U,V} : \mathscr{O}_X(U) \to \mathscr{O}_X(V)$ are inclusions so long as $V \neq \varnothing$. Suppose $\operatorname{Spec} A$ is any nonempty affine open subset of X (so A is an integral domain). Show that the natural map $\mathscr{O}_X(U) \to \mathscr{O}_{X,\eta} = K(A)$ (where U is any nonempty open set) is an inclusion.

Thus irreducible varieties (an important example of integral schemes defined later) have the convenient property that sections over different open sets can be considered subsets of the same ring. In particular, restriction maps (except to the empty set) are always inclusions, and gluing is easy: functions f_i on a cover U_i of U (as i runs over an index set) glue if and only if they are the same element of $K(X)$. This is one reason why (irreducible) varieties are usually introduced before schemes.

Integrality is not stalk-local (the disjoint union of two integral schemes is not integral, as $\operatorname{Spec} A \coprod \operatorname{Spec} B = \operatorname{Spec}(A \times B)$ by Exercise 3.6.A), but it almost is, see Exercise 5.3.C.

5.3 Properties of schemes that can be checked "affine-locally"

This section is intended to address something tricky in the definition of schemes. We have defined a scheme as a topological space with a sheaf of rings, that can be covered by affine schemes. Hence we have all of the affine open sets in the cover, but we don't know how to communicate between any two of them. Somewhat more explicitly, if I have an affine cover, and you have an affine cover, and we want to compare them, and I calculate something on my cover, there should be some way of us getting together, and figuring out how to translate my calculation over to your cover. The Affine Communication Lemma 5.3.2 will provide a convenient machine for doing this.

Thanks to this lemma, we can define a host of important properties of schemes. All of these are "affine-local" in that they can be checked on any affine cover, i.e., a covering by open affine sets. We like such properties because we can check them using any affine cover we like. If the scheme in question is quasicompact, then we need only check a finite number of affine open sets.

5.3.1. Proposition. — *Suppose* $\operatorname{Spec} A$ *and* $\operatorname{Spec} B$ *are affine open subschemes of a scheme* X*. Then* $\operatorname{Spec} A \cap \operatorname{Spec} B$ *is the union of open sets that are simultaneously distinguished open subschemes of* $\operatorname{Spec} A$ *and* $\operatorname{Spec} B$.

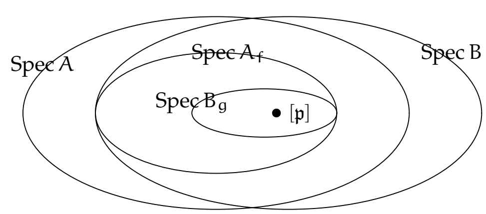


FIGURE 5.1. A trick to show that the intersection of two affine open sets may be covered by open sets that are simultaneously distinguished in both affine open sets

Proof. (See Figure 5.1.) Given any point $\mathfrak{p} \in \operatorname{Spec} A \cap \operatorname{Spec} B$, we produce an open neighborhood of $\mathfrak{p}$ in $\operatorname{Spec} A \cap \operatorname{Spec} B$ that is simultaneously distinguished in both $\operatorname{Spec} A$ and $\operatorname{Spec} B$. Let $\operatorname{Spec} A_f$ be a distinguished open subset of $\operatorname{Spec} A$ contained in $\operatorname{Spec} A \cap \operatorname{Spec} B$ and containing $\mathfrak{p}$. Let $\operatorname{Spec} B_g$ be a distinguished open subset of $\operatorname{Spec} B$ contained in $\operatorname{Spec} A_f$ and containing $\mathfrak{p}$. Then $g \in \Gamma(\operatorname{Spec} B, \mathscr{O}_X)$ restricts to an element $g' \in \Gamma(\operatorname{Spec} A_f, \mathscr{O}_X) = A_f$. The points of $\operatorname{Spec} A_f$ where g vanishes are precisely the points of $\operatorname{Spec} A_f$ where g' vanishes, so

$$\begin{aligned} \operatorname{Spec} B_g &= \operatorname{Spec} A_f \setminus \{[\mathfrak{p}] \ : \ g' \in \mathfrak{p}\} \\ &= \operatorname{Spec} (A_f)_{g'}. \end{aligned}$$

If $g' = g''/f^n$ ($g'' \in A$) then $\operatorname{Spec}(A_f)_{g'} = \operatorname{Spec} A_{fg''}$, and we are done. $\square$

The following easy result will be crucial for us.

5.3.2. Affine Communication Lemma. — *Let P be some property enjoyed by some affine open sets of a scheme X, such that*

(i) *if an affine open set* $\operatorname{Spec} A \hookrightarrow X$ *has property* P *then for any* $f \in A$, $\operatorname{Spec} A_f \hookrightarrow X$ *does too.*
(ii) *if* $(f_1, \dots, f_n) = A$, *and* $\operatorname{Spec} A_{f_i} \hookrightarrow X$ *has* P *for all* i, *then so does* $\operatorname{Spec} A \hookrightarrow X$.

Suppose that $X = \cup_{i \in I} \operatorname{Spec} A_i$ *where* $\operatorname{Spec} A_i$ *has property* P. *Then every open affine subset of* X *has* P *too.*

We say such a property is **affine-local**. Note that if U is an open subscheme of X, then U inherits any affine-local property of X. Note also that any property that is stalk-local (a scheme has property P if and only if all its stalks have property Q) is necessarily affine-local (a scheme has property P if and only if all of its affine open sets have property R, where an affine scheme has property R if and only if all its stalks have property Q). But it is sometimes not so obvious what the right definition of Q is; see for example the discussion of normality in the next section.

Proof. Let $\operatorname{Spec} A$ be an affine subscheme of X. Cover $\operatorname{Spec} A$ with a finite number of distinguished open sets $\operatorname{Spec} A_{g_j}$, each of which is distinguished in some $\operatorname{Spec} A_i$. This is possible by Proposition 5.3.1 and the quasicompactness of $\operatorname{Spec} A$ (Exercise 3.6.G(a)). By (i), each $\operatorname{Spec} A_{g_j}$ has P. By (ii), $\operatorname{Spec} A$ has P. $\square$

By choosing property P appropriately, we define some important properties of schemes.

5.3.3. Proposition. — *Suppose A is a ring, and $(f_1, \dots, f_n) = A$.*

(a) *If A is reduced, then A_{f_i} is also reduced. If each A_{f_i} is reduced, then so is A.*
(b) *If A is a Noetherian ring, then so is A_{f_i}. If each A_{f_i} is Noetherian, then so is A.*
(c) *Suppose B is a ring, and A is a B-algebra. (Hence A_g is a B-algebra for all $g \in A$.) If A is a finitely generated B-algebra, then so is A_{f_i}. If each A_{f_i} is a finitely generated B-algebra, then so is A.*

We will prove these shortly (§5.3.9). But let's first motivate you to read the proof by giving some interesting definitions and results *assuming* Proposition 5.3.3 is true.

First, the Affine Communication Lemma 5.3.2 and Proposition 5.3.3(a) implies that X is reduced if and only if X can be covered by affine open sets $\operatorname{Spec} A$ where A is reduced. (This also easily follows from the stalk-local characterization of reducedness, see Exercises 5.2.A and 5.2.B.)

5.3.4. Important Definition. Suppose X is a scheme. If X can be covered by affine open sets $\operatorname{Spec} A$ where A is Noetherian, we say that X is a **locally Noetherian scheme**. If in addition X is quasicompact, or equivalently can be covered by finitely many such affine open sets, we say that X is a **Noetherian scheme**. (We will see a number of definitions of the form "if X has this property, we say that it is locally Q; if further X is quasicompact, we say that it is Q.") By Exercise 5.1.C, the underlying topological space of a Noetherian scheme is Noetherian. Hence by Exercise 3.6.T, all open subsets of a Noetherian scheme are quasicompact.

5.3.A. EXERCISE. Show that locally Noetherian schemes are quasiseparated.

5.3.B. EXERCISE. Show that a Noetherian scheme has a finite number of irreducible components. (Hint: Proposition 3.6.15.) Show that a Noetherian scheme has a finite number of connected components, each a finite union of irreducible components.

5.3.C. EXERCISE. Show that a Noetherian scheme X is integral if and only if X is nonempty and connected and all stalks $\mathscr{O}_{X,p}$ are integral domains. Thus in "good situations", integrality is the union of local (stalks are integral domains) and global (connected) conditions. Hint: Recall that integral = irreducible + reduced (Exercise 5.2.F). If a scheme's stalks are integral domains, then it is reduced (reducedness is a stalk-local condition, Exercise 5.2.A). If a scheme X has underlying topological space that is Noetherian, then X has finitely many irreducible components (by the previous exercise); if two of them meet at a point p, then $\mathscr{O}_{X,p}$ is not an integral domain. (You can readily extend this from Noetherian schemes to locally Noetherian schemes, by showing that a connected scheme is irreducible if and only if it is nonempty and has a cover by open irreducible subsets. But some Noetherian hypotheses are necessary, see [**MO7477**].)

5.3.5. *Unimportant caution.* The ring of sections of a Noetherian scheme need not be Noetherian, see Exercise 19.11.H.

5.3.6. *Schemes over a given field* k*, or more generally over a given ring* A *(*A*-schemes).* You may be particularly interested in working over a particular field, such as $\mathbb{C}$ or $\mathbb{Q}$, or over a ring such as $\mathbb{Z}$. Motivated by this, we define the notion of A**-scheme**, or **scheme over** A, where A is a ring, as a scheme where all the rings of sections of the structure sheaf (over all open sets) are A-algebras, and all restriction maps are maps of A-algebras. (Like some earlier notions such as quasiseparatedness, this will later in Exercise 6.3.G be properly understood as a "relative notion"; it is the data of a morphism $X \to \operatorname{Spec} A$.) Suppose now X is an A-scheme. If X can be covered by affine open sets $\operatorname{Spec} B_i$ where each B_i is a *finitely generated* A-algebra, we say that X is **locally of finite type over** A, or that it is a **locally finite type** A**-scheme**. (This is admittedly cumbersome terminology; it will make more sense later, once we know about morphisms in §7.3.12.) If furthermore X is quasicompact, X is (of) **finite type over** A, or a **finite type** A**-scheme**. Note that a scheme locally of finite type over k or $\mathbb{Z}$ (or indeed any Noetherian ring) is locally Noetherian, and similarly a scheme of finite type over any Noetherian ring is Noetherian. As our key "geometric" examples: (i) $\operatorname{Spec} \mathbb{C}[x_1, \dots, x_n]/I$ is a finite type $\mathbb{C}$-scheme; and (ii) $\mathbb{P}^n_{\mathbb{C}}$ is a finite type $\mathbb{C}$-scheme. (The field $\mathbb{C}$ may be replaced by an arbitrary ring A.)

5.3.7. *Varieties.* We now make a connection to the classical language of varieties. An affine scheme that is reduced and of finite type over k is said to be an **affine variety (over** k**)**, or an **affine** k**-variety**. A reduced (quasi-)projective k-scheme is a **(quasi-)projective variety (over** k**)**, or a **(quasi-)projective** k**-variety**. Note that quasiprojective k-schemes are of finite type — do you see why? (Warning: in the literature, it is sometimes also assumed in the definition of variety that the scheme is irreducible, or that k is algebraically closed.)

5.3.D. EXERCISE.
(a) Show that $\operatorname{Spec} k[x_1, \dots, x_n]/I$ is an affine k-variety if and only if $I \subset k[x_1, \dots, x_n]$

is a radical ideal.
(b) Suppose $I \subset k[x_0, \dots, x_n]$ is a radical graded ideal. Show that $\operatorname{Proj} k[x_0, \dots, x_n]/I$ is a projective k-variety. (Caution: The example of $I = (x_0^2, x_0x_1)$ shows that $\operatorname{Proj} k[x_0, \dots, x_n]/I$ can be a projective k-variety without I being radical.)

We will not define varieties in general until §10.1.7; we will need the notion of separatedness first, to exclude abominations like the line with the doubled origin (Example 4.4.5). But many of the statements we will make in this section about affine k-varieties will automatically apply more generally to k-varieties.

5.3.E. EXERCISE. Show that a point of a locally finite type k-scheme is a closed point if and only if the residue field of the stalk of the structure sheaf at that point is a finite extension of k. Show that the closed points are dense on such a scheme (even though it needn't be quasicompact, cf. Exercise 5.1.E). Hint: §3.6.9. (Warning: closed points need not be dense even on quite reasonable schemes, see Exercise 3.6.J(b).)

5.3.F. ⋆⋆ EXERCISE (ANALYTIFICATION OF COMPLEX VARIETIES). (Warning: Any discussion of analytification will be only for readers who are familiar with the notion of complex analytic varieties, or willing to develop it on their own in parallel with our development of schemes.) Suppose X is a reduced, finite type $\mathbb{C}$-scheme. Define the corresponding complex analytic prevariety X_{an}. (The definition of an analytic prevariety is the same as the definition of a variety without the Hausdorff condition.) Caution: your definition should not depend on a choice of an affine cover of X. (Hint: First explain how to analytify reduced finite type affine $\mathbb{C}$-schemes. Then glue.) Give a bijection between the closed points of X and the points of X_{an}, using the weak Nullstellensatz 3.2.4. (In fact one may construct a continuous map of sets $X_{an} \to X$ generalizing Exercise 3.2.I.) In Exercise 6.3.K, we will see that analytification can be made into a functor. As mentioned there, two nonisomorphic complex varieties can have isomorphic analytifications, but not if they are compact.

5.3.8. *Definition.* The **degree** of a closed point p of a locally finite type k-scheme is the degree of the field extension $\kappa(p)/k$. For example, in $\mathbb{A}_k^1 = \operatorname{Spec} k[t]$, the point $[(p(t))]$ ($p(t) \in k[t]$ irreducible) is $\deg p(t)$. If k is algebraically closed, the degree of every closed point is 1.

5.3.9. *Proof of Proposition 5.3.3.* We divide each part into (i) and (ii) following the statement of the Affine Communication Lemma 5.3.2. We leave (a) for practice for you (Exercise 5.3.H) after you have read the proof of (b).

(b) (i) If $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$ is a strictly increasing chain of ideals of A_f, then we can verify that $J_1 \subsetneq J_2 \subsetneq J_3 \subsetneq \cdots$ is a strictly increasing chain of ideals of A, where

$$J_j = \{r \in A \ : \ r \in I_j\}$$

where $r \in I_j$ means "the image of r in A_f lies in I_j". (We think of this as $I_j \cap A$, except in general A needn't inject into A_{f_i}.) Clearly J_j is an ideal of A. If $x/f^n \in I_{j+1} \setminus I_j$ where $x \in A$, then $x \in J_{j+1}$, and $x \notin J_j$ (or else $x(1/f)^n \in I_j$ as well).

(ii) Suppose $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots$ is a strictly increasing chain of ideals of A. Then for each $1 \leq i \leq n$,

$$I_{i,1} \subset I_{i,2} \subset I_{i,3} \subset \cdots$$

is an increasing chain of ideals in A_{f_i}, where $I_{i,j} = I_j \otimes_A A_{f_i}$. It remains to show that for each j, $I_{i,j} \subsetneq I_{i,j+1}$ for some i; the result will then follow.

5.3.G. EXERCISE. Finish part (b). (Hint: $A \hookrightarrow \prod A_{f_i}$ by (4.1.3.1).)

5.3.H. EXERCISE. Prove (a).

(c) (i) is clear: if A is generated over B by $r_1, \dots, r_n$, then A_f is generated over B by $r_1, \dots, r_n, 1/f$.

(ii) Here is the idea. As the f_i generate A, we can write $1 = \sum c_i f_i$ for $c_i \in A$. We have generators of A_{f_i}: $r_{ij}/f_i^{k_j}$, where $r_{ij} \in A$. I claim that $\{f_i\}_i \cup \{c_i\} \cup \{r_{ij}\}_{ij}$ generate A as a B-algebra. Here is why. Suppose you have any $r \in A$. Then in A_{f_i}, we can write r as some polynomial in the r_{ij}'s and f_i, divided by some huge power of f_i. So "in each A_{f_i}, we have described r in the desired way", except for this annoying denominator. Now use a partition of unity type argument as in the proof of Theorem 4.1.2 to combine all of these into a single expression, killing the denominator. Show that the resulting expression you build still agrees with r in each of the A_{f_i}. Thus it is indeed r (by the identity axiom for the structure sheaf).

5.3.I. EXERCISE. Make this argument precise.

This concludes the proof of Proposition 5.3.3. $\square$

5.3.J. EXERCISE.
(a) If X is a quasiprojective A-scheme (Definition 4.5.9), show that X is of finite type over A. If A is furthermore assumed to be Noetherian, show that X is a Noetherian scheme, and hence has a finite number of irreducible components.
(b) Suppose U is an open subscheme of a projective A-scheme. Show that U is locally of finite type over A. If A is Noetherian, show that U is quasicompact, and hence quasiprojective over A, and hence by (a) of finite type over A. Show this need not be true if A is not Noetherian. Better: give an example of an open subscheme of a projective A-scheme that is not quasicompact, necessarily for some non-Noetherian A. (Hint: Silly example 4.5.11.)

5.4 Normality and factoriality

5.4.1. Normality.

We can now define a property of schemes that says that they are "not too far from smooth", called *normality*, which will come in very handy. We will see later that "locally Noetherian normal schemes satisfy Hartogs's Lemma" (Algebraic Hartogs's Lemma 11.3.10 for Noetherian normal schemes): functions defined away from a set of codimension 2 extend over that set. (We saw a first glimpse of this in §4.4.2.) As a consequence, rational functions that have no poles (certain sets of codimension one where the function isn't defined) are defined everywhere. We need definitions of dimension and poles to make this precise. See §12.8.7 and §26.3.5 for the fact that "smoothness" (really, "regularity") implies normality.

Recall that an integral domain A is **integrally closed** if the only zeros in $K(A)$ to any monic polynomial in $A[x]$ must lie in A itself. The basic example is $\mathbb{Z}$ (see

Exercise 5.4.F for a reason). We say a scheme X is **normal** if all of its stalks $\mathscr{O}_{X,p}$ are normal, i.e., are integral domains, and integrally closed in their fraction fields. As reducedness is a stalk-local property (Exercise 5.2.A), normal schemes are reduced.

5.4.A. EXERCISE. Show that integrally closed domains behave well under localization: if A is an integrally closed domain, and S is a multiplicative subset not containing 0, show that $S^{-1}A$ is an integrally closed domain. (Hint: assume that $x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$ where $a_i \in S^{-1}A$ has a root in the fraction field. Turn this into another equation in $A[x]$ that also has a root in the fraction field.)

It is no fun checking normality at every single point of a scheme. Thanks to this exercise, we know that if A is an integrally closed domain, then $\operatorname{Spec} A$ is normal. Also, for quasicompact schemes, normality can be checked at closed points, thanks to this exercise, and the fact that for such schemes, any point is a generization of a closed point (see Exercise 5.1.E).

It is not true that normal schemes are integral. For example, the disjoint union of two normal schemes is normal. Thus $\operatorname{Spec} k \coprod \operatorname{Spec} k \cong \operatorname{Spec}(k \times k) \cong \operatorname{Spec} k[x]/(x(x-1))$ is normal, but its ring of global sections is not an integral domain.

5.4.B. UNIMPORTANT EXERCISE. Show that a Noetherian scheme is normal if and only if it is the finite disjoint union of integral Noetherian normal schemes. (Hint: Exercise 5.3.C.)

We are close to proving a useful result in commutative algebra, so we may as well go all the way.

5.4.2. Proposition. — *If A is an integral domain, then the following are equivalent.*

(i) A *is integrally closed.*
(ii) $A_{\mathfrak{p}}$ *is integrally closed for all prime ideals* $\mathfrak{p} \subset A$.
(iii) $A_{\mathfrak{m}}$ *is integrally closed for all maximal ideals* $\mathfrak{m} \subset A$.

Proof. Exercise 5.4.A shows that integral closure is preserved by localization, so (i) implies (ii). Clearly (ii) implies (iii).

It remains to show that (iii) implies (i). This argument involves a pretty construction that we will use again. Suppose A is not integrally closed. We show that there is some $\mathfrak{m}$ such that $A_{\mathfrak{m}}$ is also not integrally closed. Suppose

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0 \tag{5.4.2.1}$$

(with $a_i \in A$) has a solution s in $K(A) \setminus A$. Let I be the **ideal of denominators of** s:

$$I := \{r \in A \ : \ rs \in A\}.$$

(Note that I is clearly an ideal of A.) Now $I \neq A$, as $1 \notin I$. Thus there is some maximal ideal $\mathfrak{m}$ containing I. Then $s \notin A_{\mathfrak{m}}$, so equation (5.4.2.1) in $A_{\mathfrak{m}}[x]$ shows that $A_{\mathfrak{m}}$ is not integrally closed as well, as desired. □

5.4.C. UNIMPORTANT EXERCISE. If A is an integral domain, show that $A = \cap A_{\mathfrak{m}}$, where the intersection runs over all maximal ideals of A. (We won't use this exercise, but it gives good practice with the ideal of denominators.)

5.4.D. UNIMPORTANT EXERCISE RELATING TO THE IDEAL OF DENOMINATORS. One might naively hope from experience with unique factorization domains that the ideal of denominators is principal. This is not true. As a counterexample, consider our new friend $A = k[w, x, y, z]/(wz - xy)$ (which we first met in Example 4.4.12, and which we will later recognize as the cone over the quadric surface), and $w/y = x/z \in K(A)$. Show that the ideal of denominators of this element of $K(A)$ is (y, z).

We will see that the I in the above exercise is not principal (Exercise 12.1.D — you may be able to show it directly, using the fact that I is a graded ideal of a graded ring). But we will also see that in good situations (Noetherian, normal), the ideal of denominators is "pure codimension 1" — this is the content of Algebraic Hartogs's Lemma 11.3.10. In its proof, §11.3.11, we give a geometric interpretation of the ideal of denominators.

5.4.3. Factoriality.

We define a notion which implies normality.

5.4.4. Definition. If all the stalks of a scheme X are unique factorization domains, we say that X is **factorial**. (Unimportant remark: This is sometimes called *locally factorial*, which may falsely suggest that this notion is affine-local, which it isn't, see Exercise 5.4.N. Another unimportant remark: the locus of points on an affine variety over an algebraically closed field that are factorial is an open subset, [**BGS**, p. 1].)

5.4.E. EXERCISE. Show that any nonzero localization of a unique factorization domain is a unique factorization domain.

5.4.5. Thus if A is a unique factorization domain, then $\operatorname{Spec} A$ is factorial. The converse need not hold, see Exercise 5.4.N. In fact, we will see that elliptic curves are factorial, yet *no* affine open set is the Spec of a unique factorization domain, §19.11.1. Hence one can show factoriality by finding an appropriate affine cover, but there need not *be* such a cover of a factorial scheme.

5.4.6. *Remark: How to check if a ring is a unique factorization domain.* There are very few means of checking that a Noetherian integral domain is a unique factorization domain. Some useful ones are: (0) elementary means: rings with a euclidean algorithm such as $\mathbb{Z}$, $k[t]$, and $\mathbb{Z}[i]$; polynomial rings over a unique factorization domain, by Gauss's Lemma (see e.g. [**Lan**, IV.2.3]). (1) Exercise 5.4.E, that the localization of a unique factorization domain is also a unique factorization domain. (2) height 1 primes are principal (Proposition 11.3.5). (3) normal and $\mathrm{Cl} = 0$ (Exercise 14.2.T). (4) Nagata's Lemma (Exercise 14.2.U). (Caution: even if A is a unique factorization domain, $A[[x]]$ need not be; see [**Mat2**, p. 165].)

5.4.7. *Factoriality implies normality.* One of the reasons we like factoriality is that it implies normality.

5.4.F. IMPORTANT EXERCISE. Show that unique factorization domains are integrally closed. Hence factorial schemes are normal, and if A is a unique factorization domain, then $\operatorname{Spec} A$ is normal. (However, rings can be integrally closed without being unique factorization domains, as we will see in Exercise 5.4.L. Another example is given without proof in Exercise 5.4.N; in that example, Spec of the ring is factorial. A variation on Exercise 5.4.L will show that schemes can be normal without being factorial, see Exercise 12.1.E.)

5.4.8. *Examples.*

5.4.G. EASY EXERCISE. Show that the following schemes are normal: $\mathbb{A}^n_k$, $\mathbb{P}^n_k$, $\operatorname{Spec}\mathbb{Z}$. (As usual, k is a field. Although it is true that if A is integrally closed then $A[x]$ is as well — see [**Bo**, Ch. 5, §1, no. 3, Cor. 2] or [**E**, Ex. 4.18] — this is not an easy fact, so do not use it here.)

5.4.H. HANDY EXERCISE (YIELDING MANY ENLIGHTENING EXAMPLES LATER). Suppose A is a unique factorization domain with 2 invertible, and $z^2 - f$ is irreducible in $A[z]$.
(a) Show that if $f \in A$ has no repeated prime factors, then $\operatorname{Spec} A[z]/(z^2 - f)$ is normal. Hint: $B := A[z]/(z^2 - f)$ is an integral domain, as $(z^2 - f)$ is prime in $A[z]$. Suppose we have monic $F(T) \in B[T]$ so that $F(T) = 0$ has a root α in $K(B) \setminus K(A)$. Then by replacing $F(T)$ by $\overline{F}(T)F(T)$, we can assume $F(T) \in A[T]$. Also, $\alpha = g + hz$ where $g, h \in K(A)$. Now α is the root of $Q(T) = 0$ for monic $Q(T) = T^2 - 2gT + (g^2 - h^2 f) \in K(A)[T]$, so we can factor $F(T) = P(T)Q(T)$ in $K(A)[T]$. By Gauss's lemma, $2g, g^2 - h^2 f \in A$. Say $g = r/2$, $h = s/t$ (s and t have no common factors, $r, s, t \in A$). Then $g^2 - h^2 f = (r^2t^2 - 4s^2 f)/4t^2$. Then t is invertible.
(b) Show that if $f \in A$ has repeated prime factors, then $\operatorname{Spec} A[z]/(z^2 - f)$ is *not* normal.

5.4.I. EXERCISE. Show that the following schemes are normal:

(a) $\operatorname{Spec}\mathbb{Z}[x]/(x^2 - n)$ where n is a square-free integer congruent to 3 modulo 4. Caution: the hypotheses of Exercise 5.4.H do not apply, so you will have to do this directly. (Your argument may also show the result when 3 is replaced by 2. A similar argument shows that $\mathbb{Z}[(1+\sqrt{n})/2]$ is integrally closed if $n \equiv 1 \pmod 4$ is square-free.)
(b) $\operatorname{Spec} k[x_1, \dots, x_n]/(x_1^2 + x_2^2 + \cdots + x_m^2)$ where $\operatorname{char} k \neq 2$, $n \geq m \geq 3$.
(c) $\operatorname{Spec} k[w, x, y, z]/(wz - xy)$ where $\operatorname{char} k \neq 2$. This is our cone over a quadric surface example from Example 4.4.12 and Exercise 5.4.D. Hint: Exercise 5.4.J may help. (The result also holds for $\operatorname{char} k = 2$, but don't worry about this.)

5.4.J. EXERCISE (DIAGONALIZING QUADRICS). Suppose k is an algebraically closed field of characteristic not 2.
(a) Show that any quadratic form in n variables can be "diagonalized" by changing coordinates to be a sum of at most n squares (e.g. $uw - v^2 = ((u + w)/2)^2 + (i(u - w)/2)^2 + (iv)^2$), *where the linear forms appearing in the squares are linearly independent*. (Hint: use induction on the number of variables, by "completing the square" at each step.)
(b) Show that the number of squares appearing depends only on the quadric. For

example, $x^2 + y^2 + z^2$ cannot be written as a sum of two squares. (Possible approach: given a basis $x_1, \dots, x_n$ of the linear forms, write the quadratic form as

$$\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

where M is a symmetric matrix. Determine how M transforms under a change of basis, and show that the rank of M is independent of the choice of basis.)

The **rank** of the quadratic form is the number of ("linearly independent") squares needed. If the number of squares equals the number of variables, the quadratic form is said to be **full rank** or (of) **maximal rank**.

5.4.K. EASY EXERCISE (RINGS CAN BE INTEGRALLY CLOSED BUT NOT UNIQUE FACTORIZATION DOMAINS, ARITHMETIC VERSION). Show that $\mathbb{Z}[\sqrt{-5}]$ is normal but not a unique factorization domain. (Hints: Exercise 5.4.I(a) and $2 \times 3 = (1+\sqrt{-5})(1-\sqrt{-5})$.)

5.4.L. EASY EXERCISE (RINGS CAN BE INTEGRALLY CLOSED BUT NOT UNIQUE FACTORIZATION DOMAINS, GEOMETRIC VERSION). Suppose $\operatorname{char} k \neq 2$. Let $A = k[w,x,y,z]/(wz - xy)$, so $\operatorname{Spec} A$ is the cone over the smooth quadric surface (cf. Exercises 4.4.12 and 5.4.D).
(a) Show that A is integrally closed. (Hint: Exercises 5.4.I(c) and 5.4.J.)
(b) Show that A is not a unique factorization domain. (Clearly $wz = xy$. But why are w, x, y, and z irreducible? Hint: A is a graded integral domain. Show that if a homogeneous element factors, the factors must be homogeneous.)

The previous two exercises look similar, but there is a difference. Thus the cone over the quadric surface is normal (by Exercise 5.4.L) but not factorial; see Exercise 12.1.E. On the other hand, $\operatorname{Spec} \mathbb{Z}[\sqrt{-5}]$ *is* factorial — all of its stalks are unique factorization domains. (You will later be able to show this by showing that $\mathbb{Z}[\sqrt{-5}]$ is a Dedekind domain, §12.5.14, whose stalks are necessarily unique factorization domains by Theorem 12.5.8(f).)

5.4.M. EXERCISE. Suppose A is a k-algebra, and l/k is a finite field extension. (Most likely your proof will not use finiteness; this hypothesis is included to avoid prevent distraction by infinite-dimensional vector spaces.) Show that if $A \otimes_k l$ is a normal integral domain, then A is a normal integral domain as well. (Although we won't need this, a version of the converse is true if l/k is separable, [**Gr-EGA**, IV$_2$.6.14.2].) Hint: fix a k-basis for l, $b_1 = 1, \dots, b_d$. Explain why $1 \otimes b_1, \dots, 1 \otimes b_d$ forms a free A-basis for $A \otimes_k l$. Explain why we have injections

$$\begin{tikzcd} A \arrow[r] \arrow[d] & K(A) \arrow[d] \\ A \otimes_k l \arrow[r] & K(A) \otimes_k l. \end{tikzcd}$$

Show that $K(A) \otimes_k l = K(A \otimes_k l)$. (Idea: $A \otimes_k l \subset K(A) \otimes_k l \subset K(A \otimes_k l)$. Why is $K(A) \otimes_k l$ a field?) Show that $(A \otimes_k l) \cap K(A) = A$. Now assume $P(T) \in A[T]$ is monic and has a root $\alpha \in K(A)$, and proceed from there.

5.4.N. EXERCISE (UFD-NESS IS NOT AFFINE-LOCAL). Let $A = (\mathbb{Q}[x,y]_{x^2+y^2})_0$ denote the homogeneous degree 0 part of the ring $\mathbb{Q}[x,y]_{x^2+y^2}$. In other words, it consists of quotients $f(x,y)/(x^2+y^2)^n$, where f has pure degree $2n$. Show that the distinguished open sets $D(\frac{x^2}{x^2+y^2})$ and $D(\frac{y^2}{x^2+y^2})$ cover $\operatorname{Spec} A$. (Hint: the sum of those two fractions is 1.) Show that $A_{\frac{x^2}{x^2+y^2}}$ and $A_{\frac{y^2}{x^2+y^2}}$ are unique factorization domains. (Hint for the first: show that both rings are isomorphic to $\mathbb{Q}[t]_{t^2+1}$; this is a localization of the unique factorization domain $\mathbb{Q}[t]$.) Finally, show that A is not a unique factorization domain. Possible hint:

$$\left(\frac{xy}{x^2+y^2}\right)^2 = \left(\frac{x^2}{x^2+y^2}\right)\left(\frac{y^2}{x^2+y^2}\right).$$

(This is generalized in Exercise 14.2.L . It is also related to Exercise 14.2.Q.)

Number theorists may prefer the example of Exercise 5.4.K: $\mathbb{Z}[\sqrt{-5}]$ is not a unique factorization domain, but it turns out that you can cover it with two affine open subsets $D(2)$ and $D(3)$, each corresponding to unique factorization domains. (For number theorists: to show that $\mathbb{Z}[\sqrt{-5}]_2$ and $\mathbb{Z}[\sqrt{-5}]_3$ are unique factorization domains, first show that the class group of $\mathbb{Z}[\sqrt{-5}]$ is $\mathbb{Z}/2$ using the geometry of numbers, as in [**Ar4**, Ch. 11, Thm. 7.9]. Then show that the ideals $(1+\sqrt{-5}, 2)$ and $(1+\sqrt{-5}, 3)$ are not principal, using the usual norm in $\mathbb{C}$.) The ring $\mathbb{Z}[\sqrt{-5}]$ is an example of a Dedekind domain, as we will discuss in §12.5.14.

5.4.9. *Remark.* For an example of k-algebra A that is not a unique factorization domain, but becomes one after a certain field extension, see Exercise 14.2.N.

5.5 Where functions are supported: Associated points of schemes

The associated points of a scheme are the few crucial points of the scheme that capture essential information about its (sheaf of) functions. There are several quite different ways of describing them, most of which are quite algebraic. We will take a nonstandard approach, beginning with geometric motivation. Because they involve both nilpotents and generic points — two concepts not part of your prior geometric intuition — it can take some time to make them "geometric" in your head. We will first meet them in a motivating example in two ways. We will then discuss their key properties. Finally, we give proper (algebraic) definitions and proofs. As is almost always the case in mathematics, it is much more important to remember the properties than it is to remember their proofs.

There are other approaches to associated points. Most notably, the algebraically most central view is via a vitally important algebraic construction, primary decomposition, mentioned only briefly in Aside 5.5.13.

5.5.1. *Associated points as "fuzz attractors".* Recall Figure 4.4, our "fuzzy" picture of the nonreduced scheme $\operatorname{Spec} k[x,y]/(y^2, xy)$. When this picture was introduced, we mentioned that the "fuzz" at the origin indicated that the nonreduced behavior was concentrated there. This was justified in Exercise 5.2.C: the origin is the only point where the stalk of the structure sheaf is nonreduced. Thus the different levels of reducedness are concentrated along two irreducible closed subsets — the origin, and the entire x-axis. Since irreducible closed subsets are in bijection with points, we may as well say that the two key points with respect to "levels of nonreducedness" were the generic point $[(y)]$, and the origin $[(x, y)]$. These will be the associated points of this scheme.

5.5.2. *Better: associated points as generic points of irreducible components of the support of sections.*

We now give a seemingly unrelated exercise about the same scheme. Recall that the support of a function on a scheme (Definition 2.4.2) is a *closed* subset.

5.5.A. EXERCISE. Suppose f is a function on $\operatorname{Spec} k[x, y]/(y^2, xy)$ (i.e., $f \in k[x, y]/(y^2, xy)$). Show that $\operatorname{Supp} f$ is either the empty set, or the origin, or the entire space.

The fact that the same closed subsets arise in two different ways is no coincidence — their generic points are the associated points of $\operatorname{Spec} k[x, y]/(y^2, xy)$.

We discuss associated points first in the affine case $\operatorname{Spec} A$. We assume that A is Noetherian, and we take this as a standing assumption when discussing associated points. More generally, we will discuss associated points of M where M is a finitely generated A module (and A is Noetherian). When speaking of rings rather than schemes, we speak of *associated primes* rather than *associated points*. Associated primes and associated points can be defined more generally, and we discuss one easy case (the integral case) in Exercise 5.5.Q.

We now state three essential properties, to be justified later. The first is the most important.

(A) *The associated primes/points of M are precisely the generic points of irreducible components of the support of some element of M (on $\operatorname{Spec} A$).*

For example, by Exercise 5.5.A, $\operatorname{Spec} k[x, y]/(y^2, xy)$ has two associated points. As another example:

5.5.B. EXERCISE (ASSUMING **(A)**). Suppose A is an integral domain. Show that the generic point is the only associated point of $\operatorname{Spec} A$.

(*Important note:* Exercises 5.5.B–5.5.H require you to work directly from some axioms, not from our later definitions. If this troubles you, feel free to work through the definitions, and use the later exercises rather than the geometric axioms **(A)–(C)** to solve these problems. But you may be surprised at how short the arguments actually are, assuming the geometric axioms.)

We could take **(A)** as the definition, although in our rigorous development below, we will take a different (but logically equivalent) starting point. (Unimportant aside: if A is a ring that is not necessarily Noetherian, then **(A)** is the definition of a *weakly associated prime*, see [**Stacks**, tag 0547].)

The next property makes **(A)** more striking.

(B) M *has a finite number of associated primes/points.*

In other words, there are only a *finite* number of irreducible closed subsets of $\operatorname{Spec} A$, such that the only possible supports of functions of $\operatorname{Spec} A$ are unions of these. You may find this unexpected, although the examples above may have prepared you for it. You should interpret this as another example of Noetherianness forcing some sort of finiteness. (For example, we will see that this generalizes "finiteness of irreducible components", cf. Proposition 3.6.15. And to get an idea of what can go wrong without Noetherian hypotheses, you can ponder the scheme of Remark 5.2.2.) This gives some meaning to the statement that their generic points are the few crucial points of the scheme.

We will see (in Exercise 5.5.O) that we can completely describe which subsets of $\operatorname{Spec} A$ are the support of an element of M: precisely those subsets which are the closure of a subset of the associated points.

5.5.3. We immediately see from **(A)** that *if* $M = A$, *the generic points of the irreducible components of* $\operatorname{Spec} A$ *are associated points* of $M = A$, by considering the function 1. The other associated points of $\operatorname{Spec} A$ are called **embedded points**. Thus in the case of $\operatorname{Spec} k[x, y]/(y^2, xy)$ (Figure 4.4), the origin is the only embedded point (by Exercise 5.5.A).

5.5.C. EXERCISE (ASSUMING **(A)**). Show that if A is reduced, $\operatorname{Spec} A$ has no embedded points. Hints: (i) first deal with the case where A is integral, i.e., where $\operatorname{Spec} A$ is irreducible. (ii) Then deal with the general case. If f is a nonzero function on a reduced affine scheme, show that $\operatorname{Supp} f = \overline{D(f)}$: the support is the closure of the locus where f doesn't vanish. Show that $\overline{D(f)}$ is the union of the irreducible components meeting $D(f)$, using (i).

Furthermore, *the natural map*

$$M \to \prod_{\text{associated } \mathfrak{p}} M_{\mathfrak{p}} \tag{5.5.3.1}$$

is an injection. (This is an important property. Once again, the associated points are "where all the action happens".) We show this by showing that the kernel is zero. Suppose a function f has a germ of zero at each associated point, so its support contains no associated points. It is supported on a closed subset, which by **(A)** must be the union of closures of associated points. Thus it must be supported nowhere, and thus be the zero function.

5.5.4. *Side Remark.* An open subscheme U of a scheme X is said to be **schematically dense** if any function on any open set V is 0 if it restricts to 0 on $U \cap V$. (We won't use this phrase.) If X is locally Noetherian, then you can use the injection (5.5.3.1) (with $M = A$) to show that an open subscheme $U \subset X$ is schematically dense if and only if it contains all the associated points of X.

5.5.D. EXERCISE (ASSUMING **(A)**). Suppose $m \in M$. Show that $\operatorname{Supp} m$ is the closure of those associated points of M where m has nonzero germ. (Hint: $\operatorname{Supp} m$ is a closed set containing the points described, and thus their closure. Why does it contain no other points?)

5.5.E. EXERCISE (ASSUMING **(A)** AND **(B)**). Show that the locus on $\operatorname{Spec} A$ of points $[\mathfrak{p}]$ where $\mathscr{O}_{\operatorname{Spec} A, [\mathfrak{p}]} = A_{\mathfrak{p}}$ is nonreduced is the closure of those associated points of $\operatorname{Spec} A$ whose stalks are nonreduced. (Hint: why do points in the closure of these associated points all have nonreduced stalks? Why can't any other point have a nonreduced stalk?)

5.5.5. *Remark.* Exercise 5.5.E partially explains the link between associated points and fuzzy pictures. (Primary decomposition, see Aside 5.5.13, gives a more explicit connection, but we won't discuss it properly.) Note for future reference that once we establish these properties (later in this section), we will have shown that if Y is a locally Noetherian scheme, the "reduced locus" of Y is an open subset of Y. This fulfills a promise made in Remark 5.2.2.

(C) *An element f of A is a zerodivisor of M (i.e., there exists $m \neq 0$ with $fm = 0$) if and only if it vanishes at some associated point of M (i.e., is contained in some associated prime of M).*

One direction is clear from the previous properties. (Do you see which?)

The next property allows us to globalize the construction of associated points to arbitrary (locally Noetherian) schemes.

5.5.F. IMPORTANT EXERCISE (ASSUMING **(A)**). Show that *the definition in* **(A)** *of associated primes/points behaves well with respect to localizing:* if S is a multiplicative subset of A, then the associated primes/points of $S^{-1}M$ are precisely those associated primes/points of M that lie in $\operatorname{Spec} S^{-1}A$, i.e., associated primes of M that do not meet S.

Thus the associated primes/points can be "determined locally". For example, associated points of A can be checked by looking at stalks of the structure sheaf (the notion is "stalk-local"). As another example, the associated primes of M may be determined by working on a distinguished open cover of $\operatorname{Spec} A$. Thanks to Exercise 5.5.F, we can (and do) define the **associated points** of a locally Noetherian scheme X to be those points $p \in X$ such that, on any affine open set $\operatorname{Spec} A$ containing p, p corresponds to an associated prime of A. This notion is independent of choice of affine neighborhood $\operatorname{Spec} A$: if p has two affine open neighborhoods $\operatorname{Spec} A$ and $\operatorname{Spec} B$ (say corresponding to primes $\mathfrak{p} \subset A$ and $\mathfrak{q} \subset B$ respectively), then p corresponds to an associated prime of A if and only if it corresponds to an associated prime of $A_{\mathfrak{p}} = \mathscr{O}_{X,p} = B_{\mathfrak{q}}$ if and only if it corresponds to an associated prime of B, by Exercise 5.5.F.

(Here we are "globalizing" only the special case $M = A$. Once we define quasicoherent sheaves, we will be able to globalize the case of a general M, see §13.6.5.)

By combining the above properties, we immediately have a number of facts, including the following. (i) A Noetherian scheme has finitely many associated points. (ii) Each of the irreducible components of the support of any function on a locally Noetherian scheme is the union of the closures of some subset of the associated points. (iii) The generic points of the irreducible components of a locally Noetherian scheme are associated points. (The remaining associated points are still called **embedded points**.) (iv) A reduced locally Noetherian scheme has no embedded points. (v) The nonreduced locus of a locally Noetherian scheme (the locus of points $p \in X$ where $\mathscr{O}_{X,p}$ is nonreduced) is the closure of those associated points that have nonreduced stalk.

Furthermore, recall that one nice property of integral schemes X (such as irreducible affine varieties) not shared by all schemes is that for any nonempty open $U \subset X$, the natural map $\Gamma(U, \mathscr{O}_X) \to K(X)$ is an inclusion (Exercise 5.2.I). Thus all sections over any nonempty open set, and elements of all stalks, can be thought of as lying in a single field $K(X)$, which is the stalk at the generic point. Associated points allow us to generalize this idea.

5.5.G. EXERCISE. Assuming the above properties of associated points, show that if X is a locally Noetherian scheme, then for any open subset $U \subset X$, the natural map

$$\Gamma(U, \mathscr{O}_X) \to \prod_{\text{associated } p \text{ in } U} \mathscr{O}_{X,p} \qquad (5.5.5.1)$$

is an injection.

We can use these properties to refine our ability to visualize schemes in a way that captures precise mathematical information. As a first check, you should be able to understand Figure 5.2. As a second, you should be able to do the following exercise.

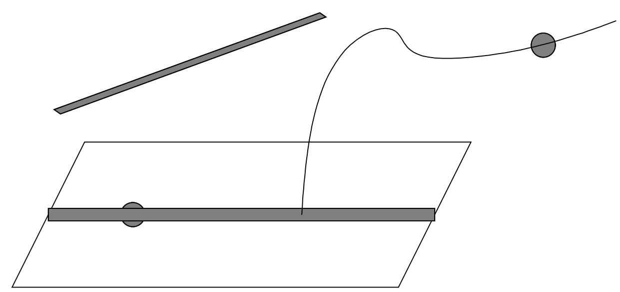


FIGURE 5.2. This scheme has 6 associated points, of which 3 are embedded points. A function is a zerodivisor if it vanishes at any of these six points.

5.5.H. EXERCISE (PRACTICE WITH FUZZY PICTURES). Assume the properties **(A)–(C)** of associated points. Suppose $X = \operatorname{Spec} \mathbb{C}[x, y]/I$, and that the associated points of X are $[(y - x^2)]$, $[(x - 1, y - 1)]$, and $[(x - 2, y - 2)]$.
(a) Sketch X as a subset of $\mathbb{A}^2_{\mathbb{C}} = \operatorname{Spec} \mathbb{C}[x, y]$, including fuzz.
(b) Do you have enough information to know if X is reduced?
(c) Do you have enough information to know if $x + y - 2$ is a zerodivisor? How about $x + y - 3$? How about $y - x^2$? (Exercise 5.5.R will verify that such an X actually exists.)

The following exercise shows that "hypersurfaces have no embedded points". (Of course, thanks to Exercise 5.5.C, this is interesting only when the hypersurface is nonreduced.)

5.5.I. EXERCISE. Assume the properties **(A)–(C)** of associated points. If $f \in k[x_1, \ldots, x_n]$ is nonzero, show that $A := k[x_1, \ldots, x_n]/(f)$ has no embedded points. Hint: suppose $\overline{g} \in A$ is a zerodivisor, and choose a lift $g \in k[x_1, \ldots, x_n]$ of $\overline{g}$. Show that g has a common factor with f. (We will use this exercise in §18.6.3. All you should use is that $k[x_1, \ldots, x_n]$ is a Noetherian unique factorization domain. We will generalize this in §26.2.7.)

5.5.6. *Definitions: Rational functions.* A **rational function** on a locally Noetherian scheme is an element of the image of $\Gamma(U, \mathscr{O}_U)$ in (5.5.5.1) for some U containing all the associated points. Equivalently, the set of rational functions is the colimit of $\mathscr{O}_X(U)$ over all open sets containing the associated points. Or if you prefer, a rational function is a function defined on an open set containing all associated points, i.e., an ordered pair (U, f), where U is an open set containing all associated points, and $f \in \Gamma(U, \mathscr{O}_X)$. Two such data (U, f) and (U', f') define the same open rational function if and only if the restrictions of f and f' to $U \cap U'$ are the same. If X is reduced, this is the same as requiring that they are defined on an open set of each of the irreducible components.

For example, on $\operatorname{Spec} k[x, y]/(y^2, xy)$ (Figure 4.4), $\frac{x-2}{(x-1)(x-3)}$ is a rational function, but $\frac{x-2}{x(x-1)}$ is not.

A rational function has a maximal **domain of definition**, because any two actual functions on an open set (i.e., sections of the structure sheaf over that open set) that agree as "rational functions" (i.e., on small enough open sets containing associated points) must be the same function, by the injectivity of (5.5.5.1). We say that a rational function f is **regular** at a point p if p is contained in this maximal domain of definition (or equivalently, if there is some open set containing p where f is defined). For example, on $\operatorname{Spec} k[x, y]/(y^2, xy)$, the rational function $\frac{x-2}{(x-1)(x-3)}$ has domain of definition consisting of everything but 1 and 3 (i.e., $[(x-1)]$ and $[(x-3)]$), and is regular away from those two points. A rational function is **regular** if it is regular at all points. (Unfortunately, "regular" is an overused word in mathematics, and in algebraic geometry in particular.)

5.5.7. The rational functions form a ring, called the **total fraction ring** or **total quotient ring** of X. If $X = \operatorname{Spec} A$ is affine, then this ring is called the **total fraction (or quotient) ring** of A. If X is integral, the total fraction ring is the function field $K(X)$ — the stalk at the generic point — so this extends our earlier Definition 5.2.H of $K(\cdot)$.

5.5.8. Definition and proofs.

We finally define associated points, and show that they have the desired properties **(A)–(C)** (and their consequences) for locally Noetherian schemes. Because the definition is a useful property to remember (on the same level as **(A)–(C)**), we dignify it with a letter. We make the definition in more generality than we will use. Suppose M is an A-module, and A is an arbitrary ring.

(D) *A prime $\mathfrak{p} \subset A$ is said to be* **associated** *to M if $\mathfrak{p}$ is the annihilator of an element m of M ($\mathfrak{p} = \{a \in A : am = 0\}$).*

5.5.9. Equivalently, $\mathfrak{p}$ is associated to M if and only if M has a submodule isomorphic to $A/\mathfrak{p}$. The set of primes associated to M is denoted $\operatorname{Ass} M$ (or $\operatorname{Ass}_A M$). Awkwardly, if I is an ideal of A, the associated primes of the module A/I are said to be the associated primes of I. This is not my fault.

5.5.10. Theorem (properties of associated primes). — *Suppose A is a Noetherian ring, and $M \neq 0$ is finitely generated.*

(a) *The set* $\operatorname{Ass} M$ *is finite (property* **(B)***) and nonempty.*
(b) *The natural map* $M \to \prod_{\mathfrak{p} \in \operatorname{Ass} M} M_{\mathfrak{p}}$ *is an injection (cf. (5.5.3.1)).*
(c) *The set of zerodivisors of* M *is* $\cup_{\mathfrak{p} \in \operatorname{Ass} M} \mathfrak{p}$ *(property* **(C)***).*
(d) *(association commutes with localization, cf. Exercise 5.5.F) If* S *is a multiplicative set, then*

$$\operatorname{Ass}_{S^{-1}A} S^{-1}M = \operatorname{Ass}_A M \cap \operatorname{Spec} S^{-1}A$$

$(= \{\mathfrak{p} \in \operatorname{Ass}_A M : \mathfrak{p} \cap S = \varnothing\})$.

We prove Theorem 5.5.10 in a series of exercises.

5.5.J. IMPORTANT EXERCISE. Suppose $M \neq 0$ is an A-module. Show that if $I \subset A$ is maximal among all proper ideals that are annihilators of elements of M, then I is prime, and hence $I \in \operatorname{Ass} M$. Thus if A is Noetherian, then $\operatorname{Ass} M$ is nonempty (part of Theorem 5.5.10(a)). (This is a good excuse to state a general philosophy: "Quite generally, proper ideals maximal with respect to some property have an uncanny tendency to be prime," [E, p. 70].)

5.5.K. EXERCISE. Suppose that M is a module over a Noetherian ring A. Show that $m = 0$ if and only if m is 0 in $M_{\mathfrak{p}}$ for each of the maximal associated primes $\mathfrak{p}$ of M. (Hint: use the previous exercise.)

This immediately implies Theorem 5.5.10(b). It also implies Theorem 5.5.10(c): Any nonzero element of $\cup_{\mathfrak{p} \in \operatorname{Ass} M} \mathfrak{p}$ is clearly a zerodivisor. Conversely, if a annihilates a nonzero element of M, then a is contained in a maximal annihilator ideal.

5.5.L. EXERCISE. If $0 \to M' \to M \to M'' \to 0$ is a short exact sequence of A-modules, show that

$$\operatorname{Ass} M' \subset \operatorname{Ass} M \subset \operatorname{Ass} M' \cup \operatorname{Ass} M''.$$

(Possible hint for the second containment: if $m \in M$ has annihilator $\mathfrak{p}$, then $Am \cong A/\mathfrak{p}$.)

5.5.M. EXERCISE.
(a) If M is a finitely generated module over Noetherian A, show that M has a filtration

$$0 = M_0 \subset M_1 \subset \cdots \subset M_n = M$$

where $M_{i+1}/M_i \cong A/\mathfrak{p}_i$ for some prime ideal $\mathfrak{p}_i$. (If the $\mathfrak{p}_i$ are all maximal, the filtration is called a *composition series*, see Definition 18.4.7.)
(b) Show that the associated primes are among the $\mathfrak{p}_i$, and thus prove Theorem 5.5.10(a).
(c) Show that for each i, $\operatorname{Supp} A/\mathfrak{p}_i$ is contained in $\operatorname{Supp} M$, or equivalently, that every $\mathfrak{p}_i$ contains an associated prime. Hint: if $\mathfrak{p}_i$ does not contain a minimal prime, then localize at $\mathfrak{p}_i$ to "make M disappear". (Caution: non-associated primes may be among the $\mathfrak{p}_i$: take $M = A = \mathbb{Z}$, and witness $0 \subset 2\mathbb{Z} \subset \mathbb{Z}$.)

5.5.N. EXERCISE. Prove Theorem 5.5.10(d) as follows.
(a) Show that

$$\operatorname{Ass}_A M \cap \operatorname{Spec} S^{-1}A \subset \operatorname{Ass}_{S^{-1}A} S^{-1}M.$$

(Hint: suppose $\mathfrak{p} \in \operatorname{Ass}_A M \cap \operatorname{Spec} S^{-1}A$, with $\mathfrak{p} = \operatorname{Ann} m$ for $m \in M$.)
(b) Suppose $\mathfrak{q} \in \operatorname{Ass}_{S^{-1}A} S^{-1}M$, which corresponds to $\mathfrak{p} \in \operatorname{Spec} A$ (i.e., $\mathfrak{q} =$

$\mathfrak{p}(S^{-1}A))$. Then $\mathfrak{q} = \operatorname{Ann}_{S^{-1}A} m$ $(m \in S^{-1}M)$, which yields a nonzero element of

$$\operatorname{Hom}_{S^{-1}A}(S^{-1}A/\mathfrak{q}, S^{-1}M).$$

Argue that this group is isomorphic to $S^{-1}\operatorname{Hom}_A(A/\mathfrak{p}, M)$ (see Exercise 1.6.G), and hence $\operatorname{Hom}_A(A/\mathfrak{p}, M) \neq 0$.

This concludes the proof of Theorem 5.5.10. The remaining important loose end is to understand associated points in terms of support.

5.5.O. EXERCISE. Show that those subsets of $\operatorname{Spec} A$ which are the support of an element of M are precisely those subsets which are the closure of a subset of the associated points. Hint: show that for any associated point $\mathfrak{p}$, there is a section supported precisely on $\overline{\mathfrak{p}}$. Remark: This can be used to solve Exercise 5.5.P, but some people prefer to do Exercise 5.5.P first, and obtain this as a consequence.

5.5.P. IMPORTANT EXERCISE. Suppose A is a Noetherian ring, and M is a finitely generated A-module. Show that associated points/primes of M satisfy property **(A)** as follows.

(a) Show that every associated point is the generic point of an irreducible component of $\operatorname{Supp} m$ for some $m \in M$. Hint: if $\mathfrak{p} \in A$ is associated, then $\mathfrak{p} = \operatorname{Ann} m$ for some $m \in M$; this is useful in Exercise 5.5.O as well.

(b) If $m \in M$, show that the support of m is the closure of those associated points at which m has nonzero germ (cf. Exercise 5.5.D, which relied on **(A)** and **(B)**). Hint: if $\mathfrak{p}$ is in the closure of such an associated point, show that m has nonzero germ at $\mathfrak{p}$. If $\mathfrak{p}$ is *not* in the closure of such an associated point, show that m is 0 in $M_{\mathfrak{p}}$ by localizing at $\mathfrak{p}$, and using Theorem 5.5.10(b) in the *localized* ring $A_{\mathfrak{p}}$ (using Theorem 5.5.10(d)).

5.5.11. Loose ends.

We can easily extend the theory of associated points of schemes to a (very special) setting without Noetherian hypotheses: integral domains, and integral schemes.

5.5.Q. EXERCISE (EASY VARIATION: ASSOCIATED POINTS OF INTEGRAL SCHEMES). Define the notion of associated points for integral domains and integral schemes. More precisely, take **(A)** as the definition, and establish **(B)** and **(C)**. (Hint: the unique associated prime of an integral domain is (0), and the unique associated point of an integral scheme is its generic point.) In particular, rational functions on an integral scheme X are precisely elements of the function field $K(X)$ (Definition 5.2.H).

Now that we have defined associated points, we can verify that there is an example of the form described in Exercise 5.5.H

5.5.R. EXERCISE. Let $I = (y - x^2)^3 \cap (x-1, y-1)^{15} \cap (x-2, y-2)$. Show that $X = \operatorname{Spec} \mathbb{C}[x,y]/I$ satisfies the hypotheses of Exercise 5.5.H. (Rhetorical question: Is there a "smaller" example? Is there a "smallest"?)

5.5.12. *A non-Noetherian remark.* By combining §5.5.3 with **(C)**, we see that if A is a Noetherian ring, then any element of any minimal prime $\mathfrak{p}$ is a zerodivisor. This is true without Noetherian hypotheses: suppose $s \in \mathfrak{p}$. Then by minimality of $\mathfrak{p}$, $\mathfrak{p}A_{\mathfrak{p}}$

is the unique prime ideal in $A_{\mathfrak{p}}$, so the element $s/1$ of $A_{\mathfrak{p}}$ is nilpotent (because it is contained in all primes of $A_{\mathfrak{p}}$, Theorem 3.2.12). Thus for some $t \in A \setminus \mathfrak{p}$, $ts^n = 0$, so s is a zerodivisor. We will use this in Exercise 11.1.G.

5.5.13. *Aside: Primary ideals.* The notion of primary ideals and primary decomposition is important, although we won't use it. (An ideal $I \subset A$ in a ring is **primary** if $I \neq A$ and if $xy \in I$ implies either $x \in I$ or $y^n \in I$ for some $n > 0$.) The associated primes of an ideal turn out to be precisely the radicals of ideals in any primary decomposition. Primary decomposition was first introduced by the world chess champion Lasker in 1905, and later axiomatized by Noether in the 1920's. See [**E**, §3.3], for example, for more on this topic.

Part III

Morphisms

CHAPTER 6

Morphisms of schemes

6.1 Introduction

We now describe the morphisms between schemes. We will define some easy-to-state properties of morphisms, but leave more subtle properties for later.

Recall that a scheme is (i) a set, (ii) with a topology, (iii) and a (structure) sheaf of rings, and that it is sometimes helpful to think of the definition as having three steps. In the same way, the notion of morphism of schemes $X \to Y$ may be defined (i) as a map of sets, (ii) that is continuous, and (iii) with some further information involving the sheaves of functions. In the case of affine schemes, we have already seen the map as sets (§3.2.9) and later saw that this map is continuous (Exercise 3.4.H).

Here are two motivations for how morphisms should behave. The first is algebraic, and the second is geometric.

6.1.1. *Algebraic motivation.* We will want morphisms of affine schemes $\operatorname{Spec} A \to \operatorname{Spec} B$ to be precisely the ring maps $B \to A$. We have already seen that ring maps $B \to A$ induce maps of topological spaces in the opposite direction (Exercise 3.4.H); the main new ingredient will be to see how to add the structure sheaf of functions into the mix. Then a morphism of schemes should be something that "on the level of affine open sets, looks like this".

6.1.2. *Geometric motivation.* Motivated by the theory of differentiable manifolds (§3.1.1), which like schemes are ringed spaces, we want morphisms of schemes at the very least to be morphisms of ringed spaces; we now motivate what these are. (We will formalize this in the next section.) Notice that if $\pi : X \to Y$ is a map of differentiable manifolds, then a differentiable function on Y pulls back to a differentiable function on X. More precisely, given an open subset $U \subset Y$, there is a natural map $\Gamma(U, \mathscr{O}_Y) \to \Gamma(\pi^{-1}(U), \mathscr{O}_X)$. This behaves well with respect to restriction (restricting a function to a smaller open set and pulling back yields the same result as pulling back and then restricting), so in fact we have a map of sheaves on Y: $\mathscr{O}_Y \to \pi_*\mathscr{O}_X$. Similarly a morphism of schemes $\pi : X \to Y$ should induce a map $\mathscr{O}_Y \to \pi_*\mathscr{O}_X$. But in fact in the category of differentiable manifolds a continuous map $X \to Y$ is a map of differentiable manifolds precisely when differentiable functions on Y pull back to differentiable functions on X (i.e., the pullback map from differentiable functions on Y to *functions* on X in fact lies in the subset of *differentiable functions*, i.e., the continuous map $X \to Y$ induces a pullback of differential functions, which can be interpreted as a map $\mathscr{O}_Y \to \pi_*\mathscr{O}_X$), so this map of sheaves *characterizes* morphisms in the differentiable category. So we could use this as the *definition* of morphism in the differentiable category (see Exercise 3.1.A).

But how do we apply this to the category of schemes? In the category of differentiable manifolds, a continuous map $\pi : X \to Y$ *induces* a pullback of (the sheaf of) functions, and we can ask when this induces a pullback of *differentiable* functions. However, functions are odder on schemes, and we can't recover the pullback map just from the map of topological spaces. The right patch is to hardwire this into the definition of morphism, i.e., to have a continuous map $\pi : X \to Y$, along with a pullback map $\pi^\sharp : \mathscr{O}_Y \to \pi_*\mathscr{O}_X$. This leads to the definition of the *category* of ringed spaces.

One might hope to define morphisms of schemes as morphisms of ringed spaces. This isn't quite right, as then Motivation 6.1.1 isn't satisfied: as desired, to each morphism $A \to B$ there is a morphism $\operatorname{Spec} B \to \operatorname{Spec} A$, but there can be additional morphisms of ringed spaces $\operatorname{Spec} B \to \operatorname{Spec} A$ not arising in this way (see Exercise 6.2.E). A revised definition as morphisms of ringed spaces that locally look of this form will work, but this is awkward to work with, and we take a different approach. However, we will check that our eventual definition actually is equivalent to this (Exercise 6.3.C).

We begin by formally defining morphisms of ringed spaces.

6.2 Morphisms of ringed spaces

6.2.1. Definition. A **morphism of ringed spaces** $\pi : X \to Y$ is a continuous map of topological spaces (which we unfortunately also call π) along with a map $\mathscr{O}_Y \to \pi_*\mathscr{O}_X$, which we think of as a "pullback map". By adjointness (§2.6.1), this is the same as a map $\pi^{-1}\mathscr{O}_Y \to \mathscr{O}_X$. (It can be convenient to package this information as in the diagram (2.6.2.1).) There is an obvious notion of composition of morphisms, so ringed spaces form a category. Hence we have notion of automorphisms and isomorphisms. You can easily verify that an isomorphism of ringed spaces means the same thing as it did before (Definition 4.3.1).

If $U \subset Y$ is an open subset, then there is a natural morphism of ringed spaces $(U, \mathscr{O}_Y|_U) \to (Y, \mathscr{O}_Y)$ (which implicitly appeared earlier in Exercise 2.6.G). More precisely, if $U \to Y$ is an isomorphism of U with an open subset V of Y, and we are given an isomorphism $(U, \mathscr{O}_U) \cong (V, \mathscr{O}_Y|_V)$ (via the isomorphism $U \cong V$), then the resulting map of ringed spaces is called an **open embedding** (or **open immersion**) of ringed spaces, and the morphism $U \to Y$ is often written $U \hookrightarrow Y$.

6.2.A. EXERCISE (MORPHISMS OF RINGED SPACES GLUE). Suppose $(X, \mathscr{O}_X)$ and $(Y, \mathscr{O}_Y)$ are ringed spaces, $X = \cup_i U_i$ is an open cover of X, and we have morphisms of ringed spaces $\pi_i : U_i \to Y$ that "agree on the overlaps", i.e., $\pi_i|_{U_i \cap U_j} = \pi_j|_{U_i \cap U_j}$. Show that there is a unique morphism of ringed spaces $\pi : X \to Y$ such that $\pi|_{U_i} = \pi_i$. (Exercise 2.2.F essentially showed this for topological spaces.)

6.2.B. EASY IMPORTANT EXERCISE: $\mathscr{O}$-MODULES PUSH FORWARD. Given a morphism of ringed spaces $\pi : X \to Y$, show that sheaf pushforward induces a functor $\mathit{Mod}_{\mathscr{O}_X} \to \mathit{Mod}_{\mathscr{O}_Y}$.

6.2.C. EASY IMPORTANT EXERCISE. Given a morphism of ringed spaces $\pi : X \to Y$ with $\pi(p) = q$, show that there is a map of stalks $(\mathscr{O}_Y)_q \to (\mathscr{O}_X)_p$.

6.2.D. KEY EXERCISE. Suppose $\pi^\sharp : B \to A$ is a morphism of rings. Define a morphism of ringed spaces $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ as follows. The map of topological spaces was given in Exercise 3.4.H. To describe a morphism of sheaves $\mathscr{O}_{\operatorname{Spec} B} \to \pi_* \mathscr{O}_{\operatorname{Spec} A}$ on $\operatorname{Spec} B$, it suffices to describe a morphism of sheaves on the distinguished base of $\operatorname{Spec} B$. On $D(g) \subset \operatorname{Spec} B$, we define

$$\mathscr{O}_{\operatorname{Spec} B}(D(g)) \to \mathscr{O}_{\operatorname{Spec} A}(\pi^{-1} D(g)) = \mathscr{O}_{\operatorname{Spec} A}(D(\pi^\sharp g))$$

by $B_g \to A_{\pi^\sharp g}$. Verify that this makes sense (e.g. is independent of g), and that this describes a morphism of sheaves on the distinguished base. (This is the third in a series of exercises. We saw that a morphism of rings induces a map of sets in §3.2.9, a map of topological spaces in Exercise 3.4.H, and now a map of ringed spaces here.)

The map of ringed spaces of Key Exercise 6.2.D is really not complicated. Here is an example. Consider the ring map $\mathbb{C}[y] \to \mathbb{C}[x]$ given by $y \mapsto x^2$ (see Figure 3.6). We are mapping the affine line with coordinate x to the affine line with coordinate y. The map is (on closed points) $a \mapsto a^2$. For example, where does $[(x-3)]$ go to? Answer: $[(y-9)]$, i.e., $3 \mapsto 9$. What is the preimage of $[(y-4)]$? Answer: those prime ideals in $\mathbb{C}[x]$ containing $[(x^2-4)]$, i.e., $[(x-2)]$ and $[(x+2)]$, so the preimage of 4 is indeed ± 2. This is just about the map of sets, which is old news (§3.2.9), so let's now think about functions pulling back. What is the pullback of the function $3/(y-4)$ on $D([(y-4)]) = \mathbb{A}^1 - \{4\}$? Of course it is $3/(x^2-4)$ on $\mathbb{A}^1 - \{-2, 2\}$.

The construction of Key Exercise 6.2.D will soon be an example of morphism of schemes! In fact we could make that definition right now. Before we do, we point out (via the next exercise) that not every morphism of ringed spaces between affine schemes is of the form of Key Exercise 6.2.D. (In the language of §6.3, this morphism of ringed spaces is not a morphism of locally ringed spaces.)

6.2.E. UNIMPORTANT EXERCISE. Recall (Exercise 3.4.K) that $\operatorname{Spec} k[y]_{(y)}$ has two points, $[(0)]$ and $[(y)]$, where the second point is closed, and the first is not. Describe a map *of ringed spaces* $\operatorname{Spec} k(x) \to \operatorname{Spec} k[y]_{(y)}$ sending the unique point of $\operatorname{Spec} k(x)$ to the closed point $[(y)]$, where the pullback map on global sections sends k to k by the identity, and sends y to x. Show that this map of ringed spaces is not of the form described in Key Exercise 6.2.D.

6.2.2. Tentative Definition we won't use (cf. Motivation 6.1.1 in §6.1). A morphism of schemes $\pi : (X, \mathscr{O}_X) \to (Y, \mathscr{O}_Y)$ is a morphism of ringed spaces that "locally looks like" the maps of affine schemes described in Key Exercise 6.2.D. Precisely, for each choice of affine open sets $\operatorname{Spec} A \subset X$, $\operatorname{Spec} B \subset Y$, such that $\pi(\operatorname{Spec} A) \subset \operatorname{Spec} B$, the induced map of ringed spaces should be of the form shown in Key Exercise 6.2.D.

We would like this definition to be checkable on an affine cover, and we might hope to use the Affine Communication Lemma to develop the theory in this way. This works, but it will be more convenient to use a clever trick: in the next section, we will use the notion of locally ringed spaces, and then once we have used it, we will discard it like yesterday's garbage.

6.3 From locally ringed spaces to morphisms of schemes

In order to prove that morphisms behave in a way we hope, we will use the notion of a *locally ringed space*. It will not be used later, although it is useful elsewhere in geometry. The notion of locally ringed spaces (and maps between them) is inspired by what we know about manifolds (see Exercise 3.1.B). If $\pi : X \to Y$ is a morphism of manifolds, with $\pi(p) = q$, and f is a function on Y vanishing at q, then the pulled back function $\pi^\sharp(f)$ on X should vanish on p. Put differently: germs of functions (at $q \in Y$) vanishing at q should pull back to germs of functions (at $p \in X$) vanishing at p.

6.3.1. *Definition.* Recall (Definition 4.3.6) that a *locally ringed space* is a ringed space $(X, \mathscr{O}_X)$ such that the stalks $\mathscr{O}_{X,p}$ are all local rings. A **morphism of locally ringed spaces** $\pi : X \to Y$ is a morphism of ringed spaces such that the induced map of stalks $\mathscr{O}_{Y,q} \to \mathscr{O}_{X,p}$ (Exercise 6.2.C) sends the maximal ideal of the former into the maximal ideal of the latter (a **"morphism of local rings"**). This means something rather concrete and intuitive: "if $p \mapsto q$, and g is a function vanishing at q, then it will pull back to a function vanishing at p." (Side Remark: you would also want: "if $p \mapsto q$, and g is a function *not* vanishing at q, then it will pull back to a function *not* vanishing at p." This follows from our definition — can you see why?) Note that locally ringed spaces form a category.

To summarize: we use the notion of locally ringed space only to define morphisms of schemes, and to show that morphisms have reasonable properties. The main things you need to remember about locally ringed spaces are (i) that the functions have values at points, and (ii) that given a map of locally ringed spaces, the pullback of where a function vanishes is precisely where the pulled back function vanishes.

6.3.A. EXERCISE. Show that morphisms of locally ringed spaces glue (cf. Exercise 6.2.A). (Hint: your solution to Exercise 6.2.A may work without change.)

6.3.B. EASY IMPORTANT EXERCISE.
(a) Show that $\operatorname{Spec} A$ is a locally ringed space. (Hint: Exercise 4.3.F.)
(b) Show that the morphism of ringed spaces $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ defined by a ring morphism $\pi^\sharp : B \to A$ (Exercise 6.2.D) is a morphism of locally ringed spaces.

6.3.2. Key Proposition. — *If $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ is a morphism of locally ringed spaces then it is the morphism of locally ringed spaces induced by the map $\pi^\sharp : B = \Gamma(\operatorname{Spec} B, \mathscr{O}_{\operatorname{Spec} B}) \to \Gamma(\operatorname{Spec} A, \mathscr{O}_{\operatorname{Spec} A}) = A$ as in Exercise 6.3.B(b).*

(Aside: Exercise 4.3.A is a special case of Key Proposition 6.3.2. You should look back at your solution to Exercise 4.3.A, and see where you implicitly used ideas about locally ringed spaces.)

Proof. Suppose $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ is a morphism of locally ringed spaces. We wish to show that it is determined by its map on global sections $\pi^\sharp : B \to A$. We first need to check that the map of points is determined by global sections. Now a point p of $\operatorname{Spec} A$ can be identified with the prime ideal of global functions vanishing on it. The image point $\pi(p)$ in $\operatorname{Spec} B$ can be interpreted as the unique point q of $\operatorname{Spec} B$, where the functions vanishing at q are precisely those that pull back to functions vanishing at $\mathfrak{p}$. (Here we use the fact that π is a map of locally ringed spaces.) This is precisely the way in which the map of sets $\operatorname{Spec} A \to \operatorname{Spec} B$ induced by a ring map $B \to A$ was defined (§3.2.9).

Note in particular that if $b \in B$, $\pi^{-1}(D(b)) = D(\pi^\sharp b)$, again using the hypothesis that π is a morphism of locally ringed spaces.

It remains to show that $\pi^\sharp : \mathscr{O}_{\operatorname{Spec} B} \to \pi_* \mathscr{O}_{\operatorname{Spec} A}$ is the morphism of sheaves given by Exercise 6.2.D (cf. Exercise 6.3.B(b)). It suffices to check this on the distinguished base (Exercise 2.7.C(a)). We now want to check that for any map of locally ringed spaces inducing the map of sheaves $\mathscr{O}_{\operatorname{Spec} B} \to \pi_* \mathscr{O}_{\operatorname{Spec} A}$, the map of sections on any distinguished open set $D(b) \subset \operatorname{Spec} B$ is determined by the map of global sections $B \to A$.

Consider the commutative diagram

$$\begin{tikzcd}
B \arrow[r, equal] & \Gamma(\operatorname{Spec} B, \mathscr{O}_{\operatorname{Spec} B}) \arrow[r, "\pi^\sharp_{\operatorname{Spec} B}"] \arrow[d, "\operatorname{res}_{\operatorname{Spec} B, D(b)}"'] & \Gamma(\operatorname{Spec} A, \mathscr{O}_{\operatorname{Spec} A}) \arrow[r, equal] \arrow[d, "\operatorname{res}_{\operatorname{Spec} A, D(\pi^\sharp b)}"] & A \\
B_b \arrow[r, equal] & \Gamma(D(b), \mathscr{O}_{\operatorname{Spec} B}) \arrow[r, "\pi^\sharp_{D(b)}"] & \Gamma(D(\pi^\sharp b), \mathscr{O}_{\operatorname{Spec} A}) \arrow[r, equal] & A_{\pi^\sharp b} = A \otimes_B B_b.
\end{tikzcd}$$

The vertical arrows (restrictions to distinguished open sets) are localizations by b, so the lower horizontal map $\pi^\sharp_{D(b)}$ is determined by the upper map (it is just localization by b). □

We are ready for our definition.

6.3.3. Definition. If X and Y are schemes, then a morphism $\pi : X \to Y$ as locally ringed spaces is called a **morphism of schemes**. We have thus defined the **category of schemes**, which we denote *Sch*. (We then have notions of **isomorphism** — just the same as before, §4.3.6 — and **automorphism**. The *target* Y of π is sometimes called the **base scheme** or the **base**, when we are interpreting π as a family of schemes parametrized by Y — this may become clearer once we have defined the fibers of morphisms in §9.3.2.)

The definition in terms of locally ringed spaces easily implies Tentative Definition 6.2.2:

6.3.C. IMPORTANT EXERCISE. Show that a morphism of schemes $\pi : X \to Y$ is a morphism of ringed spaces that looks locally like morphisms of affine schemes. Precisely, if $\operatorname{Spec} A$ is an affine open subset of X and $\operatorname{Spec} B$ is an affine open subset of Y, and $\pi(\operatorname{Spec} A) \subset \operatorname{Spec} B$, then the induced morphism of ringed spaces is a morphism of affine schemes. (In case it helps, note: if $W \subset X$ and $Y \subset Z$ are both open embeddings of ringed spaces, then any morphism of ringed spaces $X \to Y$ induces a morphism of ringed spaces $W \to Z$, by composition $W \to X \to Y \to Z$.) Show that it suffices to check on a set $(\operatorname{Spec} A_i, \operatorname{Spec} B_i)$ where the $\operatorname{Spec} A_i$ form an open cover of X.

In practice, we will use the affine cover interpretation, and forget completely about locally ringed spaces. In particular, put imprecisely, the category of affine schemes is the category of rings with the arrows reversed. More precisely:

6.3.D. Exercise. Show that the category of rings and the opposite category of affine schemes are equivalent (see §1.2.21 to read about equivalence of categories).

In particular, here is something surprising: there can be many different maps from one point to another. For example, here are two different maps from the point $\operatorname{Spec} \mathbb{C}$ to the point $\operatorname{Spec} \mathbb{C}$: the identity (corresponding to the identity $\mathbb{C} \to \mathbb{C}$), and complex conjugation. (There are even more such maps!)

It is clear (from the corresponding facts about locally ringed spaces) that morphisms glue (Exercise 6.3.A), and the composition of two morphisms is a morphism. Isomorphisms in this category are precisely what we defined them to be earlier (§4.3.6).

6.3.4. The category of complex schemes (or more generally the category of k-schemes where k is a field, or more generally the category of A-schemes where A is a ring, or more generally the category of S-schemes where S is a scheme). The category of S-schemes Sch_S (where S is a scheme) is defined as follows. The objects (S-schemes) are morphisms of the form

$$\begin{tikzcd} X \arrow[d] \\ S \end{tikzcd}$$

(The morphism to S is called the **structure morphism**. A motivation for this terminology is the fact that if $S = \operatorname{Spec} A$, the structure morphism gives the functions on each open set of X the structure of an A-algebra, cf. §5.3.6.) The morphisms in the category of S-schemes are defined to be commutative diagrams

$$\begin{tikzcd} X \arrow[r] \arrow[d] & Y \arrow[d] \\ S \arrow[r, "="] & S \end{tikzcd}$$

which is more conveniently written as a commutative diagram

$$\begin{tikzcd} X \arrow[rr] \arrow[dr] & & Y \arrow[dl] \\ & S. & \end{tikzcd}$$

When there is no confusion (if the base scheme is clear), simply the top row of the diagram is given. In the case where $S = \operatorname{Spec} A$, where A is a ring, we get the notion of an A-scheme, which is the same as the same definition as in §5.3.6 (Exercise 6.3.G), but in a more satisfactory form. For example, complex geometers may consider the category of $\mathbb{C}$-schemes.

The next two examples are important. The first will show you that you can work with these notions in a straightforward, hands-on way. The second will show that you can work with these notions in a formal way.

6.3.E. Important exercise. (This exercise can give you some practice with understanding morphisms of schemes by cutting up into affine open sets.) Make sense of the following sentence: "$\mathbb{A}_k^{n+1} \setminus \{\vec{0}\} \to \mathbb{P}_k^n$ given by

$$(x_0, x_1, \dots, x_n) \mapsto [x_0, x_1, \dots, x_n]$$

is a morphism of schemes." Caution: you can't just say where points go; you have to say where functions go. So you may have to divide these up into affines, and describe the maps, and check that they glue. (Can you generalize to the case where k is replaced by a general ring B? See Exercise 6.3.N for an answer.)

6.3.F. ESSENTIAL EXERCISE. Show that morphisms $X \to \operatorname{Spec} A$ are in natural bijection with ring morphisms $A \to \Gamma(X, \mathscr{O}_X)$. Hint: Show that this is true when X is affine. Use the fact that morphisms glue, Exercise 6.3.A. (This is even true in the category of locally ringed spaces. You are free to prove it in this generality, but it is easier in the category of schemes.)

In particular, there is a canonical morphism from a scheme to Spec of its ring of global sections. (Warning: Even if X is a finite type k-scheme, the ring of global sections might be nasty! In particular, it might not be finitely generated, see 19.11.13.) The canonical morphism $X \to \operatorname{Spec} \Gamma(X, \mathscr{O}_X)$ is an isomorphism if and only if X is affine (i.e., isomorphic to $\operatorname{Spec} A$ for some ring A), and in this case it is the isomorphism hinted at in Remark 4.3.2.

6.3.G. EASY EXERCISE. Show that this definition of A-scheme given in §6.3.4 agrees with the earlier definition of §5.3.6.

6.3.5. ⋆ *Side fact for experts:* Γ *and* Spec *are adjoints.* We have a contravariant functor Spec from rings to locally ringed spaces, and a contravariant functor Γ from locally ringed spaces to rings. In fact $(\Gamma, \operatorname{Spec})$ is an adjoint pair! (Caution: we have only discussed adjoints for covariant functors; if you care, you will have to figure out how to define adjoints for contravariant functors.) Thus we could have defined Spec by requiring it to be right-adjoint to Γ. (Fun but irrelevant side question: if you used ringed spaces rather than locally ringed spaces, Γ again has a right adjoint. What is it?)

6.3.H. EASY EXERCISE. If $S_\bullet$ is a finitely generated graded A-algebra, describe a natural "structure morphism" $\operatorname{Proj} S_\bullet \to \operatorname{Spec} A$.

6.3.I. EASY EXERCISE. Show that $\operatorname{Spec} \mathbb{Z}$ is the final object in the category of schemes. In other words, if X is any scheme, there exists a unique morphism to $\operatorname{Spec} \mathbb{Z}$. (Hence the category of schemes is isomorphic to the category of $\mathbb{Z}$-schemes.) If k is a field, show that $\operatorname{Spec} k$ is the final object in the category of k-schemes.

6.3.J. EXERCISE.
(a) Suppose p is a point of a scheme X. Describe a canonical (choice-free) morphism $\operatorname{Spec} \mathscr{O}_{X,p} \to X$. (Hint: do this for affine X first. But then for general X be sure to show that your morphism is independent of choice.)
(b) Define a canonical morphism $\operatorname{Spec} \kappa(p) \to X$. (This is often written $p \to X$; one gives p the obvious interpretation as a scheme.)

6.3.6. *Remark.* From Essential Exercise 6.3.F, it is one small step to show that some products of schemes exist: if A and B are rings, then $\operatorname{Spec} A \times \operatorname{Spec} B = \operatorname{Spec}(A \otimes_{\mathbb{Z}}$

B); and if A and B are C-algebras, then $\operatorname{Spec} A \times_{\operatorname{Spec} C} \operatorname{Spec} B = \operatorname{Spec}(A \otimes_C B)$. But we are in no hurry, so we wait until Exercise 9.1.B to discuss this properly.

6.3.K. ⋆⋆ EXERCISE FOR THOSE WITH APPROPRIATE BACKGROUND: THE ANALYTIFICATION FUNCTOR. Recall the analytification construction of Exercise 5.3.F. For each morphism of reduced finite type $\mathbb{C}$-schemes $\pi : X \to Y$ (over $\mathbb{C}$), define a morphism of complex analytic prevarieties $\pi_{an} : X_{an} \to Y_{an}$ (the **analytification of** π). Show that analytification gives a functor from the category of reduced finite type $\mathbb{C}$-schemes to the category of complex analytic prevarieties. (Remark: Two nonisomorphic varieties can have isomorphic analytifications. For example, Serre described two different algebraic structures on the complex manifold $\mathbb{C}^* \times \mathbb{C}^*$, see [**Ha2**, p. 232] and [**MO68421**]; one is "the obvious one", and the other is a $\mathbb{P}^1$-bundle over an elliptic curve, with a section removed. For an example of a smooth complex surface with infinitely many algebraic structures, see §19.11.3. On the other hand, a compact complex variety can have only one algebraic structure (see [**Se2**, §19]).)

6.3.7. Definition: The functor of points, and scheme-valued points (and ring-valued points, and field-valued points) of a scheme. If Z is a scheme, then Z**-valued points** of a scheme X, denoted $X(Z)$, are defined to be maps $Z \to X$. If A is a ring, then A**-valued points** of a scheme X, denoted $X(A)$, are defined to be the $(\operatorname{Spec} A)$-valued points of the scheme. (The most common case of this is when A is a field.) We denote Z-valued points of X by $X(Z)$ and A-valued points of X by $X(A)$.

If you are working over a base scheme B — for example, complex algebraic geometers will consider only schemes and morphisms over $B = \operatorname{Spec} \mathbb{C}$ — then in the above definition, there is an implicit structure map $Z \to B$ (or $\operatorname{Spec} A \to B$ in the case of $X(A)$). For example, for a complex geometer, if X is a scheme over $\mathbb{C}$, the $\mathbb{C}(t)$-valued points of X correspond to commutative diagrams of the form

$$\begin{tikzcd} \operatorname{Spec} \mathbb{C}(t) \arrow[rr] \arrow[dr, "\xi"'] & & X \arrow[dl, "\pi"] \\ & \operatorname{Spec} \mathbb{C} & \end{tikzcd}$$

where $\pi : X \to \operatorname{Spec} \mathbb{C}$ is the structure map for X, and ξ corresponds to the obvious inclusion of rings $\mathbb{C} \to \mathbb{C}(t)$. (Warning: a k-valued point of a k-scheme X is sometimes called a "rational point" of X, which is dangerous, as for most of the world, "rational" refers to $\mathbb{Q}$. We will use the safer phrase "k-valued point" of X.)

The terminology "Z-valued point" (and A-valued point) is unfortunate, because we earlier defined the notion of points of a scheme, and Z-valued points (and A-valued points) are not (necessarily) points! But these usages is well-established in the literature. (Look in the index under "point" to see even more inconsistent use of adjectives that modify this word.)

6.3.L. EXERCISE.
(a) (easy) Show that a morphism of schemes $X \to Y$ induces a map of Z-valued points $X(Z) \to Y(Z)$.
(b) Note that morphisms of schemes $X \to Y$ are not determined by their "underlying" map of points. (What is an example?) Show that they *are* determined by their

induced maps of Z-valued points, as Z varies over all schemes. (Hint: pick $Z = X$. In the course of doing this exercise, you will largely prove Yoneda's Lemma in the guise of Exercise 9.1.C.)

6.3.8. Furthermore, we will see that "products of Z-valued points" behave as you might hope (§9.1.3). A related reason this language is suggestive: the notation $X(Z)$ suggests the interpretation of X as a (contravariant) functor h_X from schemes to sets — the **functor of** (scheme-valued) **points** of the scheme X (cf. Example 1.2.20).

Here is a low-brow reason A-valued points are a useful notion: *the A-valued points of an affine scheme* $\operatorname{Spec} \mathbb{Z}[x_1, \dots, x_n]/(f_1, \dots, f_r)$ *(where $f_i \in \mathbb{Z}[x_1, \dots, x_n]$ are relations) are precisely the solutions to the equations*

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

in the ring A. For example, the rational solutions to $x^2 + y^2 = 16$ are precisely the $\mathbb{Q}$-valued points of $\operatorname{Spec} \mathbb{Z}[x, y]/(x^2 + y^2 - 16)$. The integral solutions are precisely the $\mathbb{Z}$-valued points. So A-valued points of an affine scheme (finite type over $\mathbb{Z}$) can be interpreted simply. In the special case where A is local, A-valued points of a general scheme have a good interpretation too:

6.3.M. EXERCISE (MORPHISMS FROM Spec OF A LOCAL RING TO X). Suppose X is a scheme, and $(A, \mathfrak{m})$ is a local ring. Suppose we have a scheme morphism $\pi : \operatorname{Spec} A \to X$ sending $[\mathfrak{m}]$ to p. Show that any open set containing p contains the image of π. Show that there is a bijection between $\operatorname{Mor}(\operatorname{Spec} A, X)$ and $\{p \in X, \text{local homomorphisms } \mathscr{O}_{X,p} \to A\}$. (Possible hint: Exercise 6.3.J(a).)

On the other hand, Z-valued points of projective space can be subtle. There are some maps we can write down easily, as shown by applying the next exercise in the case $X = \operatorname{Spec} A$, where A is a B-algebra.

6.3.N. EASY (BUT SURPRISINGLY ENLIGHTENING) EXERCISE (CF. EXERCISE 6.3.E).
(a) Suppose B is a ring. If X is a B-scheme, and $f_0, \dots, f_n$ are $n + 1$ functions on X with no common zeros, then show that $[f_0, \dots, f_n]$ gives a morphism of B-schemes $X \to \mathbb{P}^n_B$.
(b) Suppose g is a nowhere vanishing function on X, and f_i are as in part (a). Show that the morphisms $[f_0, \dots, f_n]$ and $[gf_0, \dots, gf_n]$ to $\mathbb{P}^n_B$ are the same.

6.3.9. *Example.* Consider the $n+1$ functions $x_0, \dots, x_n$ on $\mathbb{A}^{n+1}$ (otherwise known as $n + 1$ sections of the trivial bundle). They have no common zeros on $\mathbb{A}^{n+1} - 0$. Hence they determine a morphism $\mathbb{A}^{n+1} - 0 \to \mathbb{P}^n$. (We discussed this morphism in Exercise 6.3.E, but now we don't need tedious gluing arguments.)

6.3.10. You might hope that Exercise 6.3.N(a) gives all morphisms to projective space (over B). But this isn't the case. Indeed, even the identity morphism $X = \mathbb{P}^1_k \to \mathbb{P}^1_k$ isn't of this form, as the source $\mathbb{P}^1$ has no nonconstant global functions with which to build this map. (There are similar examples with an affine source.) However, there is a correct generalization (characterizing *all* maps from schemes to projective schemes) in Theorem 16.4.1. This result roughly states that this works, so long as the f_i are not quite functions, but sections of a line bundle. Our desire to understand maps to projective schemes in a clean way will be one important motivation for understanding line bundles.

We will see more ways to describe maps to projective space in the next section. A different description directly generalizing Exercise 6.3.N(a) will be given in Exercise 15.3.F, which will turn out (in Theorem 16.4.1) to be a "universal" description.

Incidentally, before Grothendieck, it was considered a real problem to figure out the right way to interpret points of projective space with "coordinates" in a ring. These difficulties were due to a lack of functorial reasoning. And the clues to the right answer already existed (the same problems arise for maps from a smooth real manifold to $\mathbb{RP}^n$) — if you ask such a geometric question (for projective space is geometric), the answer is necessarily geometric, not purely algebraic!

6.3.11. *Visualizing schemes III: picturing maps of schemes when nilpotents are present.* You now know how to visualize the points of schemes (§3.3), and nilpotents (§4.2 and §5.5). The following imprecise exercise will give you some sense of how to visualize maps of schemes when nilpotents are involved. Suppose $a \in \mathbb{C}$. Consider the map of rings $\mathbb{C}[x] \to \mathbb{C}[\epsilon]/\epsilon^2$ given by $x \mapsto a\epsilon$. Recall that $\operatorname{Spec} \mathbb{C}[\epsilon]/(\epsilon^2)$ may be pictured as a point with a tangent vector (§4.2). How would you picture this map if $a \neq 0$? How does your picture change if $a = 0$? (The tangent vector should be "crushed" in this case.)

Exercise 12.1.I will extend this considerably; you may enjoy reading its statement now.

6.4 Maps of graded rings and maps of projective schemes

As maps of rings correspond to maps of affine schemes in the opposite direction, maps of graded rings (over a base ring A) sometimes give maps of projective schemes in the opposite direction. This is an imperfect generalization: not every map of graded rings gives a map of projective schemes (§6.4.2); not every map of projective schemes comes from a map of graded rings (§19.11.9); and different maps of graded rings can yield the same map of schemes (Exercise 6.4.C).

You may find it helpful to think through Examples 6.4.1 and 6.4.2 while working through the following exercise.

6.4.A. ESSENTIAL EXERCISE. Suppose that $\phi : S_\bullet \longrightarrow R_\bullet$ is a morphism of ($\mathbb{Z}^{\geq 0}$-)graded rings. (By **map of graded rings**, we mean a map of rings that preserves the grading as a map of "graded semigroups". In other words, there is a $d > 0$ such that S_n maps to R_{dn} for all n.) Show that this induces a morphism of schemes $\operatorname{Proj} R_\bullet \setminus V(\phi(S_+)) \to \operatorname{Proj} S_\bullet$. (Hint: Suppose x is a homogeneous element of S_+. Define a map $D(\phi(x)) \to D(x)$. Show that they glue together (as x runs over all homogeneous elements of S_+). Show that this defines a map from all of $\operatorname{Proj} R_\bullet \setminus V(\phi(S_+))$.) In particular, if

$$V(\phi(S_+)) = \varnothing, \tag{6.4.0.1}$$

then we have a morphism $\operatorname{Proj} R_\bullet \to \operatorname{Proj} S_\bullet$. From your solution, it will be clear that if ϕ is furthermore a morphism of A-algebras, then the induced morphism $\operatorname{Proj} R_\bullet \setminus V(\phi(S_\bullet)) \to \operatorname{Proj} S_\bullet$ is a morphism of A-schemes.

6.4.1. *Example.* Let's see Exercise 6.4.A in action. We will scheme-theoretically interpret the map of complex projective manifolds $\mathbb{CP}^1$ to $\mathbb{CP}^2$ given by

$$\begin{array}{ccc} \mathbb{CP}^1 & \longrightarrow & \mathbb{CP}^2 \\ & & \\ [s,t] & \longmapsto & [s^{20}, s^9t^{11}, t^{20}] \end{array}$$

Notice first that this is well-defined: $[\lambda s, \lambda t]$ is sent to the same point of $\mathbb{CP}^2$ as $[s,t]$. The reason for it to be well-defined is that the three polynomials s^{20}, s^9t^{11}, and t^{20} are all homogeneous of degree 20.

Algebraically, this corresponds to a map of graded rings in the opposite direction

$$\mathbb{C}[x,y,z] \to \mathbb{C}[s,t]$$

given by $x \mapsto s^{20}$, $y \mapsto s^9t^{11}$, $z \mapsto t^{20}$. You should interpret this in light of your solution to Exercise 6.4.A, and compare this to the affine example of §3.2.10.

6.4.2. *Example.* Notice that there is no map of complex manifolds $\mathbb{CP}^2 \to \mathbb{CP}^1$ given by $[x,y,z] \mapsto [x,y]$, because the map is not defined when $x = y = 0$. This corresponds to the fact that the map of graded rings $\mathbb{C}[s,t] \to \mathbb{C}[x,y,z]$ given by $s \mapsto x$ and $t \mapsto y$, doesn't satisfy hypothesis (6.4.0.1).

6.4.B. EXERCISE. Show that if $\phi : S_\bullet \to R_\bullet$ satisfies $\sqrt{(\phi(S_+))} = R_+$, then hypothesis (6.4.0.1) is satisfied. (Hint: Exercise 4.5.I.) This algebraic formulation of the more geometric hypothesis can sometimes be easier to verify.

6.4.C. UNIMPORTANT EXERCISE. This exercise shows that different maps of graded rings can give the same map of schemes. Let $R_\bullet = k[x,y,z]/(xz, yz, z^2)$ and $S_\bullet = k[a,b,c]/(ac, bc, c^2)$, where every variable has degree 1. Show that $\operatorname{Proj} R_\bullet \cong \operatorname{Proj} S_\bullet \cong \mathbb{P}^1_k$. Show that the maps $S_\bullet \to R_\bullet$ given by $(a,b,c) \mapsto (x,y,z)$ and $(a,b,c) \mapsto (x,y,0)$ give the same (iso)morphism $\operatorname{Proj} R_\bullet \to \operatorname{Proj} S_\bullet$. (The real reason is that all of these constructions are insensitive to what happens in a finite number of degrees. This will be made precise in a number of ways later, most immediately in Exercise 6.4.F.)

6.4.3. *Unimportant remark.* Exercise 16.4.G shows that not every morphism of schemes $\operatorname{Proj} R_\bullet \to \operatorname{Proj} S_\bullet$ comes from a map of graded rings $S_\bullet \to R_\bullet$, even in quite reasonable circumstances.

6.4.4. Veronese subrings.

Here is a useful construction. Suppose $S_\bullet$ is a finitely generated graded ring. Define the n**th Veronese subring** of $S_\bullet$ by $S_{n\bullet} = \oplus_{j=0}^{\infty} S_{nj}$. (The "old degree" n is "new degree" 1.)

6.4.D. EXERCISE. Show that the map of graded rings $S_{n\bullet} \hookrightarrow S_\bullet$ induces an *isomorphism* $\operatorname{Proj} S_\bullet \to \operatorname{Proj} S_{n\bullet}$. (Hint: if $f \in S_+$ is homogeneous of degree divisible by n, identify $D(f)$ on $\operatorname{Proj} S_\bullet$ with $D(f)$ on $\operatorname{Proj} S_{n\bullet}$. Why do such distinguished open sets cover $\operatorname{Proj} S_\bullet$?)

6.4.E. EXERCISE. If $S_\bullet$ is generated in degree 1, show that $S_{n\bullet}$ is also generated in degree 1. (You may want to consider the case of the polynomial ring first.)

6.4.F. EXERCISE. Show that if $R_\bullet$ and $S_\bullet$ are the same finitely generated graded rings except in a finite number of nonzero degrees (make this precise!), then $\operatorname{Proj} R_\bullet \cong \operatorname{Proj} S_\bullet$.

6.4.G. EXERCISE. Suppose $S_\bullet$ is generated over S_0 by $f_1, \dots, f_n$. Find a d such that $S_{d\bullet}$ is finitely generated in "new" degree 1 (= "old" degree d). (This is surprisingly tricky, so here is a hint. Suppose there are generators $x_1, \dots, x_n$ of degrees $d_1, \dots, d_n$ respectively. Show that any monomial $x_1^{a_1} \cdots x_n^{a_n}$ of degree at least $nd_1 \dots d_n$ has $a_i \geq (\prod_j d_j)/d_i$ for some i. Show that the $nd_1 \dots d_n$th Veronese subring is generated by elements in "new" degree 1.)

Exercise 6.4.G, in combination with Exercise 6.4.D, shows that there is little harm in assuming that finitely generated graded rings are generated in degree 1, as after a regrading (or more precisely, keeping only terms of degree a multiple of d, then dividing the degree by d), this is indeed the case. This is handy, as it means that, using Exercise 6.4.D, we can assume that any finitely generated graded ring is generated in degree 1. Exercise 8.2.G will later imply as a consequence that we can embed every Proj in some projective space.

6.4.H. LESS IMPORTANT EXERCISE. Suppose $S_\bullet$ is a finitely generated ring. Show that $S_{n\bullet}$ is a finitely generated graded ring. (Possible approach: use the previous exercise, or something similar, to show there is some N such that $S_{nN\bullet}$ is generated in degree 1, so the graded ring $S_{nN\bullet}$ is finitely generated. Then show that for each $0 < j < N$, $S_{nN\bullet+nj}$ is a finitely generated module over $S_{nN\bullet}$.)

6.5 Rational maps from reduced schemes

Informally speaking, a "rational map" is "a morphism defined almost everywhere", much as a rational function (Definition 5.5.6) is a name for a function defined almost everywhere. We will later see that in good situations, just as with rational functions, where a rational map is defined, it is uniquely defined (the Reduced-to-Separated Theorem 10.2.2), and has a largest *"domain of definition"* (§10.2.3). For this section only, *we assume* X *to be reduced*. A key example will be irreducible varieties (§6.5.6), and the language of rational maps is most often used in this case.

6.5.1. *Definition.* A **rational map** from X to Y, denoted $X \dashrightarrow Y$, is a morphism on a dense open set, with the equivalence relation $(f : U \to Y) \sim (g : V \to Y)$ if there is a dense open set $Z \subset U \cap V$ such that $f|_Z = g|_Z$. (In §10.2.3, we will improve this to: if $f|_{U\cap V} = g|_{U\cap V}$ in good circumstances — when Y is separated.) People often use the word "map" for "morphism", which is quite reasonable, except that a rational map need not be a map. So to avoid confusion, when one means "rational map", one should never just say "map".

We will also talk about rational maps of S-schemes for a scheme S. The definition is the same, except now X and Y are S-schemes, and $f : U \to Y$ is a morphism of S-schemes.

6.5.2. ⋆ *Rational maps more generally.* Just as with rational functions, Definition 6.5.1 can be extended to where X is not reduced, as is (using the same name, "rational map"), or in a version that imposes some control over what happens over the nonreduced locus (*pseudo-morphisms*, [**Stacks**, tag 01RX]). We will see in §10.2 that rational maps from reduced schemes to separated schemes behave particularly well, which is why they are usually considered in this context. The reason for the definition of pseudo-morphisms is to extend these results to when X is nonreduced. We will not use the notion of pseudo-morphism.

6.5.3. An obvious example of a rational map is a morphism. Another important example is the projection $\mathbb{P}^n_A \dashrightarrow \mathbb{P}^{n-1}_A$ given by $[x_0, \cdots, x_n] \mapsto [x_0, \cdots, x_{n-1}]$. (How precisely is this a rational map in the sense of Definition 6.5.1? What is its domain of definition?)

A rational map $\pi : X \dashrightarrow Y$ is **dominant** (or in some sources, *dominating*) if for some (and hence every) representative $U \to Y$, the image is dense in Y.

6.5.A. Exercise. Show that a rational map $\pi : X \dashrightarrow Y$ of irreducible schemes is dominant if and only if π sends the generic point of X to the generic point of Y.

A little thought will convince you that you can compose (in a well-defined way) a dominant map $\pi : X \dashrightarrow Y$ from an irreducible scheme X with a rational map $\rho : Y \dashrightarrow Z$. Furthermore, the composition $\rho \circ \pi$ will be dominant if ρ is. Integral schemes and dominant rational maps between them form a category which is geometrically interesting.

6.5.B. Easy Exercise. Show that dominant rational maps of integral schemes give morphisms of function fields in the opposite direction.

It is not true that morphisms of function fields always give dominant rational maps, or even rational maps. For example, $\operatorname{Spec} k[x]$ and $\operatorname{Spec} k(x)$ have the same function field ($k(x)$), but there is no corresponding rational map $\operatorname{Spec} k[x] \dashrightarrow \operatorname{Spec} k(x)$ of k-schemes. Reason: that would correspond to a morphism from an open subset U of $\operatorname{Spec} k[x]$, say $\operatorname{Spec} k[x, 1/f(x)]$, to $\operatorname{Spec} k(x)$. But there is no map of rings $k(x) \to k[x, 1/f(x)]$ (sending k identically to k and x to x) for any one $f(x)$. However, maps of function fields indeed give dominant rational maps of integral finite type k-schemes (and in particular, irreducible varieties, to be defined in §10.1.7), see Proposition 6.5.7 below.

(If you want more evidence that the topologically-defined notion of dominance is simultaneously algebraic, you can show that if $\phi : A \to B$ is a ring morphism, then the corresponding morphism $\operatorname{Spec} B \to \operatorname{Spec} A$ is dominant if and only if ϕ has kernel contained in the nilradical of A.)

6.5.4. *Definition.* A rational map $\pi : X \dashrightarrow Y$ is said to be **birational** if it is dominant, and there is another rational map (a "rational inverse") ψ that is also dominant, such that $\pi \circ \psi$ is (in the same equivalence class as) the identity on Y, and $\psi \circ \pi$ is (in the same equivalence class as) the identity on X. This is the notion of isomorphism in the category of integral schemes and dominant rational maps. (We note in passing that in the differentiable category, this is a useless notion — any connected manifolds of dimension d have isomorphic dense open sets. To show this, use the fact that every smooth manifold admits a triangulation, see [**Wh**, p. 124-135].)

We say X and Y are **birational** (to each other) if there exists a birational map $X \dashrightarrow Y$. If X and Y are irreducible, then birational maps induce isomorphisms of function fields. The fact that maps of function fields correspond to rational maps in the opposite direction for integral finite type k-schemes, to be proved in Proposition 6.5.7, shows that a map between integral finite type k-schemes that induces an isomorphism of function fields is birational. An integral finite type k-scheme is said to be **rational** if it is birational to $\mathbb{A}^n_k$ for some k. A *morphism* is **birational** if it is birational as a rational map.

6.5.5. Proposition. — *Suppose X and Y are reduced schemes. Then X and Y are birational if and only if there is a dense open subscheme U of X and a dense open subscheme V of Y such that $U \cong V$.*

Proposition 6.5.5 tells you how to think of birational maps. Just as a rational map is a "mostly defined function", two birational reduced schemes are "mostly isomorphic". For example, a reduced finite type k-scheme (such as a reduced affine variety over k) is rational if it has a dense open subscheme isomorphic to an open subscheme of $\mathbb{A}^n$.

Proof. The "if" direction is trivial, so we prove the "only if" direction.

Step 1. Because X and Y are birational, we can find some dense open subschemes $X_1 \subset X$ and $Y_1 \subset Y$, along with $F : X_1 \to Y$ and $G : Y_1 \to X$ whose composition in either order is the identity morphism on *some* dense open subscheme where it makes sense. Replace X_1 and Y_1 by those dense open subschemes.

We have thus found dense open subschemes $X_1 \subset X$ and $Y_1 \subset Y$, along with morphisms $F : X_1 \to Y$ and $G : Y_1 \to X$, whose composition in either order is the identity on the open subset where it is defined. (More precisely, if $X_2 = F^{-1}(Y_1)$, and $Y_2 = G^{-1}(X_1)$, then $G \circ F|_{X_2} = \mathrm{id}_{X_2}$, and $F \circ G|_{Y_2} = \mathrm{id}_{Y_2}$.)

Step 2. For $n > 1$, inductively define $X_{n+1} = F^{-1}(Y_n)$ and $Y_{n+1} = G^{-1}(X_n)$. Informally, X_n is the (dense) open subset of points of X that can be mapped n times by F and G alternately, and analogously for Y_n. Define $X_\infty = \cap_{n \geq 1} X_n$, and $Y_\infty = \cap_{n \geq 1} Y_n$. Then $X_\infty = X_2$, as $G \circ F$ is the identity on X_2 (so any point of X_2 can be acted on by F and G alternately any number of times), and similarly $Y_\infty = Y_2$. Thus F and G define maps between X_2 and Y_2, and these are inverse maps by assumption. $\square$

6.5.6. Rational maps of irreducible varieties.

6.5.7. Proposition. — *Suppose X is an integral k-scheme and Y is an integral finite type k-scheme, and we are given an extension of function fields $\phi^\sharp : K(Y) \hookrightarrow K(X)$ preserving k. Then there exists a dominant rational map of k-schemes $\phi : X \dashrightarrow Y$ inducing $\phi^\sharp$.*

Proof. By replacing Y with an open subset, we may assume that Y is affine, say $\operatorname{Spec} B$, where B is generated over k by finitely many elements $y_1, \ldots, y_n$. Since we only need to define ϕ on an open subset of X, we may similarly assume that $X = \operatorname{Spec} A$ is affine. Then $\phi^\sharp$ gives an inclusion $\phi^\sharp : B \hookrightarrow K(A)$. Write the product of the images of $y_1, \ldots, y_n$ as f/g, with $f, g \in A$. Then $\phi^\sharp$ further induces an inclusion $B \hookrightarrow A_g$. Therefore $\phi : \operatorname{Spec} A_g \to \operatorname{Spec} B$ induces $\phi^\sharp$. The morphism

ϕ is dominant because the inverse image of the zero ideal under the inclusion $B \hookrightarrow A_g$ is the zero ideal, so ϕ takes the generic point of X to the generic point of Y. $\square$

6.5.C. EXERCISE. Let K be a finitely generated field extension of k. (Recall that a field extension K over k is **finitely generated** if there is a finite "generating set" x_1, ..., x_n in K such that every element of K can be written as a rational function in x_1, ..., x_n with coefficients in k.) Show that there exists an irreducible affine k-variety with function field K. (Hint: Consider the map $k[t_1, \dots, t_n] \to K$ given by $t_i \mapsto x_i$, and show that the kernel is a prime ideal $\mathfrak{p}$, and that $k[t_1, \dots, t_n]/\mathfrak{p}$ has fraction field K. Interpreted geometrically: consider the map $\operatorname{Spec} K \to \operatorname{Spec} k[t_1, \dots, t_n]$ given by the ring map $t_i \mapsto x_i$, and take the closure of the one-point image.)

6.5.D. EXERCISE. Describe equivalences of categories among the following.

(a) the category with objects "integral k-varieties", and morphisms "dominant rational maps defined over k";
(b) the category with objects "integral *affine* k-varieties", and morphisms "dominant rational maps defined over k"; and
(c) the opposite ("arrows-reversed") category with objects "finitely generated field extensions of k", and morphisms "inclusions extending the identity on k".

In particular, an integral affine k-variety X is rational if its function field $K(X)$ is a purely transcendent extension of k, i.e., $K(X) \cong k(x_1, \dots, x_n)$ for some n. (This needs to be said more precisely: the map $k \hookrightarrow K(X)$ induced by $X \to \operatorname{Spec} k$ should agree with the "obvious" map $k \hookrightarrow k(x_1, \dots, x_n)$ under this isomorphism.)

6.5.8. More examples of rational maps.
A recurring theme in these examples is that domains of definition of rational maps to projective schemes extend over regular codimension one points. We will make this precise in the Curve-to-Projective Extension Theorem 16.5.1, when we discuss curves.

The first example is the classical formula for Pythagorean triples. Suppose you are looking for rational points on the circle C given by $x^2+y^2=1$ (Figure 6.1). One rational point is $p=(1,0)$. If q is another rational point, then pq is a line of rational (non-infinite) slope. This gives a rational map from the conic C (now interpreted as $\operatorname{Spec} \mathbb{Q}[x,y]/(x^2+y^2-1)$) to $\mathbb{A}^1_{\mathbb{Q}}$, given by $(x,y) \mapsto y/(x-1)$. (Something subtle just happened: we were talking about $\mathbb{Q}$-points on a circle, and ended up with a rational map of schemes.) Conversely, given a line of slope m through p, where m is rational, we can recover q by solving the equations $y=m(x-1)$, $x^2+y^2=1$. We substitute the first equation into the second, to get a quadratic equation in x. We know that we will have a solution $x=1$ (because the line meets the circle at $(x,y)=(1,0)$), so we expect to be able to factor this out, and find the other factor. This indeed works:

$$\begin{aligned} x^2+(m(x-1))^2 &= 1 \\ \Longrightarrow \quad (m^2+1)x^2+(-2m^2)x+(m^2-1) &= 0 \\ \Longrightarrow \quad (x-1)((m^2+1)x-(m^2-1)) &= 0 \end{aligned}$$

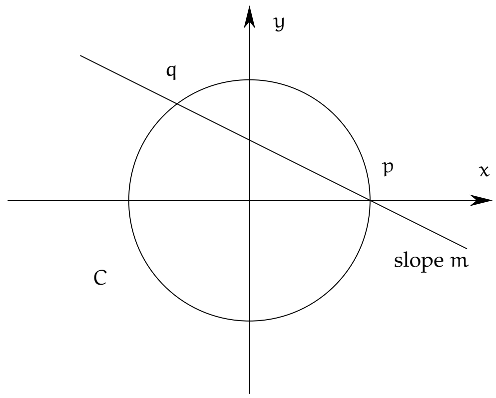


FIGURE 6.1. Finding primitive Pythagorean triples using geometry

The other solution is $x = (m^2-1)/(m^2+1)$, which gives $y = -2m/(m^2+1)$. Thus we get a birational map between the conic C and $\mathbb{A}^1$ with coordinate m, given by $f : (x, y) \mapsto y/(x-1)$ (which is defined for $x \neq 1$), and with inverse rational map given by $m \mapsto ((m^2-1)/(m^2+1), -2m/(m^2+1))$ (which is defined away from $m^2+1=0$).

We can extend this to a rational map $C \dashrightarrow \mathbb{P}^1_{\mathbb{Q}}$ via the "inclusion" $\mathbb{A}^1_{\mathbb{Q}} \to \mathbb{P}^1_{\mathbb{Q}}$ (which we later call an open embedding). Then f is given by $(x, y) \mapsto [y, x-1]$. We then have an interesting question: what is the domain of definition of f? It appears to be defined everywhere except for where $y = x-1 = 0$, i.e., everywhere but p. But in fact it can be extended over p! Note that $(x, y) \mapsto [x+1, -y]$ (where $(x, y) \neq (-1, 0)$) agrees with f on their common domains of definition, as $[x+1, -y] = [y, x-1]$. Hence this rational map can be extended farther than we at first thought. This will be a special case of the Curve-to-Projective Extension Theorem 16.5.1.

6.5.E. EXERCISE. Use the above to find a "formula" yielding all Pythagorean triples.

6.5.F. EXERCISE. Show that the conic $x^2 + y^2 = z^2$ in $\mathbb{P}^2_k$ is isomorphic to $\mathbb{P}^1_k$ for any field k of characteristic not 2. (Aside: What happens in characteristic 2?)

6.5.9. In fact, any conic in $\mathbb{P}^2_k$ with a k-valued point (i.e., a point with residue field k) of rank 3 (after base change to $\overline{k}$, so "rank" makes sense, see Exercise 5.4.J) is isomorphic to $\mathbb{P}^1_k$. (The hypothesis of having a k-valued point is certainly necessary: $x^2 + y^2 + z^2 = 0$ over $k = \mathbb{R}$ is a conic that is not isomorphic to $\mathbb{P}^1_k$.)

6.5.G. EXERCISE. Find all rational solutions to $y^2 = x^3 + x^2$, by finding a birational map to $\mathbb{A}^1_{\mathbb{Q}}$, mimicking what worked with the conic. Hint: what point should you project from? (In Exercise 19.10.F, we will see that these points form a group, and that this is a degenerate elliptic curve.)

You will obtain a rational map to $\mathbb{P}^1_{\mathbb{Q}}$ that is not defined over the node $x = y = 0$, and *cannot* be extended over this codimension 1 set. This is an example of the limits of our future result, the Curve-to-Projective Extension Theorem 16.5.1, showing how to extend rational maps to projective space over codimension 1 sets: the codimension 1 sets have to be regular.

6.5.H. EXERCISE. Use a similar idea to find a birational map from the quadric $Q = \{x^2 + y^2 = w^2 + z^2\} \subset \mathbb{P}^3_{\mathbb{Q}}$ to $\mathbb{P}^2_{\mathbb{Q}}$. Use this to find all rational points on Q. (This illustrates a good way of solving Diophantine equations. You will find a dense open subset of Q that is isomorphic to a dense open subset of $\mathbb{P}^2$, where you can easily find all the rational points. There will be a closed subset of Q where the rational map is not defined, or not an isomorphism, but you can deal with this subset in an ad hoc fashion.)

6.5.I. EXERCISE (THE CREMONA TRANSFORMATION, A USEFUL CLASSICAL CONSTRUCTION). Consider the rational map $\mathbb{P}^2_k \dashrightarrow \mathbb{P}^2_k$, given by $[x, y, z] \to [1/x, 1/y, 1/z]$. What is the the domain of definition? (It is bigger than the locus where $xyz \neq 0$!) You will observe that you can extend it over "codimension 1 sets" (ignoring the fact that we don't yet know what codimension means). This again foreshadows the Curve-to-Projective Extension Theorem 16.5.1.

6.5.10. $\star$ **Complex curves that are not rational (fun but inessential).**

We now describe two examples of curves C that do not admit a nonconstant rational map from $\mathbb{P}^1_{\mathbb{C}}$. Both proofs are by Fermat's method of *infinite descent*. These results can be interpreted (as you will later be able to check using Theorem 17.4.3) as the fact that these curves have no "nontrivial" $\mathbb{C}(t)$-valued points, where by this, we mean that any $\mathbb{C}(t)$-valued point is secretly a $\mathbb{C}$-valued point. You may notice that if you consider the same examples with $\mathbb{C}(t)$ replaced by $\mathbb{Q}$ (and where C is a curve over $\mathbb{Q}$ rather than $\mathbb{C}$), you get two fundamental questions in number theory and geometry. The analog of Exercise 6.5.L is the question of rational points on elliptic curves, and you may realize that the analog of Exercise 6.5.J is even more famous. Also, the arithmetic analog of Exercise 6.5.L(a) is the "four squares theorem" (there are not four integer squares in arithmetic progression), first stated by Fermat. These examples will give you a glimpse of how and why facts over number fields are often parallelled by facts over function fields of curves. This parallelism is a recurring deep theme in the subject.

6.5.J. EXERCISE. If $n > 2$, show that $\mathbb{P}^1_{\mathbb{C}}$ has no dominant rational maps to the "Fermat curve" $x^n + y^n = z^n$ in $\mathbb{P}^2_{\mathbb{C}}$. Hint: reduce this to showing that there is no "nonconstant" solution $(f(t), g(t), h(t))$ to $f(t)^n + g(t)^n = h(t)^n$, where $f(t)$, $g(t)$, and $h(t)$ are rational functions in t. By clearing denominators, reduce this to showing that there is no nonconstant solution where $f(t)$, $g(t)$, and $h(t)$ are relatively prime polynomials. For this, assume there is a solution, and consider one of the lowest positive degree. Then use the fact that $\mathbb{C}[t]$ is a unique factorization domain, and $h(t)^n - g(t)^n = \prod_{i=1}^n (h(t) - \zeta^i g(t))$, where ζ is a primitive nth root of unity. Argue that each $h(t) - \zeta^i g(t)$ is an nth power. Then use

$$(h(t) - g(t)) + \alpha (h(t) - \zeta g(t)) = \beta \left(h(t) - \zeta^2 g(t)\right)$$

for suitably chosen α and β to get a solution of smaller degree. (How does this argument fail for $n = 2$?)

6.5.K. EXERCISE. Give two smooth complex curves X and Y so that no nonempty open subset of X is isomorphic to a nonempty open subset of Y. Hint: Exercise 6.5.J.

6.5.L. EXERCISE. Suppose a, b, and c are distinct complex numbers. By the following steps, show that if $x(t)$ and $y(t)$ are two rational functions of t (elements of $\mathbb{C}(t)$) such that

$$y(t)^2 = (x(t) - a)(x(t) - b)(x(t) - c), \tag{6.5.10.1}$$

then $x(t)$ and $y(t)$ are constants ($x(t), y(t) \in \mathbb{C}$). (Here $\mathbb{C}$ may be replaced by any field K of characteristic not 2; slight extra care is needed if K is not algebraically closed.)
(a) Suppose $P, Q \in \mathbb{C}[t]$ are relatively prime polynomials such that four linear combinations of them are perfect squares, no two of which are constant multiples of each other. Show that P and Q are constant (i.e., $P, Q \in \mathbb{C}$). Hint: By renaming P and Q, show that you may assume that the perfect squares are P, Q, $P - Q$, $P - \lambda Q$ (for some $\lambda \in \mathbb{C}$). Define u and v to be square roots of P and Q respectively. Show that $u - v$, $u + v$, $u - \sqrt{\lambda} v$, $u + \sqrt{\lambda} v$ are perfect squares, and that u and v are relatively prime. If P and Q are not both constant, note that $0 < \max(\deg u, \deg v) < \max(\deg P, \deg Q)$. Assume from the start that P and Q were chosen as a counterexample with minimal $\max(\deg P, \deg Q)$ to obtain a contradiction. (Aside: It is possible to have *three* distinct linear combinations that are perfect squares. Such examples essentially correspond to primitive Pythagorean triples in $\mathbb{C}(t)$ — can you see how?)
(b) Suppose $(x, y) = (p/q, r/s)$ is a solution to (6.5.10.1), where $p, q, r, s \in \mathbb{C}[t]$, and p/q and r/s are in lowest terms. Clear denominators to show that $r^2 q^3 = s^2(p - aq)(p - bq)(p - cq)$. Show that $s^2 | q^3$ and $q^3 | s^2$, and hence that $s^2 = \delta q^3$ for some $\delta \in \mathbb{C}$. From $r^2 = \delta(p - aq)(p - bq)(p - cq)$, show that $(p - aq)$, $(p - bq)$, $(p - cq)$ are perfect squares. Show that q is also a perfect square, and then apply part (a).

A much better geometric approach to Exercises 6.5.J and 6.5.L is given in Exercise 21.7.H.

6.6 ⋆ Representable functors and group schemes

6.6.1. Maps to $\mathbb{A}^1$ correspond to functions. If X is a scheme, there is a bijection between the maps $X \to \mathbb{A}^1$ and global sections of the structure sheaf: by Exercise 6.3.F, maps $\pi : X \to \mathbb{A}^1_{\mathbb{Z}}$ correspond to maps to ring maps $\pi^\sharp : \mathbb{Z}[t] \to \Gamma(X, \mathscr{O}_X)$, and $\pi^\sharp(t)$ is a function on X; this is reversible.

This map is very natural in an informal sense: you can even picture this map to $\mathbb{A}^1$ as being *given* by the function. (By analogy, a function on a manifold is a map to $\mathbb{R}$.) But it is natural in a more precise sense: this bijection is functorial in X. We will ponder this example at length, and see that it leads us to two important sophisticated notions: representable functors and group schemes.

6.6.A. EASY EXERCISE. Suppose X is a $\mathbb{C}$-scheme. Verify that there is a natural bijection between maps $X \to \mathbb{A}^1_{\mathbb{C}}$ in the category of $\mathbb{C}$-schemes and functions on X. (Here the base ring $\mathbb{C}$ can be replaced by any ring A.)

This interpretation can be extended to rational maps, as follows.

6.6.B. UNIMPORTANT EXERCISE. Interpret rational functions on an integral scheme (Exercise 5.5.Q, see also Definition 5.5.6) as rational maps to $\mathbb{A}^1_{\mathbb{Z}}$.

6.6.2. Representable functors. We restate the bijection of §6.6.1 as follows. We have two different contravariant functors from *Sch* to *Sets*: maps to $\mathbb{A}^1$ (i.e., $H : X \mapsto \mathrm{Mor}(X, \mathbb{A}^1_{\mathbb{Z}})$), and functions on X ($F : X \mapsto \Gamma(X, \mathscr{O}_X)$). The "naturality" of the bijection — the functoriality in X — is precisely the statement that the bijection gives a natural isomorphism of functors (§1.2.21): given any $\pi : X \to X'$, the diagram

$$\begin{tikzcd} H(X') \arrow[r, "H(\pi)"] \arrow[d] & H(X) \arrow[d] \\ F(X') \arrow[r, "F(\pi)"] & F(X) \end{tikzcd}$$

(where the vertical maps are the bijections given in §6.6.1) commutes.

More generally, if Y is an element of a category $\mathscr{C}$ (we care about the special case $\mathscr{C} = Sch$), recall the contravariant functor $h_Y : \mathscr{C} \to Sets$ defined by $h_Y(X) = \mathrm{Mor}(X, Y)$ (Example 1.2.20). We say a contravariant functor from $\mathscr{C}$ to *Sets* is **represented by** Y if it is naturally isomorphic to the functor h_Y. We say it is **representable** if it is represented by *some* Y.

The bijection of §6.6.1 may now be stated as: *the global section functor is represented by* $\mathbb{A}^1$.

6.6.C. IMPORTANT EASY EXERCISE (REPRESENTING OBJECTS ARE UNIQUE UP TO UNIQUE ISOMORPHISM). Show that if a contravariant functor F is represented by Y and by Z, then we have a unique isomorphism $Y \to Z$ induced by the natural isomorphism of functors $h_Y \to h_Z$. Hint: this is a version of the universal property arguments of §1.3: once again, we are recognizing an object (up to unique isomorphism) by maps to that object. This exercise is essentially Exercise 1.3.Z(b). (This extends readily to Yoneda's Lemma in this setting, Exercise 9.1.C. You are welcome to try that now.)

You have implicitly seen this notion before: you can interpret the existence of products and fibered products in a category as examples of representable functors. (You may wish to work out how a natural isomorphism $h_{Y \times Z} \cong h_Y \times h_Z$ induces the projection maps $Y \times Z \to Y$ and $Y \times Z \to Z$.)

6.6.D. EXERCISE (WARM-UP). Suppose F is the contravariant functor $Sch \to Sets$ defined by $F(X) = \{\text{Grothendieck}\}$ for all schemes X. Show that F is representable. (What is it representable by?)

6.6.E. EXERCISE. In this exercise, $\mathbb{Z}$ may be replaced by any ring.
(a) *(Affine* n*-space represents the functor of* n *functions.)* Show that the contravariant

functor from ($\mathbb{Z}$-)schemes to *Sets*

$$X \mapsto \{(f_1, \dots, f_n) : f_i \in \Gamma(X, \mathscr{O}_X)\}$$

is represented by $\mathbb{A}^n_{\mathbb{Z}}$. Show that $\mathbb{A}^1_{\mathbb{Z}} \times_{\mathbb{Z}} \mathbb{A}^1_{\mathbb{Z}} \cong \mathbb{A}^2_{\mathbb{Z}}$, i.e., that $\mathbb{A}^2$ satisfies the universal property of $\mathbb{A}^1 \times \mathbb{A}^1$. (You will undoubtedly be able to immediately show that $\prod \mathbb{A}^{m_i}_{\mathbb{Z}} \cong \mathbb{A}^{\sum m_i}_{\mathbb{Z}}$.)
(b) *(The functor of invertible functions is representable.)* Show that the contravariant functor from ($\mathbb{Z}$-)schemes to *Sets* taking X to invertible functions on X is representable by $\operatorname{Spec} \mathbb{Z}[t, t^{-1}]$.

6.6.3. *Definition.* The scheme defined in Exercise 6.6.E(b) is called the **multiplicative group** $\mathbb{G}_m$. "$\mathbb{G}_m$ over a field k" ("the multiplicative group over k") means $\operatorname{Spec} k[t, t^{-1}]$, with the same group operations. Better: it represents the group of invertible functions in the category of k-schemes. We can similarly define $\mathbb{G}_m$ over an arbitrary ring or even arbitrary scheme.

6.6.F. LESS IMPORTANT EXERCISE. Fix a ring A. Consider the functor H from the category of locally ringed spaces to *Sets* given by $H(X) = \{A \to \Gamma(X, \mathscr{O}_X)\}$. Show that this functor is representable (by $\operatorname{Spec} A$). This gives another (admittedly odd) motivation for the definition of $\operatorname{Spec} A$, closely related to that of §6.3.5.

6.6.4. ⋆⋆ **Group schemes (or more generally, group objects in a category).**

(The rest of §6.6 should be read only for entertainment.) We return again to Example 6.6.1. Functions on X are better than a set: they form a group. (Indeed they even form a ring, but we will worry about this later.) Given a morphism $X \to Y$, pullback of functions $\Gamma(Y, \mathscr{O}_Y) \to \Gamma(X, \mathscr{O}_X)$ is a group homomorphism. So we should expect $\mathbb{A}^1$ to have some group-like structure. This leads us to the notion of *group scheme*, or more generally a *group object* in a category, which we now define.

Suppose $\mathscr{C}$ is a category with a final object Z and with products. (We know that *Sch* has a final object $Z = \operatorname{Spec} \mathbb{Z}$, by Exercise 6.3.I. We will later see that it has products, §9.1. But in Exercise 6.6.K we will give an alternative characterization of group objects that applies in any category, so we won't worry about this.)

A **group object** in $\mathscr{C}$ is an element X along with three morphisms:

- *multiplication:* $m : X \times X \to X$
- *inverse:* $i : X \to X$
- *identity element:* $e : Z \to X$ (*not* the identity map)

These morphisms are required to satisfy several conditions.

(i) associativity axiom:

$$\begin{tikzcd} X \times X \times X \arrow[r, "(m, id)"] \arrow[d, "(id, m)"'] & X \times X \arrow[d, "m"] \\ X \times X \arrow[r, "m"] & X \end{tikzcd}$$

commutes. (Here id means the equality $X \to X$.)

(ii) identity axiom:

$$X \xrightarrow{\sim} Z \times X \xrightarrow{e \times id} X \times X \xrightarrow{m} X$$

and

$$X \xrightarrow{\sim} X \times Z \xrightarrow{\mathrm{id} \times e} X \times X \xrightarrow{m} X$$

are both the identity map $X \to X$. (This corresponds to the group axiom: "multiplication by the identity element is the identity map".)

(iii) inverse axiom: $X \xrightarrow{i, \mathrm{id}} X \times X \xrightarrow{m} X$ and $X \xrightarrow{\mathrm{id}, i} X \times X \xrightarrow{m} X$ are both the map that is the composition $X \longrightarrow Z \xrightarrow{e} X$.

As motivation, you can check that a group object in the category of sets is in fact the same thing as a group. (This is symptomatic of how you take some notion and make it categorical. You write down its axioms in a categorical way, and if all goes well, if you specialize to the category of sets, you get your original notion. You can apply this to the notion of "rings" in an exercise below.)

A **group scheme** is defined to be a group object in the category of schemes. A **group scheme** over a ring A (or a scheme S) is defined to be a group object in the category of A-schemes (or S-schemes).

6.6.G. EXERCISE. Give $\mathbb{A}^1_{\mathbb{Z}}$ the structure of a group scheme, by describing the three structural morphisms, and showing that they satisfy the axioms. (Hint: the morphisms should not be surprising. For example, inverse is given by $t \mapsto -t$. Note that we know that the product $\mathbb{A}^1_{\mathbb{Z}} \times \mathbb{A}^1_{\mathbb{Z}}$ exists, by Exercise 6.6.E(a).)

6.6.H. EXERCISE. Show that if G is a group object in a category $\mathscr{C}$, then for any $X \in \mathscr{C}$, $\mathrm{Mor}(X, G)$ has the structure of a group, and the group structure is preserved by pullback (i.e., $\mathrm{Mor}(\cdot, G)$ is a contravariant functor to *Groups*).

6.6.I. EXERCISE. Show that the group structure described by the previous exercise translates the group scheme structure on $\mathbb{A}^1_{\mathbb{Z}}$ to the group structure on $\Gamma(X, \mathscr{O}_X)$, via the bijection of §6.6.1.

6.6.J. EXERCISE. Define the notion of **abelian group scheme**, and **ring scheme**. (You will undoubtedly at the same time figure out how to define the notion of abelian group object and ring object in any category C. You may discover a more efficient approach to such questions after reading §6.6.5.)

The language of scheme-valued points (Definition 6.3.7) has the following advantage: notice that the *points* of a group scheme need not themselves form a group (consider $\mathbb{A}^1_{\mathbb{Z}}$). But Exercise 6.6.H shows that the Z-*valued points* of a group scheme (where Z is any given scheme) indeed form a group.

6.6.5. *Group schemes, more functorially.* There was something unsatisfactory about our discussion of the "group-respecting" nature of the bijection in §6.6.1: we observed that the right side (functions on X) formed a group, then we developed the axioms of a group scheme, then we cleverly figured out the maps that made $\mathbb{A}^1_{\mathbb{Z}}$ into a group scheme, then we showed that this induced a group structure on the left side of the bijection ($\mathrm{Mor}(X, \mathbb{A}^1)$) that precisely corresponded to the group structure on the right side (functions on X).

The picture is more cleanly explained as follows.

6.6.K. EXERCISE. Suppose we have a contravariant functor F from *Sch* (or indeed any category) to *Groups*. Suppose further that F composed with the forgetful functor *Groups* $\to$ *Sets* is represented by an object Y. Show that the group operations on $F(X)$ (as X varies through *Sch*) uniquely determine $m : Y \times Y \to Y$, $i : Y \to Y$, $e : Z \to Y$ satisfying the axioms defining a group scheme, such that the group operation on $\mathrm{Mor}(X, Y)$ is the same as that on $F(X)$.

In particular, the definition of a group object in a category was forced upon us by the definition of group. More generally, you should expect that any class of objects that can be interpreted as sets with additional structure should fit into this picture.

You should apply this exercise to $\mathbb{A}^1_{\mathbb{Z}}$, and see how the explicit formulas you found in Exercise 6.6.G are forced on you.

6.6.L. EXERCISE. Work out the maps m, i, and e in the group schemes of Exercise 6.6.E.

6.6.M. EXERCISE. Explain why the product of group objects in a category can be naturally interpreted as a group object in that category.

6.6.N. EXERCISE.
(a) Define **morphism of group schemes**.
(b) Recall that if A is a ring, then $GL_n(A)$ (the **general linear group over** A) is the group of invertible $n \times n$ matrices with entries in the ring A. Figure out the right definition of **the group scheme** GL_n **(over a ring** A**)**, and describe the **determinant map** $\det : GL_n \to \mathbb{G}_m$.
(c) Make sense of the statement: "$(\cdot^n) : \mathbb{G}_m \to \mathbb{G}_m$ given by $t \mapsto t^n$ is a morphism of group schemes."

The language of Exercise 6.6.N(a) suggests that group schemes form a category; feel free to prove this if you want. In fact, the category of group schemes has a zero object. What is it?

6.6.O. EXERCISE (KERNELS OF MAPS OF GROUP SCHEMES). Suppose $F : G_1 \to G_2$ is a morphism of group schemes. Consider the contravariant functor *Sch* $\to$ *Groups* given by $X \mapsto \ker(\mathrm{Mor}(X, G_1) \to \mathrm{Mor}(X, G_2))$. If this is representable, by a group scheme G_0, say, show that $G_0 \to G_1$ is the kernel of F in the category of group schemes.

6.6.P. EXERCISE. Show that the kernel of $(\cdot^n)$ (Exercise 6.6.N) is representable. If $n > 0$, show that over a field k of characteristic p dividing n, this group scheme is nonreduced. (This group scheme, denoted μ_n, is important, although we will not use it.)

6.6.Q. EXERCISE. Show that the kernel of $\det : GL_n \to \mathbb{G}_m$ is representable. This is the group scheme SL_n. (You can do this over $\mathbb{Z}$, or over a field k, or even over an arbitrary ring A; the algebra is the same.)

6.6.R. EXERCISE. Show (as easily as possible) that $\mathbb{A}^1_k$ is a ring k-scheme. (Here k can be replaced by any ring.)

6.6.S. EXERCISE.
(a) Define the notion of a (left) **group scheme action** (of a group scheme on a scheme).
(b) Suppose A is a ring. Show that specifying an integer-valued grading on A is equivalent to specifying an action of $\mathbb{G}_m$ on $\operatorname{Spec} A$. (This interpretation of a grading is surprisingly enlightening. Caution: there are two possible choices of the integer-valued grading, and there are reasons for both. Both are used in the literature.)

6.6.6. *Aside: Hopf algebras.* Here is a notion that we won't use, but it is easy enough to define now. Suppose $G = \operatorname{Spec} A$ is an affine group scheme, i.e., a group scheme that is an affine scheme. The categorical definition of group scheme can be restated in terms of the ring A. (This requires thinking through Remark 6.3.6; see Exercise 9.1.B.) Then these axioms define a **Hopf algebra**. For example, we have a "comultiplication map" $A \to A \otimes A$.

6.6.T. EXERCISE. As $\mathbb{A}^1_{\mathbb{Z}}$ is a group scheme, $\mathbb{Z}[t]$ has a Hopf algebra structure. Describe the comultiplication map $\mathbb{Z}[t] \to \mathbb{Z}[t] \otimes_{\mathbb{Z}} \mathbb{Z}[t]$.

6.7 ⋆⋆ The Grassmannian (initial construction)

The Grassmannian is a useful geometric construction that is "the geometric object underlying linear algebra". In (classical) geometry over a field $K = \mathbb{R}$ or $\mathbb{C}$, just as projective space parametrizes one-dimensional subspaces of a given n-dimensional vector space, the Grassmannian parametrizes k-dimensional subspaces of n-dimensional space. The Grassmannian $G(k, n)$ is a manifold of dimension $k(n-k)$ (over the field). The manifold structure is given as follows. Given a basis $(v_1, \ldots, v_n)$ of n-space, "most" k-planes can be described as the span of the k vectors

$$\left\langle v_1 + \sum_{i=k+1}^{n} a_{1i} v_i, v_2 + \sum_{i=k+1}^{n} a_{2i} v_i, \ldots, v_k + \sum_{i=k+1}^{n} a_{ki} v_i \right\rangle . \tag{6.7.0.1}$$

(Can you describe which k-planes are *not* of this form? Hint: row reduced echelon form. Aside: the stratification of $G(k, n)$ by normal form is the decomposition of the Grassmannian into *Schubert cells*. You may be able to show using the normal form that each Schubert cell is isomorphic to an affine space.) Any k-plane of this form can be described in such a way uniquely. We use this to identify those k-planes of this form with the manifold $K^{k(n-k)}$ (with coordinates a_{ji}). This is a large affine patch on the Grassmannian (called the "open Schubert cell" with respect to this basis). As the v_i vary, these patches cover the Grassmannian (why?), and the manifold structures agree (a harder fact).

We now *define* the Grassmannian in algebraic geometry, over a ring A. Suppose $v = (v_1, \ldots, v_n)$ is a basis for $A^{\oplus n}$. More precisely: $v_i \in A^{\oplus n}$, and the map $A^{\oplus n} \to A^{\oplus n}$ given by $(a_1, \ldots, a_n) \mapsto a_1 v_1 + \cdots + a_n v_n$ is an isomorphism.

6.7.A. EXERCISE. Show that any two bases are related by an invertible $n \times n$ matrix over A — a matrix with entries in A whose determinant is an invertible element of A.

For each such basis v, we consider the scheme $U_v \cong \mathbb{A}_A^{k(n-k)}$, with coordinates a_{ji} ($k+1 \leq i \leq n$, $1 \leq j \leq k$), which we imagine as corresponding to the k-plane spanned by the vectors (6.7.0.1).

6.7.B. EXERCISE. Given two bases v and w, explain how to glue U_v to U_w along appropriate open sets. You may find it convenient to work with coordinates a_{ji} where i runs from 1 to n, not just $k+1$ to n, but imposing $a_{ji} = \delta_{ji}$ (i.e., 1 when $i = j$ and 0 otherwise) when $i \leq k$. This convention is analogous to coordinates $x_{i/j}$ on the patches of projective space (§4.4.9). Hint: the relevant open subset of U_v will be where a certain determinant doesn't vanish.

6.7.C. EXERCISE/DEFINITION. By checking triple intersections, verify that these patches (over all possible bases) glue together to a single scheme (Exercise 4.4.A). This is the **Grassmannian** $G(k,n)$ over the ring A. Because it can be interpreted as a space of linear "$\mathbb{P}_A^{k-1}$'s" in $\mathbb{P}_A^{n-1}$, it is often also written $\mathbb{G}(k-1, n-1)$. (You will see that this is wise notation in Exercise 11.2.J, for example.)

Although this definition is pleasantly explicit (it is immediate that the Grassmannian is covered by $\mathbb{A}^{k(n-k)}$'s), and perhaps more "natural" than our original definition of projective space in §4.4.9 (we aren't making a choice of basis; we use *all* bases), there are several things unsatisfactory about this definition of the Grassmannian. In fact the Grassmannian is always projective; this isn't obvious with this definition. Furthermore, the Grassmannian comes with a natural closed embedding into $\mathbb{P}^{\binom{n}{k}-1}$ (the *Plücker embedding*). Finally, there is an action of GL_n on the space of k-planes in n-space, so we should be able to see this in our algebraic incarnation. We will address these issues in §16.7, by giving a better description, as a moduli space.

6.7.1. *(Partial) flag varieties.* Just as the Grassmannian "parametrizes" k-planes in n-space, the **flag variety** parametrizes "flags" nested sequences of subspaces of n-space

$$F_0 \subset F_1 \subset \cdots \subset F_n$$

where $\dim F_i = i$. Generalizing both of these is the notion of a **partial flag variety** associated to some data $0 \leq a_1 < \cdots < a_\ell \leq n$, which parametrizes nested sequences of subspaces of n-space

$$F_{a_1} \subset \cdots \subset F_{a_\ell}$$

where $\dim F_{a_i} = a_i$. You should be able to generalize all of the discussion in §6.7 to this setting.

CHAPTER 7

Useful classes of morphisms of schemes

We now define an excessive number of types of morphisms. Some (often finiteness properties) are useful because *every* "reasonable" morphism has such properties, and they will be used in proofs in obvious ways. Others correspond to geometric behavior, and you should have a picture of what each means.

7.0.2. One of Grothendieck's lessons is that things that we often think of as properties of *objects* are better understood as properties of *morphisms*. One way of turning properties of objects into properties of morphisms is as follows. If P is a property of schemes, we often (but not always) say that a *morphism* $\pi : X \to Y$ has P if for every affine open subset $U \subset Y$, $\pi^{-1}(U)$ has P. We will see this for P = quasicompact, quasiseparated, affine, and more. (As you might hope, in good circumstances, P will satisfy the hypotheses of the Affine Communication Lemma 5.3.2, so we don't have to check *every* affine open subset.) Informally, you can think of such a morphism as one where all the fibers have P. (You can quickly define the fiber of a morphism as a topological space, but once we define fiber product, we will define the *scheme-theoretic* fiber, and then this discussion will make sense.) But it means more than that: it means that "being P" is really not just fiber-by-fiber, but behaves well as the fiber varies. (For comparison, a smooth morphism of manifolds means more than that the fibers are smooth.)

7.1 An example of a reasonable class of morphisms: Open embeddings

7.1.1. What to expect of any "reasonable" type of morphism. You will notice that essentially all classes of morphisms have three properties.

(i) They are **local on the target**. In other words, (a) if $\pi : X \to Y$ is in the class, then for any open subset V of Y, the restricted morphism $\pi^{-1}(V) \to V$ is in the class; and (b) for a morphism $\pi : X \to Y$, if there is an open cover $\{V_i\}$ of Y for which each restricted morphism $\pi^{-1}(V_i) \to V_i$ is in the class, then π is in the class. In particular, as schemes are built out of affine schemes, it should be possible to check on an affine cover, as described in §7.0.2.

(ii) They are closed under composition: if $\pi : X \to Y$ and $\rho : Y \to Z$ are both in this class, then so is $\rho \circ \pi$.

(iii) They are closed under "base change" or "pullback" or "fibered product". We will discuss fibered product of schemes in Chapter 9.1.

When anyone tells you a new class of morphism, you should immediately ask yourself (or them) whether these three properties hold. And it is essentially true that a class of morphism is "reasonable" if and only if it satisfies these three properties. Here is a first example.

An **open embedding** (or **open immersion**) of schemes is defined to be an open embedding as ringed spaces (§6.2.1). In other words, a morphism $\pi : (X, \mathscr{O}_X) \to (Y, \mathscr{O}_Y)$ of schemes is an open embedding if π factors as

$$(X, \mathscr{O}_X) \xrightarrow[\sim]{\rho} (U, \mathscr{O}_Y|_U) \xhookrightarrow{\tau} (Y, \mathscr{O}_Y)$$

where ρ is an isomorphism, and $\tau : U \hookrightarrow Y$ is an inclusion of an open set. It is immediate that isomorphisms are open embeddings. We often sloppily say that $(X, \mathscr{O}_X)$ is an *open subscheme* of $(Y, \mathscr{O}_Y)$. The symbol $\hookrightarrow$ is often used to indicate that a morphism is an open embedding (or more generally, a locally closed embedding, see §8.1.2). This is a bit confusing, and not too important: at the level of sets, open subschemes are subsets, while open embeddings are bijections onto subsets.

7.1.A. Exercise (properties (i) and (ii)). Verify that the class of open embeddings satisfies properties (i) and (ii) of §7.1.1.

7.1.B. Important but easy exercise (property (iii)). Verify that the class of open embeddings satisfies property (iii) of §7.1.1. More specifically: suppose $i : U \to Z$ is an open embedding, and $\rho : Y \to Z$ is any morphism. Show that $U \times_Z Y$ exists and $U \times_Z Y \to Y$ is an open embedding. (Hint: I'll even tell you what $U \times_Z Y$ is: $(\rho^{-1}(U), \mathscr{O}_Y|_{\rho^{-1}(U)})$.) In particular, if $U \hookrightarrow Z$ and $V \hookrightarrow Z$ are open embeddings, $U \times_Z V \cong U \cap V$.

7.1.C. Easy Exercise. Suppose $\rho : X \to Y$ is an open embedding. Show that if Y is locally Noetherian, then X is too. Show that if Y is Noetherian, then X is too. However, show that if Y is quasicompact, X need not be. (Hint: let Y be affine but not Noetherian, see Exercise 3.6.G(b).)

"Open embeddings" are scheme-theoretic analogs of open subsets. "Closed embeddings" are scheme-theoretic analogs of closed subsets, but they have a surprisingly different flavor, as we will see in §8.1.

7.1.2. *Definition.* In analogy with "local on the target" (§7.1.1), we define what it means for a property P of morphisms to be **local on the source**: to check if a morphism $\pi : X \to Y$ has P, it suffices to check on any open cover $\{U_i\}$ of X. We then define **affine-local on the source** (and **affine-local on the target**) in the obvious way: it suffices to check on any affine cover of the source (resp. target). The use of "affine-local" rather than "local" is to emphasize that the criterion on affine schemes is simple to describe.

7.1.D. Exercise (practice with the concept). Show that the notion of "open embedding" is not local on the source.

7.2 Algebraic interlude: Lying Over and Nakayama

Algebra is the offer made by the devil to the mathematician. The devil says: I will give you this powerful machine, it will answer any question you like. All you need to do is give me your soul: give up geometry and you will have this marvelous machine.
— *M. Atiyah,* [**At2**, p. 659]; *but see the Atiyah quote at the start of* §0.3

To set up our discussion in the next section on integral morphisms, we develop some algebraic preliminaries. A clever trick we use can also be used to show Nakayama's Lemma, so we discuss this as well.

Suppose $\phi : B \to A$ is a ring morphism. We say $a \in A$ is **integral** over B if a satisfies some monic polynomial

$$a^n + ?a^{n-1} + \cdots + ? = 0$$

where the coefficients lie in $\phi(B)$. A ring *homomorphism* $\phi : B \to A$ is **integral** if every element of A is integral over $\phi(B)$. An integral ring morphism ϕ is an **integral** *extension* if ϕ is an *inclusion* of rings. You should think of integral homomorphisms and integral extensions as ring-theoretic generalizations of the notion of algebraic extensions of fields.

7.2.A. EXERCISE. Show that if $\phi : B \to A$ is a ring morphism, $(b_1, \ldots, b_n) = 1$ in B, and $B_{b_i} \to A_{\phi(b_i)}$ is integral for all i, then ϕ is integral. Hint: replace B by $\phi(B)$ to reduce to the case where B is a subring of A. Suppose $a \in A$. Show that there is some t and m such that $b_i^t a^m \in B + Ba + Ba^2 + \cdots + Ba^{m-1}$ for some t and m independent of i. Use a "partition of unity" argument as in the proof of Theorem 4.1.2 to show that $a^m \in B + Ba + Ba^2 + \cdots + Ba^{m-1}$.

7.2.B. EXERCISE.
(a) Show that the property of a *homomorphism* $\phi : B \to A$ being integral is always preserved by localization and quotient of B, and quotient of A, but not localization of A. More precisely: suppose ϕ is integral. Show that the induced maps $T^{-1}B \to \phi(T)^{-1}A$, $B/J \to A/\phi(J)A$, and $B \to A/I$ are integral (where T is a multiplicative subset of B, J is an ideal of B, and I is an ideal of A), but $B \to S^{-1}A$ need not be integral (where S is a multiplicative subset of A). (Hint for the latter: show that $k[t] \to k[t]$ is an integral homomorphism, but $k[t] \to k[t]_{(t)}$ is not.)
(b) Show that the property of ϕ being an integral *extension* is preserved by localization of B, but not localization or quotient of A. (Hint for the latter: $k[t] \to k[t]$ is an integral extension, but $k[t] \to k[t]/(t)$ is not.)
(c) In fact the property of ϕ being an integral *extension* (as opposed to integral *homomorphism*) is not preserved by taking quotients of B either. (Let $B = k[x,y]/(y^2)$ and $A = k[x,y,z]/(z^2, xz - y)$. Then B injects into A, but $B/(x)$ doesn't inject into $A/(x)$.) But it is in some cases. Suppose $\phi : B \to A$ is an integral extension, $J \subset B$ is the restriction of an ideal $I \subset A$. (Side Remark: you can show that this holds if J is prime.) Show that the induced map $B/J \to A/JA$ is an integral extension. (Hint: show that the composition $B/J \to A/JA \to A/I$ is an injection.)

The following lemma uses a useful but sneaky trick.

7.2.1. Lemma. — *Suppose $\phi : B \to A$ is a ring homomorphism. Then $a \in A$ is integral over B if and only if it is contained in a subalgebra of A that is a finitely generated B-module.*

Proof. If a satisfies a monic polynomial equation of degree n, then the B-submodule of A generated by 1, a, …, a^{n-1} is closed under multiplication, and hence a subalgebra of A.

Assume conversely that a is contained in a subalgebra A' of A that is a finitely generated B-module. Choose a finite generating set m_1, …, m_n of A' (as a B-module). Then $a m_i = \sum b_{ij} m_j$, for some $b_{ij} \in B$. Thus

$$(7.2.1.1) \qquad \left(a \mathrm{Id}_{n \times n} - [b_{ij}]_{ij}\right) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

where Id_n is the $n \times n$ identity matrix in A. We can't invert the matrix $(a\mathrm{Id}_{n\times n} - [b_{ij}]_{ij})$, but we almost can. Recall that an $n \times n$ matrix M has an *adjugate matrix* $\mathrm{adj}(M)$ such that $\mathrm{adj}(M)M = \det(M)\mathrm{Id}_n$. (The (i,j)th entry of $\mathrm{adj}(M)$ is the determinant of the matrix obtained from M by deleting the ith column and jth row, times $(-1)^{i+j}$. You have likely seen this in the form of a formula for M^{-1} when there *is* an inverse; see for example [**DF**, p. 440].) The coefficients of $\mathrm{adj}(M)$ are polynomials in the coefficients of M. Multiplying (7.2.1.1) by $\mathrm{adj}(a\mathrm{Id}_{n\times n} - [b_{ij}]_{ij})$, we get

$$\det(a\mathrm{Id}_{n\times n} - [b_{ij}]_{ij}) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

So $\det(aI - [b_{ij}])$ annihilates the generating elements m_i, and hence every element of A', i.e., $\det(aI - [b_{ij}]) = 0$. But expanding the determinant yields an integral equation for a with coefficients in B. □

7.2.2. Corollary (finite implies integral). — *If A is a finite B-algebra (a finitely generated B-module), then ϕ is an integral homomorphism.*

The converse is false: integral does not imply finite, as $\mathbb{Q} \hookrightarrow \overline{\mathbb{Q}}$ is an integral homomorphism, but $\overline{\mathbb{Q}}$ is not a finite $\mathbb{Q}$-module. (A field extension is integral if it is algebraic.)

7.2.C. EXERCISE. Show that if $C \to B$ and $B \to A$ are both integral homomorphisms, then so is their composition.

7.2.D. EXERCISE. Suppose $\phi : B \to A$ is a ring morphism. Show that the elements of A integral over B form a subalgebra of A.

7.2.3. *Remark: Transcendence theory.* These ideas lead to the main facts about transcendence theory we will need for a discussion of dimension of varieties, see Exercise/Definition 11.2.A.

7.2.4. The Lying Over and Going-Up Theorems. The Lying Over Theorem is a useful property of integral extensions.

7.2.5. The Lying Over Theorem. — *Suppose $\phi : B \to A$ is an integral* extension. *Then for any prime ideal $\mathfrak{q} \subset B$, there is a prime ideal $\mathfrak{p} \subset A$ such that $\mathfrak{p} \cap B = \mathfrak{q}$.*

To be clear on how weak the hypotheses are: B need not be Noetherian, and A need not be finitely generated over B.

7.2.6. Geometric translation: $\operatorname{Spec} A \to \operatorname{Spec} B$ is surjective. (A map of schemes is **surjective** if the underlying map of sets is surjective.)

Although this is a theorem in algebra, the name can be interpreted geometrically: the theorem asserts that the corresponding morphism of schemes is surjective, and that "above" every prime $\mathfrak{q}$ "downstairs", there is a prime $\mathfrak{p}$ "upstairs", see Figure 7.1. (For this reason, it is often said that $\mathfrak{p}$ "lies over" $\mathfrak{q}$ if $\mathfrak{p} \cap B = \mathfrak{q}$.) The following exercise sets up the proof.

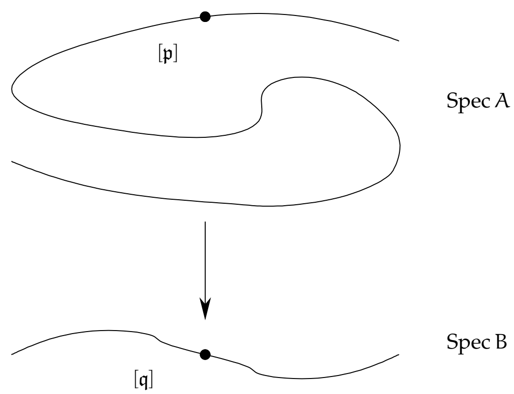


FIGURE 7.1. A picture of the Lying Over Theorem 7.2.5: if $\phi : B \to A$ is an integral extension, then $\operatorname{Spec} A \to \operatorname{Spec} B$ is surjective

7.2.E. ⋆ EXERCISE. Show that the special case where A is a field translates to: if $B \subset A$ is a subring with A integral over B, then B is a field. Prove this. (Hint: you must show that all nonzero elements in B have inverses in B. Here is the start: If $b \in B$, then $1/b \in A$, and this satisfies some integral equation over B.)

⋆ *Proof of the Lying Over Theorem 7.2.5.* We first make a reduction: by localizing at $\mathfrak{q}$ (preserving integrality by Exercise 7.2.B(b)), we can assume that $(B, \mathfrak{q})$ is a local ring. Then let $\mathfrak{p}$ be any *maximal* ideal of A. Consider the following diagram.

$$\begin{tikzcd} A \arrow[r, two heads] & A/\mathfrak{p} & \text{field} \\ B \arrow[u, hook] \arrow[r, two heads] & B/(\mathfrak{p} \cap B) \arrow[u, hook] & \end{tikzcd}$$

The right vertical arrow is an integral extension by Exercise 7.2.B(c). By Exercise 7.2.E, $B/(\mathfrak{p} \cap B)$ is a field too, so $\mathfrak{p} \cap B$ is a maximal ideal, hence it is $\mathfrak{q}$. □

7.2.F. IMPORTANT EXERCISE (THE GOING-UP THEOREM).
(a) Suppose $\phi : B \to A$ is an integral *homomorphism* (not necessarily an integral extension). Show that if $\mathfrak{q}_1 \subset \mathfrak{q}_2 \subset \cdots \subset \mathfrak{q}_n$ is a chain of prime ideals of B, and

$\mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_m$ is a chain of prime ideals of A such that $\mathfrak{p}_i$ "lies over" $\mathfrak{q}_i$ (and $1 \leq m < n$), then the second chain can be extended to $\mathfrak{p}_1 \subset \cdots \subset \mathfrak{p}_n$ so that this remains true. (Hint: reduce to the case $m = 1, n = 2$; reduce to the case where $\mathfrak{q}_1 = (0)$ and $\mathfrak{p}_1 = (0)$; use the Lying Over Theorem.)
(b) Draw a picture of this theorem.

There are analogous "Going-Down" results (requiring quite different hypotheses); see for example Theorem 11.2.12 and Exercise 24.5.E.

7.2.7. Nakayama's Lemma.
The trick in the proof of Lemma 7.2.1 can be used to quickly prove Nakayama's Lemma, which we will use repeatedly in the future. This name is used for several different but related results, which we discuss here. (A geometric interpretation will be given in Exercise 13.7.E.) We may as well prove it while the trick is fresh in our minds.

7.2.8. Nakayama's Lemma version 1. — *Suppose A is a ring, I is an ideal of A, and M is a finitely generated A-module, such that $M = IM$. Then there exists an $a \in A$ with $a \equiv 1 \pmod{I}$ with $aM = 0$.*

Proof. Say M is generated by $m_1, \ldots, m_n$. Then as $M = IM$, we have $m_i = \sum_j a_{ij} m_j$ for some $a_{ij} \in I$. Thus

$$(\mathrm{Id}_n - Z) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = 0 \qquad (7.2.8.1)$$

where $Z = (a_{ij})$. Multiplying both sides of (7.2.8.1) on the left by $\operatorname{adj}(\mathrm{Id}_n - Z)$, we obtain

$$\det(\mathrm{Id}_n - Z) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = 0.$$

But when you expand out $\det(\mathrm{Id}_n - Z)$, as Z has entries in I, you get something that is $1 \pmod{I}$. □

Here is why you care. Suppose I is contained in all maximal ideals of A. (The intersection of all the maximal ideals is called the *Jacobson radical*, but we won't use this phrase. For comparison, recall that the nilradical was the intersection of the *prime ideals* of A.) Then any $a \equiv 1 \pmod{I}$ is invertible. (We are not using Nakayama yet!) Reason: otherwise $(a) \neq A$, so the ideal (a) is contained in some maximal ideal $\mathfrak{m}$ — but $a \equiv 1 \pmod{\mathfrak{m}}$, contradiction. As a is invertible, we have the following.

7.2.9. Nakayama's Lemma version 2. — *Suppose A is a ring, I is an ideal of A contained in all maximal ideals, and M is a finitely generated A-module. (The most interesting case is when A is a local ring, and I is the maximal ideal.) Suppose $M = IM$. Then $M = 0$.*

7.2.G. EXERCISE (NAKAYAMA'S LEMMA VERSION 3). Suppose A is a ring, and I is an ideal of A contained in all maximal ideals. Suppose M is a *finitely generated*

A-module, and $N \subset M$ is a submodule. If $N/IN \to M/IM$ is surjective, then $M = N$.

7.2.H. IMPORTANT EXERCISE (NAKAYAMA'S LEMMA VERSION 4: GENERATORS OF $M/\mathfrak{m}M$ LIFT TO GENERATORS OF M). Suppose $(A, \mathfrak{m})$ is a local ring. Suppose M is a finitely generated A-module, and $f_1, \dots, f_n \in M$, with (the images of) $f_1, \dots, f_n$ generating $M/\mathfrak{m}M$. Then $f_1, \dots, f_n$ generate M. (In particular, taking $M = \mathfrak{m}$, if we have generators of $\mathfrak{m}/\mathfrak{m}^2$, they also generate $\mathfrak{m}$.)

7.2.I. IMPORTANT EXERCISE GENERALIZING LEMMA 7.2.1. Recall that a B-module N is said to be **faithful** if the only element of B acting on N by the identity is 1 (or equivalently, if the only element of B acting as the 0-map on N is 0). Suppose S is a subring of a ring A, and $r \in A$. Suppose there is a faithful $S[r]$-module M that is finitely generated as an S-module. Show that r is integral over S. (Hint: change a few words in the proof of version 1 of Nakayama, Lemma 7.2.8.)

7.2.J. EXERCISE. Suppose A is an integral domain, and $\widetilde{A}$ is the integral closure of A in $K(A)$, i.e., those elements of $K(A)$ integral over A, which form a subalgebra by Exercise 7.2.D. Show that $\widetilde{A}$ is integrally closed in $K(\widetilde{A}) = K(A)$.

7.3 A gazillion finiteness conditions on morphisms

By the end of this section, you will have seen the following types of morphisms: quasicompact, quasiseparated, affine, finite, integral, closed, (locally) of finite type, quasifinite — and possibly, (locally) of finite presentation.

7.3.1. Quasicompact and quasiseparated morphisms.

A morphism $\pi : X \to Y$ of schemes is **quasicompact** if for every open affine subset U of Y, $\pi^{-1}(U)$ is quasicompact. (Equivalently, the preimage of any quasicompact open subset is quasicompact. This is the right definition in other parts of geometry.)

We will like this notion because (i) finite sets have advantages over infinite sets (e.g. a finite set of integers has a maximum; also, things can be proved inductively), and (ii) most reasonable schemes will be quasicompact.

Along with quasicompactness comes the weird notion of quasiseparatedness. A morphism $\pi : X \to Y$ is **quasiseparated** if for every affine open subset U of Y, $\pi^{-1}(U)$ is a quasiseparated scheme (§5.1.1). This will be a useful hypothesis in theorems, usually in conjunction with quasicompactness. (For this reason, "quasicompact and quasiseparated" are often abbreviated as *qcqs*, as we do in the name of the Qcqs Lemma 7.3.5.) Various interesting kinds of morphisms (locally Noetherian source, affine, separated, see Exercises 7.3.B(b), 7.3.D, and 10.1.H resp.) are quasiseparated, and having the word "quasiseparated" will allow us to state theorems more succinctly.

7.3.A. EASY EXERCISE. Show that the composition of two quasicompact morphisms is quasicompact. (It is also true that the composition of two quasiseparated morphisms is quasiseparated. This is not impossible to show directly, but will in any case follow easily once we understand it in a more sophisticated way, see Proposition 10.1.13(b).)

7.3.B. EASY EXERCISE.
(a) Show that any morphism from a Noetherian scheme is quasicompact.
(b) Show that any morphism from a quasiseparated scheme is quasiseparated. Thus by Exercise 5.3.A, any morphism from a locally Noetherian scheme is quasiseparated. Thus those readers working only with locally Noetherian schemes may take quasiseparatedness as a standing hypothesis.

7.3.2. *Caution.* The two parts of the Exercise 7.3.B may lead you to suspect that any morphism $\pi : X \to Y$ with quasicompact source and target is necessarily quasicompact. This is false, and you may verify that the following is a counterexample. Let Z be the nonquasiseparated scheme constructed in Exercise 5.1.J, and let $X = \operatorname{Spec} k[x_1, x_2, \dots]$ as in Exercise 5.1.J. The obvious projection $\pi : Z \to X$ is not quasicompact.

7.3.C. EXERCISE. (Obvious hint for both parts: the Affine Communication Lemma 5.3.2.)
(a) *(quasicompactness is affine-local on the target)* Show that a morphism $\pi : X \to Y$ is quasicompact if there is a cover of Y by open affine sets U_i such that $\pi^{-1}(U_i)$ is quasicompact.
(b) *(quasiseparatedness is affine-local on the target)* Show that a morphism $\pi : X \to Y$ is quasiseparated if there is a cover of Y by open affine sets U_i such that $\pi^{-1}(U_i)$ is quasiseparated.

Following Grothendieck's philosophy of thinking that the important notions are properties of morphisms, not of objects (§7.0.2), we can restate the definition of quasicompact (resp. quasiseparated) scheme as a scheme that is quasicompact (resp. quasiseparated) over the final object $\operatorname{Spec} \mathbb{Z}$ in the category of schemes (Exercise 6.3.I).

7.3.3. Affine morphisms.
A morphism $\pi : X \to Y$ is **affine** if for every affine open set U of Y, $\pi^{-1}(U)$ (interpreted as an open subscheme of X) is an affine scheme.

7.3.D. FAST EXERCISE. Show that affine morphisms are quasicompact and quasiseparated. (Hint for the second: Exercise 5.1.G.)

7.3.4. Proposition (the property of "affineness" is affine-local on the target). — *A morphism $\pi : X \to Y$ is affine if there is a cover of Y by affine open sets U such that $\pi^{-1}(U)$ is affine.*

This proof is the hardest part of this section. For part of the proof (which will start in §7.3.7), it will be handy to have a lemma.

7.3.5. Qcqs Lemma. — *If X is a quasicompact quasiseparated scheme and $s \in \Gamma(X, \mathscr{O}_X)$, then the natural map $\Gamma(X, \mathscr{O}_X)_s \to \Gamma(X_s, \mathscr{O}_X)$ is an isomorphism.*

Here X_s means the locus on X where s doesn't vanish. (By Exercise 4.3.G(a), X_s is open.) We avoid the notation $D(s)$ to avoid any suggestion that X is affine. (If X is affine, then X_s is $D(s)$, and we already know the theorem is true — do you see why?)

7.3.E. EXERCISE (REALITY CHECK). What is the natural map $\Gamma(X, \mathscr{O}_X)_s \to \Gamma(X_s, \mathscr{O}_X)$ of the Qcqs Lemma 7.3.5? (Hint: the universal property of localization, Exercise 1.3.D.)

To repeat the earlier reassuring comment on the "quasicompact quasiseparated" hypothesis: this just means that X can be covered by a finite number of affine open subsets, any two of which have intersection also covered by a finite number of affine open subsets (Exercise 5.1.H). The hypothesis applies in lots of interesting situations, such as if X is affine (Exercise 5.1.G) or Noetherian (Exercise 5.3.A). And conversely, whenever you see quasicompact quasiseparated hypotheses (e.g. Exercises 13.3.E, 13.3.H), they are most likely there because of this lemma. To remind ourselves of this fact, we call it the Qcqs Lemma.

7.3.6. *Proof of the Qcqs Lemma 7.3.5.* Cover X with finitely many affine open sets $U_i = \operatorname{Spec} A_i$. Let $U_{ij} = U_i \cap U_j$. Then

$$0 \to \Gamma(X, \mathscr{O}_X) \to \prod_i A_i \to \prod_{\{i,j\}} \Gamma(U_{ij}, \mathscr{O}_X)$$

is exact. (See the discussion after (4.1.3.1) for the signs arising in the last map.) By the quasiseparated hypotheses, we can cover each U_{ij} with a finite number of affine open sets $U_{ijk} = \operatorname{Spec} A_{ijk}$, so we have that

$$0 \to \Gamma(X, \mathscr{O}_X) \to \prod_i A_i \to \prod_{\{i,j\},k} A_{ijk}$$

is exact. Localizing at s (an exact functor, Exercise 1.6.F(a)) gives

$$0 \to \Gamma(X, \mathscr{O}_X)_s \to \left(\prod_i A_i\right)_s \to \left(\prod_{\{i,j\},k} A_{ijk}\right)_s$$

As localization commutes with *finite* products (Exercise 1.3.F(a)),

$$0 \to \Gamma(X, \mathscr{O}_X)_s \to \prod_i (A_i)_{s_i} \to \prod_{\{i,j\},k} (A_{ijk})_{s_{ijk}} \tag{7.3.6.1}$$

is exact, where the global function s induces functions $s_i \in A_i$ and $s_{ijk} \in A_{ijk}$.

But similarly, the scheme X_s can be covered by affine opens $\operatorname{Spec}(A_i)_{s_i}$, and $\operatorname{Spec}(A_i)_{s_i} \cap \operatorname{Spec}(A_j)_{s_j}$ are covered by a finite number of affine opens $\operatorname{Spec}(A_{ijk})_{s_{ijk}}$, so we have

$$0 \to \Gamma(X_s, \mathscr{O}_X) \to \prod_i (A_i)_{s_i} \to \prod_{\{i,j\},k} (A_{ijk})_{s_{ijk}}. \tag{7.3.6.2}$$

Notice that the maps $\prod_i (A_i)_{s_i} \to \prod_{\{i,j\},k} (A_{ijk})_{s_{ijk}}$ in (7.3.6.1) and (7.3.6.2) are the same, and we have described the kernel of the map in two ways, so $\Gamma(X, \mathscr{O}_X)_s \to \Gamma(X_s, \mathscr{O}_X)$ is indeed an isomorphism. (Notice how the quasicompact and quasiseparated hypotheses were used in an easy way: to obtain finite products, which would commute with localization.) $\square$

7.3.7. *Proof of Proposition 7.3.4.* As usual, we use the Affine Communication Lemma 5.3.2. (We apply it to the condition "π is affine over".) We check our two criteria. First, suppose $\pi : X \to Y$ is affine over $\operatorname{Spec} B$, i.e., $\pi^{-1}(\operatorname{Spec} B) = \operatorname{Spec} A$. Then for any $s \in B$, $\pi^{-1}(\operatorname{Spec} B_s) = \operatorname{Spec} A_{\pi^\sharp s}$.

Second, suppose we are given $\pi : X \to \operatorname{Spec} B$ and $(s_1, \ldots, s_n) = B$ with $X_{\pi^\sharp s_i}$ affine ($\operatorname{Spec} A_i$, say). (As in the statement of the Qcqs Lemma 7.3.5, $X_{\pi^\sharp s_i}$ is the subset of X where $\pi^\sharp s_i$ doesn't vanish.) We wish to show that X is affine too. Let $A = \Gamma(X, \mathscr{O}_X)$. Then $X \to \operatorname{Spec} B$ factors through the tautological map $\alpha : X \to \operatorname{Spec} A$ (arising from the (iso)morphism $A \to \Gamma(X, \mathscr{O}_X)$, Exercise 6.3.F).

$$\begin{tikzcd} \cup_i X_{\pi^\sharp s_i} = X \arrow[rr, "\alpha"] \arrow[dr, "\pi"'] & & \operatorname{Spec} A \arrow[dl, "\beta"] \\ & \cup_i D(s_i) = \operatorname{Spec} B & \end{tikzcd}$$

Then $\beta^{-1}(D(s_i)) = D(\beta^\sharp s_i) \cong \operatorname{Spec} A_{\beta^\sharp s_i}$ (the preimage of a distinguished open set is a distinguished open set), and $\pi^{-1}(D(s_i)) = \operatorname{Spec} A_i$. Now X is quasicompact and quasiseparated by the affine-locality of these notions (Exercise 7.3.C), so the hypotheses of the Qcqs Lemma 7.3.5 are satisfied. Hence we have an induced isomorphism of $A_{\beta^\sharp s_i} = \Gamma(X, \mathscr{O}_X)_{\beta^\sharp s_i} \cong \Gamma(X_{\beta^\sharp s_i}, \mathscr{O}_X) = A_i$. Thus α induces an isomorphism $\operatorname{Spec} A_i \to \operatorname{Spec} A_{\beta^\sharp s_i}$ (an isomorphism of rings induces an isomorphism of affine schemes, Exercise 4.3.A). Thus α is an isomorphism over each $\operatorname{Spec} A_{\beta^\sharp s_i}$, which cover $\operatorname{Spec} A$, and thus α is an isomorphism. Hence $X \cong \operatorname{Spec} A$, so is affine as desired. $\square$

The affine-locality of affine morphisms (Proposition 7.3.4) has some nonobvious consequences, as shown in the next exercise.

7.3.F. USEFUL EXERCISE. Suppose Z is a closed subset of an affine scheme $\operatorname{Spec} A$ locally cut out by one equation. (In other words, $\operatorname{Spec} A$ can be covered by smaller open sets, and on each such set Z is cut out by one equation.) Show that the complement Y of Z is affine. (This is clear if Z is globally cut out by one equation f, even set-theoretically; then $Y = \operatorname{Spec} A_f$. However, Z is not always of this form, see §19.11.10.)

7.3.8. Finite and integral morphisms.

Before defining finite and integral morphisms, we give an example to keep in mind. If L/K is a field extension, then $\operatorname{Spec} L \to \operatorname{Spec} K$ (i) is always affine; (ii) is integral if L/K is algebraic; and (iii) is finite if L/K is finite.

A morphism $\pi : X \to Y$ is **finite** if for every affine open set $\operatorname{Spec} B$ of Y, $\pi^{-1}(\operatorname{Spec} B)$ is the spectrum of a B-algebra that is a finitely generated B-*module*. By definition, finite morphisms are affine. Warning about terminology (finite vs. finitely generated): Recall that if we have a ring morphism $B \to A$ such that A is a finitely generated B-*module* then we say that A is a **finite** B-algebra. This is stronger than being a finitely generated B-*algebra*.

7.3.G. EXERCISE (THE PROPERTY OF FINITENESS IS AFFINE-LOCAL ON THE TARGET). Show that a morphism $\pi : X \to Y$ is finite if there is a cover of Y by affine open sets $\operatorname{Spec} A$ such that $\pi^{-1}(\operatorname{Spec} A)$ is the spectrum of a finite A-algebra.

The following four examples will give you some feeling for finite morphisms. In each example, you will notice two things. In each case, the maps are always finite-to-one (as maps of sets). We will verify this in general in Exercise 7.3.K. You will also notice that the morphisms are **closed** as maps of topological spaces, i.e., the images of closed sets are closed. We will show that finite morphisms are always closed in Exercise 7.3.M (and give a second proof in §8.2.5). Intuitively, you should think of finite as being closed plus finite fibers, although this isn't quite true. We will make this precise in Theorem 29.6.2.

Example 1: Branched covers. Consider the morphism $\operatorname{Spec} k[t] \to \operatorname{Spec} k[u]$ given by $u \mapsto p(t)$, where $p(t) \in k[t]$ is a degree n polynomial (see Figure 7.2). This is finite: $k[t]$ is generated as a $k[u]$-module by $1, t, t^2, \dots, t^{n-1}$.

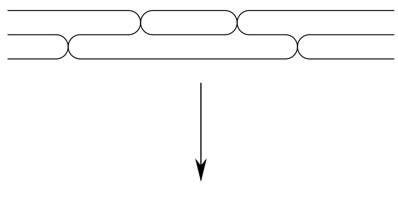


FIGURE 7.2. The "branched cover" $\mathbb{A}^1_k \to \mathbb{A}^1_k$ of the "u-line" by the "t-line" given by $u \mapsto p(t)$ is finite

Example 2: Closed embeddings (to be defined soon, in §8.1.1). If I is an ideal of a ring A, consider the morphism $\operatorname{Spec} A/I \to \operatorname{Spec} A$ given by the obvious map $A \to A/I$ (see Figure 7.3 for an example, with $A = k[t]$, $I = (t)$). This is a finite morphism (A/I is generated as a A-module by the element $1 \in A/I$).

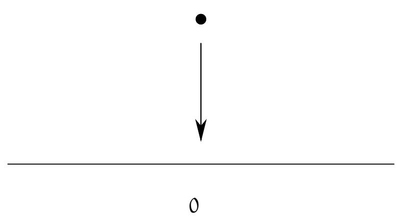


FIGURE 7.3. The "closed embedding" $\operatorname{Spec} k \to \operatorname{Spec} k[t]$ given by $t \mapsto 0$ is finite

Example 3: Normalization (to be defined in §9.7). Consider the morphism $\operatorname{Spec} k[t] \to \operatorname{Spec} k[x, y]/(y^2 - x^2 - x^3)$ corresponding to $k[x, y]/(y^2 - x^2 - x^3) \to k[t]$ given by $x \mapsto t^2 - 1$, $y \mapsto t^3 - t$ (check that this is a well-defined ring map!), see Figure 7.4. This is a finite morphism, as $k[t]$ is generated as a $(k[x, y]/(y^2 - x^2 - x^3))$-module by 1 and t. (The figure suggests that this is an isomorphism away from the "node" of the target. You can verify this, by checking that it induces an isomorphism between $D(t^2 - 1)$ in the source and $D(x)$ in the target. We will meet this example again!)

7.3.H. IMPORTANT EXERCISE (EXAMPLE 4, FINITE MORPHISMS TO $\operatorname{Spec} k$). Show that if $X \to \operatorname{Spec} k$ is a finite morphism, then X is a finite union of points with

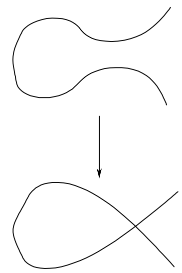


FIGURE 7.4. The "normalization" $\operatorname{Spec} k[t] \to \operatorname{Spec} k[x,y]/(y^2 - x^2 - x^3)$ given by $(x,y) \mapsto (t^2-1, t^3-t)$ is finite

the discrete topology, each point with residue field a finite extension of k, see Figure 7.5. (An example is $\operatorname{Spec} \mathbb{F}_8 \times \mathbb{F}_4[x,y]/(x^2,y^4) \times \mathbb{F}_4[t]/(t^9) \times \mathbb{F}_2 \to \operatorname{Spec} \mathbb{F}_2$.) Do *not* just quote some fancy theorem! Possible approach: By Exercise 3.2.G, any integral domain which is a finite k-algebra must be a field. If $X = \operatorname{Spec} A$, show that every prime $\mathfrak{p}$ of A is maximal. Show that the irreducible components of $\operatorname{Spec} A$ are closed points. Show $\operatorname{Spec} A$ is discrete and hence finite. Show that the residue fields $K(A/\mathfrak{p})$ of A are finite field extensions of k. (See Exercise 7.4.D for an extension to quasifinite morphisms.)

FIGURE 7.5. A picture of a finite morphism to $\operatorname{Spec} k$. Bigger fields are depicted as bigger points.

7.3.I. EASY EXERCISE (CF. EXERCISE 7.2.C). Show that the composition of two finite morphisms is also finite.

7.3.J. EXERCISE ("FINITE MORPHISMS TO $\operatorname{Spec} A$ ARE PROJECTIVE"). If R is an A-algebra, define a graded ring $S_\bullet$ by $S_0 = A$, and $S_n = R$ for $n > 0$. (What is the multiplicative structure? Hint: you know how to multiply elements of R together, and how to multiply elements of A with elements of R.) Describe an isomorphism

$\operatorname{Proj} S_\bullet \cong \operatorname{Spec} R$. Show that if R is a *finite* A-algebra (finitely generated as an A-module) then $S_\bullet$ is a finitely generated graded ring over A, and hence that $\operatorname{Spec} R$ is a projective A-scheme (§4.5.9).

7.3.K. IMPORTANT EXERCISE. Show that finite morphisms have finite fibers. (This is a useful exercise, because you will have to figure out how to get at points in a fiber of a morphism: given $\pi : X \to Y$, and $q \in Y$, what are the points of $\pi^{-1}(q)$? This will be easier to do once we discuss fibers in greater detail, see Remark 9.3.4, but it will be enlightening to do it now.) Hint: if $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$ are both affine, and $q = [\mathfrak{q}]$, then we can throw out everything in B outside $\overline{q}$ by modding out by $\mathfrak{q}$; show that the preimage is $\operatorname{Spec}(A/\pi^\sharp \mathfrak{q} A)$. Then you have reduced to the case where Y is the Spec of an integral domain B, and $[\mathfrak{q}] = [(0)]$ is the generic point. We can throw out the rest of the points of B by localizing at (0). Show that the preimage is Spec of A localized at $\pi^\sharp B^\times$. Show that the condition of finiteness is preserved by the constructions you have done, and thus reduce the problem to Exercise 7.3.H.

There is more to finiteness than finite fibers, as is shown by the following two examples.

7.3.9. *Example.* The open embedding $\mathbb{A}^2 - \{(0,0)\} \to \mathbb{A}^2$ has finite fibers, but is not affine (as $\mathbb{A}^2 - \{(0,0)\}$ isn't affine, §4.4.1) and hence not finite.

7.3.L. EASY EXERCISE. Show that the open embedding $\mathbb{A}^1_{\mathbb{C}} - \{0\} \to \mathbb{A}^1_{\mathbb{C}}$ has finite fibers and is affine, but is not finite.

7.3.10. *Definition.* A morphism $\pi : X \to Y$ of schemes is **integral** if π is affine, and for every affine open subset $\operatorname{Spec} B \subset Y$, with $\pi^{-1}(\operatorname{Spec} B) = \operatorname{Spec} A$, the induced map $B \to A$ is an integral ring morphism. This is an affine-local condition by Exercises 7.2.A and 7.2.B, and the Affine Communication Lemma 5.3.2. It is closed under composition by Exercise 7.2.C. Integral morphisms are mostly useful because finite morphisms are integral by Corollary 7.2.2. Note that the converse implication doesn't hold (witness $\operatorname{Spec} \overline{\mathbb{Q}} \to \operatorname{Spec} \mathbb{Q}$, as discussed after the statement of Corollary 7.2.2).

7.3.M. EXERCISE. Prove that integral morphisms are closed, i.e., that the image of closed subsets are closed. (Hence finite morphisms are closed. A second proof will be given in §8.2.5.) Hint: Reduce to the affine case. If $\pi^\sharp : B \to A$ is a ring map, inducing integral $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$, then suppose $I \subset A$ cuts out a closed set of $\operatorname{Spec} A$, and $J = (\pi^\sharp)^{-1}(I)$, then note that $B/J \subset A/I$, and apply the Lying Over Theorem 7.2.5 here.

7.3.N. UNIMPORTANT EXERCISE. Suppose $B \to A$ is integral. Show that for any ring homomorphism $B \to C$, the induced map $C \to A \otimes_B C$ is integral. (Hint: We wish to show that any $\sum_{i=1}^n a_i \otimes c_i \in A \otimes_B C$ is integral over C. Use the fact that each of the finitely many a_i are integral over B, and then Exercise 7.2.D.) Once we know what "base change" is, this will imply that the property of integrality of a morphism is preserved by base change, Exercise 9.4.B(e).

7.3.11. *Fibers of integral morphisms.* Unlike finite morphisms (Exercise 7.3.K), integral morphisms don't always have finite fibers. (Can you think of an example?)

However, once we make sense of fibers as topological spaces (or even schemes) in §9.3.2, you can check (Exercise 11.1.D) that the fibers have the property that no point is in the closure of any other point.

7.3.12. Morphisms (locally) of finite type.

A morphism $\pi : X \to Y$ is **locally of finite type** if for every affine open set $\operatorname{Spec} B$ of Y, and every affine open subset $\operatorname{Spec} A$ of $\pi^{-1}(\operatorname{Spec} B)$, the induced morphism $B \to A$ expresses A as a finitely generated B-algebra. By the affine-locality of finite-typeness of B-schemes (Proposition 5.3.3(c)), this is equivalent to: $\pi^{-1}(\operatorname{Spec} B)$ can be covered by affine open subsets $\operatorname{Spec} A_i$ so that each A_i is a finitely generated B-algebra.

A morphism π is **of finite type** if it is locally of finite type and quasicompact. Translation: for every affine open set $\operatorname{Spec} B$ of Y, $\pi^{-1}(\operatorname{Spec} B)$ can be covered with *a finite number of* open sets $\operatorname{Spec} A_i$ so that the induced morphism $B \to A_i$ expresses A_i as a finitely generated B-algebra.

7.3.13. *Linguistic side remark.* It is a common practice to name properties as follows: P equals locally P plus quasicompact. Two exceptions are "ringed space" (§6.3) and "finite presentation" (§7.3.17).

7.3.O. EXERCISE (THE NOTIONS "LOCALLY FINITE TYPE" AND "FINITE TYPE" ARE AFFINE-LOCAL ON THE TARGET). Show that a morphism $\pi : X \to Y$ is locally of finite type if there is a cover of Y by affine open sets $\operatorname{Spec} B_i$ such that $\pi^{-1}(\operatorname{Spec} B_i)$ is locally finite type over B_i.

Example: the "structure morphism" $\mathbb{P}^n_A \to \operatorname{Spec} A$ is of finite type, as $\mathbb{P}^n_A$ is covered by $n+1$ open sets of the form $\operatorname{Spec} A[x_1, \dots, x_n]$.

Our earlier definition of schemes of "finite type over k" (or "finite type k-schemes") from §5.3.6 is now a special case of this more general notion: the phrase "a scheme X is of finite type over k" means that we are given a morphism $X \to \operatorname{Spec} k$ (the "structure morphism") that is of finite type.

Here are some properties enjoyed by morphisms of finite type.

7.3.P. EXERCISE (FINITE = INTEGRAL + FINITE TYPE).
(a) (easier) Show that finite morphisms are of finite type.
(b) Show that a morphism is finite if and only if it is integral and of finite type.

7.3.Q. EXERCISES (NOT HARD, BUT IMPORTANT).
(a) Show that every open embedding is locally of finite type, and hence that every quasicompact open embedding is of finite type. Show that every open embedding into a locally Noetherian scheme is of finite type.
(b) Show that the composition of two morphisms locally of finite type is locally of finite type. (Hence as the composition of two quasicompact morphisms is quasicompact, Easy Exercise 7.3.A, the composition of two morphisms of finite type is of finite type.)
(c) Suppose $X \to Y$ is locally of finite type, and Y is locally Noetherian. Show that X is also locally Noetherian. If $X \to Y$ is a morphism of finite type, and Y is Noetherian, show that X is Noetherian.

7.3.14. *Definition.* A morphism π is **quasifinite** if it is of finite type, and for all $q \in Y$, $\pi^{-1}(q)$ is a finite set. The main point of this definition is the "finite fiber" part; the "finite type" hypothesis will ensure that this notion is "preserved by fibered product," Exercise 9.4.C.

Combining Exercise 7.3.K with Exercise 7.3.P(a), we see that finite morphisms are quasifinite. There are quasifinite morphisms which are not finite, such as $\mathbb{A}^2 - \{(0,0)\} \to \mathbb{A}^2$ (Example 7.3.9). However, we will soon see that quasifinite morphisms to $\operatorname{Spec} k$ are finite (Exercise 7.4.D). A key example of a morphism with finite fibers that is not quasifinite is $\operatorname{Spec} \mathbb{C}(t) \to \operatorname{Spec} \mathbb{C}$. Another is $\operatorname{Spec} \overline{\mathbb{Q}} \to \operatorname{Spec} \mathbb{Q}$. (For interesting behavior caused by the fact that $\operatorname{Spec} \overline{\mathbb{Q}} \to \operatorname{Spec} \mathbb{Q}$ is not of finite type, see Warning 9.1.4.)

7.3.15. *How to picture quasifinite morphisms.* If $\pi : X \to Y$ is a finite morphism, then for any quasi-compact open subset $U \subset X$, the induced morphism $\pi|_U : U \to Y$ is quasi-finite. In fact *every* reasonable quasifinite morphism arises in this way. (This simple-sounding statement is in fact a deep and important result — a form of Zariski's Main Theorem, see Exercise 29.6.E.) Thus the right way to visualize quasifiniteness is as a finite map with some (closed locus of) points removed.

7.3.16. Frobenius.

7.3.R. EXERCISE. Suppose p is prime and $r \in \mathbb{Z}^+$. Let $q = p^r$, and $k = \mathbb{F}_q$. Define $\phi : k[x_1, \dots, x_n] \to k[x_1, \dots, x_n]$ by $\phi(x_i) = x_i^p$ for each i, and let $F : \mathbb{A}_k^n \to \mathbb{A}_k^n$ be the map of schemes corresponding to ϕ.
(a) Show that F^r is the identity on the level of sets, but is not the identity morphism.
(b) Show that F is a bijection, but is not an isomorphism of schemes.
(c) If $K = \overline{\mathbb{F}}_p$, show that the morphism $F : \mathbb{A}_K^n \to \mathbb{A}_K^n$ of K-schemes corresponding to $x_i \mapsto x_i^p$ is a bijection, but no power of F is the identity on the level of sets.

7.3.S. EXERCISE. Suppose X is a scheme over $\mathbb{F}_p$. Explain how to define (without choice) an endomorphism $F : X \to X$ such that for each affine open subset $\operatorname{Spec} A \subset X$, F corresponds to the map $A \to A$ given by $f \mapsto f^p$ for all $f \in A$. (The morphism F is called the **absolute Frobenius morphism**.)

7.3.17. $\star$ **Morphisms (locally) of finite presentation.**

There is a variant often useful to non-Noetherian people. A ring A is a **finitely presented** B-algebra (or $B \to A$ is **finitely presented**) if

$$A \cong B[x_1, \dots, x_n]/(r_1(x_1, \dots, x_n), \dots, r_j(x_1, \dots, x_n))$$

("A has a finite number of generators and a finite number of relations over B"). If A is Noetherian, then finitely presented is the same as finite type, as the "finite number of relations" comes for free, so most of you will not care. A morphism $\pi : X \to Y$ is **locally of finite presentation** (or **locally finitely presented**) if for each affine open set $\operatorname{Spec} B$ of Y, $\pi^{-1}(\operatorname{Spec} B) = \cup_i \operatorname{Spec} A_i$ with $B \to A_i$ finitely presented. A morphism is of **finite presentation** (or **finitely presented**) if it is locally of finite presentation and quasiseparated and quasicompact. If X is locally Noetherian, then locally of finite presentation is the same as locally of finite type, and finite presentation is the same as finite type. So if you are a Noetherian person, you don't need to worry about this notion.

This definition is a violation of the general principle that erasing "locally" is the same as adding "quasicompact and" (Remark 7.3.13). But it is well motivated: finite presentation means "finite in all possible ways" (the algebra corresponding to each affine open set has a finite number of generators, and a finite number of relations, and a finite number of such affine open sets cover, and their intersections are also covered by a finite number affine open sets) — it is all you would hope for in a scheme without it actually being Noetherian. Exercise 9.3.H makes this precise, and explains how this notion often arises in practice.

7.3.T. EXERCISE. Show that the notion of "locally of finite presentation" is affine-local on the target.

7.3.U. ⋆⋆ HARD EXERCISE (BUT REASSURING TO KNOW). Show that if $\pi : X \to Y$ is locally of finite presentation, then for *any* affine open subscheme $\operatorname{Spec} B$ of Y and *any* affine open subscheme $\operatorname{Spec} A$ of X with $\pi(\operatorname{Spec} A) \subset \operatorname{Spec} B$, A is a finitely presented B-algebra. In particular, the notion of "locally of finite presentation" is affine-local on the source.

7.3.V. EXERCISE. Show that open embeddings are locally finitely presented.

7.3.W. EXERCISE. Show that the composition of two locally finitely presented morphisms is locally finitely presented. (Then once we show that the composition of two quasiseparated morphisms is quasiseparated in Proposition 10.1.13(b), we will know that the composition of two finitely presented morphisms is finitely presented — recall that the composition of two quasicompact morphisms is quasicompact, by Easy Exercise 7.3.A.)

7.3.18. ⋆⋆ *Remark.* A morphism $\pi : X \to Y$ is locally of finite presentation if and only if for every projective system of Y-schemes $\{S_\lambda\}_{\lambda \in I}$ with each S_λ an affine scheme, the natural map

$$\varinjlim_{\lambda} \operatorname{Hom}(S_\lambda, X) \to \operatorname{Hom}_Y(\varprojlim_{\lambda} S_\lambda, X)$$

is a bijection (see [**Gr-EGA**, IV_3.8.14.2]). This characterization of locally finitely presented morphisms as "limit-preserving" can be useful.

7.4 Images of morphisms: Chevalley's theorem and elimination theory

In this section, we will answer a question that you may have wondered about long before hearing the phrase "algebraic geometry". If you have a number of polynomial equations in a number of variables with indeterminate coefficients, you would reasonably ask what conditions there are on the coefficients for a (common) solution to exist. Given the algebraic nature of the problem, you might hope that the answer should be purely algebraic in nature — it shouldn't be "random", or involve bizarre functions like exponentials or cosines. You should expect the answer to be given by "algebraic conditions". This is indeed the case, and it can be profitably interpreted as a question about images of maps of varieties or schemes, in which guise it is answered by Chevalley's Theorem 7.4.2 (see 7.4.5 for a more precise proof). Chevalley's Theorem will give an immediate proof of the Nullstellensatz 3.2.5 (§7.4.3).

In special cases, the image is nicer still. For example, we have seen that finite morphisms are closed (the image of closed subsets under finite morphisms are closed, Exercise 7.3.M). We will prove a classical result, the Fundamental Theorem of Elimination Theory 7.4.7, which essentially generalizes this (as explained in §8.2.5) to maps from projective space. We will use it repeatedly. In a different direction, in the distant future we will see that in certain good circumstances ("flat" plus a bit more, see Exercise 24.5.G), morphisms are open (the image of open subsets is open); one example (which you can try to show directly) is $\mathbb{A}^n_B \to \operatorname{Spec} B$.

7.4.1. Chevalley's theorem.

If $\pi : X \to Y$ is a morphism of schemes, the notion of the image of π as *sets* is clear: we just take the points in Y that are the image of points in X. We know that the image can be open (open embeddings), and we have seen examples where it is closed, and more generally, locally closed. But it can be weirder still: consider the morphism $\mathbb{A}^2_k \to \mathbb{A}^2_k$ given by $(x, y) \mapsto (x, xy)$. The image is the plane, with the y-axis removed, but the origin put back in (see Figure 7.6). This isn't so horrible. We make a definition to capture this phenomenon. A **constructible subset** of a Noetherian topological space is a subset which belongs to the smallest family of subsets such that (i) every open set is in the family, (ii) a finite intersection of family members is in the family, and (iii) the complement of a family member is also in the family. For example the image of $(x, y) \mapsto (x, xy)$ is constructible. (An extension of the notion of constructibility to more general topological spaces is mentioned in Exercise 9.3.I.)

[to make]

FIGURE 7.6. The image of $(x, y) \mapsto (x, xy)$.

7.4.A. EXERCISE: CONSTRUCTIBLE SUBSETS ARE FINITE DISJOINT UNIONS OF LOCALLY CLOSED SUBSETS. Recall that a subset of a topological space X is *locally closed* if it is the intersection of an open subset and a closed subset. (Equivalently, it is an open subset of a closed subset, or a closed subset of an open subset. We will later have trouble extending this to open and closed and locally closed subschemes, see Exercise 8.1.M.) Show that a subset of a Noetherian topological space X is constructible if and only if it is the finite disjoint union of locally closed subsets. As a consequence, if $X \to Y$ is a continuous map of Noetherian topological spaces, then the preimage of a constructible set is a constructible set.

7.4.B. EXERCISE (USED IN PROOF 7.4.3 OF THE NULLSTELLENSATZ 3.2.5). Show that the generic point of $\mathbb{A}^1_k$ does not form a constructible subset of $\mathbb{A}^1_k$ (where k is a field).

7.4.C. EXERCISE (USED IN EXERCISE 24.5.G).

(a) Show that a constructible subset of a Noetherian scheme is closed if and only if it is "stable under specialization". More precisely, if Z is a constructible subset of a Noetherian scheme X, then Z is closed if and only if for every pair of points y_1 and y_2 with $y_1 \in \overline{y_2}$, if $y_2 \in Z$, then $y_1 \in Z$. Hint for the "if" implication: show that Z

can be written as $\coprod_{i=1}^n U_i \cap Z_i$ where $U_i \subset X$ is open and $Z_i \subset X$ is closed. Show that Z can be written as $\coprod_{i=1}^n U_i \cap Z_i$ (with possibly different n, U_i, Z_i) where each Z_i is irreducible and meets U_i. Now use "stability under specialization" and the generic point of Z_i to show that $Z_i \subset Z$ for all i, so $Z = \cup Z_i$.
(b) Show that a constructible subset of a Noetherian scheme is open if and only if it is "stable under generization". (Hint: this follows in one line from (a).)

The image of a morphism of schemes can be stranger than a constructible set. Indeed if S is *any* subset of a scheme Y, it can be the image of a morphism: let X be the disjoint union of spectra of the residue fields of all the points of S, and let $\pi : X \to Y$ be the natural map. This is quite pathological, but in any reasonable situation, the image is essentially no worse than arose in the previous example of $(x, y) \mapsto (x, xy)$. This is made precise by Chevalley's theorem.

7.4.2. Chevalley's Theorem. — *If $\pi : X \to Y$ is a finite type morphism of Noetherian schemes, the image of any constructible set is constructible. In particular, the image of π is constructible.*

(For the minority who might care: see §9.3.7 for an extension to locally finitely presented morphisms.) We discuss the proof after giving some important consequences that may seem surprising, in that they are algebraic corollaries of a seemingly quite geometric and topological theorem. The first is a proof of the Nullstellensatz.

7.4.3. *Proof of the Nullstellensatz 3.2.5.* We wish to show that if K is a field extension of k that is finitely generated as a k-algebra, say by $x_1, \dots, x_n$, then it is a finite field extension. It suffices to show that each x_i is algebraic over k. But if x_i is not algebraic over k, then we have an inclusion of rings $k[x] \to K$, corresponding to a dominant morphism $\pi : \operatorname{Spec} K \to \mathbb{A}^1_k$ of finite type k-schemes. Of course $\operatorname{Spec} K$ is a single point, so the image of π is one point. By Chevalley's Theorem 7.4.2 and Exercise 7.4.B, the image of π is not the generic point of $\mathbb{A}^1_k$, so $\operatorname{im}(\pi)$ is a closed point of $\mathbb{A}^1_k$, and thus π is not dominant. $\square$

A similar idea can be used in the following exercise.

7.4.D. EXERCISE (QUASIFINITE MORPHISMS TO A FIELD ARE FINITE). Suppose $\pi : X \to \operatorname{Spec} k$ is a quasifinite morphism. Show that π is finite. (Hint: deal first with the affine case, $X = \operatorname{Spec} A$, where A is finitely generated over k. Suppose A contains an element x that is not algebraic over k, i.e., we have an inclusion $k[x] \hookrightarrow A$. Exercise 7.3.H may help.)

7.4.E. EXERCISE ("FOR MAPS OF VARIETIES, SURJECTIVITY CAN BE CHECKED ON CLOSED POINTS"). Assume Chevalley's Theorem 7.4.2. Show that a morphism of affine k-varieties $\pi : X \to Y$ is surjective if and only if it is surjective on closed points (i.e., if every closed point of Y is the image of a closed point of X). (Once we define varieties in general, in Definition 10.1.7, you will see that your argument works without change with the adjective "affine" removed.)

In order to prove Chevalley's Theorem 7.4.2 (in Exercise 7.4.O), we introduce a useful idea of Grothendieck's. For the purposes of this discussion only, we say *a*

B*-algebra* A *satisfies* $(\dagger)$ *if for each finitely generated* A*-module* M*, there exists a nonzero* $f \in B$ *such that* M_f *is a free* B_f*-module.*

7.4.4. Grothendieck's Generic Freeness Lemma. — *Suppose* B *is a Noetherian integral domain. Then every finitely generated* B*-algebra satisfies* $(\dagger)$*.*

Proof. We prove the Generic Freeness Lemma 7.4.4 in a series of exercises. We assume that B is a Noetherian integral domain until Lemma 7.4.4 is proved, at the end of Exercise 7.4.K.

7.4.F. EXERCISE. Show that B itself satisfies $(\dagger)$.

7.4.G. EXERCISE. Reduce the proof of Lemma 7.4.4 to the following statement: if A is a finitely-generated B-algebra satisfying $(\dagger)$, then $A[T]$ does too. (Hint: induct on the number of generators of A as an B-algebra.)

We now prove this statement. Suppose A satisfies $(\dagger)$, and let M be a finitely generated $A[T]$-module, generated by the finite set S. Let $M_0 = 0$, and let M_1 be the sub-A-module of M generated by S. For $n > 0$, inductively define

$$M_{n+1} = M_n + TM_n,$$

a sub-A-module of M. Note that M is the increasing union of the A-modules M_n.

7.4.H. EXERCISE. Show that multiplication by T induces a surjection

$$\psi_n : M_n/M_{n-1} \to M_{n+1}/M_n.$$

7.4.I. EXERCISE. Show that for $n \gg 0$, ψ_n is an isomorphism. Hint: use the ascending chain condition on M_1.

7.4.J. EXERCISE. Show that there is a nonzero $f \in B$ such that $(M_{i+1}/M_i)_f$ is free as an B_f-module, for all i. Hint: as i varies, M_{i+1}/M_i passes through only finitely many isomorphism classes.

The following result concludes the proof of the Generic Freeness Lemma 7.4.4.

7.4.K. EXERCISE (NOT REQUIRING NOETHERIAN HYPOTHESES). Suppose M is an B-module that is an increasing union of submodules M_i, with $M_0 = 0$, and that every M_{i+1}/M_i is free. Show that M is free. Hint: first construct compatible isomorphisms $\phi_n : \oplus_{i=0}^{n-1} M_{i+1}/M_i \to M_n$ by induction on n. Then show that the colimit $\phi := \varinjlim \phi_n : \oplus_{i=0}^{\infty} M_{i+1}/M_i \to M$ is an isomorphism. Side Remark: More generally, your argument will show that if the M_{i+1}/M_i are all projective modules (to be defined in §23.2.1), then M is (non-naturally) isomorphic to their direct sum.

$\square$

We now set up the proof of Chevalley's Theorem 7.4.2.

7.4.L. EXERCISE. Suppose $\pi : X \to Y$ is a finite type morphism of Noetherian schemes, and Y is irreducible. Show that there is a dense open subset U of Y such that the image of π either contains U or else does not meet U. (Hint: suppose $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ is such a morphism. Then by the Generic Freeness Lemma 7.4.4, there is a nonzero $f \in B$ such that A_f is a free B_f-module. It must have zero rank or positive rank. In the first case, show that the image of π does not meet $D(f) \subset \operatorname{Spec} B$. In the second case, show that the image of π contains $D(f)$.)

There are more direct ways of showing the content of the above hint. For example, another proof in the case of varieties will turn up in the proof of Proposition 11.4.1. We only use the Generic Freeness Lemma because we will use it again in the future (§24.5.8).

7.4.M. EXERCISE. Show that to prove Chevalley's Theorem, it suffices to prove that if $\pi : X \to Y$ is a finite type morphism of Noetherian schemes, the image of π is constructible.

7.4.N. EXERCISE. Reduce further to the case where Y is affine, say $Y = \operatorname{Spec} B$. Reduce further to the case where X is affine.

We now give the rest of the proof by waving our hands, and leave it to you to make it precise. The idea is to use Noetherian induction, and to reduce the problem to Exercise 7.4.L.

We can deal with each of the components of Y separately, so we may assume that Y is irreducible. We can then take B to be an integral domain. By Exercise 7.4.L, there is a dense open subset U of Y where either the image of π includes it, or is disjoint from it. If $U = Y$, we are done. Otherwise, it suffices to deal with the complement of U. Renaming this complement Y, we return to the start of the paragraph.

7.4.O. EXERCISE. Complete the proof of Chevalley's Theorem 7.4.2, by making the above argument precise.

7.4.5. $\star$ **Elimination of quantifiers.** A basic sort of question that arises in any number of contexts is when a system of equations has a solution. Suppose for example you have some polynomials in variables $x_1, \dots, x_n$ over an algebraically closed field $\overline{k}$, some of which you set to be zero, and some of which you set to be nonzero. (This question is of fundamental interest even before you know any scheme theory!) Then there is an algebraic condition on the coefficients which will tell you if there is a solution. Define the **Zariski topology on** $\overline{k}^n$ in the obvious way: closed subsets are cut out by equations.

7.4.P. EXERCISE (ELIMINATION OF QUANTIFIERS, OVER AN ALGEBRAICALLY CLOSED FIELD). Fix an algebraically closed field $\overline{k}$. Suppose

$$f_1, \dots, f_p, g_1, \dots, g_q \in \overline{k}[W_1, \dots, W_m, X_1, \dots X_n]$$

are given. Show that there is a (Zariski-)constructible subset Y of $\overline{k}^m$ such that

$$f_1(w_1, \dots, w_m, X_1, \dots, X_n) = \cdots = f_p(w_1, \dots, w_m, X_1, \dots, X_n) = 0 \tag{7.4.5.1}$$

and

$$g_1(w_1, \dots, w_m, X_1, \dots, X_n) \neq 0 \quad \cdots \quad g_q(w_1, \dots, w_m, X_1, \dots, X_n) \neq 0 \tag{7.4.5.2}$$

has a solution $(X_1, \dots, X_n) = (x_1, \dots, x_n) \in \overline{k}^n$ if and only if $(w_1, \dots, w_m) \in Y$. Hints: if Z is a finite type scheme over $\overline{k}$, and the closed points are denoted Z^{cl} ("cl" is for either "closed" or "classical"), then under the inclusion of topological spaces $Z^{cl} \hookrightarrow Z$, the Zariski topology on Z induces the Zariski topology on Z^{cl}. Note that we can identify $(\mathbb{A}^p_{\overline{k}})^{\mathrm{cl}}$ with $\overline{k}^p$ by the Nullstellensatz (Exercise 5.3.E). If X is the locally closed subset of $\mathbb{A}^{m+n}$ cut out by the equalities and inequalities (7.4.5.1) and (7.4.5.2), we have the diagram

$$\begin{tikzcd} X^{\mathrm{cl}} \arrow[r, hook] \arrow[d, "\pi^{\mathrm{cl}}"'] & X \arrow[r, hook, "\text{loc. cl.}"] \arrow[d, "\pi"] & \mathbb{A}^{m+n} \arrow[dl] \\ \overline{k}^m \arrow[r, hook] & \mathbb{A}^m & \end{tikzcd}$$

where $Y = \operatorname{im} \pi^{\mathrm{cl}}$. By Chevalley's theorem 7.4.2, $\operatorname{im}\pi$ is constructible, and hence so is $(\operatorname{im}\pi) \cap \overline{k}^m$. It remains to show that $(\operatorname{im}\pi) \cap \overline{k}^m = Y$ ($= \operatorname{im}\pi^{\mathrm{cl}}$). You might use the Nullstellensatz.

This is called "elimination of quantifiers" because it gets rid of the quantifier "there exists a solution". The analogous statement for real numbers, where inequalities are also allowed, is a special case of Tarski's celebrated theorem of elimination of quantifiers for real closed fields (see, for example, [T]).

7.4.6. The Fundamental Theorem of Elimination Theory.
In the case of projective space (and later, projective morphisms), one can do better than Chevalley.

7.4.7. Theorem (Fundamental Theorem of Elimination Theory). — *The morphism* $\pi : \mathbb{P}^n_A \to \operatorname{Spec} A$ *is* **closed** *(sends closed sets to closed sets).*

Note that *no* Noetherian hypotheses are needed.

A great deal of classical algebra and geometry is contained in this theorem as special cases. Here are some examples.

First, let $A = k[a, b, c, \dots, i]$, and consider the closed subset of $\mathbb{P}^2_A$ (taken with coordinates x, y, z) corresponding to $ax + by + cz = 0$, $dx + ey + fz = 0$, $gx + hy + iz = 0$. Then we are looking for the locus in $\operatorname{Spec} A$ where these equations have a nontrivial solution. This indeed corresponds to a Zariski-closed set — where

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = 0.$$

Thus the idea of the determinant is embedded in elimination theory.

As a second example, let $A = k[a_0, a_1, \dots, a_m, b_0, b_1, \dots, b_n]$. Now consider the closed subset of $\mathbb{P}^1_A$ (taken with coordinates x and y) corresponding to $a_0 x^m + a_1 x^{m-1} y + \cdots + a_m y^m = 0$ and $b_0 x^n + b_1 x^{n-1} y + \cdots + b_n y^n = 0$. Then there is a polynomial in the coefficients $a_0, \dots, b_n$ (an element of A) which vanishes if and only if these two polynomials have a common nonzero root — this polynomial is called the *resultant*.

More generally, given a number of homogeneous equations in $n + 1$ variables with indeterminate coefficients, Theorem 7.4.7 implies that one can write down equations in the coefficients that precisely determine when the equations have a nontrivial solution.

7.4.8. *Proof of the Fundamental Theorem of Elimination Theory 7.4.7.* Suppose $Z \hookrightarrow \mathbb{P}^n_A$ is a closed subset. We wish to show that $\pi(Z)$ is closed.

By the definition of the Zariski topology on $\operatorname{Proj} A[x_0, \dots, x_n]$ (§4.5.7), Z is cut out (set-theoretically) by some homogeneous elements $f_1, f_2, \dots \in A[x_0, \dots, x_n]$. We wish to show that the points $\mathfrak{p} \in \operatorname{Spec} A$ that are in $\pi(Z)$ form a closed subset. Equivalently, we want to show that those $\mathfrak{p}$ for which $f_1, f_2, \dots$ have a common zero in $\operatorname{Proj} \kappa(\mathfrak{p})[x_0, \dots, x_n]$ form a closed subset of $\operatorname{Spec} A$.

To motivate our argument, we consider a related question. Suppose that $S_\bullet := k[x_0, \dots, x_n]$, and that $g_1, g_2, \dots \in k[x_0, \dots, x_n]$ are homogeneous polynomials. How can we tell if $g_1, g_2, \dots$ have a common zero in $\operatorname{Proj} S_\bullet = \mathbb{P}^n_k$?

They would have a common zero if and only if in

$$\mathbb{A}^{n+1}_k = \operatorname{Spec} S_\bullet,$$

they cut out (set-theoretically) more than the origin, i.e., if set-theoretically

$$V(g_1, g_2, \dots) \not\subset V(x_0, \dots, x_n).$$

This is true if and only if

$$(x_0, \dots, x_n)^N \not\subset (g_1, g_2, \dots) \quad \text{for all } N.$$

(Do you see why?) This is equivalent to

$$S_N \not\subset (g_1, g_2, \dots) \quad \text{for all } N,$$

which may be rewritten as

$$S_N \not\subset g_1 S_{N-\deg g_1} \oplus g_2 S_{N-\deg g_2} \oplus \cdots \tag{7.4.8.1}$$

for all N. In other words, this is equivalent to the statement that the k-linear map

$$S_{N-\deg g_1} \oplus S_{N-\deg g_2} \oplus \cdots \to S_N$$

is not surjective. This map is given by a matrix with $\dim S_N$ rows. (It may have an infinite number of columns, but this will not bother us.) To check that this linear map is not surjective, we need only check that all the "maximal" ($\dim S_N \times \dim S_N$) determinants are zero. (Of course, we need to check this for all N.) Thus the condition that $g_1, g_2, \dots$ have no common zeros in $\mathbb{P}^n_k$ is the same as checking some (admittedly infinite) number of equations. In other words, it is a Zariski-closed condition on the coefficients of the polynomials $g_1, g_2, \dots$.

7.4.Q. EXERCISE. Complete the proof of the Fundamental Theorem of Elimination Theory 7.4.7. (Hint: follow precisely the same argument, with k replaced by A, and the g_i replaced by the f_i. How and why does this prove the theorem?) □

Notice that projectivity was crucial to the proof: we used graded rings in an essential way. Notice also that the proof is essentially just linear algebra.

CHAPTER 8

Closed embeddings and related notions

8.1 Closed embeddings and closed subschemes

The scheme-theoretic analog of closed subsets has a surprisingly different flavor from the analog of open sets (open embeddings). However, just as open embeddings (the scheme-theoretic version of open set) are locally modeled on open sets $U \subset Y$, the analog of closed subsets also has a local model. This was foreshadowed by our understanding of closed subsets of $\operatorname{Spec} B$ as roughly corresponding to ideals. If $I \subset B$ is an ideal, then $\operatorname{Spec} B/I \hookrightarrow \operatorname{Spec} B$ is a morphism of schemes, and we have checked that on the level of topological spaces, this describes $\operatorname{Spec} B/I$ as a closed subset of $\operatorname{Spec} B$, with the subspace topology (Exercise 3.4.I). This morphism is our "local model" of a closed embedding.

8.1.1. *Definition.* A morphism $\pi : X \to Y$ is a **closed embedding** (or **closed immersion**) if it is an affine morphism, and for every affine open subset $\operatorname{Spec} B \subset Y$, with $\pi^{-1}(\operatorname{Spec} B) \cong \operatorname{Spec} A$, the map $B \to A$ is surjective (i.e., of the form $B \to B/I$, our desired local model). If X is a *subset* of Y (and π on the level of sets is the inclusion), we say that X is a **closed subscheme** of Y. The difference between a closed embedding and a closed subscheme is confusing and unimportant; the same issue for open embeddings/subschemes was discussed in §7.1.1. The symbol $\hookrightarrow$ often is used to indicate that a morphism is a closed embedding (or more generally, a locally closed embedding, §8.1.2.)

8.1.A. EXERCISE. Show that a closed embedding identifies the topological space of X with a closed subset of the topological space of Y. (Caution: The closed embeddings $\operatorname{Spec} k[x]/(x) \hookrightarrow k[x]/(x^2)$ and $\operatorname{Spec} k[x]/(x^2) \hookrightarrow k[x]/(x^2)$ show that the closed subset does not determine the closed subscheme. The "infinitesimal" information, or "fuzz", is lost.)

8.1.B. EASY EXERCISE. Show that closed embeddings are finite, hence of finite type.

8.1.C. EASY EXERCISE. Show that the composition of two closed embeddings is a closed embedding.

8.1.D. EXERCISE. Show that the property of being a closed embedding is affine-local on the target.

8.1.E. EXERCISE. Suppose $B \to A$ is a surjection of rings. Show that the induced morphism $\operatorname{Spec} A \to \operatorname{Spec} B$ is a closed embedding. (Our definition would be a terrible one if this were not true!)

A closed embedding $\pi : X \hookrightarrow Y$ determines an *ideal sheaf* on Y, as the kernel $\mathscr{I}_{X/Y}$ of the map of $\mathscr{O}_Y$-modules

$$\mathscr{O}_Y \to \pi_* \mathscr{O}_X.$$

An **ideal sheaf** on Y is what it sounds like: it is a sheaf of ideals. It is a sub-$\mathscr{O}_Y$-module $\mathscr{I}$ of $\mathscr{O}_Y$. On each open subset, it gives an ideal $\mathscr{I}(U)$ of the ring $\mathscr{O}_Y(U)$. We thus have an exact sequence (of $\mathscr{O}_Y$-modules) $0 \to \mathscr{I}_{X/Y} \to \mathscr{O}_Y \to \pi_* \mathscr{O}_X \to 0$. (On $\operatorname{Spec} B$, the epimorphism $\mathscr{O}_Y \to \pi_* \mathscr{O}_X$ is the surjection $B \to A$ of Definition 8.1.1.)

Thus for each affine open subset $\operatorname{Spec} B \hookrightarrow Y$, we have an ideal $I(B) \subset B$, and we can recover X from this information: the $I(B)$ (as $\operatorname{Spec} B \hookrightarrow Y$ varies over the affine open subsets) defines an $\mathscr{O}$-module on the base, hence an $\mathscr{O}_Y$-module on Y, and the cokernel of $\mathscr{I} \hookrightarrow \mathscr{O}_Y$ is $\mathscr{O}_X$. It will be useful to understand when the information of the $I(B)$ (for all affine opens $\operatorname{Spec} B \hookrightarrow Y$) actually determines a closed subscheme. Our life is complicated by the fact that the answer is "not always", as shown by the following example.

8.1.F. UNIMPORTANT EXERCISE. Let $X = \operatorname{Spec} k[x]_{(x)}$, the germ of the affine line at the origin, which has two points, the closed point and the generic point η. Define $\mathscr{I}(X) = \{0\} \subset \mathscr{O}_X(X) = k[x]_{(x)}$, and $\mathscr{I}(\eta) = k(x) = \mathscr{O}_X(\eta)$. Show that this sheaf of ideals does not correspond to a closed subscheme. (Possible approach: do the next exercise first.)

The next exercise gives a necessary condition.

8.1.G. EXERCISE. Suppose $\mathscr{I}_{X/Y}$ is a sheaf of ideals corresponding to a closed embedding $X \hookrightarrow Y$. Suppose $\operatorname{Spec} B \hookrightarrow Y$ is an affine open subscheme, and $f \in B$. Show that the natural map $I(B)_f \to I(B_f)$ is an isomorphism. (First state what the "natural map" is!)

It is an important and useful fact that this is sufficient:

8.1.H. ESSENTIAL (HARD) EXERCISE: A USEFUL CRITERION FOR WHEN IDEALS IN AFFINE OPEN SETS DEFINE A CLOSED SUBSCHEME. Suppose Y is a scheme, and for each affine open subset $\operatorname{Spec} B$ of Y, $I(B) \subset B$ is an ideal. Suppose further that for each affine open subset $\operatorname{Spec} B \hookrightarrow Y$ and each $f \in B$, restriction of functions from $B \to B_f$ induces an isomorphism $I(B_f) \cong I(B)_f$. Show that these data arise from a (unique) closed subscheme $X \hookrightarrow Y$ by the above construction. In other words, the closed embeddings $\operatorname{Spec} B/I \hookrightarrow \operatorname{Spec} B$ glue together in a well-defined manner to obtain a closed embedding $X \hookrightarrow Y$.

This is a hard exercise, so as a hint, here are three different ways of proceeding; some combination of them may work for you. *Approach 1.* For each affine open $\operatorname{Spec} B$, we have a closed subscheme $\operatorname{Spec} B/I \hookrightarrow \operatorname{Spec} B$. (i) For any two affine open subschemes $\operatorname{Spec} A$ and $\operatorname{Spec} B$, show that the two closed subschemes

$\operatorname{Spec} A/I(A) \hookrightarrow \operatorname{Spec} A$ and $\operatorname{Spec} B/I(B) \hookrightarrow \operatorname{Spec} B$ restrict to the *same* closed subscheme of their intersection. (Hint: cover their intersection with open sets simultaneously distinguished in both affine open sets, Proposition 5.3.1.) Thus for example we can glue these two closed subschemes together to get a closed subscheme of $\operatorname{Spec} A \cup \operatorname{Spec} B$. (ii) Use Exercise 4.4.A on gluing schemes (or the ideas therein) to glue together the closed embeddings in all affine open subschemes simultaneously. You will only need to worry about triple intersections. *Approach 2.* (i) Use the data of the ideals $I(B)$ to define a sheaf of ideals $\mathscr{I} \hookrightarrow \mathscr{O}$. (ii) For each affine open subscheme $\operatorname{Spec} B$, show that $\mathscr{I}(\operatorname{Spec} B)$ is indeed $I(B)$, and $(\mathscr{O}/\mathscr{I})(\operatorname{Spec} B)$ is indeed $B/I(B)$, so the data of $\mathscr{I}$ recovers the closed subscheme on each $\operatorname{Spec} B$ as desired. *Approach 3.* (i) Describe X first as a subset of Y. (ii) Check that X is closed. (iii) Define the sheaf of functions $\mathscr{O}_X$ on this subset, perhaps using compatible stalks. (iv) Check that this resulting ringed space is indeed locally the closed subscheme given by $\operatorname{Spec} B/I \hookrightarrow \operatorname{Spec} B$.)

We will see later (§13.5.6) that closed subschemes correspond to *quasicoherent* sheaves of ideals; the mathematical content of this statement will turn out to be precisely Exercise 8.1.H.

8.1.I. IMPORTANT EXERCISE/DEFINITION: THE VANISHING SCHEME.
(a) Suppose Y is a scheme, and $s \in \Gamma(\mathscr{O}_Y, Y)$. Define the closed scheme **cut out by** s. We call this the **vanishing scheme** $V(s)$ of s, as it is the scheme-theoretic version of our earlier (set-theoretical) version of $V(s)$ (§3.4). (Hint: on affine open $\operatorname{Spec} B$, we just take $\operatorname{Spec} B/(s_B)$, where s_B is the restriction of s to $\operatorname{Spec} B$. Use Exercise 8.1.H to show that this yields a well-defined closed subscheme.)
(b) If u is an invertible function, show that $V(s) = V(su)$.
(c) If S is a set of functions, define $V(S)$.

8.1.J. IMPORTANT EXERCISE.
(a) In analogy with closed subsets, define the notion of a **finite union of closed subschemes** of X, and an arbitrary (not necessarily finite) **intersection of closed subschemes** of X. (Exercise 8.1.H may help.) Hint: If X is affine, then you might expect that the union of closed subschemes corresponding to I_1 and I_2 would be the closed subscheme corresponding to either $I_1 \cap I_2$ or $I_1 I_2$ — but which one? We would want the union of a closed subscheme with itself to *be* itself, so the right choice is $I_1 \cap I_2$.
(b) Describe the scheme-theoretic intersection of $V(y - x^2)$ and $V(y)$ in $\mathbb{A}^2$. See Figure 4.5 for a picture. (For example, explain informally how this corresponds to two curves meeting at a single point with multiplicity 2 — notice how the 2 is visible in your answer. Alternatively, what is the nonreducedness telling you — both its "size" and its "direction"?) Describe their scheme-theoretic union.
(c) Show that the underlying set of a finite union of closed subschemes is the finite union of the underlying sets, and similarly for arbitrary intersections.
(d) Describe the scheme-theoretic intersection of $V(y^2 - x^2)$ and $V(y)$ in $\mathbb{A}^2$. Draw a picture. (Did you expect the intersection to have multiplicity one or multiplicity two?) Hence show that if X, Y, and Z are closed subschemes of W, then $(X \cap Z) \cup (Y \cap Z) \neq (X \cup Y) \cap Z$ in general. In particular, not all properties of intersection and union carry over from sets to schemes.

8.1.K. $\star$ HARD EXERCISE (NOT USED LATER). In the literature, the usual definition of a closed embedding is a morphism $\pi : X \to Y$ such that π induces a homeomorphism of the underlying topological space of X onto a closed subset of the topological space of Y, and the induced map $\pi^\sharp : \mathscr{O}_Y \to \pi_* \mathscr{O}_X$ of sheaves on Y is surjective. (By "surjective" we mean that the ring homomorphism on stalks is surjective.) Show that this definition agrees with the one given above. (To show that our definition involving surjectivity on the level of affine open sets implies this definition, you can use the fact that surjectivity of a morphism of sheaves can be checked on a base, Exercise 2.7.E.)

We have now defined the analog of open subsets and closed subsets in the land of schemes. Their definition is slightly less "symmetric" than in the classical topological setting: the "complement" of a closed subscheme is a unique open subscheme, but there are many "complementary" closed subschemes to a given open subscheme in general. (We will soon define one that is "best", that has a reduced structure, §8.3.9.)

8.1.2. Locally closed embeddings and locally closed subschemes.

Now that we have defined analogs of open and closed subsets, it is natural to define the analog of locally closed subsets. Recall that locally closed subsets are intersections of open subsets and closed subsets. Hence they are closed subsets of open subsets, or equivalently open subsets of closed subsets. The analog of these equivalences will be a little problematic in the land of schemes.

We say a morphism $\pi : X \to Y$ is a **locally closed embedding** (or **locally closed immersion**) if π can factored into

$$X \xrightarrow{\rho} Z \xrightarrow{\tau} Y$$

where ρ is a closed embedding and τ is an open embedding. If X is a subset of Y (and π on the level of sets is the inclusion), we say X is a **locally closed subscheme** of Y. (Warning: The term *immersion* is often used instead of *locally closed embedding* or *locally closed immersion*, but this is unwise terminology. The differential geometric notion of immersion is closer to what algebraic geometers call unramified, which we will define in §21.6. The naked term *embedding* should be avoided, because it is not precise.) The symbol $\hookrightarrow$ is often used to indicate that a morphism is a locally closed embedding.

For example, the morphism $\operatorname{Spec} k[t, t^{-1}] \to \operatorname{Spec} k[x, y]$ given by $(x, y) \mapsto (t, 0)$ is a locally closed embedding (Figure 8.1).

8.1.L. EASY EXERCISE. Show that locally closed embeddings are locally of finite type.

At this point, you could define the intersection of two locally closed embeddings in a scheme X (which will also be a locally closed embedding in X). But it would be awkward, as you would have to show that your construction is independent of the factorizations of each locally closed embedding into a closed embedding and an open embedding. Instead, we wait until Exercise 9.2.C, when recognizing the intersection as a fibered product will make this easier.

Clearly an open subscheme U of a closed subscheme V of X can be interpreted as a closed subscheme of an open subscheme: as the topology on V is induced from the topology on X, the underlying set of U is the intersection of some open

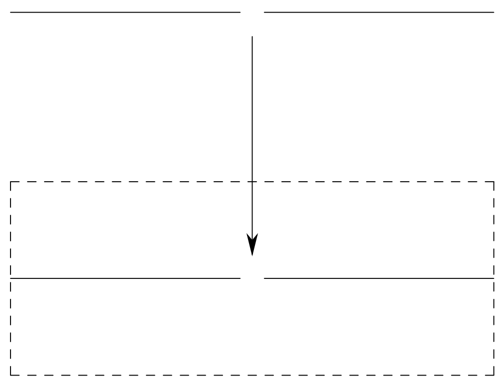


FIGURE 8.1. The locally closed embedding $\operatorname{Spec} k[t, t^{-1}] \rightarrow \operatorname{Spec} k[x, y]$ ($t \mapsto (t, 0) = (x, y)$, i.e., $(x, y) \rightarrow (t, 0)$)

subset U' on X with V. We can take $V' = V \cap U'$, and then $V' \rightarrow U'$ is a closed embedding, and $U' \rightarrow X$ is an open embedding.

It is not clear that a closed subscheme V' of an open subscheme U' can be expressed as an open subscheme of a closed subscheme V. In the category of topological spaces, we would take V as the closure of V', so we are now motivated to define the analogous construction, which will give us an excuse to introduce several related ideas, in §8.3. We will then resolve this issue in good cases (e.g. if X is Noetherian) in Exercise 8.3.C.

We formalize our discussion in an exercise.

8.1.M. EXERCISE. Suppose $V \rightarrow X$ is a morphism. Consider three conditions:

(i) V is the intersection of an open subscheme of X and a closed subscheme of X (you will have to define the meaning of "intersection" here, see Exercise 7.1.B, or else see the hint below).

(ii) V is an open subscheme of a closed subscheme of X, i.e., it factors into an open embedding followed by a closed embedding.

(iii) V is a closed subscheme of an open subscheme of X, i.e., V is a locally closed embedding.

Show that (i) and (ii) are equivalent, and both imply (iii). (Remark: (iii) does *not* always imply (i) and (ii), see the pathological example [**Stacks**, tag 01QW].) Hint: It may be helpful to think of the problem as follows. You might hope to think of a locally closed embedding as a fibered diagram

$$\begin{tikzcd} V \arrow[r, hook, "\text{open emb.}"] \arrow[d, hook, "\text{closed emb.}"'] & K \arrow[d, hook, "\text{closed emb.}"] \\ U \arrow[r, hook, "\text{open emb.}"'] & X. \end{tikzcd}$$

Interpret (i) as the existence of the diagram. Interpret (ii) as this diagram minus the lower left corner. Interpret (iii) as the diagram minus the upper right corner.

8.1.N. EXERCISE. Show that the composition of two locally closed embeddings is a locally closed embedding. (Hint: you might use (ii) implies (iii) in the previous exercise.)

8.1.3. *Unimportant remark.* It may feel odd that in the definition of a locally closed embedding, we had to make a choice (as a composition of a closed embedding followed by an open embedding, rather than vice versa), but this type of issue comes up earlier: a subquotient of a group can be defined as the quotient of a subgroup, or a subgroup of a quotient. Which is the right definition? Or are they the same? (Hint: compositions of two subquotients should certainly be a subquotient, cf. Exercise 8.1.N.)

8.2 More projective geometry

We now interpret closed embeddings in terms of graded rings. Don't worry; most of the annoying foundational discussion of graded rings is complete, and we now just take advantage of our earlier work.

8.2.1. Example: Closed embeddings in projective space $\mathbb{P}^n_A$. Recall the definition of projective space $\mathbb{P}^n_A$ given in §4.4.10 (and the terminology defined there). Any *homogeneous* polynomial f in $x_0, \dots, x_n$ defines a closed subscheme. (Thus even if f doesn't make sense as a function, its vanishing scheme still makes sense.) On the open set U_i, the closed subscheme is $V(f(x_{0/i}, \dots, x_{n/i}))$, which we privately think of as $V(f(x_0, \dots, x_n)/x_i^{\deg f})$. On the overlap

$$U_i \cap U_j = \operatorname{Spec} A[x_{0/i}, \dots, x_{n/i}, x_{j/i}^{-1}]/(x_{i/i} - 1),$$

these functions on U_i and U_j don't exactly agree, but they agree up to a nonvanishing scalar, and hence cut out the same closed subscheme of $U_i \cap U_j$ (Exercise 8.1.I(b)):

$$f(x_{0/i}, \dots, x_{n/i}) = x_{j/i}^{\deg f} f(x_{0/j}, \dots, x_{n/j}).$$

Similarly, a collection of homogeneous polynomials in $A[x_0, \dots, x_n]$ cuts out a closed subscheme of $\mathbb{P}^n_A$. (Exercise 8.2.C will show that all closed subschemes of $\mathbb{P}^n_A$ are of this form.)

8.2.2. *Definition.* A closed subscheme cut out by a single (homogeneous) equation is called a **hypersurface** in $\mathbb{P}^n_A$. Of course, a hypersurface is not in general cut out by a single global function on $\mathbb{P}^n_A$: if $A = k$, there *are* no nonconstant global functions (Exercise 4.4.E). The **degree of a hypersurface** is the degree of the polynomial. (Implicit in this is that this notion can be determined from the subscheme itself. You may have the tools to prove this now, but we won't formally prove it until Exercise 18.6.H.) A hypersurface of degree 1 (resp. degree 2, 3, ...) is called a **hyperplane** (resp. **quadric, cubic, quartic, quintic, sextic, septic, octic, ... hypersurface**). If $n = 2$, a degree 1 hypersurface is called a **line**, and a degree 2 hypersurface is called a **conic curve**, or a **conic** for short. If $n = 3$, a hypersurface is called a **surface**. (In Chapter 11, we will justify the terms *curve* and *surface*.)

8.2.A. EXERCISE.
(a) Show that $wz = xy, x^2 = wy, y^2 = xz$ describes an irreducible subscheme in $\mathbb{P}^3_k$. In fact it is a curve, a notion we will define once we know what dimension is. This curve is called the **twisted cubic**. (The twisted cubic is a good nontrivial example of many things, so you should make friends with it as soon as possible. It implicitly appeared earlier in Exercise 3.6.F.)
(b) Show that the twisted cubic is isomorphic to $\mathbb{P}^1_k$.

We now extend this discussion to projective schemes in general.

8.2.B. EXERCISE. Suppose that $S_\bullet \twoheadrightarrow R_\bullet$ is a surjection of graded rings. Show that the domain of the induced morphism (Exercise 6.4.A) is $\operatorname{Proj} R_\bullet$, and that the induced morphism $\operatorname{Proj} R_\bullet \to \operatorname{Proj} S_\bullet$ is a closed embedding. (Exercise 15.4.H will show that every closed embedding in $\operatorname{Proj} S_\bullet$ is of this form.)

8.2.C. EXERCISE (CONVERSE TO EXERCISE 8.2.B). Suppose that $X \hookrightarrow \operatorname{Proj} S_\bullet$ is a closed embedding in a projective A-scheme (where $S_\bullet$ is a finitely generated graded A-algebra). Show that X is projective by describing it as $\operatorname{Proj}(S_\bullet/I)$, where I is a homogeneous ideal, of "projective functions" vanishing on X. (Many find this easier if $S_\bullet$ is generated in degree 1, and this case is the most important, so you may wish to deal only with this case.)

8.2.D. EXERCISE. Show that an injective linear map of k-vector spaces $V \hookrightarrow W$ induces a closed embedding $\mathbb{P}V \hookrightarrow \mathbb{P}W$. (This is another justification for the definition of $\mathbb{P}V$ in Example 4.5.12 in terms of the *dual* of V.)

8.2.3. *Definition.* The closed subscheme defined in Exercise 8.2.D is called a **linear space**. Once we know about dimension, we will call this closed subscheme a linear space of dimension $\dim V - 1 = \dim \mathbb{P}V$. More explicitly, a linear space of dimension n in $\mathbb{P}^N$ is any closed subscheme cut out by $N - n$ k-linearly independent homogeneous linear polynomials in $x_0, \dots, x_N$. A linear space of dimension 1 (resp. 2, n, $\dim \mathbb{P}W - 1$) is called a **line** (resp. **plane**, **n-plane**, **hyperplane**). (If the linear map in the previous exercise is not injective, then the hypothesis (6.4.0.1) of Exercise 6.4.A fails.)

8.2.E. EXERCISE (A SPECIAL CASE OF BÉZOUT'S THEOREM). Suppose $X \subset \mathbb{P}^n_k$ is a degree d hypersurface cut out by $f = 0$, and ℓ is a line not contained in X. A very special case of Bézout's theorem (Exercise 18.6.K) implies that X and ℓ meet with multiplicity d, "counted correctly". Make sense of this, by restricting the homogeneous degree d polynomial f to the line ℓ, and using the fact that a degree d polynomial in $k[x]$ has d roots, counted properly. (If it makes you feel better, assume $k = \overline{k}$.)

8.2.F. EXERCISE. Show that the map of graded rings $k[w, x, y, z] \to k[s, t]$ given by $(w, x, y, z) \mapsto (s^3, s^2t, st^2, t^3)$ induces a closed embedding $\mathbb{P}^1_k \hookrightarrow \mathbb{P}^3_k$, which yields an isomorphism of $\mathbb{P}^1_k$ with the twisted cubic (defined in Exercise 8.2.A — in fact, this will solve Exercise 8.2.A(b)). Doing this in a hands-on way will set you up well for the general Veronese construction of §8.2.6; see Exercise 8.2.J for a generalization.

8.2.4. A particularly nice case: when $S_\bullet$ is generated in degree 1.

Suppose $S_\bullet$ is a finitely generated graded ring generated in degree 1. Then S_1 is a finitely generated S_0-module, and the irrelevant ideal S_+ is generated in degree 1 (cf. Exercise 4.5.D(a)).

8.2.G. EXERCISE. Show that if $S_\bullet$ is generated (as an A-algebra) in degree 1 by $n+1$ elements $x_0, \dots, x_n$, then $\operatorname{Proj} S_\bullet$ may be described as a closed subscheme of $\mathbb{P}^n_A$ as follows. Consider $A^{\oplus(n+1)}$ as a free module with generators $t_0, \dots, t_n$ associated to $x_0, \dots, x_n$. The surjection of

$$\begin{array}{ccc} \operatorname{Sym}^\bullet A^{\oplus(n+1)} = A[t_0, t_1, \dots, t_n] & \twoheadrightarrow & S_\bullet \\ & & \\ t_i & \longmapsto & x_i \end{array}$$

implies $S_\bullet = A[t_0, t_1, \dots t_n]/I$, where I is a homogeneous ideal. (In particular, $\operatorname{Proj} S_\bullet$ can always be interpreted as a closed subscheme of some $\mathbb{P}^n_A$ if $S_\bullet$ is finitely generated in degree 1. Then using Exercises 6.4.D and 6.4.G, you can remove the hypothesis of generation in degree 1.)

This is analogous to the fact that if R is a finitely generated A-algebra, then choosing n generators of R as an algebra is the same as describing $\operatorname{Spec} R$ as a closed subscheme of $\mathbb{A}^n_A$. In the affine case this is "choosing coordinates"; in the projective case this is "choosing projective coordinates".

Recall (Exercise 4.4.F) that if k is algebraically closed, then we can interpret the closed points of $\mathbb{P}^n$ as the lines through the origin in $(n+1)$-space. The following exercise states this more generally.

8.2.H. EXERCISE. Suppose $S_\bullet$ is a finitely generated graded ring over an algebraically closed field k, generated in degree 1 by $x_0, \dots, x_n$, inducing closed embeddings $\operatorname{Proj} S_\bullet \hookrightarrow \mathbb{P}^n$ and $\operatorname{Spec} S_\bullet \hookrightarrow \mathbb{A}^{n+1}$. Give a bijection between the closed points of $\operatorname{Proj} S_\bullet$ and the "lines through the origin" in $\operatorname{Spec} S_\bullet \subset \mathbb{A}^{n+1}$.

8.2.5. *A second proof that finite morphisms are closed.* This interpretation of $\operatorname{Proj} S_\bullet$ as a closed subscheme of projective space (when it is generated in degree 1) yields the following second proof of the fact (shown in Exercise 7.3.M) that finite morphisms are closed. Suppose $\pi : X \to Y$ is a finite morphism. The question is local on the target, so it suffices to consider the affine case $Y = \operatorname{Spec} B$. It suffices to show that $\pi(X)$ is closed. Then by Exercise 7.3.J, X is a projective B-scheme, and hence by the Fundamental Theorem of Elimination Theory 7.4.7, its image is closed.

8.2.6. Important classical construction: The Veronese embedding.
Suppose $S_\bullet = k[x, y]$, so $\operatorname{Proj} S_\bullet = \mathbb{P}^1_k$. Then $S_{2\bullet} = k[x^2, xy, y^2] \subset k[x, y]$ (see §6.4.4 on the Veronese subring). We identify this subring as follows.

8.2.I. EXERCISE. Let $u = x^2$, $v = xy$, $w = y^2$. Show that $S_{2\bullet} \cong k[u, v, w]/(uw - v^2)$, by mapping u, v, w to x^2, xy, y^2, respectively.

We have a graded ring generated by three elements in degree 1. Thus we think of it as sitting "in" $\mathbb{P}^2$, via the construction of §8.2.G. This can be interpreted as "$\mathbb{P}^1$ as a conic in $\mathbb{P}^2$".

8.2.7. Thus if k is algebraically closed of characteristic not 2, using the fact that we can diagonalize quadrics (Exercise 5.4.J), the conics in $\mathbb{P}^2$, up to change of coordinates, come in only a few flavors: sums of 3 squares (e.g. our conic of the previous exercise), sums of 2 squares (e.g. $y^2 - x^2 = 0$, the union of 2 lines), a single square (e.g. $x^2 = 0$, which looks set-theoretically like a line, and is nonreduced), and 0 (perhaps not a conic at all). Thus we have proved: any plane conic (over an algebraically closed field of characteristic not 2) that can be written as the sum of three nonzero squares is isomorphic to $\mathbb{P}^1$. (See Exercise 6.5.F for a closely related fact.)

We now soup up this example.

8.2.J. EXERCISE. We continue to take $S_\bullet = k[x,y]$. Show that $\operatorname{Proj} S_{d\bullet}$ is given by the equations that

$$\begin{pmatrix} y_0 & y_1 & \cdots & y_{d-1} \\ y_1 & y_2 & \cdots & y_d \end{pmatrix}$$

is rank 1 (i.e., that all the 2×2 minors vanish). This is called the **degree d rational normal curve** "in" $\mathbb{P}^d$. You did the *twisted cubic* case $d = 3$ in Exercises 8.2.A and 8.2.F.

8.2.8. *Definition.* More generally, if $S_\bullet = k[x_0, \dots, x_n]$, then $\operatorname{Proj} S_{d\bullet} \subset \mathbb{P}^{N-1}$ (where N is the dimension of the vector space of homogeneous degree d polynomials in $x_0, \dots, x_n$) is called the **d-uple embedding** or **d-uple Veronese embedding**. The reason for the word "embedding" is historical; we really mean closed embedding. (Combining Exercise 6.4.E with Exercise 8.2.G shows that $\operatorname{Proj} S_\bullet \to \mathbb{P}^{N-1}$ is a closed embedding.)

8.2.K. COMBINATORIAL EXERCISE (CF. REMARK 4.5.3). Show that $N = \binom{n+d}{d}$.

8.2.L. UNIMPORTANT EXERCISE. Find six linearly independent quadratic equations vanishing on the **Veronese surface** $\operatorname{Proj} S_{2\bullet}$ where $S_\bullet = k[x_0, x_1, x_2]$, which sits naturally in $\mathbb{P}^5$. (You needn't show that these equations generate all the equations cutting out the Veronese surface, although this is in fact true.) Possible hint: use the identity

$$\det \begin{pmatrix} x_0x_0 & x_0x_1 & x_0x_2 \\ x_1x_0 & x_1x_1 & x_1x_2 \\ x_2x_0 & x_2x_1 & x_2x_2 \end{pmatrix} = 0.$$

8.2.9. Rulings on the quadric surface. We return to rulings on the quadric surface, which first appeared in the optional (starred) section §4.4.12.

8.2.M. USEFUL GEOMETRIC EXERCISE: THE RULINGS ON THE QUADRIC SURFACE $wz = xy$. This exercise is about the lines on the quadric surface X given by $wz - xy = 0$ in $\mathbb{P}^3_k$ (where the projective coordinates on $\mathbb{P}^3_k$ are ordered w, x, y, z). This construction arises all over the place in nature.

(a) Suppose a_0 and b_0 are elements of k, not both zero. Make sense of the statement: as $[c, d]$ varies in $\mathbb{P}^1$, $[a_0c, b_0c, a_0d, b_0d]$ is a line in the quadric surface. (This describes "a family of lines parametrized by $\mathbb{P}^1$", although we can't yet make this precise.) Find another family of lines. These are the two **rulings** of the smooth quadric surface.

(b) Show that through every k-valued point of the quadric surface X, there passes

one line from each ruling.
(c) Show there are no other lines. (There are many ways of proceeding. At risk of predisposing you to one approach, here is a germ of an idea. Suppose L is a line on the quadric surface, and $[1, x, y, z]$ and $[1, x', y', z']$ are distinct points on it. Because they are both on the quadric, $z = xy$ and $z' = x'y'$. Because all of L is on the quadric, $(1+t)(z+tz') - (x+tx')(y+ty') = 0$ for all t. After some algebraic manipulation, this translates into $(x-x')(y-y') = 0$. How can this be made watertight? Another possible approach uses Bézout's theorem, in the form of Exercise 8.2.E.)

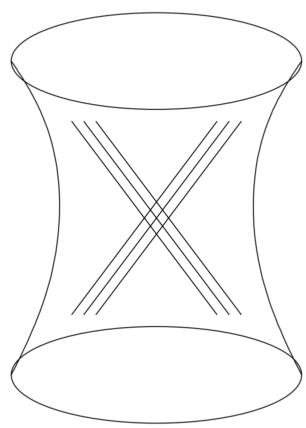


FIGURE 8.2. The two rulings on the quadric surface $V(wz - xy) \subset \mathbb{P}^3$. One ruling contains the line $V(w, x)$ and the other contains the line $V(w, y)$.

Hence by Exercise 5.4.J, if we are working over an algebraically closed field of characteristic not 2, we have shown that all rank 4 quadric surfaces have two rulings of lines. (In Example 9.6.2, we will recognize this quadric as $\mathbb{P}^1 \times \mathbb{P}^1$.)

8.2.10. *Side Remark.* The existence of these two rulings is the first chapter of a number of important and beautiful stories. The second chapter is often the following. If k is an algebraically closed field, then a "maximal rank" (Exercise 5.4.J) quadric hypersurface X of dimension m contains no linear spaces of dimension greater than $m/2$. (We will see in Exercise 12.3.D that the maximal rank quadric hypersurfaces are the "smooth" quadrics.) If $m = 2a+1$, then X contains an irreducible $\binom{a+2}{2}$-dimensional family of a-planes. If $m = 2a$, then X contains two irreducible $\binom{a+1}{2}$-dimensional families of a-planes, and furthermore two a-planes Λ and Λ' are in the same family if and only if $\dim(\Lambda \cap \Lambda') \equiv a \pmod 2$. These families of linear spaces are also called **rulings**. (For more information, see [**GH1**, §6.1, p. 735, Prop.].) You already know enough to think through the examples of $m = 0$, 1, and 2. The case of 3 is discussed in Exercise 16.7.K.

8.2.11. Weighted projective space. If we put a nonstandard weighting on the variables of $k[x_1, \dots, x_n]$ — say we give x_i degree d_i — then $\operatorname{Proj} k[x_1, \dots, x_n]$ is called **weighted projective space** $\mathbb{P}(d_1, d_2, \dots, d_n)$.

8.2.N. EXERCISE. Show that $\mathbb{P}(m, n)$ is isomorphic to $\mathbb{P}^1$. Show that $\mathbb{P}(1,1,2) \cong \operatorname{Proj} k[u, v, w, z]/(uw - v^2)$. Hint: do this by looking at the even-graded parts of $k[x_0, x_1, x_2]$, cf. Exercise 6.4.D. (This is a projective cone over a conic curve. Over a field of characteristic not 2, it is isomorphic to the traditional cone $x^2 + y^2 = z^2$ in $\mathbb{P}^3$, see Figure 8.3.)

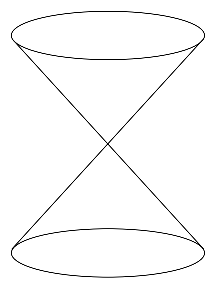


FIGURE 8.3. The cone $\operatorname{Spec} k[x, y, z]/(z^2 - x^2 - y^2)$.

8.2.12. Affine and projective cones.

If $S_\bullet$ is a finitely generated graded ring, then the **affine cone** of $\operatorname{Proj} S_\bullet$ is $\operatorname{Spec} S_\bullet$. Caution: this terminology is not ideal, as this construction depends on $S_\bullet$, not just on $\operatorname{Proj} S_\bullet$. As motivation, consider the graded ring $S_\bullet = \mathbb{C}[x, y, z]/(z^2 - x^2 - y^2)$. Figure 8.3 is a sketch of $\operatorname{Spec} S_\bullet$. (Here we draw the "real picture" of $z^2 = x^2 + y^2$ in $\mathbb{R}^3$.) It is a cone in the traditional sense; the origin $(0, 0, 0)$ is the "cone point".

This gives a useful way of picturing Proj (even over arbitrary rings, not just $\mathbb{C}$). Intuitively, you could imagine that if you discarded the origin, you would get something that would project onto $\operatorname{Proj} S_\bullet$. The following exercise makes that precise.

8.2.O. EXERCISE. If $\operatorname{Proj} S_\bullet$ is a projective scheme over a field k, describe a natural morphism $\operatorname{Spec} S_\bullet \setminus V(S_+) \to \operatorname{Proj} S_\bullet$. (Can you see why $V(S_+)$ is a single point, and should reasonably be called the origin?)

This readily generalizes to the following exercise, which again motivates the terminology "irrelevant".

8.2.P. EASY EXERCISE. If $S_\bullet$ is a finitely generated graded ring, describe a natural morphism $\operatorname{Spec} S_\bullet \setminus V(S_+) \to \operatorname{Proj} S_\bullet$.

In fact, it can be made precise that $\operatorname{Proj} S_\bullet$ is the quotient (by the multiplicative group of scalars) of the affine cone minus the origin.

8.2.13. *Definition.* The **projective cone** of $\operatorname{Proj} S_\bullet$ is $\operatorname{Proj} S_\bullet[T]$, where T is a new variable of degree 1. For example, the cone corresponding to the conic $\operatorname{Proj} k[x,y,z]/(z^2-x^2-y^2)$ is $\operatorname{Proj} k[x,y,z,T]/(z^2-x^2-y^2)$. The projective cone is sometimes called the **projective completion** of $\operatorname{Spec} S_\bullet$.

8.2.Q. LESS IMPORTANT EXERCISE (CF. §4.5.1). Show that the "projective cone" $\operatorname{Proj} S_\bullet[T]$ of $\operatorname{Proj} S_\bullet$ has a closed subscheme isomorphic to $\operatorname{Proj} S_\bullet$ (informally, corresponding to $T=0$), whose complement (the distinguished open set $D(T)$) is isomorphic to the affine cone $\operatorname{Spec} S_\bullet$.

This construction can be usefully pictured as the affine cone union some points "at infinity", and the points at infinity form the Proj. (The reader may wish to start with Figure 8.3, and try to visualize the conic curve "at infinity", and then compare this visualization to Figure 4.9.)

We have thus completely described the algebraic analog of the classical picture of 4.5.1.

8.3 Smallest closed subschemes such that ...

We now define a series of notions that are all of the form "the smallest closed subscheme such that something or other is true". One example will be the notion of scheme-theoretic closure of a locally closed embedding, which will allow us to interpret locally closed embeddings in three equivalent ways (open subscheme intersect closed subscheme; open subscheme of closed subscheme; and closed subscheme of open subscheme — cf. Exercise 8.1.M).

8.3.1. Scheme-theoretic image.

We start with the notion of scheme-theoretic image. Set-theoretic images are badly behaved in general (§7.4.1), and even with reasonable hypotheses such as those in Chevalley's theorem 7.4.2, things can be confusing. For example, there is no reasonable way to impose a scheme structure on the image of $\mathbb{A}^2_k \to \mathbb{A}^2_k$ given by $(x,y) \mapsto (x,xy)$. It will be useful (e.g. Exercise 8.3.C) to define a notion of a closed subscheme of the target that "best approximates" the image. This will incorporate the notion that the image of something with nonreduced structure ("fuzz") can also have nonreduced structure. As usual, we will need to impose reasonable hypotheses to make this notion behave well (see Theorem 8.3.4 and Corollary 8.3.5).

8.3.2. *Definition.* Suppose $i : Z \hookrightarrow Y$ is a closed subscheme, giving an exact sequence $0 \to \mathscr{I}_{Z/Y} \to \mathscr{O}_Y \to i_*\mathscr{O}_Z \to 0$. We say that **the image of** $\pi : X \to Y$ **lies in** Z if the composition $\mathscr{I}_{Z/Y} \to \mathscr{O}_Y \to \pi_*\mathscr{O}_X$ is zero. Informally, locally, functions vanishing on Z pull back to the zero function on X. If the image of π lies in some subschemes Z_j (as j runs over some index set), it clearly lies in their intersection (cf. Exercise 8.1.J(a) on intersections of closed subschemes). We then define the **scheme-theoretic image of** π, a closed subscheme of Y, as the "smallest closed subscheme containing the image", i.e., the intersection of all closed subschemes containing the image. In particular (and in our first examples), if Y is affine, the scheme-theoretic image is cut out by functions on Y that are 0 when pulled back to X.

Example 1. Consider $\pi : \operatorname{Spec} k[\epsilon]/(\epsilon^2) \to \operatorname{Spec} k[x] = \mathbb{A}^1_k$ given by $x \mapsto \epsilon$. Then the scheme-theoretic image of π is given by $\operatorname{Spec} k[x]/(x^2)$ (the polynomials pulling back to 0 are precisely multiples of x^2). Thus the image of the fuzzy point still has some fuzz.

Example 2. Consider $\pi : \operatorname{Spec} k[\epsilon]/(\epsilon^2) \to \operatorname{Spec} k[x] = \mathbb{A}^1_k$ given by $x \mapsto 0$. Then the scheme-theoretic image is given by $k[x]/x$: the image is reduced. In this picture, the fuzz is "collapsed" by π.

Example 3. Consider $\pi : \operatorname{Spec} k[t, t^{-1}] = \mathbb{A}^1 - \{0\} \to \mathbb{A}^1 = \operatorname{Spec} k[u]$ given by $u \mapsto t$. Any function $g(u)$ which pulls back to 0 as a function of t must be the zero-function. Thus the scheme-theoretic image is everything. The set-theoretic image, on the other hand, is the distinguished open set $\mathbb{A}^1 - \{0\}$. Thus in not-too-pathological cases, the underlying set of the scheme-theoretic image is not the set-theoretic image. But the situation isn't terrible: the underlying set of the scheme-theoretic image must be closed, and indeed it is the closure of the set-theoretic image. We might imagine that in reasonable cases this will be true, and in even nicer cases, the underlying set of the scheme-theoretic image will be set-theoretic image. We will later see that this is indeed the case (§8.3.6).

But sadly pathologies can sometimes happen in, well, pathological situations.

Example 4 (see Figure 8.4). Let $X = \coprod \operatorname{Spec} k[\epsilon_n]/((\epsilon_n)^n)$ (a scheme which appeared in the hint to Exercise 5.2.E) and $Y = \operatorname{Spec} k[x]$, and define $X \to Y$ by $x \to \epsilon_n$ on the nth component of X. Then if a function $g(x)$ on Y pulls back to 0 on X, then its Taylor expansion is 0 to order n (by examining the pullback to the nth component of X) for all n, so $g(x)$ must be 0. (This argument will be vastly generalized in Exercise 12.9.A(b).) Thus the scheme-theoretic image is $V(0)$ on Y, i.e., Y itself, while the set-theoretic image is easily seen to be just the origin (the closed point 0). (This morphism implicitly arises in Caution/Example 8.3.11.)

[to be made]

FIGURE 8.4. A pathological morphism (Example 4 of Definition 8.3.2)

8.3.3. *Criteria for computing scheme-theoretic images affine-locally.* Example 4 clearly is weird though, and we can show that in "reasonable circumstances" such pathology doesn't occur. It would be great to compute the scheme-theoretic image affine-locally. On the affine open set $\operatorname{Spec} B \subset Y$, define the ideal $I(B) \subset B$ of functions which pull back to 0 on X. Formally, $I(B) := \ker(B \to \Gamma(\operatorname{Spec} B, \pi_*(\mathscr{O}_X))$. Then if for each such B, and each $g \in B$, $I(B) \otimes_B B_g \to I(B_g)$ is an isomorphism, then we will have defined the scheme-theoretic image as a closed subscheme (see Exercise 8.1.H). Clearly each function on $\operatorname{Spec} B$ that vanishes when pulled back to $\pi^{-1}(\operatorname{Spec} B)$ also vanishes when restricted to $D(g)$ and then pulled back to $\pi^{-1}(D(g))$. So the question is: given a function r/g^n on $D(g)$ that pulls back to zero on $\pi^{-1}(D(g))$, is it true that for some m, $rg^m = 0$ when pulled back to

$\pi^{-1}(\operatorname{Spec} B)$? Here are three cases where the answer is "yes". (I would like to add a picture here, but I can't think of one that would enlighten more people than it would confuse. So you should try to draw one that suits you.) For each affine in the source, there is some m which works. There is one that works for all affines in a cover (i) if $m = 1$ always works, or (ii) if there are only a finite number of affines in the cover.

(i) The answer is yes if $\pi^{-1}(\operatorname{Spec} B)$ is reduced: we simply take $m = 1$ (as r vanishes on $\operatorname{Spec} B_g$ and g vanishes on $V(g)$, so rg vanishes on $\operatorname{Spec} B = \operatorname{Spec} B_g \cup V(g)$.)

(ii) The answer is also yes if $\pi^{-1}(\operatorname{Spec} B)$ is affine, say $\operatorname{Spec} A$: if $r' = \pi^\sharp r$ and $g' = \pi^\sharp g$ in A, then if $r' = 0$ on $D(g')$, then there is an m such that $r'(g')^m = 0$ (as the statement $r' = 0$ in $D(g')$ means precisely this fact — the functions on $D(g')$ are $A_{g'}$).

(ii)′ More generally, the answer is yes if $\pi^{-1}(\operatorname{Spec} B)$ is quasicompact: cover $\pi^{-1}(\operatorname{Spec} B)$ with finitely many affine open sets. For each one there will be some m_i so that $rg^{m_i} = 0$ when pulled back to this open set. Then let $m = \max(m_i)$. (We see again that quasicompactness is our friend!)

In conclusion, we have proved the following (subtle) theorem.

8.3.4. Theorem. — *Suppose $\pi : X \to Y$ is a morphism of schemes. If X is reduced or π is quasicompact, then the scheme-theoretic image of π may be computed affine-locally: on* $\operatorname{Spec} A \subset Y$, *it is cut out by the functions that pull back to 0.*

8.3.5. Corollary. — *Under the hypotheses of Theorem 8.3.4, the closure of the set-theoretic image of π is the underlying set of the scheme-theoretic image.*

(Example 4 above shows that we cannot excise these hypotheses.)

8.3.6. In particular, if the set-theoretic image is closed (e.g. if π is finite or projective), the set-theoretic image is the underlying set of the scheme-theoretic image, as promised in Example 3 above.

8.3.7. *Proof of Corollary 8.3.5.* The set-theoretic image is in the underlying set of the scheme-theoretic image. (Check this!) The underlying set of the scheme-theoretic image is closed, so the closure of the set-theoretic image is contained in the underlying set of the scheme-theoretic image. On the other hand, if U is the complement of the closure of the set-theoretic image, $\pi^{-1}(U) = \varnothing$. Under these hypotheses, the scheme theoretic image can be computed locally, so the scheme-theoretic image is the empty set on U. $\square$

We conclude with a few stray remarks.

8.3.A. EASY EXERCISE. If X is reduced, show that the scheme-theoretic image of $\pi : X \to Y$ is also reduced.

More generally, you might expect there to be no unnecessary nonreduced structure on the image not forced by nonreduced structure on the source. We make this precise in the locally Noetherian case, when we can talk about associated points.

8.3.B. ⋆ UNIMPORTANT EXERCISE. If $\pi : X \to Y$ is a *quasicompact* morphism of locally Noetherian schemes, show that the associated points of the image subscheme are a subset of the image of the associated points of X. (The example of $\coprod_{a \in \mathbb{C}} \operatorname{Spec} \mathbb{C}[t]/(t-a) \to \operatorname{Spec} \mathbb{C}[t]$ shows what can go wrong if you give up quasicompactness — note that reducedness of the source doesn't help.) Hint: reduce to the case where X and Y are affine. (Can you develop your geometric intuition so that this becomes plausible to you?)

8.3.8. Scheme-theoretic closure of a locally closed subscheme.

We define the **scheme-theoretic closure** of a locally closed embedding $\pi : X \to Y$ as the scheme-theoretic image of X. (A shorter phrase for this is **schematic closure**, although this nomenclature has not fully caught on.)

8.3.C. TRICKY EXERCISE. If a locally closed embedding $V \to X$ is quasicompact (e.g. if V is Noetherian, Exercise 7.3.B(a)), or if V is reduced, show that (iii) implies (i) and (ii) in Exercise 8.1.M. Thus in this fortunate situation, a locally closed embedding can be thought of in three different ways, whichever is convenient.

8.3.D. UNIMPORTANT EXERCISE, USEFUL FOR INTUITION. If $\pi : X \to Y$ is a locally closed embedding into a locally Noetherian scheme (so X is also locally Noetherian), then the associated points of the scheme-theoretic closure are (naturally in bijection with) the associated points of X. (Hint: Exercise 8.3.B.) Informally, we get no nonreduced structure on the scheme-theoretic closure not "forced by" that on X.

8.3.9. The (reduced) subscheme structure on a closed subset.

Suppose X^{set} is a closed sub*set* of a scheme Y. Then we can define a canonical scheme structure X on X^{set} that is reduced. We could describe it as being cut out by those functions whose values are zero at all the points of X^{set}. On the affine open set $\operatorname{Spec} B$ of Y, if the set X^{set} corresponds to the radical ideal $I = I(X^{\text{set}})$ (recall the $I(\cdot)$ function from §3.7), the scheme X corresponds to $\operatorname{Spec} B/I$. You can quickly check that this behaves well with respect to any distinguished inclusion $\operatorname{Spec} B_f \hookrightarrow \operatorname{Spec} B$. We could also consider this construction as an example of a scheme-theoretic image in the following crazy way: let W be the scheme that is a disjoint union of all the points of X^{set}, where the point corresponding to p in X^{set} is Spec of the residue field of $\mathscr{O}_{Y,p}$. Let $\rho : W \to Y$ be the "canonical" map sending "p to p", and giving an isomorphism on residue fields. Then the scheme structure on X is the scheme-theoretic image of ρ. A third definition: it is the smallest closed subscheme whose underlying set contains X^{set}.

This construction is called the (induced) **reduced subscheme structure** on the closed subset X^{set}. (Vague exercise: Make a definition of the reduced subscheme structure precise and rigorous to your satisfaction.)

8.3.E. EXERCISE. Show that the underlying set of the induced reduced subscheme $X \to Y$ is indeed the closed subset X^{set}. Show that X is reduced.

8.3.10. Reduced version of a scheme.

In the main interesting case where X^{set} is all of Y, we obtain a *reduced closed subscheme* $Y^{\text{red}} \to Y$, called the **reduction** of Y. On the affine open subset $\operatorname{Spec} B \hookrightarrow Y$, $Y^{\text{red}} \hookrightarrow Y$ corresponds to the nilradical $\mathfrak{N}(B)$ of B. The *reduction* of a scheme is the "reduced version" of the scheme, and informally corresponds to "shearing off the fuzz".

An alternative equivalent definition: on the affine open subset $\operatorname{Spec} B \hookrightarrow Y$, the reduction of Y corresponds to the ideal $\mathfrak{N}(B) \subset B$ of nilpotents. As for any $f \in B$, $\mathfrak{N}(B)_f = \mathfrak{N}(B_f)$, by Exercise 8.1.H this defines a closed subscheme.

8.3.11. $\star$ *Caution/Example.* It is not true that for *every* open subset $U \subset Y$, $\Gamma(U, \mathscr{O}_{Y^{\text{red}}})$ is $\Gamma(U, \mathscr{O}_Y)$ modulo its nilpotents. For example, on $Y = \coprod \operatorname{Spec} k[x]/(x^n)$, the function x is not nilpotent, but is 0 on Y^{red}, as it is "locally nilpotent". This may remind you of Example 4 after Definition 8.3.2.

8.3.12. Scheme-theoretic support of a quasicoherent sheaf. Similar ideas are used in the definition of the scheme-theoretic support of a quasicoherent sheaf, see Exercise 18.9.B.

8.4 Effective Cartier divisors, regular sequences and regular embeddings

We now introduce regular embeddings, an important class of locally closed embeddings. Locally closed embeddings of regular schemes in regular schemes are one important example of regular embeddings (Exercise 12.2.K). Effective Cartier divisors, basically the codimension 1 case, will turn out to be repeatedly useful as well — to see *how* useful, notice how often it appears in the index. We begin with this case.

8.4.1. Locally principal closed subschemes, and effective Cartier divisors.

A closed subscheme is **locally principal** if on each open set in a small enough open cover it is cut out by a single equation (i.e. by a principal ideal, hence the terminology). More specifically, a locally principal closed subscheme is a closed embedding $\pi : X \to Y$ for which there is an open cover $\{U_i : i \in I\}$ of Y for which for each i, the restricted morphism $\pi^{-1}(U_i) \to U_i$ is (isomorphic as a U_i-scheme to) the closed subscheme $V(s_i)$ of U_i for some $s_i \in \Gamma(U_i, \mathscr{O}_X)$. In particular, if $\pi : X \to Y$ is a locally principal closed subscheme, then the open cover of Y can be chosen to be open affines: Simply cover each open set U_i by open affines, and restrict each s_i to each open affine.

For example, hyperplanes in $\mathbb{P}^n_A$ (Definition 8.2.2) are locally principal: each homogeneous polynomial in $n+1$ variables defines a locally principal closed subscheme of $\mathbb{P}^n_A$. (Warning: local principality is not an affine-local condition, see Exercise 5.4.N! Also, the example of a projective hypersurface, §8.2.1 shows that a locally principal closed subscheme need not be cut out by a globally-defined function.)

If the ideal sheaf is locally generated by a function that is not a zerodivisor, we call the closed subscheme an *effective Cartier divisor*. More precisely: if $\pi : X \to Y$ is a closed embedding, and there is a cover Y by *affine* open subsets $\operatorname{Spec} A_i \subset Y$, and there exist non-zerodivisors $t_i \in A_i$ with $V(t_i) = X|_{\operatorname{Spec} A_i}$ (scheme-theoretically — i.e., the ideal sheaf of X over $\operatorname{Spec} A_i$ is generated by t_i), then we say that X is an **effective Cartier divisor** on Y. (We will not explain the origin of the phrase, as it is not relevant for this point of view.)

8.4.A. EXERCISE. Suppose $t \in A$ is a non-zerodivisor. Show that t is a non-zerodivisor in $A_{\mathfrak{p}}$ for each prime $\mathfrak{p}$.

8.4.2. *Caution.* If D is an effective Cartier divisor on an affine scheme $\operatorname{Spec} A$, it is not necessarily true that $D = V(t)$ for some $t \in A$ (see Exercise 14.2.M — §19.11.10 gives a different flavor of example). In other words, the condition of a closed subscheme being an effective Cartier divisor can be verified on *an* affine cover, but cannot be checked on an *arbitrary* affine cover — it is not an affine-local condition in this obvious a way.

8.4.B. EXERCISE. Suppose X is a locally Noetherian scheme, and $t \in \Gamma(X, \mathscr{O}_X)$ is a function on it. Show that t (or more precisely the closed subscheme $V(t)$) is an effective Cartier divisor if and only if it doesn't vanish on any associated point of X.

8.4.C. UNIMPORTANT EXERCISE. Suppose $V(t) = V(t') \hookrightarrow \operatorname{Spec} A$ is an effective Cartier divisor, with t and t' non-zerodivisors in A. Show that t is an invertible function times t'.

The idea of an effective Cartier divisor leads us to the notion of regular sequences. (We will close the loop in Exercise 8.4.H, where we will interpret effective Cartier divisors on any reasonable scheme as regular embeddings of codimension 1.)

8.4.3. Regular sequences.

The definition of regular sequence is the algebraic version of the following geometric idea: locally, we take an effective Cartier divisor (a non-zerodivisor); then an effective Cartier divisor on that; then an effective Cartier divisor on that; and so on, a finite number of times. A little care is necessary; for example, we might want this to be independent of the order of the equations imposed, and this is true only when we say this in the right way.

We make the definition of regular sequence for a ring A, and more generally for an A-module M.

8.4.4. *Definition.* If M is an A-module, a sequence $x_1, \dots, x_r \in A$ is called an **M-regular sequence** (or a **regular sequence for** M) if for each i, x_i is not a zerodivisor for $M/(x_1, \dots, x_{i-1})M$. (The case $i = 1$ should be interpreted as: "x_1 is not a zerodivisor of M.")

In the case most relevant to us, when $M = A$, this should be seen as a reasonable approximation of a "complete intersection", and indeed we will use this as the definition (§8.4.7). An A-regular sequence is just called a **regular sequence**.

8.4.D. EXERCISE. If M is an A-module, show that an M-regular sequence remains regular upon any localization. (More generally, your argument will likely show that sequences remain regular upon any flat ring extension, but we will not need this, and you may not know what this means.)

8.4.E. EXERCISE. If x, y is an M-regular sequence, show that x^N, y is an M-regular sequence. (More generally, if $x_1, \dots, x_n$ is a regular sequence, and $a_1, \dots, a_n \in \mathbb{Z}^+$, then $x_1^{a_1}, \dots, x_n^{a_n}$ is a regular sequence, see [E, Ex. 17.5], [Mat2, Thm. 16.1], or

[**Mat1**, Thm. 26]. We give this easier special case as an exercise because we will use it.)

8.4.5. *Interesting example.* We now give an example ([**E**, Example 17.3]) showing that the order of a regular sequence matters. Suppose $A = k[x, y, z]/(x-1)z$, so $X = \operatorname{Spec} A$ is the union of the $z = 0$ plane and the $x = 1$ plane — X is reduced and has two components (see Figure 8.5). You can readily verify that x is a non-zerodivisor for A ($x = 0$ misses one component of X, and doesn't vanish entirely on the other), and that the effective Cartier divisor, $X' = \operatorname{Spec} k[x, y, z]/(x, z)$ is integral. Then $(x-1)y$ gives an effective Cartier divisor on X' (it doesn't vanish entirely on X'), so x, $(x-1)y$ is a regular sequence for A. However, $(x-1)y$ is *not* a non-zerodivisor of A, as it *does* vanish entirely on one of the two components. Thus $(x-1)y$, x is *not* a regular sequence. The reason that reordering the regular sequence x, $(x-1)y$ ruins regularity is clear: there is a locus on which $(x-1)y$ isn't effective Cartier, but it disappears if we enforce $x = 0$ first. The problem is one of "nonlocality" — "near" $x = y = z = 0$ there is no problem. This may motivate the fact that in the (Noetherian) local situation, this problem disappears. We now make this precise.

[to be made]

FIGURE 8.5. Order matters in a regular sequence (in the "non-local" situation)

8.4.6. Theorem. — *Suppose $(A, \mathfrak{m})$ is a Noetherian local ring, and M is a finitely generated A-module. Then any M-regular sequence $(x_1, \dots, x_r)$ in $\mathfrak{m}$ remains a regular sequence upon any reordering.*

(Dieudonné showed that Noetherian hypotheses are necessary in Theorem 8.4.6, [**Di**].)

Before proving Theorem 8.4.6 (in Exercise 8.4.F), we prove the first nontrivial case, when $r = 2$. This discussion is secretly a baby case of the Koszul complex. We will gratuitously use the language of spectral sequences, in order to give you practice.

Suppose $x, y \in \mathfrak{m}$, and x, y is an M-regular sequence. Translation: $x \in \mathfrak{m}$ is a non-zerodivisor on M, and $y \in \mathfrak{m}$ is a non-zerodivisor on M/xM.

Consider the double complex

(8.4.6.1)
$$\begin{tikzcd} M \arrow[r, "\times(-x)"] & M \\ M \arrow[u, "\times y"] \arrow[r, "\times x"] & M \arrow[u, "\times y"'] \end{tikzcd}$$

where the bottom left is considered to be in position $(0, 0)$. (The only reason for the minus sign in the top row is solely our arbitrary preference for anti-commuting rather than commuting squares in §1.7.1, but it really doesn't matter.)

We compute the cohomology of the total complex using a (simple) spectral sequence, beginning with the rightward orientation. (The use of spectral sequences here, as in many of our other applications, is overkill; we do this partially in order to get practice with the machine.) On the first page, we have

$$\begin{tikzcd} (0:(x)) & M/xM \\ (0:(x)) \arrow[u, "\times y"] & M/xM \arrow[u, "\times y"'] \end{tikzcd}$$

The entries $(0:(x))$ in the first column are 0, as x is a non-zerodivisor on M. Taking homology in the vertical direction to obtain the second page, we find

(8.4.6.2)
$$\begin{matrix} 0 & M/(x,y)M \\ & \\ 0 & 0 \end{matrix}$$

using the fact that y is a non-zerodivisor on M/xM. The sequence clearly converges here. Thus the original double complex (8.4.6.1) only has nonzero cohomology in degree 2, where it is $M/(x,y)M$.

Now we run the spectral sequence on (8.4.6.1) using the upward orientation. The first page of the sequence is:

$$\begin{tikzcd} M/yM \arrow[r, "\times(-x)"] & M/yM \\ (0:(y)) \arrow[r, "\times x"] & (0:(y)) \end{tikzcd}$$

The sequence must converge to (8.4.6.2) after the next step. From the top row, we see that multiplication by x must be injective on M/yM, so x is a non-zerodivisor on M/yM. From the bottom row, multiplication by x gives an isomorphism of $(0:(y))$ with itself. As $x \in \mathfrak{m}$, by version 2 of Nakayama's Lemma (Lemma 7.2.9), this implies that $(0:(y)) = 0$, so y is a non-zerodivisor on M. Thus we have shown that y, x is a regular sequence on M — the $n = 2$ case of Theorem 8.4.6.

8.4.F. EASY EXERCISE. Prove Theorem 8.4.6. (Hint: show it first in the case of a reordering where only two adjacent x_i are swapped, using the $n = 2$ case just discussed.) Where are the Noetherian hypotheses used?

8.4.7. Regular embeddings.

Suppose $\pi : X \hookrightarrow Y$ is a locally closed embedding. We say that π is a **regular embedding (of codimension** r**) at a point** $p \in X$ if in the local ring $\mathscr{O}_{Y,p}$, the ideal of X is generated by a regular sequence (of length r). We say that π is a **regular embedding (of codimension** r**)** if it is a regular embedding (of codimension r) at all $p \in X$. (Another reasonable name for a regular embedding might be "local complete intersection". Unfortunately, "local complete intersection morphism", or "lci morphism", is already used for a related notion, see [**Stacks**, tag 068E].)

Our terminology uses the word "codimension", which we have not defined as a word on its own. The reason for using this word will become clearer once you meet Krull's Principal Ideal Theorem 11.3.3 and Krull's Height Theorem 11.3.7.

8.4.G. EXERCISE (THE CONDITION OF A LOCALLY CLOSED EMBEDDING BEING A REGULAR EMBEDDING IS OPEN). Show that if a locally closed embedding $\pi : X \hookrightarrow Y$ of locally Noetherian schemes is a regular embedding at p, then it is a regular embedding in some neighborhood of p in X. Hint: reduce to the case where π is a closed embedding, and then where Y (hence X) is affine — say $Y = \operatorname{Spec} B$, $X = \operatorname{Spec} B/I$, and $p = [\mathfrak{p}]$ — and there are $f_1, \dots, f_r$ such that in $\mathscr{O}_{Y,p}$, the images of the f_i are a regular sequence generating $I_\mathfrak{p}$. We wish to show that $(f_1, \dots, f_r) = I$ "in a neighborhood of p". Prove the following fact in algebra: if I and J are ideals of a Noetherian ring A, and $\mathfrak{p} \subset A$ is a prime ideal such that $I_\mathfrak{p} = J_\mathfrak{p}$, show that there exists $a \in A \setminus \mathfrak{p}$ such that $I_a = J_a$ in A_a. To do this, show that it suffices to consider the special case $I \subset J$, by considering $I \cap J$ and J instead of I and J. To show this special case, let $K = J/I$, a finitely generated module, and show that if $K_\mathfrak{p} = 0$ then $K_a = 0$ for some $a \in A \setminus \mathfrak{p}$.

Hence if X is quasicompact, then to check that a closed embedding π is a regular embedding it suffices to check at closed points of X.

Exercise 12.1.F(b) will show that not all closed embeddings are regular embeddings.

8.4.H. EXERCISE. Show that a closed embedding $X \hookrightarrow Y$ of locally Noetherian schemes is a regular embedding of codimension 1 if and only if X is an effective Cartier divisor on Y. Unimportant remark: the Noetherian hypotheses can be replaced by requiring $\mathscr{O}_Y$ to be coherent, and essentially the same argument shows it. It is interesting to note that "effective Cartier divisor" implies "regular embedding of codimension 1" always, but that the converse argument requires Noetherian(-like) assumptions. (See [**MO129242**] for a counterexample to the converse.)

8.4.8. *Definition.* A **codimension** r **complete intersection** in a scheme Y is a closed subscheme X that can be written as the scheme-theoretic intersection of r effective Cartier divisors $D_1, \dots, D_r$, such that at every point $p \in X$, the equations corresponding to $D_1, \dots, D_r$ form a regular sequence. The phrase **complete intersection** means "codimension r complete intersection for some r".

CHAPTER 9

Fibered products of schemes, and base change

9.1 They exist

Before we get to products, we note that coproducts exist in the category of schemes: just as with the category of sets (Exercise 1.3.T), coproduct is disjoint union. The next exercise makes this precise (and directly extends to coproducts of an infinite number of schemes).

9.1.A. EASY EXERCISE. Suppose X and Y are schemes. Let $X \coprod Y$ be the scheme whose underlying topological space is the disjoint union of the topological spaces of X and Y, and with structure sheaf on (the part corresponding to) X given by $\mathscr{O}_X$, and similarly for Y. Show that $X \coprod Y$ is the coproduct of X and Y (justifying the use of the coproduct symbol $\coprod$).

We will now construct the fibered product in the category of schemes.

9.1.1. Theorem: Fibered products exist. — *Suppose* $\alpha : X \to Z$ *and* $\beta : Y \to Z$ *are morphisms of schemes. Then the fibered product*

$$\begin{tikzcd} X \times_Z Y \arrow[r, "\alpha'"] \arrow[d, "\beta'"'] & Y \arrow[d, "\beta"] \\ X \arrow[r, "\alpha"] & Z \end{tikzcd}$$

exists in the category of schemes.

Note: if A is a ring, people often sloppily write $\times_A$ for $\times_{\operatorname{Spec} A}$. If B is an A-algebra, and X is an A-scheme, people often write X_B or $X \times_A B$ for $X \times_{\operatorname{Spec} A} \operatorname{Spec} B$.

9.1.2. *Warning: products of schemes aren't products of sets.* Before showing existence, here is a warning: the product of schemes isn't a product of sets (and more generally for fibered products). We have made a big deal about schemes being *sets*, endowed with a *topology*, upon which we have a *structure sheaf*. So you might think that we will construct the product in this order. But we won't, because products behave oddly on the level of sets. You may have checked (Exercise 6.6.E(a)) that the product of two affine lines over your favorite algebraically closed field $\overline{k}$ is the affine plane: $\mathbb{A}^1_{\overline{k}} \times_{\overline{k}} \mathbb{A}^1_{\overline{k}} \cong \mathbb{A}^2_{\overline{k}}$. But the underlying set of the latter is *not* the underlying set of the former —- we get additional points, corresponding to curves in $\mathbb{A}^2$ that are not lines parallel to the axes!

9.1.3. On the other hand, W-valued points (where W is a scheme, Definition 6.3.7) *do* behave well under (fibered) products (as mentioned in §6.3.8). This is just the universal property *definition* of fibered product: an W-valued point of a scheme X is defined as an element of $\operatorname{Hom}(W, X)$, and the fibered product is defined by

$$\operatorname{Hom}(W, X \times_Z Y) = \operatorname{Hom}(W, X) \times_{\operatorname{Hom}(W,Z)} \operatorname{Hom}(W, Y). \tag{9.1.3.1}$$

This is one justification for making the definition of scheme-valued point. For this reason, those classical people preferring to think only about varieties over an algebraically closed field $\overline{k}$ (or more generally, finite type schemes over $\overline{k}$), and preferring to understand them through their closed points — or equivalently, the $\overline{k}$-valued points, by the Nullstellensatz (Exercise 5.3.E) — needn't worry: the closed points of the product of two finite type $\overline{k}$-schemes over $\overline{k}$ are (naturally identified with) the product of the closed points of the factors. This will follow from the fact that the product is also finite type over $\overline{k}$, which we verify in Exercise 9.2.D. This is one of the reasons that varieties over algebraically closed fields can be easier to work with. But over a nonalgebraically closed field, things become even more interesting; Example 9.2.2 is a first glimpse.

(Fancy remark: You may feel that (i) "products of topological spaces are products on the underlying sets" is natural, while (ii) "products of schemes are not necessarily products on the underlying sets" is weird. But really (i) is the lucky consequence of the fact that the underlying set of a topological space can be interpreted as set of p-valued points, where p is a point, so it is best seen as a consequence of paragraph 9.1.3, which is the "more correct" — i.e., more general — fact.)

9.1.4. *Warning on Noetherianness.* The fibered product of Noetherian schemes need not be Noetherian. You will later be able to verify that Exercise 9.2.E gives an example, i.e., that $A := \overline{\mathbb{Q}} \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}$ is not Noetherian, as follows. By Exercise 11.1.G(a), $\dim A = 0$. A Noetherian dimension 0 scheme has a finite number of points (Exercise 11.1.C). But by Exercise 9.2.E, $\operatorname{Spec} A$ has an infinite number of points.

On the other hand, the fibered product of finite type k-schemes over finite type k-schemes is a finite type k-scheme (Exercise 9.2.D), so this pathology does not arise for varieties.

9.1.5. Philosophy behind the proof of Theorem 9.1.1. The proof of Theorem 9.1.1 can be confusing. The following comments may help a little.

We already basically know existence of fibered products in two cases: the case where X, Y, and Z are affine (stated explicitly below), and the case where $\beta : Y \to Z$ is an open embedding (Exercise 7.1.B).

9.1.B. EXERCISE (PROMISED IN REMARK 6.3.6). Use Exercise 6.3.F ($\operatorname{Hom}_{Sch}(W, \operatorname{Spec} A) = \operatorname{Hom}_{Rings}(A, \Gamma(W, \mathscr{O}_W))$) to show that given ring maps $C \to A$ and $C \to B$,

$$\operatorname{Spec}(A \otimes_C B) \cong \operatorname{Spec} A \times_{\operatorname{Spec} C} \operatorname{Spec} B.$$

(Interpret tensor product as the "cofibered product" in the category of rings.) Hence the fibered product of affine schemes exists (in the category of schemes). (This generalizes the fact that the product of affine lines exist, Exercise 6.6.E(a).)

The main theme of the proof of Theorem 9.1.1 is that because schemes are built by gluing affine schemes along open subsets, these two special cases will be all that we need. The argument will repeatedly use the same ideas — roughly, that schemes glue (Exercise 4.4.A), and that morphisms of schemes glue (Exercise 6.3.A). This is a sign that something more structural is going on; §9.1.6 describes this for experts.

Proof of Theorem 9.1.1. The key idea is this: we cut everything up into affine open sets, do fibered products there, and show that everything glues nicely. The conceptually difficult part of the proof comes from the gluing, and the realization that we have to check almost nothing. We divide the proof up into a number of bite-sized pieces.

Step 1: fibered products of affine with almost-affine over affine. We begin by combining the affine case with the open embedding case as follows. Suppose X and Z are affine, and $\beta : Y \to Z$ factors as $Y \overset{\iota}{\hookrightarrow} Y' \longrightarrow Z$ where ι is an open embedding and Y' is affine. Then $X \times_Z Y$ exists. This is because if the two small squares of

$$\begin{tikzcd} W \arrow[r] \arrow[d, hook] & Y \arrow[d, hook] \\ W' \arrow[r] \arrow[d] & Y' \arrow[d] \\ X \arrow[r] & Z \end{tikzcd}$$

are fibered diagrams, then the "outside rectangle" is also a fibered diagram. (This was Exercise 1.3.Q, although you should be able to see this on the spot.) It will be important to remember (from Important Exercise 7.1.B) that "open embeddings" are "preserved by fibered product": the fact that $Y \to Y'$ is an open embedding implies that $W \to W'$ is an open embedding.

Key Step 2: fibered product of affine with arbitrary over affine exists. We now come to the key part of the argument: if X and Z are affine, and Y is arbitrary. This is confusing when you first see it, so we first deal with a special case, when Y is the union of two affine open sets $Y_1 \cup Y_2$. Let $Y_{12} = Y_1 \cap Y_2$.

Now for $i = 1$ and 2, $X \times_Z Y_i$ exists by the affine case, Exercise 9.1.B. Call this W_i. Also, $X \times_Z Y_{12}$ exists by Step 1 (call it W_{12}), and comes with *canonical* open embeddings into W_1 and W_2 (by construction of fibered products with open embeddings, see the last sentence of Step 1). Thus we can glue W_1 to W_2 along W_{12}; call this resulting scheme W.

We check that the result is the fibered product by verifying that it satisfies the universal property. Suppose we have maps $\alpha'' : V \to X$, $\beta'' : V \to Y$ that compose (with α and β respectively) to the same map $V \to Z$. We need to construct a unique map $\gamma : V \to W$, so that $\alpha' \circ \gamma = \beta''$ and $\beta' \circ \gamma = \alpha''$.

(9.1.5.1)
$$\begin{tikzcd} V \arrow[dr, "\exists!\gamma?"] \arrow[drr, bend left, "\beta''"] \arrow[ddr, bend right, "\alpha''"'] & & \\ & W \arrow[r, "\alpha'"'] \arrow[d, "\beta'"] & Y \arrow[d, "g"] \\ & X \arrow[r, "\alpha"'] & Z \end{tikzcd}$$

For $i = 1, 2$, define $V_i := (\beta'')^{-1}(Y_i)$. Define $V_{12} := (\beta'')^{-1}(Y_{12}) = V_1 \cap V_2$. Then there is a unique map $V_i \to W_i$ such that the composed maps $V_i \to X$ and $V_i \to Y_i$ are as desired (by the universal product of the fibered product $X \times_Z Y_i = W_i$), hence a unique map $\gamma_i : V_i \to W$. Similarly, there is a unique map $\gamma_{12} : V_{12} \to W$ such that the composed maps $V_{12} \to X$ and $V_{12} \to Y$ are as desired. But the restriction of γ_i to V_{12} is one such map, so it must be γ_{12}. Thus the maps γ_1 and γ_2 agree on V_{12}, and glue together to a unique map $\gamma : V \to W$. We have shown existence and uniqueness of the desired γ.

We have thus shown that if Y is the union of two affine open sets, and X and Z are affine, then $X \times_Z Y$ exists.

We now tackle the general case. (You may prefer to first think through the case where "two" is replaced by "three".) We now cover Y with open sets Y_i, as i runs over some index set (not necessarily finite!). As before, we define W_i and W_{ij}. We can glue these together to produce a scheme W along with open sets we identify with W_i (Exercise 4.4.A — you should check the triple intersection "cocycle" condition).

As in the two-affine case, we show that W is the fibered product by showing that it satisfies the universal property. Suppose we have maps $\alpha'' : V \to X$, $\beta'' : V \to Y$ that compose to the same map $V \to Z$. We construct a unique map $\gamma : V \to W$, so that $\alpha' \circ \gamma = \beta''$ and $\beta' \circ \gamma = \alpha''$. Define $V_i = (\beta'')^{-1}(Y_i)$ and $V_{ij} := (\beta'')^{-1}(Y_{ij}) = V_i \cap V_j$. Then there is a unique map $V_i \to W_i$ such that the composed maps $V_i \to X$ and $V_i \to Y_i$ are as desired, hence a unique map $\gamma_i : V_i \to W$. Similarly, there is a unique map $\gamma_{ij} : V_{ij} \to W$ such that the composed maps $V_{ij} \to X$ and $V_{ij} \to Y$ are as desired. But the restriction of γ_i to V_{ij} is one such map, so it must be γ_{ij}. Thus the maps γ_i and γ_j agree on V_{ij}. Thus the γ_i glue together to a unique map $\gamma : V \to W$. We have shown existence and uniqueness of the desired γ, completing this step.

Step 3: Z affine, X and Y arbitrary. We next show that if Z is affine, and X and Y are arbitrary schemes, then $X \times_Z Y$ exists. We just follow Step 2, with the roles of X and Y reversed, using the fact that by the previous step, we can assume that the fibered product of an affine scheme with an arbitrary scheme over an affine scheme exists.

Step 4: Z admits an open embedding into an affine scheme Z', X and Y arbitrary. This is akin to Step 1: $X \times_Z Y$ satisfies the universal property of $X \times_{Z'} Y$.

Step 5: the general case. We employ the same trick yet again. Suppose $\alpha : X \to Z$, $\beta : Y \to Z$ are two morphisms of schemes. Cover Z with affine open subschemes Z_i, and let $X_i = \alpha^{-1}(Z_i)$ and $Y_i = \beta^{-1}(Z_i)$. Define $Z_{ij} := Z_i \cap Z_j$, $X_{ij} := \alpha^{-1}(Z_{ij})$, and $Y_{ij} := \beta^{-1}(Z_{ij})$. Then $W_i := X_i \times_{Z_i} Y_i$ exists for all i (Step 3), and $W_{ij} := X_{ij} \times_{Z_{ij}} Y_{ij}$ exists for all i, j (Step 4), and for each i and j, W_{ij} comes with a canonically open immersion into both W_i and W_j (see the last sentence in Step 1). As W_i satisfies the universal property of $X \times_Z Y_i$ (do you see why?), we may canonically identify W_i (which we know to exist by Step 3) with with $X \times_Z Y_i$. Similarly, we identify W_{ij} with $X \times_Z Y_{ij}$.

We then proceed exactly as in Step 2: the W_i's can be glued together along the W_{ij} (the cocycle condition can be readily checked to be satisfied), and W can be checked to satisfy the universal property of $X \times_Z Y$ (again, exactly as in Step 2). $\square$

9.1.6. $\star\star$ **Describing the existence of fibered products using the high-falutin' language of representable functors.** The proof above can be described more cleanly in the language of representable functors (§6.6). This will be enlightening only after you have absorbed the above argument and meditated on it for a long time. It may be most useful to shed light on representable functors, rather than on the existence of the fibered product.

Until the end of §9.1 only, by functor*, we mean contravariant functor from the category Sch of schemes to the category of Sets.* For each scheme X, we have a functor h_X, taking a scheme Y to the set $\mathrm{Mor}(Y, X)$ (§1.2.20). Recall (§1.3.10, §6.6) that a functor is *representable* if it is naturally isomorphic to some h_X. If a functor is representable, then the representing scheme is unique up to unique isomorphism (Exercise 6.6.C). This can be usefully extended as follows:

9.1.C. EXERCISE (YONEDA'S LEMMA). If X and Y are schemes, describe a bijection between morphisms of schemes $X \to Y$ and natural transformations of functors $h_X \to h_Y$. Hence show that the category of schemes is a fully faithful subcategory (§1.2.15) of the "functor category" of all functors (contravariant, *Sch* $\to$ *Sets*). Hint: this has nothing to do with schemes; your argument will work in any category. This is the contravariant version of Exercise 1.3.Z(c).

One of Grothendieck's insights is that we should try to treat such functors as "geometric spaces", without worrying about representability. Many notions carry over to this more general setting without change, and some notions are easier. For example, fibered products of functors always exist: $h \times_{h''} h'$ may be defined by

$$(h \times_{h''} h')(W) = h(W) \times_{h''(W)} h'(W),$$

where the fibered product on the right is a fibered product of sets, which always exists. (This isn't quite enough to define a functor; we have only described where objects go. You should work out where morphisms go too.) We didn't use anything about schemes; this works with *Sch* replaced by any category.

Then "$X \times_Z Y$ exists" translates to "$h_X \times_{h_Z} h_Y$ is representable".

9.1.7. *Representable functors are Zariski sheaves.* Because "morphisms to schemes glue" (Exercise 6.3.A), we have a necessary condition for a functor to be representable. We know that if $\{U_i\}$ is an open cover of Y, a morphism $Y \to X$ is determined by its restrictions $U_i \to X$, and given morphisms $U_i \to X$ that agree on the overlap $U_i \cap U_j \to X$, we can glue them together to get a morphism $Y \to X$. In the language of equalizer exact sequences (§2.2.7),

$$\cdot \longrightarrow h_X(Y) \longrightarrow \prod h_X(U_i) \rightrightarrows \prod h_X(U_i \cap U_j)$$

is exact. Thus morphisms to X (i.e., the functor h_X) form a sheaf on every scheme Y. If this holds, we say that the functor is a **Zariski sheaf**. (You can impress your friends by telling them that this is a *sheaf on the big Zariski site*.) We can repeat this discussion with *Sch* replaced by the category Sch_S of schemes over a given base scheme S. We have proved (or observed) that *in order for a functor to be representable, it is necessary for it to be a Zariski sheaf.*

The fiber product passes this test:

9.1.D. EXERCISE. If $X, Y \to Z$ are schemes, show that $h_X \times_{h_Z} h_Y$ is a Zariski sheaf. (Do not use the fact that $X \times_Z Y$ is representable! The point of this section is to recover representability from a more sophisticated perspective.)

We can make some other definitions that extend notions from schemes to functors. We say that a map (i.e., natural transformation) of functors $h' \to h$ expresses h' as an **open subfunctor** of h if for all representable functors h_X and maps $h_X \to h$, the fibered product $h_X \times_h h'$ is representable, by U say, and $h_U \to h_X$ corresponds to an open embedding of schemes $U \to X$. The following fibered square may help.

$$\begin{tikzcd} h_U \arrow[r, "\text{open}"] \arrow[d] & h_X \arrow[d] \\ h' \arrow[r] & h \end{tikzcd}$$

9.1.E. EXERCISE. Show that a map of representable functors $h_W \to h_Z$ is an open subfunctor if and only if $W \to Z$ is an open embedding, so this indeed extends the notion of open embedding to (contravariant) functors (*Sch* $\to$ *Sets*).

9.1.F. EXERCISE (THE GEOMETRIC NATURE OF THE NOTION OF "OPEN SUBFUNCTOR").
(a) Show that an open subfunctor of an open subfunctor is also an open subfunctor.
(b) Suppose $h' \to h$ and $h'' \to h$ are two open subfunctors of h. Define the intersection of these two open subfunctors, which should also be an open subfunctor of h.
(c) Suppose U and V are two open subschemes of a scheme X, so $h_U \to h_X$ and $h_V \to h_X$ are open subfunctors. Show that the intersection of these two open subfunctors is, as you would expect, $h_{U \cap V}$.

9.1.G. EXERCISE. Suppose $\alpha : X \to Z$ and $\beta : Y \to Z$ are morphisms of schemes, and $U \subset X$, $V \subset Y$, $W \subset Z$ are open embeddings, where U and V map to W. Interpret $h_U \times_{h_W} h_V$ as an open subfunctor of $h_X \times_{h_Z} h_Y$. (Hint: given a map $h_T \to h_{X \times_Z Y}$, what open subset of T should correspond to $U \times_W V$?)

A collection h_i of open subfunctors of h is said to **cover** h if for *every* map $h_X \to h$ from a representable subfunctor, the corresponding open subsets $U_i \hookrightarrow X$ cover X.

Given that functors do not have an obvious underlying set (let alone a topology), it is rather amazing that we are talking about when one is an "open subset" of another, or when some functors "cover" another!

9.1.H. EXERCISE. Suppose $\{Z_i\}_i$ is an affine cover of Z, $\{X_{ij}\}_j$ is an affine cover of the preimage of Z_i in X, and $\{Y_{ik}\}_k$ is an affine cover of the preimage of Z_i in Y. Show that $\{h_{X_{ij}} \times_{h_{Z_i}} h_{Y_{ik}}\}_{ijk}$ is an open cover of the functor $h_X \times_{h_Z} h_Y$. (Hint: consider a map $h_T \to h_X \times_{h_Z} h_Y$, and extend your solution to Exercise 9.1.G.)

We now come to a key point: a Zariski sheaf that is "locally representable" must be representable:

9.1.I. KEY EXERCISE. If a functor h is a Zariski sheaf that has an open cover by representable functors ("is covered by schemes"), then h is representable. (Hint: use Exercise 4.4.A to glue together the schemes representing the open subfunctors.)

This immediately leads to the existence of fibered products as follows. Exercise 9.1.D shows that $h_X \times_{h_Z} h_Y$ is a Zariski sheaf. But $h_{X_{ij}} \times_{h_{Z_i}} h_{Y_{ik}}$ is representable for each i, j, k (fibered products of affines over an affine exist, Exercise 9.1.B), and these functors are an open cover of $h_X \times_{h_Z} h_Y$ by Exercise 9.1.H, so by Key Exercise 9.1.I we are done.

9.2 Computing fibered products in practice

Before giving some examples, we first see how to compute fibered products in practice. There are four types of morphisms **(1)–(4)** that it is particularly easy to take fibered products with, and all morphisms can be built from these atomic components. More precisely, **(1)** will imply that we can compute fibered products locally on the source and target. Thus to understand fibered products in general, it suffices to understand them on the level of affine sets, i.e., to be able to compute $A \otimes_B C$ given ring maps $B \to A$ and $B \to C$. Any map $B \to A$ (and similarly $B \to C$) may be expressed as $B \to B[t_1, \dots]/I$, so if we know how to base change by "adding variables" **(2)** and "taking quotients" **(3)**, we can "compute" any fibered product (at least in theory). The fourth type of morphism **(4)**, corresponding to localization, is useful to understand explicitly as well.

(1) Base change by open embeddings.

We have already done this (Exercise 7.1.B), and we used it in the proof that fibered products of schemes exist.

(2) Adding an extra variable.

9.2.A. EASY ALGEBRA EXERCISE. Show that $A \otimes_B B[t] \cong A[t]$, so the following is a fibered diagram. (Your argument might naturally extend to allow the addition of infinitely many variables, but we won't need this generality.) Hint: show that $A[t]$ satisfies an appropriate universal property.

$$\begin{tikzcd}
\operatorname{Spec} A[t] \arrow[r] \arrow[d] & \operatorname{Spec} B[t] \arrow[d] \\
\operatorname{Spec} A \arrow[r] & \operatorname{Spec} B
\end{tikzcd}$$

(3) Base change by closed embeddings

9.2.B. EXERCISE. Suppose $\phi : B \to A$ is a ring morphism, and $I \subset B$ is an ideal. Let $I^e := \langle \phi(i) \rangle_{i \in I} \subset A$ be the **extension of** I **to** A. Describe a natural isomorphism $A/I^e \cong A \otimes_B (B/I)$. (Hint: consider $I \to B \to B/I \to 0$, and use the right-exactness of $\otimes_B A$, Exercise 1.3.H.)

9.2.1. As an immediate consequence: the fibered product with a closed subscheme is a closed subscheme of the fibered product in the obvious way. We say that "closed embeddings are preserved by base change".

9.2.C. EXERCISE.
(a) Interpret the intersection of two closed embeddings into X (cf. Exercise 8.1.J) as their fibered product over X.
(b) Show that "locally closed embeddings" are preserved by base change.
(c) Define the **intersection of n locally closed embeddings** $X_i \hookrightarrow Z$ ($1 \leq i \leq n$) by the fibered product of the X_i over Z (mapping to Z). Show that the intersection of (a finite number of) locally closed embeddings is also a locally closed embedding.

As an application of Exercise 9.2.B, we can compute tensor products of finitely generated k algebras over k. For example, we have

$$k[x_1, x_2]/(x_1^2 - x_2) \otimes_k k[y_1, y_2]/(y_1^3 + y_2^3) \cong k[x_1, x_2, y_1, y_2]/(x_1^2 - x_2, y_1^3 + y_2^3).$$

9.2.D. EXERCISE. Suppose X and Y are locally of finite type A-schemes. Show that $X \times_A Y$ is also locally of finite type over A. Prove the same thing with "locally" removed from both the hypothesis and conclusion.

9.2.2. *Example.* We can these ideas to compute $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}$:

$$\begin{aligned}
\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} &\cong \mathbb{C} \otimes_{\mathbb{R}} (\mathbb{R}[x]/(x^2+1)) \\
&\cong (\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}[x])/(x^2+1) \qquad \text{by 9.2\textbf{(3)}} \\
&\cong \mathbb{C}[x]/(x^2+1) \qquad \text{by 9.2\textbf{(2)}} \\
&\cong \mathbb{C}[x]/((x-i)(x+i)) \\
&\cong \mathbb{C}[x]/(x-i) \times \mathbb{C}[x]/(x+i) \qquad \text{by the Chinese Remainder Theorem} \\
&\cong \mathbb{C} \times \mathbb{C}
\end{aligned}$$

Thus $\operatorname{Spec} \mathbb{C} \times_{\mathbb{R}} \operatorname{Spec} \mathbb{C} \cong \operatorname{Spec} \mathbb{C} \coprod \operatorname{Spec} \mathbb{C}$. This example is the first example of many different behaviors. Notice for example that two points somehow correspond to the Galois group of $\mathbb{C}$ over $\mathbb{R}$; for one of them, x (the "i" in one of the copies of $\mathbb{C}$) equals i (the "i" in the other copy of $\mathbb{C}$), and in the other, $x = -i$.

9.2.3. ⋆ *Remark.* Here is a clue that there is something deep going on behind Example 9.2.2. If L/K is a (finite) Galois extension with Galois group G, then $L \otimes_K L$ is isomorphic to L^G (the product of $|G|$ copies of L). This turns out to be a restatement of the classical form of linear independence of characters! In the language of schemes, $\operatorname{Spec} L \times_K \operatorname{Spec} L$ is a union of a number of copies of $\operatorname{Spec} L$ that naturally form a torsor over the Galois group G; but we will not define torsor here.

9.2.E. ⋆ HARD BUT FASCINATING EXERCISE FOR THOSE FAMILIAR WITH $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Show that the points of $\operatorname{Spec} \overline{\mathbb{Q}} \otimes_{\mathbb{Q}} \overline{\mathbb{Q}}$ are in natural bijection with $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and the Zariski topology on the former agrees with the profinite topology on the latter. (Some hints: first do the case of finite Galois extensions. Relate the topology on Spec of a direct limit of rings to the inverse limit of Specs. Can you see which point corresponds to the identity of the Galois group?)

At this point, we can compute any $A \otimes_B C$ (where A and C are B-algebras): any map of rings $\phi : B \to A$ can be interpreted by adding variables (perhaps infinitely many) to B, and then imposing relations. But in practice **(4)** is useful, as we will see in examples.

(4) Base change of affine schemes by localization.

9.2.F. EXERCISE. Suppose $\phi : B \to A$ is a ring morphism, and $S \subset B$ is a multiplicative subset of B, which implies that $\phi(S)$ is a multiplicative subset of A. Describe a natural isomorphism $\phi(S)^{-1}A \cong A \otimes_B (S^{-1}B)$.

Translation: the fibered product with a localization is the localization of the fibered product in the obvious way. We say that "localizations are preserved by base change". This is handy if the localization is of the form $B \hookrightarrow B_f$ (corresponding to taking distinguished open sets) or $B \hookrightarrow K(B)$ (from B to the fraction field of B, corresponding to taking generic points), and various things in between.

9.2.4. Examples. These four facts let you calculate lots of things in practice, and we will use them freely.

9.2.G. EXERCISE: THE THREE IMPORTANT TYPES OF MONOMORPHISMS OF SCHEMES. Show that the following are monomorphisms (Definition 1.3.9): open embeddings, closed embeddings, and localization of affine schemes. As monomorphisms are closed under composition, Exercise 1.3.V, compositions of the above are also monomorphisms — for example, locally closed embeddings, or maps from "Spec of stalks at points of X" to X. (Caution: if p is a point of a scheme X, the natural morphism $\operatorname{Spec} \mathscr{O}_{X,p} \to X$, cf. Exercise 6.3.M, is a monomorphism but is not in general an open embedding.)

9.2.H. EXERCISE. Prove that $\mathbb{A}^n_A \cong \mathbb{A}^n_{\mathbb{Z}} \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} A$. Prove that $\mathbb{P}^n_A \cong \mathbb{P}^n_{\mathbb{Z}} \times_{\operatorname{Spec} \mathbb{Z}} \operatorname{Spec} A$. Thus affine space and projective space are pulled back from their "universal manifestation" over the final object $\operatorname{Spec} \mathbb{Z}$.

9.2.5. *Extending the base field.* One special case of base change is called **extending the base field**: if X is a k-scheme, and ℓ is a field extension (often ℓ is the algebraic closure of k), then $X \times_{\operatorname{Spec} k} \operatorname{Spec} \ell$ (sometimes informally written $X \times_k \ell$ or X_ℓ) is an ℓ-scheme. Often properties of X can be checked by verifying them instead on X_ℓ. This is the subject of *descent* — certain properties "descend" from X_ℓ to X. We have already seen that the property of being the Spec of a normal integral domain descends in this way (Exercise 5.4.M). Exercises 9.2.I and 9.2.J give other examples of properties which descend: the property of two morphisms being equal, and the property of a(n affine) morphism being a closed embedding, both descend in this way. Those interested in schemes over non-algebraically closed fields will use this repeatedly, to reduce results to the algebraically closed case.

9.2.I. EXERCISE. Suppose $\pi : X \to Y$ and $\rho : X \to Y$ are morphisms of k-schemes, ℓ/k is a field extension, and $\pi_\ell : X \times_{\operatorname{Spec} k} \operatorname{Spec} \ell \to Y \times_{\operatorname{Spec} k} \operatorname{Spec} \ell$ and $\rho_\ell : X \times_{\operatorname{Spec} k} \operatorname{Spec} \ell \to Y \times_{\operatorname{Spec} k} \operatorname{Spec} \ell$ are the induced maps of ℓ-schemes. (Be sure you understand what this means!) Show that if $\pi_\ell = \rho_\ell$ then $\pi = \rho$. (Hint: show that π and ρ are the same on the level of sets. To do this, you may use that

$X \times_{\operatorname{Spec} k} \operatorname{Spec} \ell \to X$ is surjective, which we will soon prove in Exercise 9.4.D. Then reduce to the case where X and Y are affine.)

9.2.J. EASY EXERCISE. Suppose $\pi : X \to Y$ is an affine morphism over k. Show that π is a closed embedding if and only if $\pi \times_k \overline{k} : X \times_k \overline{k} \to Y \times_k \overline{k}$ is. (The affine hypothesis is not necessary for this result, but it makes the proof easier, and this is the situation in which we will most need it.)

9.2.K. UNIMPORTANT BUT FUN EXERCISE. Show that $\operatorname{Spec} \mathbb{Q}(t) \otimes_{\mathbb{Q}} \mathbb{C}$ has closed points in natural correspondence with the transcendental complex numbers. (If the description $\operatorname{Spec} \mathbb{Q}(t) \otimes_{\mathbb{Q}[t]} \mathbb{C}[t]$ is more striking, you can use that instead.) This scheme doesn't come up in nature, but it is certainly neat! A related idea comes up in Remark 11.2.16.

9.3 Interpretations: Pulling back families, and fibers of morphisms

9.3.1. Pulling back families.
We can informally interpret fibered product in the following geometric way. Suppose $Y \to Z$ is a morphism. We interpret this as a "family of schemes parametrized by a **base scheme** (or just plain **base**) Z." Then if we have another morphism $\psi : X \to Z$, we interpret the induced map $X \times_Z Y \to X$ as the "pulled back family" (see Figure 9.1).

$$\begin{tikzcd} X \times_Z Y \arrow[r] \arrow[d, "\text{pulled back family}"'] & Y \arrow[d, "\text{family}"] \\ X \arrow[r, "\psi"] & Z \end{tikzcd}$$

We sometimes say that $X \times_Z Y$ is the **scheme-theoretic pullback of** Y, **scheme-theoretic inverse image**, or **inverse image scheme of** Y. (Our forthcoming discussion of fibers may give some motivation for this.) For this reason, fibered product is often called **base change** or **change of base** or **pullback**. In addition to the various names for a Cartesian diagram given in §1.3.6, in algebraic geometry it is often called a **base change diagram** or a **pullback diagram**, and $X \times_Z Y \to X$ is called the **pullback** of $Y \to Z$ by ψ, and $X \times_Z Y$ is called the **pullback** of Y by ψ. One often uses the phrase "over X" or "above X" to when discussing $X \times_Z Y$, especially if X is a locally closed subscheme of Z. (Random side remark: scheme-theoretic pullback always makes sense, while the notion of scheme-theoretic image is somehow problematic, as discussed in §8.3.1.)

Before making any definitions, we give a motivating informal example. Consider the "family of curves" $y^2 = x^3 + tx$ in the xy-plane parametrized by t. Translation: consider $\operatorname{Spec} k[x,y,t]/(y^2 - x^3 - tx) \to \operatorname{Spec} k[t]$. If we pull back to a family parametrized by the uv-plane via $uv = t$ (i.e., $\operatorname{Spec} k[u,v] \to \operatorname{Spec} k[t]$ given by $t \mapsto uv$), we get $y^2 = x^3 + uvx$, i.e., $\operatorname{Spec} k[x,y,u,v]/(y^2 - x^3 - uvx) \to \operatorname{Spec} k[u,v]$. If instead we set t to 3 (i.e., pull back by $\operatorname{Spec} k[t]/(t-3) \to \operatorname{Spec} k[t]$, we get the curve $y^2 = x^3 + 3x$ (i.e., $\operatorname{Spec} k[x,y]/(y^2 - x^3 - 3x) \to \operatorname{Spec} k$), which we interpret

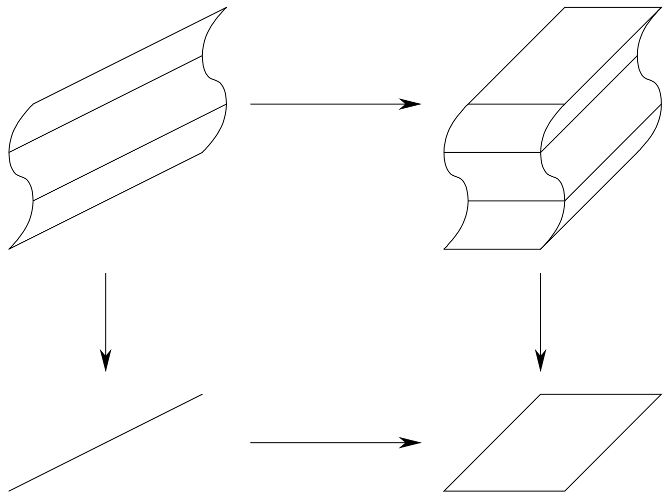


FIGURE 9.1. A picture of a pulled back family

as the fiber of the original family above $t = 3$. We will soon be able to interpret these constructions in terms of fiber products.

9.3.2. Fibers of morphisms.

(If you did Exercise 7.3.K, that finite morphisms have finite fibers, you will not find this discussion surprising.) A special case of pullback is the notion of a fiber of a morphism. We motivate this with the notion of fiber in the category of topological spaces.

9.3.A. EXERCISE. Show that if $Y \to Z$ is a continuous map of topological spaces, and X is a point p of Z, then the fiber of Y over p (the set-theoretic fiber, with the induced topology) is naturally identified with $X \times_Z Y$.

More generally, for any $\pi : X \to Z$, the fiber of $X \times_Z Y \to X$ over a point p of X is naturally identified with the fiber of $Y \to Z$ over $\pi(p)$.

Motivated by topology, we return to the category of schemes. Suppose $p \to Z$ is the inclusion of a point (not necessarily closed). More precisely, if p is a point with residue field K, consider the map $\operatorname{Spec} K \to Z$ sending $\operatorname{Spec} K$ to p, with the natural isomorphism of residue fields. Then if $g : Y \to Z$ is any morphism, the base change with $p \to Z$ is called the (scheme-theoretic) **fiber** (or *fibre*) **of** g **above** p or the (scheme-theoretic) **preimage of** p, and is denoted $g^{-1}(p)$. If Z is irreducible, the fiber above the generic point of Z is called the **generic fiber** (of g). In an affine open subscheme $\operatorname{Spec} A$ containing p, p corresponds to some prime ideal $\mathfrak{p}$, and the morphism $\operatorname{Spec} K \to Z$ corresponds to the ring map $A \to A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}}$. This is the composition of localization and closed embedding, and thus can be computed by the tricks above. (Note that $p \to Z$ is a monomorphism, by Exercise 9.2.G.)

9.3.B. EXERCISE. Show that the underlying topological space of the (scheme-theoretic) fiber of $X \to Y$ above a point p is naturally identified with the topological fiber of $X \to Y$ above p.

9.3.C. EXERCISE (ANALOG OF EXERCISE 9.3.A). Suppose that $\pi : Y \to Z$ and $\tau : X \to Z$ are morphisms, and $p \in X$ is a point. Show that the fiber of $X \times_Z Y \to X$ over p is (isomorphic to) the base change to p of the fiber of $\pi : Y \to Z$ over $\tau(p)$.

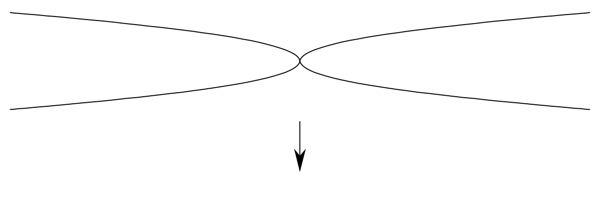


FIGURE 9.2. The map $\mathbb{C} \to \mathbb{C}$ given by $y \mapsto y^2$

9.3.3. *Example (enlightening in several ways).* Consider the projection of the parabola $y^2 = x$ to the x-axis over $\mathbb{Q}$, corresponding to the map of rings $\mathbb{Q}[x] \to \mathbb{Q}[y]$, with $x \mapsto y^2$. If $\mathbb{Q}$ alarms you, replace it with your favorite field and see what happens. (You should look at Figure 9.2, which is a flipped version of the parabola of Figure 3.6, and figure out how to edit it to reflect what we glean here.) Writing $\mathbb{Q}[y]$ as $\mathbb{Q}[x,y]/(y^2 - x)$ helps us interpret the morphism conveniently.

(i) Then the preimage of 1 is two points:

$$\begin{aligned} \operatorname{Spec} \mathbb{Q}[x,y]/(y^2 - x) \otimes_{\mathbb{Q}[x]} \mathbb{Q}[x]/(x-1) &\cong \operatorname{Spec} \mathbb{Q}[x,y]/(y^2 - x, x-1) \\ &\cong \operatorname{Spec} \mathbb{Q}[y]/(y^2 - 1) \\ &\cong \operatorname{Spec} \mathbb{Q}[y]/(y-1) \coprod \operatorname{Spec} \mathbb{Q}[y]/(y+1). \end{aligned}$$

(ii) The preimage of 0 is one nonreduced point:

$$\operatorname{Spec} \mathbb{Q}[x,y]/(y^2 - x, x) \cong \operatorname{Spec} \mathbb{Q}[y]/(y^2).$$

(iii) The preimage of -1 is one reduced point, but of "size 2 over the base field".

$$\operatorname{Spec} \mathbb{Q}[x,y]/(y^2 - x, x+1) \cong \operatorname{Spec} \mathbb{Q}[y]/(y^2+1) \cong \operatorname{Spec} \mathbb{Q}[i] = \operatorname{Spec} \mathbb{Q}(i).$$

(iv) The preimage of the generic point is again one reduced point, but of "size 2 over the residue field", as we verify now.

$$\operatorname{Spec} \mathbb{Q}[x,y]/(y^2 - x) \otimes_{\mathbb{Q}[x]} \mathbb{Q}(x) \cong \operatorname{Spec} \mathbb{Q}[y] \otimes_{\mathbb{Q}[y^2]} \mathbb{Q}(y^2)$$

i.e., (informally) the Spec of the ring of polynomials in y divided by polynomials in y^2. A little thought shows you that in this ring you may invert *any* polynomial in y, as if $f(y)$ is any polynomial in y, then

$$\frac{1}{f(y)} = \frac{f(-y)}{f(y)f(-y)},$$

and the latter denominator is a polynomial in y^2. Thus

$$\mathbb{Q}[x,y]/(y^2-x)\otimes_{\mathbb{Q}[x]}\mathbb{Q}(x)\cong\mathbb{Q}(y)$$

which is a degree 2 field extension of $\mathbb{Q}(x)$ (note that $\mathbb{Q}(x)=\mathbb{Q}(y^2)$).

Notice the following interesting fact: in each of the four cases, the number of preimages can be interpreted as 2, where you count to two in several ways: you can count points (as in the case of the preimage of 1); you can get nonreduced behavior (as in the case of the preimage of 0); or you can have a field extension of degree 2 (as in the case of the preimage of -1 or the generic point). In each case, the fiber is an affine scheme whose dimension as a vector space over the residue field of the point is 2. Number theoretic readers may have seen this behavior before. We will discuss this example again in §17.4.8. This is going to be symptomatic of a very important kind of morphism (a finite flat morphism, see Exercise 24.4.G and §24.4.11).

Try to draw a picture of this morphism if you can, so you can develop a pictorial shorthand for what is going on. A good first approximation is the parabola of Figure 9.2, but you will want to somehow depict the peculiarities of (iii) and (iv).

9.3.4. *Remark: Finite morphisms have finite fibers.* If you haven't done Exercise 7.3.K, that finite morphisms have finite fibers, now would be a good time to do it, as you will find it more straightforward given what you know now.

9.3.D. EXERCISE (IMPORTANT FOR THOSE WITH MORE ARITHMETIC BACKGROUND). What is the scheme-theoretic fiber of $\operatorname{Spec}\mathbb{Z}[i]\to\operatorname{Spec}\mathbb{Z}$ over the prime (p)? Your answer will depend on p, and there are four cases, corresponding to the four cases of Example 9.3.3. (Can you draw a picture?)

9.3.E. EXERCISE. (This exercise will give you practice in computing a fibered product over something that is not a field.) Consider the morphism of schemes $X=\operatorname{Spec}k[t]\to Y=\operatorname{Spec}k[u]$ corresponding to $k[u]\to k[t]$, $u\mapsto t^2$, where $\operatorname{char}k\neq 2$. Show that $X\times_Y X$ has two irreducible components. (What happens if $\operatorname{char}k=2$? See Exercise 9.5.A for a clue.)

9.3.5. *A first view of a blow-up.*

9.3.F. IMPORTANT CONCRETE EXERCISE. (The discussion here immediately generalizes to $\mathbb{A}^n_A$.) Define a closed subscheme $\operatorname{Bl}_{(0,0)}\mathbb{A}^2_k$ of $\mathbb{A}^2_k\times_k\mathbb{P}^1_k$ as follows (see Figure 9.3). If the coordinates on $\mathbb{A}^2_k$ are x,y, and the projective coordinates on $\mathbb{P}^1_k$ are u,v, this subscheme is cut out in $\mathbb{A}^2_k\times_k\mathbb{P}^1_k$ by the single equation $xv=yu$. (You may wish to interpret $\operatorname{Bl}_{(0,0)}\mathbb{A}^2_k$ as follows. The $\mathbb{P}^1_k$ parametrizes lines through the origin. The blow-up corresponds to ordered pairs of (point p, line ℓ) such that $(0,0),p\in\ell$.) Describe the fiber of the morphism $\operatorname{Bl}_{(0,0)}\mathbb{A}^2_k\to\mathbb{P}^1_k$ over each closed point of $\mathbb{P}^1_k$. Show that the morphism $\operatorname{Bl}_{(0,0)}\mathbb{A}^2_k\to\mathbb{A}^2_k$ is an isomorphism away from $(0,0)\in\mathbb{A}^2_k$. Show that the fiber over $(0,0)$ is an effective Cartier divisor (§8.4.1, a closed subscheme that is locally cut out by a single equation, which is not a zerodivisor). It is called the **exceptional divisor**. We will discuss blow-ups in Chapter 22. This particular example will come up in the motivating example of §22.1, and in Exercise 20.2.D.

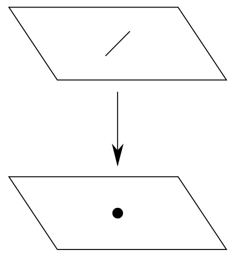


FIGURE 9.3. A first example of a blow-up

We haven't yet discussed regularity, but here is a hand-waving argument suggesting that the $\mathrm{Bl}_{(0,0)} \mathbb{A}^2_k$ is "smooth": the preimage above either standard open set $U_i \subset \mathbb{P}^1$ is isomorphic to $\mathbb{A}^2$. Thus "the blow-up is a surgery that takes the smooth surface $\mathbb{A}^2_k$, cuts out a point, and glues back in a $\mathbb{P}^1$, in such a way that the outcome is another smooth surface."

9.3.6. General fibers, generic fibers, generically finite morphisms.
The phrases "generic fiber" and "general fiber" parallel the phrases "generic point" and "general point" (Definition 3.6.11). Suppose $\pi : X \to Y$ is a morphism of schemes. When one says the **general fiber** (or *a* general fiber) of π a has a certain property, this means that there exists a dense open subset $U \subset Y$ such that the fibers above any point in U have that property.

When one says **the generic fiber** of $\pi : X \to Y$, this implicitly means that Y is irreducible, and the phrase refers to the fiber over the generic point. *General fiber* and *generic fiber* are not the same thing! Clearly if something holds for the general fiber, then it holds for the generic fiber, but the converse is not always true. However, in good circumstances, it can be — properties of the generic fiber extend to an honest neighborhood. For example, if Y is irreducible and Noetherian, and π is finite type, then if the generic fiber of π is empty (resp. nonempty), then the general fiber is empty (resp. nonempty), by Chevalley's theorem (or more simply, by Exercise 7.4.L).

If $\pi : X \to Y$ is finite type, we say π is **generically finite** if π is finite after base change to the generic point of each irreducible component (or equivalently, by Exercise 7.4.D, if the preimage of the generic point of each irreducible component of Y is finite). (The notion of generic finiteness can be defined in more general circumstances, see [**Stacks**, tag 073A].)

9.3.G. EXERCISE ("GENERICALLY FINITE" MEANS "GENERALLY FINITE" IN GOOD CIRCUMSTANCES). Suppose $\pi : X \to Y$ is an affine finite type morphism of locally Noetherian schemes, and Y is reduced. Show that there is an open neighborhood of each generic point of Y over which π is actually finite. (The hypotheses can be weakened considerably, see [**Stacks**, tag 02NW].) Hint: reduce to the case where Y is $\operatorname{Spec} B$, where B is an integral domain. Then X is affine, say $X = \operatorname{Spec} A$. Write $A = B[x_1, \dots, x_n]/I$. Now $A \otimes_B K(B)$ is a finite $K(B)$-module (finite-dimensional vector space) by hypothesis, so there are monic polynomials $f_i(t) \in K(B)[t]$ such that $f_i(x_i) = 0$ in $A \otimes_B K(B)$. Let b be the product of the (finite number of) denominators appearing in the coefficients in the $f_i(x)$. By replacing B by B_b, argue that you can assume that $f_i(t) \in B[t]$. Then $f_i(x_i) = 0$ in $A \otimes_B K(B)$, meaning that $f_i(x_i)$ is annihilated by some nonzero element of B. By replacing B by its localization at the product of these n nonzero elements ("shrinking Spec B further"), argue that $f_i(x_i) = 0$ in A. Then conclude.

9.3.7. $\star\star$ **Finitely presented families (morphisms) are locally pullbacks of particularly nice families.** If you are macho and are embarrassed by Noetherian rings, the following exercise can be used to extend results from the Noetherian case to finitely presented situations. Exercise 9.3.I, an extension of Chevalley's Theorem 7.4.2, is a good example.

9.3.H. EXERCISE. Suppose $\pi : X \to \operatorname{Spec} B$ is a finitely presented morphism. Show that there exists a base change diagram of the form

$$\begin{tikzcd} X \arrow[r] \arrow[d, "\pi"'] & X' \arrow[d, "\pi'"] \\ \operatorname{Spec} B \arrow[r, "\rho"] & \operatorname{Spec} \mathbb{Z}[x_1, \dots, x_N] \end{tikzcd}$$

where N is some integer, and π' is finitely presented (= finite type as the target is Noetherian, see §7.3.17). Thus each finitely presented morphism is locally (on the base) a pullback of a finite type morphism to a Noetherian scheme. Hence any result proved for Noetherian schemes and stable under base change is automatically proved for finitely presented morphisms to arbitrary schemes. Hint: think about the case where X is affine first. If $X = \operatorname{Spec} A$, then $A = B[y_1, \dots, y_n]/(f_1, \dots, f_r)$. Choose one variable x_i for each coefficient of $f_i \in B[y_1, \dots, y_n]$. What is X' in this case? Then consider the case where X is the union of two affine open sets, that intersect in an affine open set. Then consider more general cases until you solve the full problem. You will need to use every part of the definition of finite presentation. (Exercise 28.2.L extends this result.)

9.3.I. EXERCISE (CHEVALLEY'S THEOREM FOR LOCALLY FINITELY PRESENTED MORPHISMS).
(a) Suppose that A is a finitely presented B-algebra (B not necessarily Noetherian), so $A = B[x_1, \dots, x_n]/(f_1, \dots, f_r)$. Show that the image of $\operatorname{Spec} A \to \operatorname{Spec} B$ is a finite union of locally closed subsets of $\operatorname{Spec} B$. Hint: Exercise 9.3.H (the simpler affine case).
(b) Show that if $\pi : X \to Y$ is a quasicompact locally finitely presented morphism, and Y is quasicompact, then $\pi(X)$ is a finite union of locally closed subsets. (For hardened experts only: [**Gr-EGA**, 0_{III}.9.1] gives a definition of *local constructibility*, and of constructibility in more generality. The general form of Chevalley's constructibility theorem [**Gr-EGA**, IV_1.1.8.4] is that the image of a locally constructible set, under a finitely presented map, is also locally constructible.)

9.4 Properties preserved by base change

All reasonable properties of morphisms are preserved under base change. (In fact, one might say that a property of morphisms cannot be reasonable if it is not preserved by base change, cf. §7.1.1.) We discuss this, and in §9.5 we will explain how to fix those that don't fit this pattern.

We have already shown that the notion of "open embedding" is preserved by base change (Exercise 7.1.B). We did this by explicitly describing what the fibered product of an open embedding is: if $Y \hookrightarrow Z$ is an open embedding, and $\psi : X \to Z$ is any morphism, then we checked that the open subscheme $\psi^{-1}(Y)$ of X satisfies the universal property of fibered products.

We have also shown that the notion of "closed embedding" is preserved by base change (§9.2 **(3)**). In other words, given a fiber diagram

$$\begin{tikzcd} W \arrow[r] \arrow[d] & Y \arrow[d, hook, "\text{cl. emb.}"] \\ X \arrow[r] & Z \end{tikzcd}$$

where $Y \hookrightarrow Z$ is a closed embedding, $W \to X$ is as well.

9.4.A. EASY EXERCISE. Show that locally principal closed subschemes (Definition 8.4.1) pull back to locally principal closed subschemes.

Similarly, other important properties are preserved by base change.

9.4.B. EXERCISE. Show that the following properties of morphisms are preserved by base change.

(a) quasicompact
(b) quasiseparated
(c) affine morphism
(d) finite
(e) integral
(f) locally of finite type
(g) finite type
⋆⋆ (h) locally of finite presentation
⋆⋆ (i) finite presentation

9.4.C. ⋆ EXERCISE. Show that the notion of "quasifinite morphism" (finite type + finite fibers, Definition 7.3.14) is preserved by base change. (Warning: the notion of "finite fibers" is not preserved by base change. $\operatorname{Spec} \overline{\mathbb{Q}} \to \operatorname{Spec} \mathbb{Q}$ has finite fibers, but $\operatorname{Spec} \overline{\mathbb{Q}} \otimes_{\mathbb{Q}} \overline{\mathbb{Q}} \to \operatorname{Spec} \overline{\mathbb{Q}}$ has one point for each element of $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, see Exercise 9.2.E.) Hint: reduce to the case $\operatorname{Spec} A \to \operatorname{Spec} B$. Reduce to the case $\phi : \operatorname{Spec} A \to \operatorname{Spec} k$. By Exercise 7.4.D, such ϕ are actually finite, and finiteness is preserved by base change.

9.4.D. EXERCISE. Show that surjectivity is preserved by base change. (**Surjectivity** has its usual meaning: surjective as a map of sets.) You may end up showing that for any fields k_1 and k_2 containing k_3, $k_1 \otimes_{k_3} k_2$ is nonzero, and using the axiom of choice to find a maximal ideal in $k_1 \otimes_{k_3} k_2$.

9.4.1. On the other hand, injectivity is not preserved by base change — witness the bijection $\operatorname{Spec}\mathbb{C} \to \operatorname{Spec}\mathbb{R}$, which loses injectivity upon base change by $\operatorname{Spec}\mathbb{C} \to \operatorname{Spec}\mathbb{R}$ (see Example 9.2.2). This can be rectified (see §9.5.24).

9.4.E. EXERCISE (CF. EXERCISE 9.2.D). Suppose X and Y are integral finite type $\overline{k}$-schemes. Show that $X \times_{\overline{k}} Y$ is an integral finite type $\overline{k}$-scheme. (Once we define "variety", this will become the important fact that the product of irreducible varieties over an algebraically closed field is an irreducible variety, Exercise 10.1.E. The fact that the base field $\overline{k}$ is algebraically closed is important, see §9.5. See Exercise 9.5.M for an improvement.) Hint: reduce to the case where X and Y are both affine, say $X = \operatorname{Spec} A$ and $Y = \operatorname{Spec} B$ with A and B integral domains. You might flip ahead to Easy Exercise 9.5.L to see how to do this. Suppose $(\sum a_i \otimes b_i)\left(\sum a'_j \otimes b'_j\right) = 0$ in $A \otimes_{\overline{k}} B$ with $a_i, a'_j \in A$, $b_i, b'_j \in B$, where both $\{b_i\}$ and $\{b'_j\}$ are linearly independent over $\overline{k}$, and a_1 and a'_1 are nonzero. Show that $D(a_1 a'_1) \subset \operatorname{Spec} A$ is nonempty. By the Weak Nullstellensatz 3.2.4, there is a maximal $\mathfrak{m} \subset A$ in $D(a_1 a'_1)$ with $A/\mathfrak{m} = \overline{k}$. By reducing modulo $\mathfrak{m}$, deduce $(\sum \overline{a}_i \otimes b_i)\left(\sum \overline{a'}_j \otimes b'_j\right) = 0$ in B, where the overline indicates residue modulo $\mathfrak{m}$. Show that this contradicts the fact that B is a domain.

9.4.F. EXERCISE. If P is a property of morphisms preserved by base change and composition, and $X \to Y$ and $X' \to Y'$ are two morphisms of S-schemes with property P, show that $X \times_S X' \to Y \times_S Y'$ has property P as well.

9.5 ⋆ Properties not preserved by base change, and how to fix them

There are some notions that you should reasonably expect to be preserved by pullback based on your geometric intuition. Given a family in the topological category, fibers pull back in reasonable ways. So for example, any pullback of a family in which all the fibers are irreducible will also have this property; ditto for connected. Unfortunately, both of these fail in algebraic geometry, as Example 9.2.2 shows:

$$\begin{tikzcd} \operatorname{Spec}\mathbb{C} \coprod \operatorname{Spec}\mathbb{C} \arrow[r] \arrow[d] & \operatorname{Spec}\mathbb{C} \arrow[d] \\ \operatorname{Spec}\mathbb{C} \arrow[r] & \operatorname{Spec}\mathbb{R} \end{tikzcd}$$

The family on the right (the vertical map) has irreducible and connected fibers, and the one on the left doesn't. The same example shows that the notion of "integral fibers" also doesn't behave well under pullback. And we used it in 9.4.1 to show that injectivity isn't preserved by Base Change.

9.5.A. EXERCISE. Suppose k is a field of characteristic p, so $k(u)/k(u^p)$ is an inseparable extension. By considering $k(u) \otimes_{k(u^p)} k(u)$, show that the notion of "reduced fibers" does not necessarily behave well under pullback. (We will soon see that this happens only in characteristic p, in the presence of inseparability.)

We rectify this problem as follows.

9.5.1. A **geometric point** of a scheme X is defined to be a morphism $\operatorname{Spec} k \to X$ where k is an algebraically closed field. Awkwardly, this is now the third kind of "point" of a scheme! There are just plain points, which are elements of the underlying set; there are Z-valued points (Z a scheme), which are maps $Z \to X$, §6.3.7; and there are geometric points. Geometric points are clearly a flavor of a scheme-valued point, but they are also an enriched version of a (plain) point: they are the data of a point with an inclusion of the residue field of the point in an algebraically closed field.

A **geometric fiber** of a morphism $X \to Y$ is defined to be the fiber over a geometric point of Y. A morphism has **connected** (resp. **irreducible**, **integral**, **reduced**) **geometric fibers** if all its geometric fibers are connected (resp. irreducible, integral, reduced). One usually says that the morphism has **geometrically connected** (resp. **geometrically irreducible**, **geometrically integral**, **geometrically reduced**) **fibers**. A k-scheme X is **geometrically connected** (resp. **geometrically irreducible**, **geometrically integral**, **geometrically reduced**) if the structure morphism $X \to \operatorname{Spec} k$ has geometrically connected (resp. irreducible, integral, reduced) fibers. We will soon see that to check any of these conditions, we need only base change to $\overline{k}$.

(Warning: in some sources, in the definition of "geometric point", "algebraically closed" is replaced by "separably closed".)

9.5.B. EXERCISE. Show that the notion of "connected (resp. irreducible, integral, reduced) geometric fibers" behaves well under base change.

9.5.C. EXERCISE FOR THE ARITHMETICALLY-MINDED. Show that for the morphism $\operatorname{Spec} \mathbb{C} \to \operatorname{Spec} \mathbb{R}$, all geometric fibers consist of two reduced points. (Cf. Example 9.2.2.) Thus $\operatorname{Spec} \mathbb{C}$ is a geometrically reduced but not geometrically irreducible $\mathbb{R}$-scheme.

9.5.D. EASY EXERCISE. Give examples of k-schemes that:

(a) are reduced but not geometrically reduced;
(b) are connected but not geometrically connected;
(c) are integral but not geometrically integral.

9.5.E. EXERCISE. Recall Example 9.3.3, the projection of the parabola $y^2 = x$ to the x-axis, corresponding to the map of rings $\mathbb{Q}[x] \to \mathbb{Q}[y]$, with $x \mapsto y^2$. Show that the geometric fibers of this map are always two points, except for those geometric fibers "over $0 = [(x)]$". (Note that $\operatorname{Spec} \mathbb{C} \to \operatorname{Spec} \mathbb{Q}[x]$ and $\operatorname{Spec} \overline{\mathbb{Q}} \to \operatorname{Spec} \mathbb{Q}[x]$, both corresponding to ring maps with $x \mapsto 0$, are both geometric points "above 0".)

Checking whether a k-scheme is geometrically connected etc. seems annoying: you need to check every single algebraically closed field containing k. However, in each of these four cases, the failure of nice behavior of geometric fibers can already be detected after a finite field extension. For example, $\operatorname{Spec} \mathbb{Q}(i) \to \operatorname{Spec} \mathbb{Q}$ is not geometrically connected, and in fact you only need to base change by $\operatorname{Spec} \mathbb{Q}(i)$ to see this. We make this precise as follows.

Suppose X is a k-scheme. If K/k is a field extension, define $X_K = X \times_k \operatorname{Spec} K$. Consider the following twelve statements.

- X_K is reduced:
  (R$_a$) for all fields K,
  (R$_b$) for all algebraically closed fields K (X is geometrically reduced),
  (R$_c$) for $K = \overline{k}$,
  (R$_d$) for $K = k^p$ (where k^p is the perfect closure of k)
- X_K is irreducible:
  (I$_a$) for all fields K,
  (I$_b$) for all algebraically closed fields K (X is geometrically irreducible),
  (I$_c$) for $K = \overline{k}$,
  (I$_d$) for $K = k^s$ (where k^s is the separable closure of k).
- X_K is connected:
  (C$_a$) for all fields K,
  (C$_b$) for all algebraically closed fields K (X is geometrically connected),
  (C$_c$) for $K = \overline{k}$,
  (C$_d$) for $K = k^s$.

Trivially *(R$_a$)* implies *(R$_b$)* implies *(R$_c$)*, and *(R$_a$)* implies *(R$_d$)*, and similarly with "reduced" replaced by "irreducible" and "connected".

9.5.F. EXERCISE.
(a) Suppose that E/F is a field extension, and A is an F-algebra. Show that A is a subalgebra of $A \otimes_F E$. (Hint: think of these as vector spaces over F.)
(b) Show that: *(R$_b$)* implies *(R$_a$)* and *(R$_c$)* implies *(R$_d$)*.
(c) Show that: *(I$_b$)* implies *(I$_a$)* and *(I$_c$)* implies *(I$_d$)*.
(d) Show that: *(C$_b$)* implies *(C$_a$)* and *(C$_c$)* implies *(C$_d$)*.
Possible hint: You may use the fact that if Y is a nonempty F-scheme, then $Y \times_F \operatorname{Spec} E$ is nonempty, cf. Exercise 9.4.D.

Thus for example a k-scheme is geometrically integral if and only if it remains integral under any field extension.

9.5.2. Hard fact. In fact, *(R$_d$)* implies *(R$_a$)*, and thus *(R$_a$)* through *(R$_d$)* are all equivalent, and similarly for the other two rows. The explanation is below. On a first reading, you may want to read only Corollary 9.5.11 on connectedness, Proposition 9.5.14 on irreducibility, Proposition 9.5.20 on reducedness, and Theorem 9.5.23 on varieties, and then to use them to solve Exercise 9.5.N. You can later come back and read the proofs, which include some useful tricks turning questions about general schemes over a field to questions about finite type schemes.

The following exercise may help even the geometrically-minded reader appreciate the utility of these notions. (There is nothing important about the dimension 2 and the degree 4 in this exercise!)

9.5.G. EXERCISE. Recall from Remark 4.5.3 that the quartic curves in $\mathbb{P}^2_k$ are parametrized by a $\mathbb{P}^{14}_k$. (This will be made much more precise in §28.3.5.) Show that the points of $\mathbb{P}^{14}_k$ corresponding to geometrically irreducible curves form an open subset. Explain the necessity of the modifier "geometrically" (even if k is algebraically closed).

9.5.3. $\star\star$ **The rest of §9.5 is double-starred.**

9.5.4. Proposition. — *Suppose* A *and* B *are finite type* k*-algebras. Then* $\operatorname{Spec} A \times_k \operatorname{Spec} B \to \operatorname{Spec} B$ *is an open map.*

This is the one fact we will not prove here. We could (it isn't too hard), but instead we leave it until Exercise 24.5.H.

9.5.5. Preliminary discussion.

9.5.6. Lemma. — *Suppose* X *is a* k*-scheme. Then* $X \to \operatorname{Spec} k$ *is universally open, i.e., remains open after any base change.*

Proof. If S is an arbitrary k-scheme, we wish to show that $X_S \to S$ is open. It suffices to consider the case $X = \operatorname{Spec} A$ and $S = \operatorname{Spec} B$. To show that $\phi : \operatorname{Spec} A \otimes_k B \to \operatorname{Spec} B$ is open, it suffices to show that the image of a distinguished open set $D(f)$ ($f \in A \otimes_k B$) is open.

We come to a trick we will use repeatedly, which we will call the tensor-finiteness trick. Write $f = \sum a_i \otimes b_i$, where the sum is *finite*. It suffices to replace A by the subring generated by the a_i. (Reason: if this ring is A', then factor ϕ through $\operatorname{Spec} A' \otimes_k B$.) Thus we may assume A is finitely generated over k. Then use Proposition 9.5.4. $\square$

9.5.7. Lemma. — *Suppose the field extension* E/F *is purely inseparable (i.e., any* $a \in E$ *has minimal polynomial over* F *with only one root, perhaps with multiplicity). Suppose* X *is any* F*-scheme. Then* $\phi : X_E \to X$ *is a homeomorphism.*

Proof. The morphism ϕ is a bijection, so we may identify the points of X and X_E. (Reason: for any point $p \in X$, the scheme-theoretic fiber $\phi^{-1}(p)$ is a single point, by the definition of pure inseparability.) The morphism ϕ is continuous (so opens in X are open in X_E), and by Lemma 9.5.6, ϕ is open (so opens in X are open in X_E). $\square$

9.5.8. Connectedness.

Recall that a connected component of a topological space is a maximal connected subset §3.6.12.

9.5.H. EXERCISE (PROMISED IN REMARK 3.6.13). Show that every point is contained in a connected component, and that connected components are closed. (Hint: see the hint for Exercise 3.6.N.)

9.5.I. TOPOLOGICAL EXERCISE. Suppose $\phi : X \to Y$ is open, and has nonempty connected fibers. Then ϕ induces a bijection of connected components.

9.5.9. Lemma. — *Suppose* X *is geometrically connected over* k*. Then for any scheme* Y/k, $X \times_k Y \to Y$ *induces a bijection of connected components.*

Proof. Combine Lemma 9.5.6 and Exercise 9.5.I. $\square$

9.5.J. EXERCISE (PROMISED IN REMARK 3.6.3). Show that a scheme X is disconnected if and only if there exists a function $e \in \Gamma(X, \mathscr{O}_X)$ that is an idempotent ($e^2 = e$) distinct from 0 and 1. (Hint: if X is the disjoint union of two open sets X_0

and X_1, let e be the function that is 0 on X_0 and 1 on X_1. Conversely, given such an idempotent, define $X_0 = V(e)$ and $X_1 = V(1-e)$.)

9.5.10. Proposition. — *Suppose k is separably closed, and A is an k-algebra with* $\operatorname{Spec} A$ *connected. Then* $\operatorname{Spec} A$ *is geometrically connected over k.*

Proof. We wish to show that $\operatorname{Spec} A \otimes_k K$ is connected for any field extension K/k. It suffices to assume that K is algebraically closed (as $\operatorname{Spec} A \otimes_k \overline{K} \to \operatorname{Spec} A \otimes_k K$ is surjective). By choosing an embedding $\overline{k} \hookrightarrow K$ and considering the diagram

$$\begin{tikzcd}
\operatorname{Spec} A \otimes_k K \arrow[r] \arrow[d] & \operatorname{Spec} A \otimes_k \overline{k} \arrow[r, "\text{homeo.}", "\text{by Lem. 9.5.7}"'] \arrow[d] & \operatorname{Spec} A \arrow[d] \\
\operatorname{Spec} K \arrow[r] & \operatorname{Spec} \overline{k} \arrow[r] & \operatorname{Spec} k
\end{tikzcd}$$

it suffices to assume k is algebraically closed.

If $\operatorname{Spec} A \otimes_k K$ is disconnected, then $A \otimes_k K$ contains an idempotent $e \neq 0, 1$ (by Exercise 9.5.J). By the tensor-finiteness trick, we may assume that A is a finitely generated algebra over k, and K is a finitely generated field extension. Write $K = K(B)$ for some integral domain B of finite type over k. Then by the tensor-finiteness trick, by considering the finite number of denominators appearing in a representative of e as a sum of decomposable tensors, $e \in A \otimes_k B[1/b]$ for some nonzero $b \in B$, so $\operatorname{Spec} A \otimes_k B[1/b]$ is disconnected, say with open subsets U and V with $U \coprod V = \operatorname{Spec} A \otimes_k B[1/b]$.

Now $\phi : \operatorname{Spec} A \otimes_k B[1/b] \to \operatorname{Spec} B[1/b]$ is an open map (Proposition 9.5.4), so $\phi(U)$ and $\phi(V)$ are nonempty open sets. As $\operatorname{Spec} B[1/b]$ is connected, the intersection $\phi(U) \cap \phi(V)$ is a nonempty open set, which has a closed point p (with residue field k, as $k = \overline{k}$). But then $\phi^{-1}(p) \cong \operatorname{Spec} A$, and we have covered $\operatorname{Spec} A$ with two disjoint open sets, yielding a contradiction. □

9.5.11. Corollary. — *If k is separably closed, and Y is a connected k-scheme, then Y is geometrically connected.*

Proof. We wish to show that for any field extension K/k, Y_K is connected. By Proposition 9.5.10, $\operatorname{Spec} K$ is geometrically connected over k. Then apply Lemma 9.5.9 with $X = \operatorname{Spec} K$. □

9.5.12. Irreducibility.

9.5.13. Proposition. — *Suppose k is separably closed, A is a k-algebra with* $\operatorname{Spec} A$ *irreducible, and K/k is a field extension. Then* $\operatorname{Spec} A \otimes_k K$ *is irreducible.*

Proof. We follow the philosophy of the proof of Proposition 9.5.10. As in the first paragraph of that proof, it suffices to assume that K and k are algebraically closed. If $A \otimes_k K$ is not irreducible, then we can find x and y with $V(x), V(y) \neq \operatorname{Spec} A \otimes_k K$ and $V(x) \cup V(y) = \operatorname{Spec} A \otimes_k K$. As in the second paragraph of the proof of Proposition 9.5.10, we may assume that A is a finitely generated algebra over k, and $K = K(B)$ for an integral domain B of finite type over k, and $x, y \in A \otimes_k B[1/b]$ for some nonzero $b \in B$. Then $D(x)$ and $D(y)$ are nonempty open subsets of

$\operatorname{Spec} A \otimes_k B[1/b]$, whose image in $\operatorname{Spec} B[1/b]$ are nonempty opens, and thus their intersection is nonempty and contains a closed point p. But then $\phi^{-1}(p) \cong \operatorname{Spec} A$, and we have covered $\operatorname{Spec} A$ with two proper closed sets (the restrictions of $V(x)$ and $V(y)$), yielding a contradiction. □

9.5.K. EXERCISE. Suppose k is separably closed, and A and B are k-algebras, both irreducible (with irreducible Spec, i.e., with one minimal prime). Show that $A \otimes_k B$ is irreducible too. (Hint: reduce to the case where A and B are finite type over k. Extend the proof of the previous proposition.)

9.5.L. EASY EXERCISE. Show that a scheme X is irreducible if and only if there exists an open cover $X = \cup U_i$ with U_i irreducible for all i, and $U_i \cap U_j \neq \varnothing$ for all i, j.

9.5.14. Proposition. — *Suppose K/k is a field extension of a separably closed field and X_k is irreducible. Then X_K is irreducible.*

Proof. Take an open cover $X = \cup U_i$ by pairwise intersecting irreducible affine open subsets. The base change of each U_i to K is irreducible by Proposition 9.5.13, and they pairwise intersect. The result then follows from Exercise 9.5.L. □

9.5.M. EXERCISE. Suppose B is a geometrically integral k-algebra, and A is an integral k-algebra. Show that $A \otimes_k B$ is integral. (Once we define "variety", this will imply that the product of a geometrically integral variety with an integral variety is an integral variety.) Hint: revisit the proof of Exercise 9.4.E.

9.5.15. Reducedness.

We recall the following fact from field theory, which is a refined version of the basics of transcendence theory developed in Exercise 11.2.A. Because this is a starred section, we content ourselves with a reference rather than a proof.

9.5.16. Algebraic Fact: finitely generated extensions of perfect fields are separably generated, see [E, Cor. 16.17(b)] or [vdW, §19.7]. — *Suppose E/F is a finitely generated extension of a perfect field. Then it can be factored into a finite separable part and a purely transcendent part: $E/F(t_1, \dots, t_n)/F$.*

9.5.17. Proposition. — *Suppose B is a geometrically reduced k-algbra, and A is a reduced k-algebra. Then $A \otimes_k B$ is reduced.*

(Compare this to Exercise 9.5.M.)

Proof. Reduce to the case where A is finitely generated over k using the tensor-finiteness trick. (Suppose we have $x \in A \otimes_k B$ with $x^n = 0$. Then $x = \sum a_i \otimes b_i$. Let A' be the finitely generated subring of A generated by the a_i. Then $A' \otimes_k B$ is a subring of $A \otimes_k B$. Replace A by A'.) Then A is a subring of the product $\prod K_i$ of the function fields of its irreducible components (from our discussion on associated points: Theorem 5.5.10(b), see also Exercise 5.5.G). So it suffices to prove it for A a product of fields. Then it suffices to prove it when A is a field. But then we are done, by the definition of geometric reducedness. □

9.5.18. Proposition. — *Suppose A is a reduced k-algebra. Then:*

(a) $A \otimes_k k(t)$ *is reduced.*
(b) *If* E/k *is a finite separable extension, then* $A \otimes_k E$ *is reduced.*

Proof. (a) Clearly $A \otimes k[t]$ is reduced, and localization preserves reducedness (as reducedness is stalk-local, Exercise 5.2.A).

(b) Working inductively, we can assume E is generated by a single element, with minimal polynomial $p(t)$. By the tenor-finiteness trick, we can assume A is finitely generated over k. Then by the same trick as in the proof of Proposition 9.5.17, we can replace A by the product of its function fields of its components, and then we can assume A is a field. But then $A[t]/p(t)$ is reduced by the definition of separability of p. □

9.5.19. Lemma. — *Suppose* E/k *is a field extension of a perfect field, and* A *is a reduced* k*-algebra. Then* $A \otimes_k E$ *is reduced.*

Proof. By the tensor product finiteness trick, we may assume E is finitely generated over k. By Algebraic Fact 9.5.16, we can factor E/k into extensions of the forms of Proposition 9.5.18 (a) and (b). We then apply Proposition 9.5.18. □

9.5.20. Proposition. — *Suppose* E/k *is an extension of a perfect field, and* X *is a reduced* k*-scheme. Then* X_E *is reduced.*

Proof. Reduce to the case where X is affine. Use Lemma 9.5.19. □

9.5.21. Corollary. — *Suppose* k *is perfect, and* A *and* B *are reduced* k*-algebras. Then* $A \otimes_k B$ *is reduced.*

Proof. By Lemma 9.5.19, A is a geometrically reduced k-algebra. Then apply Lemma 9.5.17. □

9.5.22. Varieties.

9.5.23. Theorem. —
(a) If k *is perfect, the product of affine* k*-varieties (over* $\operatorname{Spec} k$*) is an affine* k*-variety.*
(b) If k *is separably closed, the product of irreducible affine* k*-varieties is an irreducible affine* k*-variety.*
(c) If k *is separably closed, the product of connected affine* k*-varieties is a connected affine* k*-variety.*

Once we define varieties in general, in Definition 10.1.7, you will be able to remove the adjective "affine" throughout this statement of Theorem 9.5.23. (We also remark that (b) was proven in Exercise 9.4.E under the stronger hypothesis that k is *algebraically* closed.)

Proof. (a) The finite type and separated statements are straightforward, as both properties are preserved by base change and composition. For reducedness, reduce to the affine case, then use Corollary 9.5.21.

(b) It only remains to show irreducibility. Reduce to the affine case using Exercise 9.5.L (as in the proof of Proposition 9.5.14). Then use Proposition 9.5.K.

(c) This follows from Corollary 9.5.11. □

9.5.N. EXERCISE (COMPLETING HARD FACT 9.5.2). Show that *(R_d)* implies *(R_a)*, *(I_d)* implies *(I_a)*, and *(C_d)* implies *(C_a)*.

9.5.O. EXERCISE. Suppose that A and B are two integral domains that are $\overline{k}$-algebras. Show that $A \otimes_{\overline{k}} B$ is an integral domain.

9.5.24. ⋆⋆ **Universally injective (radicial) morphisms.** As remarked in §9.4.1, injectivity is not preserved by base change. A better notion is that of **universally injective** morphisms: morphisms that are injections of sets after any base change. In keeping with the traditional agricultural terminology (sheaves, germs, ..., cf. Remark 2.4.4), these morphisms were named **radicial** after one of the lesser root vegetables. As a first example: all locally closed embeddings are universally injective (as they are injective, and remain locally closed embeddings upon any base change). If you wish, you can show more generally that all monomorphisms are universally injective. (Hint: show that monomorphisms are injective, and that the class of monomorphisms is preserved by base change.)

Universal injectivity is the algebro-geometric generalization of the notion of purely inseparable extensions of fields. A field extension K/L is **purely inseparable** if L is separably closed in K, or equivalently, if for every $\alpha \in K \setminus L$, the minimal polynomial of α is not a separable polynomial. This is obviously more interesting in positive characteristic.

9.5.P. EXERCISE. If K/L is an extension of fields. Show that $\operatorname{Spec} K \to \operatorname{Spec} L$ is universally injective if and only if K/L is purely inseparable.

9.5.Q. EXERCISE. Suppose $\pi : X \to Y$ is a morphism of schemes. Show that the following are equivalent.

(i) The morphism π is universally injective.
(ii) For every field K, the induced map $\operatorname{Hom}(\operatorname{Spec} K, X) \to \operatorname{Hom}(\operatorname{Spec} K, Y)$ is injective.
(iii) For each $p \in X$, the field extension $\kappa(p)/\kappa(\pi(p))$ is purely inseparable.

Possible hint: Connect i) and ii), and connect ii) and iii). Side Remark: in ii), you can show that it suffices to take K to be algebraically closed, so "universally injective" is the same as "injective on geometric points".

9.5.R. EASY EXERCISE. Show that the class of universally injective morphisms is stable under composition, base change, and products. Show that this notion is local on the target. Thus the class of universally injective morphisms is reasonable in the sense of §7.1.1.

Exercise 10.1.P shows that universal injectivity is really a property of the diagonal morphism, and explores the consequences of this.

9.6 Products of projective schemes: The Segre embedding

We next describe products of projective A-schemes over A. (The case of greatest initial interest is if $A = k$.) To do this, we need only describe $\mathbb{P}^m_A \times_A \mathbb{P}^n_A$, because any projective A-scheme has a closed embedding in some $\mathbb{P}^m_A$, and closed embeddings behave well under base change, so if $X \hookrightarrow \mathbb{P}^m_A$ and $Y \hookrightarrow \mathbb{P}^n_A$ are closed embeddings, then $X \times_A Y \hookrightarrow \mathbb{P}^m_A \times_A \mathbb{P}^n_A$ is also a closed embedding, cut out by the equations of X and Y (§9.2**(3)**). We will describe $\mathbb{P}^m_A \times_A \mathbb{P}^n_A$, and see that it too is a projective A-scheme. (Hence if X and Y are projective A-schemes, then their product $X \times_A Y$ over A is also a projective A-scheme.)

Before we do this, we will get some motivation from classical projective spaces (nonzero vectors modulo nonzero scalars, Exercise 4.4.F) in a special case. Our map will send $[x_0, x_1, x_2] \times [y_0, y_1]$ to a point in $\mathbb{P}^5$, whose coordinates we think of as being entries in the "multiplication table"

$$\begin{bmatrix} x_0y_0, & x_1y_0, & x_2y_0, \\ x_0y_1, & x_1y_1, & x_2y_1 \end{bmatrix}.$$

This is indeed a well-defined map of sets. Notice that the resulting matrix is rank one, and from the matrix, we can read off $[x_0, x_1, x_2]$ and $[y_0, y_1]$ up to scalars. For example, to read off the point $[x_0, x_1, x_2] \in \mathbb{P}^2$, we take the first row, unless it is all zero, in which case we take the second row. (They can't both be all zero.) In conclusion: in classical projective geometry, given a point of $\mathbb{P}^m$ and $\mathbb{P}^n$, we have produced a point in $\mathbb{P}^{mn+m+n}$, and from this point in $\mathbb{P}^{mn+m+n}$, we can recover the points of $\mathbb{P}^m$ and $\mathbb{P}^n$.

Suitably motivated, we return to algebraic geometry. We define a map

$$\mathbb{P}^m_A \times_A \mathbb{P}^n_A \to \mathbb{P}^{mn+m+n}_A$$

by

$$([x_0, \dots, x_m], [y_0, \dots, y_n]) \mapsto [z_{00}, z_{01}, \dots, z_{ij}, \dots, z_{mn}]$$

$$= [x_0y_0, x_0y_1, \dots, x_iy_j, \dots x_my_n].$$

More explicitly, we consider the map from the affine open set $U_i \times V_j$ (where $U_i = D(x_i)$ and $V_j = D(y_j)$ to the affine open set $W_{ij} = D(z_{ij})$ by

$$(x_{0/i}, \dots, x_{m/i}, y_{0/j}, \dots, y_{n/j}) \mapsto (x_{0/i}y_{0/j}, \dots, x_{i/i}y_{j/j}, \dots, x_{m/i}y_{n/j})$$

or, in terms of algebras, $z_{ab/ij} \mapsto x_{a/i}y_{b/j}$.

9.6.A. EXERCISE. Check that these maps glue to give a well-defined morphism $\mathbb{P}^m_A \times_A \mathbb{P}^n_A \to \mathbb{P}^{mn+m+n}_A$.

9.6.1. We next show that this morphism is a closed embedding. We can check this on an open cover of the target (the notion of being a closed embedding is affine-local, Exercise 8.1.D). Let's check this on the open set where $z_{ij} \neq 0$. The preimage of this open set in $\mathbb{P}^m_A \times \mathbb{P}^n_A$ is the locus where $x_i \neq 0$ and $y_j \neq 0$, i.e., $U_i \times V_j$. As described above, the map of rings is given by $z_{ab/ij} \mapsto x_{a/i}y_{b/j}$; this is clearly a surjection, as $z_{aj/ij} \mapsto x_{a/i}$ and $z_{ib/ij} \mapsto y_{b/j}$. (A generalization of this ad hoc description will be given in Exercise 16.4.D.)

This map is called the **Segre morphism** or **Segre embedding**. If A is a field, the image is called the **Segre variety**.

9.6.B. EXERCISE. Show that the Segre scheme (the image of the Segre embedding) is cut out (scheme-theoretically) by the equations corresponding to

$$\operatorname{rank}\begin{pmatrix} a_{00} & \cdots & a_{0n} \\ \vdots & \ddots & \vdots \\ a_{m0} & \cdots & a_{mn} \end{pmatrix} = 1,$$

i.e., that all 2×2 minors vanish. Hint: suppose you have a polynomial in the a_{ij} that becomes zero upon the substitution $a_{ij} = x_i y_j$. Give a recipe for subtracting polynomials of the form "monomial times 2×2 minor" so that the end result is 0. (The analogous question for the Veronese embedding in special cases is the content of Exercises 8.2.J and 8.2.L.)

9.6.2. *Important Example.* Let's consider the first nontrivial example, when $m = n = 1$. We get $\mathbb{P}^1_k \times_k \mathbb{P}^1_k \hookrightarrow \mathbb{P}^3_k$ (where k is a field). We get a single equation

$$\operatorname{rank}\begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} = 1,$$

i.e., $a_{00}a_{11} - a_{01}a_{10} = 0$. We again meet our old friend, the quadric surface (§8.2.9)! Hence: the smooth quadric surface $wz - xy = 0$ (Figure 8.2) is isomorphic to $\mathbb{P}^1_k \times_k \mathbb{P}^1_k$. Recall from Exercise 8.2.M that the quadric has two families (rulings) of lines. You may wish to check that one family of lines corresponds to the image of $\{x\} \times_k \mathbb{P}^1_k$ as x varies, and the other corresponds to the image $\mathbb{P}^1_k \times_k \{y\}$ as y varies.

If k is an algebraically closed field of characteristic not 2, then by diagonalizability of quadratics (Exercise 5.4.J), all rank 4 ("full rank") quadratics in 4 variables are isomorphic, so all rank 4 quadric surfaces over an algebraically closed field of characteristic not 2 are isomorphic to $\mathbb{P}^1_k \times_k \mathbb{P}^1_k$.

Note that this is not true over a field that is not algebraically closed. For example, over $\mathbb{R}$, $w^2 + x^2 + y^2 + z^2 = 0$ (in $\mathbb{P}^3_{\mathbb{R}}$) is not isomorphic to $\mathbb{P}^1_{\mathbb{R}} \times_{\mathbb{R}} \mathbb{P}^1_{\mathbb{R}}$. Reason: the former has no real points, while the latter has lots of real points.

You may wish to do the next two exercises in either order. The second can be used to show the first, but the first may give you insight into the second.

9.6.C. EXERCISE: A COORDINATE-FREE DESCRIPTION OF THE SEGRE EMBEDDING. Show that the Segre embedding can be interpreted as $\mathbb{P}V \times \mathbb{P}W \to \mathbb{P}(V \otimes W)$ via the surjective map of graded rings

$$\operatorname{Sym}^\bullet(V^\vee \otimes W^\vee) \longrightarrow\!\!\!\!\!\rightarrow \oplus_{i=0}^{\infty}\left(\operatorname{Sym}^i V^\vee\right) \otimes \left(\operatorname{Sym}^i W^\vee\right)$$

"in the opposite direction".

9.6.D. EXERCISE: A COORDINATE-FREE DESCRIPTION OF PRODUCTS OF PROJECTIVE A-SCHEMES IN GENERAL. Suppose that $S_\bullet$ and $T_\bullet$ are finitely generated graded rings over A. Describe an isomorphism

$$(\operatorname{Proj} S_\bullet) \times_A (\operatorname{Proj} T_\bullet) \cong \operatorname{Proj} \oplus_{n=0}^{\infty} (S_n \otimes_A T_n)$$

(where hopefully the definition of multiplication in the graded ring $\oplus_{n=0}^{\infty}(S_n \otimes_A T_n)$ is clear).

9.7 Normalization

We discuss normalization now only because the central construction gives practice with the central idea behind the construction in §9.1 of the fibered product (see Exercises 9.7.B and 9.7.I).

Normalization is a means of turning a *reduced* scheme into a normal scheme. A *normalization* of a reduced scheme X is a morphism $\nu : \widetilde{X} \to X$ from a normal scheme, where ν induces a bijection of irreducible components of $\widetilde{X}$ and X, and ν gives a birational morphism on each of the irreducible components. It will satisfy a universal property, and hence it is unique up to unique isomorphism. Figure 7.4 is an example of a normalization.

We begin with the case where X is irreducible, and hence integral. (We will then deal with a more general case, and also discuss normalization in a function field extension.) In this case of irreducible X, the **normalization** $\nu : \widetilde{X} \to X$ is a dominant morphism from an irreducible normal scheme to X, such that any other such morphism factors through ν:

$$\begin{tikzcd} \text{normal} & Y \arrow[rr, dotted, "\exists!"] \arrow[dr, "\text{dominant}"'] & & \widetilde{X} \arrow[dl, "\nu \text{ dominant}"] & \text{normal} \\ & & X & & \end{tikzcd}$$

Thus if the normalization exists, then it is unique up to unique isomorphism. We now have to show that it exists, and we do this in a way that will look familiar. We deal first with the case where X is affine, say $X = \operatorname{Spec} A$, where A is an integral domain. Then let $\widetilde{A}$ be the *integral closure* of A in its fraction field $K(A)$. (Recall that the integral closure of A in its fraction field consists of those elements of $K(A)$ that are solutions to monic polynomials in $A[x]$. It is a ring extension by Exercise 7.2.D, and integrally closed by Exercise 7.2.J.)

9.7.A. EXERCISE. Show that $\nu : \operatorname{Spec} \widetilde{A} \to \operatorname{Spec} A$ satisfies the universal property of normalization. (En route, you might show that the global sections of an irreducible normal scheme are also "normal", i.e., integrally closed.)

9.7.B. IMPORTANT (BUT SURPRISINGLY EASY) EXERCISE. Show that normalizations of integral schemes exist in general. (Hint: Ideas from the existence of fiber products, §9.1, may help.)

9.7.C. EASY EXERCISE. Show that normalizations are integral and surjective. (Hint for surjectivity: the Lying Over Theorem, see §7.2.6.)

We will soon see that normalization of integral finite type k-schemes is always a birational morphism, in Exercise 9.7.N.

9.7.D. EXERCISE. Explain (by defining a universal property) how to extend the notion of normalization to the case where X is a reduced scheme, with possibly more than one component, but under the hypothesis that every affine open subset of X has finitely many irreducible components. Note that this includes all locally Noetherian schemes. (If you wish, you can show that the normalization exists in this case. See [**Stacks**, tag 035Q] for more.)

Here are some examples.

9.7.E. EXERCISE. Show that $\operatorname{Spec} k[t] \to \operatorname{Spec} k[x,y]/(y^2 - x^2(x+1))$ given by $(x,y) \mapsto (t^2-1, t(t^2-1))$ (see Figure 7.4) is a normalization. (Hint: show that $k[t]$ and $k[x,y]/(y^2-x^2(x+1))$ have the same fraction field. Show that $k[t]$ is integrally closed. Show that $k[t]$ is contained in the integral closure of $k[x,y]/(y^2-x^2(x+1))$.)

You will see from the previous exercise that once we guess what the normalization is, it isn't hard to verify that it is indeed the normalization. Perhaps a few words are in order as to where the polynomials t^2-1 and $t(t^2-1)$ arose in the previous exercise. The key idea is to guess $t = y/x$. (Then $t^2 = x+1$ and $y = xt$ quickly.) This idea comes from three possible places. We begin by sketching the curve, and noticing the "node" at the origin. ("Node" will be formally defined in §29.3.1.) (a) The function y/x is well-defined away from the node, and at the node, the two branches have "values" $y/x = 1$ and $y/x = -1$. (b) We can also note that if $t = y/x$, then t^2 is a polynomial, so we will need to adjoin t in order to obtain the normalization. (c) The curve is cubic, so we expect a general line to meet the cubic in three points, counted with multiplicity. (We will make this precise when we discuss Bézout's Theorem, Exercise 18.6.K, but in this case we have already gotten a hint of this in Exercise 6.5.G.) There is a $\mathbb{P}^1$ parametrizing lines through the origin (with coordinate equal to the slope of the line, y/x), and most such lines meet the curve with multiplicity two at the origin, and hence meet the curve at precisely one other point of the curve. So this "coordinatizes" most of the curve, and we try adding in this coordinate.

9.7.F. EXERCISE. Find the normalization of the *cusp* $y^2 = x^3$ (see Figure 9.4). ("Cusp" will be formally defined in Definition 29.3.3.)

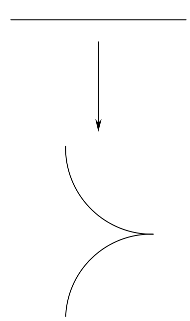


FIGURE 9.4. Normalization of a cusp

9.7.G. EXERCISE. Suppose $\operatorname{char} k \neq 2$. Find the normalization of the tacnode $y^2 = x^4$, and draw a picture analogous to Figure 9.4. ("Tacnode" will be formally defined in Definition 29.3.3.)

(Although we haven't defined "singularity", "smooth", "curve", or "dimension", you should still read this.) Notice that in the previous examples, normalization "resolves" the singularities ("non-smooth" points) of the curve. In general, it will do so in dimension one (in reasonable Noetherian circumstances, as normal Noetherian integral domains of dimension one are all discrete valuation rings, §12.5), but won't do so in higher dimension (the cone $z^2 = x^2 + y^2$ over a field k of characteristic not 2 is normal, Exercise 5.4.I(b)).

9.7.H. EXERCISE. Suppose $X = \operatorname{Spec} \mathbb{Z}[15i]$. Describe the normalization $\widetilde{X} \to X$. (Hint: $\mathbb{Z}[i]$ is a unique factorization domain, §5.4.6(0), and hence is integrally closed by Exercise 5.4.F.) Over what points of X is the normalization not an isomorphism?

Another exercise in a similar vein is the normalization of the "crumpled plane", Exercise 12.5.I.

9.7.I. EXERCISE (NORMALIZATION IN A FUNCTION FIELD EXTENSION, AN IMPORTANT GENERALIZATION). Suppose X is an integral scheme. The **normalization** $\nu : \widetilde{X} \to X$ **of** X **in a given algebraic field extension** L **of the function field** $K(X)$ **of** X is a dominant morphism from a normal scheme $\widetilde{X}$ with function field L, such that ν induces the inclusion $K(X) \hookrightarrow L$, and that is universal with respect to this property.

$$\begin{tikzcd}
\operatorname{Spec} L \arrow[r, equal] \arrow[dr, equal] & \operatorname{Spec} K(Y) \arrow[r] \arrow[d, "\sim"'] & Y \arrow[d, dotted, "\exists!"'] \arrow[dd, bend left] & \text{normal} \\
& \operatorname{Spec} K(\widetilde{X}) \arrow[r] \arrow[d] & \widetilde{X} \arrow[d] & \text{normal} \\
& \operatorname{Spec} K(X) \arrow[r] & X &
\end{tikzcd}$$

Show that the normalization in a finite field extension exists.

The following two examples, one arithmetic and one geometric, show that this is an interesting construction.

9.7.J. EXERCISE. Suppose $X = \operatorname{Spec} \mathbb{Z}$ (with function field $\mathbb{Q}$). Find its integral closure in the field extension $\mathbb{Q}(i)$. (There is no "geometric" way to do this; it is purely an algebraic problem, although the answer should be understood geometrically.)

9.7.1. *Remark: Rings of integers in number fields.* A finite extension K of $\mathbb{Q}$ is called a **number field**, and the integral closure of $\mathbb{Z}$ in K the **ring of integers in** K, denoted $\mathscr{O}_K$. (This notation is awkward given our other use of the symbol $\mathscr{O}$.)

$$\begin{tikzcd}
\operatorname{Spec} K \arrow[r] \arrow[d] & \operatorname{Spec} \mathscr{O}_K \arrow[d] \\
\operatorname{Spec} \mathbb{Q} \arrow[r] & \operatorname{Spec} \mathbb{Z}
\end{tikzcd}$$

By the previous exercises, $\mathscr{O}_K$ is a normal integral domain, and we will see soon (Theorem 9.7.3(a)) that it is Noetherian, and later (Exercise 11.1.F) that it has "dimension 1". This is an example of a *Dedekind domain*, see §12.5.14. We will think of it as a "smooth" curve as soon as we define what "smooth" (really, regular) and "curve" mean.

9.7.K. EXERCISE. Find the ring of integers in $\mathbb{Q}(\sqrt{n})$, where n is square-free and $n \equiv 3 \pmod 4$. (Hint: Exercise 5.4.I(a), where you will also be able to figure out the answer for square-free n in general.)

9.7.L. EXERCISE. Suppose $\operatorname{char} k \neq 2$ for convenience (although it isn't necessary).
(a) Suppose $X = \operatorname{Spec} k[x]$ (with function field $k(x)$). Find its normalization in the field extension $k(y)$, where $y^2 = x^2 + x$. (Again we get a Dedekind domain.) Hint: this can be done without too much pain. Show that $\operatorname{Spec} k[x, y]/(x^2 + x - y^2)$ is normal, possibly by identifying it as an open subset of $\mathbb{P}^1_k$, or possibly using Exercise 5.4.H.
(b) Suppose $X = \mathbb{P}^1$, with distinguished open $\operatorname{Spec} k[x]$. Find its normalization in the field extension $k(y)$, where $y^2 = x^2 + x$. (Part (a) involves computing the normalization over one affine open set; now figure out what happens over the "other" affine open set, and how to glue. The main lesson to draw is about how to glue — there will be two difference choices of how to glue the two pieces, corresponding to the Galois group of the function field extension, and the construction forces you to choose one of them.)

9.7.2. Fancy fact: Finiteness of integral closure.
The following fact is useful.

9.7.3. Theorem (finiteness of integral closure). — *Suppose A is a Noetherian integral domain, $K = K(A)$, L/K is a finite field extension, and B is the integral closure of A in L ("the integral closure of A in the field extension L/K", i.e., those elements of L integral over A).*
(a) If A is integrally closed and L/K is separable, then B is a finitely generated A-module.
(b) If A is a finitely generated k-algebra, then B is a finitely generated A-module.

Eisenbud gives a proof in a page and a half: (a) is [**E**, Prop. 13.14] and (b) is [**E**, Cor. 13.13]. A sketch is given in §9.7.5.

9.7.4. *Warning.* Part (b) does *not* hold for Noetherian A in general. In fact, the integral closure of a Noetherian ring need not be Noetherian (see [**E**, p. 299] for some discussion). This is alarming. The existence of such an example is a sign that Theorem 9.7.3 is not easy.

9.7.M. EXERCISE.
(a) Show that if X is an integral finite type k-scheme, then its normalization $\nu : \widetilde{X} \to X$ is a finite morphism.
(b) Suppose X is an integral scheme. Show that if X is normal, then the normalization in a *finite* separable field extension is a finite morphism. Show that if X is an integral finite type k-scheme, then the normalization in a finite field extension is a finite morphism. In particular, once we define "variety" (Definition 10.1.7), you will see that this implies that the normalization of a variety (including in a finite field extension) is a variety.

9.7.N. EXERCISE. Show that if X is an integral finite type k-scheme, then the normalization morphism is birational. (Hint: Proposition 6.5.7; or solve Exercise 9.7.O first.)

9.7.O. EXERCISE. Suppose that if X is an integral finite type k-scheme. Show that the normalization map of X is an isomorphism on an open dense subset of X. Hint: Proposition 6.5.5.

9.7.P. EXERCISE. Suppose $\rho : Z \to X$ is a finite birational morphism from an irreducible variety to an irreducible *normal* variety. Show that ρ is an isomorphism.

9.7.5. ⋆⋆ *Sketch of proof of finiteness of integral closure, Theorem 9.7.3.* Here is a sketch to show the structure of the argument. It uses commutative algebra ideas from Chapter 11, so you should only glance at this to see that nothing fancy is going on. *Part (a)*: reduce to the case where L/K is Galois, with group $\{\sigma_1, \dots, \sigma_n\}$. Choose $b_1, \dots, b_n \in B$ forming a K-vector space basis of L. Let M be the matrix (familiar from Galois theory) with ijth entry $\sigma_i b_j$, and let $d = \det M$. Show that the entries of M lie in B, and that $d^2 \in K$ (as d^2 is Galois-fixed). Show that $d \neq 0$ using linear independence of characters. Then complete the proof by showing that $B \subset d^{-2}(Ab_1 + \cdots + Ab_n)$ (submodules of finitely generated modules over Noetherian rings are also Noetherian, Exercise 3.6.X) as follows. Suppose $b \in B$, and write $b = \sum c_i b_i$ ($c_i \in K$). If c is the column vector with entries c_i, show that the ith entry of the column vector Mc is $\sigma_i b \in B$. Multiplying Mc on the left by $\operatorname{adj} M$ (see the trick of the proof of Lemma 7.2.1), show that $dc_i \in B$. Thus $d^2 c_i \in B \cap K = A$ (as A is integrally closed), as desired.

For *(b)*, use the Noether Normalization Lemma 11.2.4 to reduce to the case $A = k[x_1, \dots, x_n]$. Reduce to the case where L is normally closed over K. Let L' be the subextension of L/K so that L/L' is Galois and L'/K is purely inseparable. Use part (a) to reduce to the case $L = L'$. If $L' \neq K$, then for some q, L' is generated over K by the qth root of a finite set of rational functions. Reduce to the case $L' = k'(x_1^{1/q}, \dots, x_n^{1/q})$ where k'/k is a finite purely inseparable extension. In this case, show that $B = k'[x_1^{1/q}, \dots, x_n^{1/q}]$, which is indeed finite over $k[x_1, \dots, x_n]$. □

CHAPTER 10

Separated and proper morphisms, and (finally!) varieties

10.1 Separated morphisms (and quasiseparatedness done properly)

Separatedness is a fundamental notion. It is the analog of the Hausdorff condition for manifolds (see Exercise 10.1.A), and as with Hausdorffness, this geometrically intuitive notion ends up being just the right hypothesis to make theorems work. Although the definition initially looks odd, in retrospect it is just perfect.

10.1.1. Motivation. Let's review why we like Hausdorffness. Recall that a topological space is *Hausdorff* if for every two points x and y, there are disjoint open neighborhoods of x and y. The real line is Hausdorff, but the "real line with doubled origin" (of which Figure 4.6 may be taken as a sketch) is not. Many proofs and results about manifolds use Hausdorffness in an essential way. For example, the classification of compact one-dimensional manifolds is very simple, but if the Hausdorff condition were removed, we would have a very wild set.

So once armed with this definition, we can cheerfully exclude the line with doubled origin from civilized discussion, and we can (finally) define the notion of a *variety*, in a way that corresponds to the classical definition.

With our motivation from manifolds, we shouldn't be surprised that all of our affine and projective schemes are separated: certainly, in the land of manifolds, the Hausdorff condition comes for free for "subsets" of manifolds. (More precisely, if Y is a manifold, and X is a subset that satisfies all the hypotheses of a manifold except possibly Hausdorffness, then Hausdorffness comes for free. Similarly, we will see that locally closed embeddings in something separated are also separated: combine Exercise 10.1.B and Proposition 10.1.13(a).)

As an unexpected added bonus, a separated morphism to an affine scheme has the property that the intersection of two affine open sets in the source is affine (Proposition 10.1.8). This will make Čech cohomology work very easily on (quasicompact) schemes (Chapter 18). You might consider this an analog of the fact that in $\mathbb{R}^n$, the intersection of two convex sets is also convex. As affine schemes are trivial from the point of view of quasicoherent cohomology, just as convex sets in $\mathbb{R}^n$ have no cohomology, this metaphor is apt.

A lesson arising from the construction is the importance of the *diagonal morphism*. More precisely, given a morphism $\pi : X \to Y$, good consequences can be leveraged from good behavior of the **diagonal morphism** $\delta_\pi : X \to X \times_Y X$ (the product of the identity morphism $X \to X$ with itself), usually through fun diagram chases. This lesson applies across many fields of geometry. (Another nice gift of the diagonal morphism: it will give us a good algebraic definition of differentials, in Chapter 21.)

Grothendieck taught us that one should try to define properties of morphisms, not of objects; then we can say that an object has that property if its morphism to the final object has that property. We discussed this briefly at the start of Chapter 7. In this spirit, separatedness will be a property of morphisms, not schemes.

10.1.2. Defining separatedness. Before we define separatedness, we make an observation about all diagonal morphisms.

10.1.3. Proposition. — *Let $\pi : X \to Y$ be a morphism of schemes. Then the diagonal morphism $\delta : X \to X \times_Y X$ is a locally closed embedding.*

We will often use δ to denote a diagonal morphism. This locally closed subscheme of $X \times_Y X$ (which we also call the **diagonal**) will be denoted Δ.

Proof. We will describe a union of open subsets of $X \times_Y X$ covering the image of X, such that the image of X is a closed embedding in this union.

Say Y is covered with affine open sets V_i and X is covered with affine open sets U_{ij}, with $\pi : U_{ij} \to V_i$. Note that $U_{ij} \times_{V_i} U_{ij}$ is an affine open subscheme of the product $X \times_Y X$ (basically this is how we constructed the product, by gluing together affine building blocks). Then the diagonal is covered by these affine open subsets $U_{ij} \times_{V_i} U_{ij}$. (Any point $p \in X$ lies in some U_{ij}; then $\delta(p) \in U_{ij} \times_{V_i} U_{ij}$. Figure 10.1 may be helpful.) Note that $\delta^{-1}(U_{ij} \times_{V_i} U_{ij}) = U_{ij}$: clearly $U_{ij} \subset \delta^{-1}(U_{ij} \times_{V_i} U_{ij})$, and because $pr_1 \circ \delta = id_X$ (where pr_1 is the first projection), $\delta^{-1}(U_{ij} \times_{V_i} U_{ij}) \subset U_{ij}$. Finally, we check that $U_{ij} \to U_{ij} \times_{V_i} U_{ij}$ is a closed embedding. Say $V_i = \operatorname{Spec} B$ and $U_{ij} = \operatorname{Spec} A$. Then this corresponds to the natural ring map $A \otimes_B A \to A$ ($a_1 \otimes a_2 \mapsto a_1 a_2$), which is obviously surjective. $\square$

The open subsets we described may not cover $X \times_Y X$, so we have not shown that δ is a closed embedding.

10.1.4. *Definition.* A morphism $\pi : X \to Y$ is **separated** if the diagonal morphism $\delta_\pi : X \to X \times_Y X$ is a closed embedding. An A-scheme X is said to be **separated over** A if the structure morphism $X \to \operatorname{Spec} A$ is separated. When people say that a scheme (rather than a morphism) X is separated, they mean implicitly that some "structure morphism" is separated. For example, if they are talking about A-schemes, they mean that X is separated over A.

Thanks to Proposition 10.1.3 (and once you show that a locally closed embedding whose image is closed is actually a closed embedding), a morphism is separated if and only if the diagonal Δ is a closed subset — a purely topological condition on the diagonal. This is reminiscent of a definition of Hausdorff, as the next exercise shows.

10.1.A. UNIMPORTANT EXERCISE (FOR THOSE SEEKING TOPOLOGICAL MOTIVATION). Show that a topological space X is Hausdorff if and only if the diagonal is a closed subset of $X \times X$. (The reason separatedness of schemes doesn't give Hausdorffness — i.e., that for any two open points x and y there aren't necessarily disjoint open neighborhoods — is that in the category of schemes, the topological

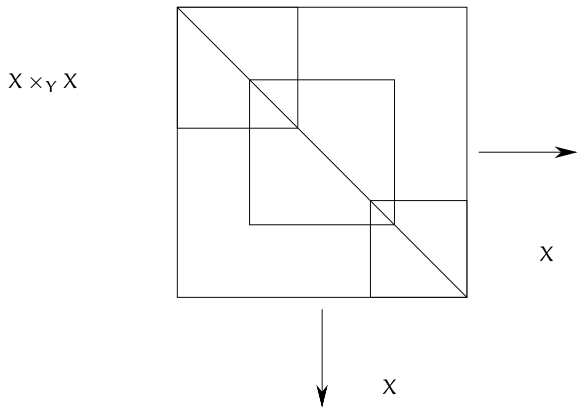


FIGURE 10.1. A neighborhood of the diagonal is covered by $U_{ij} \times_{V_i} U_{ij}$

space $X \times X$ is not in general the product of the topological space X with itself, see §9.1.2.)

10.1.B. IMPORTANT EASY EXERCISE. Show locally closed embeddings (and in particular open and closed embeddings) are separated. (Hint: Do this by hand. Alternatively, show that monomorphisms are separated. Open and closed embeddings are monomorphisms, by Exercise 9.2.G.)

10.1.C. IMPORTANT EASY EXERCISE. Show that every morphism of affine schemes is separated. (Hint: this was essentially done in the proof of Proposition 10.1.3.)

10.1.D. EXERCISE. Show that the line with doubled origin X (Example 4.4.5) is not separated, by verifying that the image of the diagonal morphism is not closed. (Another argument is given below, in Exercise 10.2.C. A fancy argument is given in Exercise 12.7.C.)

We next come to our first example of something separated but not affine. The following single calculation will imply that all quasiprojective A-schemes are separated (once we know that the composition of separated morphisms are separated, Proposition 10.1.13).

10.1.5. Proposition. — *The morphism* $\mathbb{P}^n_A \to \operatorname{Spec} A$ *is separated.*

We give two proofs. The first is by direct calculation. The second requires no calculation, and just requires that you remember some classical constructions described earlier.

Proof 1: Direct calculation. We cover $\mathbb{P}^n_A \times_A \mathbb{P}^n_A$ with open sets of the form $U_i \times_A U_j$, where $U_0, \ldots, U_n$ form the "usual" affine open cover. The case $i = j$ was taken care of before, in the proof of Proposition 10.1.3. If $i \neq j$ then

$$U_i \times_A U_j \cong \operatorname{Spec} A[x_{0/i}, \ldots, x_{n/i}, y_{0/j}, \ldots, y_{n/j}]/(x_{i/i} - 1, y_{j/j} - 1).$$

Now the restriction of the diagonal Δ is contained in U_i (as the diagonal morphism composed with projection to the first factor is the identity), and similarly is contained in U_j. Thus the diagonal morphism over $U_i \times_A U_j$ is $U_i \cap U_j \to U_i \times_A U_j$. This is a closed embedding, as the corresponding map of rings

$$A[x_{0/i}, \dots, x_{n/i}, y_{0/j}, \dots, y_{n/j}] \to A[x_{0/i}, \dots, x_{n/i}, x_{j/i}^{-1}]/(x_{i/i} - 1)$$

(given by $x_{k/i} \mapsto x_{k/i}$, $y_{k/j} \mapsto x_{k/i}/x_{j/i}$) is clearly a surjection (as each generator of the ring on the right is clearly in the image — note that $x_{j/i}^{-1}$ is the image of $y_{i/j}$). $\square$

Proof 2: Classical geometry. Note that the diagonal morphism $\delta : \mathbb{P}_A^n \to \mathbb{P}_A^n \times_A \mathbb{P}_A^n$ followed by the Segre embedding $S : \mathbb{P}_A^n \times_A \mathbb{P}_A^n \to \mathbb{P}^{n^2+2n}$ (§9.6, a closed embedding) can also be factored as the second Veronese embedding $\nu_2 : \mathbb{P}_A^n \to \mathbb{P}^{\binom{n+2}{2}-1}$ (§8.2.6) followed by a linear map $L : \mathbb{P}^{\binom{n+2}{2}-1} \to \mathbb{P}^{n^2+2n}$ (another closed embedding, Exercise 8.2.D), both of which are closed embeddings.

$$\begin{tikzcd}
& \mathbb{P}_A^n \times_A \mathbb{P}_A^n \arrow[dr, "S"', "\text{cl. emb.}"] & \\
\mathbb{P}_A^n \arrow[ur, "\delta", "\text{1. (?) cl. emb.}"'] \arrow[dr, "\text{cl. emb.}", "\nu_2"'] & & \mathbb{P}^{n^2+2n} \\
& \mathbb{P}^{\binom{n+2}{2}-1} \arrow[ur, "\text{cl. emb.}", "L"'] &
\end{tikzcd}$$

Informally, in coordinates:

$$\begin{tikzcd}
& ([x_0, x_1, \dots, x_n], [x_0, x_1, \dots, x_n]) \arrow[dr, "S"] & \\
{[x_0, x_1, \dots, x_n]} \arrow[ur, "\delta"] \arrow[dr, "\nu_2"'] & & \begin{bmatrix} x_0^2, & x_0x_1, & \cdots & x_0x_n, \\ x_1x_0, & x_1^2, & \cdots & x_1x_n, \\ \vdots & \vdots & \ddots & \vdots \\ x_nx_0, & x_nx_1, & \cdots & x_n^2 \end{bmatrix} \\
& {[x_0^2, x_0x_1, \dots, x_{n-1}x_n, x_n^2]} \arrow[ur, "L"'] &
\end{tikzcd}$$

The composed map $\mathbb{P}_A^n$ may be written as $[x_0, \dots, x_n] \mapsto [x_0^2, x_0x_1, x_0x_2, \dots, x_n^2]$, where the subscripts on the right run over all ordered pairs (i, j) where $0 \le i, j \le n$.) This forces δ to send closed sets to closed sets (or else $S \circ \delta$ won't, but $L \circ \nu_2$ does). $\square$

We note for future reference a minor result proved in the course of Proof 1.

10.1.6. Small Proposition. — *If U and V are open subsets of an A-scheme X, then $\Delta \cap (U \times_A V) \cong U \cap V$.*

Figure 10.2 may help show why this is natural. You could also interpret this statement as

$$X \times_{(X \times_A X)} (U \times_A V) \cong U \times_X V$$

which follows from the magic diagram, Exercise 1.3.S.

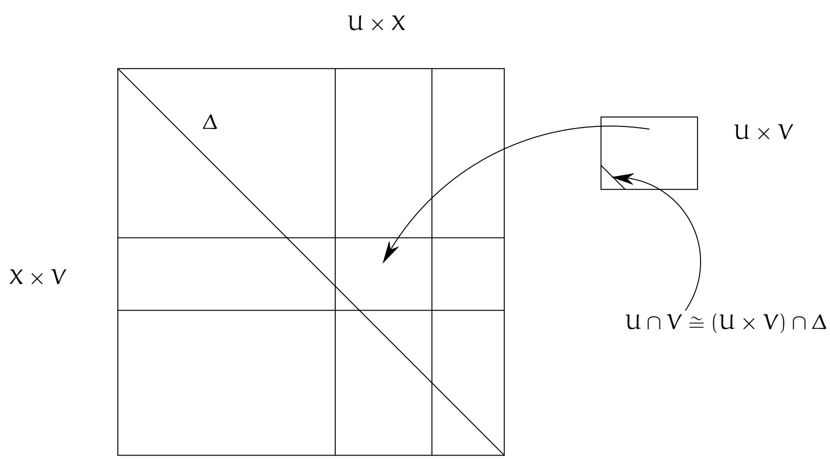


FIGURE 10.2. Small Proposition 10.1.6

10.1.7. Important Definition (finally!). A **variety** over a field k, or k-**variety**, is a reduced, separated scheme of finite type over k. For example, a reduced finite type affine k-scheme is a variety. (Do you see why separatedness holds?) The notion of variety generalizes our earlier notion of affine variety (§5.3.7) and projective variety (§5.3.D, see Proposition 10.1.14). (Notational caution: In some sources, the additional condition of irreducibility is imposed. Also, it is often assumed that k is algebraically closed.)

Varieties over k form a category: morphisms of k-varieties are just morphisms as k-schemes. We will shortly see (in Exercise 10.1.L) that such morphisms are automatically finite type and separated. (Of course, the category of varieties over an algebraically closed field k long predates modern scheme theory.)

A **subvariety** of a variety X is a *reduced* locally closed subscheme of X (which you can quickly check is a variety itself). An open subvariety of X is an open subscheme of X. (Reducedness is automatic in this case.) A closed subvariety of X is a reduced closed subscheme of X.

If you have read the double-starred section on group objects in a category (§6.6.4), you will automatically have the notion of a group variety. We will discuss this a bit more in §10.3.9.

10.1.E. EXERCISE (PRODUCTS OF IRREDUCIBLE VARIETIES OVER $\overline{k}$ ARE IRREDUCIBLE VARIETIES). Use Exercise 9.4.E and properties of separatedness to show that the product of two irreducible $\overline{k}$-varieties is an irreducible $\overline{k}$-variety.

10.1.F. ⋆⋆ EXERCISE (COMPLEX ALGEBRAIC VARIETIES YIELD COMPLEX ANALYTIC VARIETIES; FOR THOSE WITH SUFFICIENT BACKGROUND). Show that the analytification (Exercises 5.3.F and 6.3.K) of a complex algebraic variety is a complex analytic variety.

Here is a very handy consequence of separatedness.

10.1.8. Proposition. — *Suppose $X \to \operatorname{Spec} A$ is a separated morphism to an affine scheme, and U and V are affine open subsets of X. Then $U \cap V$ is an affine open subset of X.*

Before proving this, we state a consequence that is otherwise nonobvious. If $X = \operatorname{Spec} A$, then the intersection of any two affine open subsets of A is an affine open subset. This is certainly not an obvious fact! We know the intersection of two distinguished affine open sets is affine (from $D(f) \cap D(g) = D(fg)$), but we have little handle on affine open sets in general.

Warning: this property does not characterize separatedness. For example, if $A = \operatorname{Spec} k$ and X is the line with doubled origin over k, then X also has this property.

Proof. By Proposition 10.1.6, $(U \times_A V) \cap \Delta \cong U \cap V$, where Δ is the diagonal. But $U \times_A V$ is affine (the fibered product of two affine schemes over an affine scheme is affine, Step 1 of our construction of fibered products, Theorem 9.1.1), and $(U \times_A V) \cap \Delta$ is a closed subscheme of the affine scheme $U \times_A V$, and hence $U \cap V$ is affine. □

10.1.9. Redefinition: Quasiseparated morphisms.

We say a morphism $\pi : X \to Y$ is **quasiseparated** if the diagonal morphism $\delta_\pi : X \to X \times_Y X$ is quasicompact.

10.1.G. EXERCISE. Show that this agrees with our earlier definition of quasiseparated (§7.3.1): show that $\pi : X \to Y$ is quasiseparated if and only if for any affine open $\operatorname{Spec} A$ of Y, and two affine open subsets U and V of X mapping to $\operatorname{Spec} A$, $U \cap V$ is a *finite* union of affine open sets. (Possible hint: compare this to Proposition 10.1.8. Another possible hint: the magic diagram, Exercise 1.3.S.)

Here are two large classes of morphisms that are quasiseparated.

10.1.H. EASY EXERCISE. Show that separated morphisms are quasiseparated. (Hint: closed embeddings are affine, hence quasicompact.)

Second, any morphism from a locally Noetherian scheme is quasiseparated, see Exercise 7.3.B(b), so Noetherian people need never worry about this issue.

We now give four quick propositions showing that separatedness and quasiseparatedness behave well, just as many other classes of morphisms did.

10.1.10. Proposition. — *Both separatedness and quasiseparatedness are preserved by base change.*

Proof. Suppose

$$\begin{tikzcd} W \arrow[r] \arrow[d] & X \arrow[d] \\ Y \arrow[r] & Z \end{tikzcd}$$

is a fiber diagram. We will show that if $Y \to Z$ is separated or quasiseparated, then so is $W \to X$. Then you can quickly verify that

$$\begin{tikzcd} W \arrow[r, "\delta_W"] \arrow[d] & W \times_X W \arrow[d] \\ Y \arrow[r, "\delta_Y"] & Y \times_Z Y \end{tikzcd}$$

is a fiber diagram. (This is true in any category with fibered products.) As the property of being a closed embedding is preserved by base change (§9.2 **(3)**), if δ_Y is a closed embedding, so is δ_W.

The quasiseparatedness case follows in the identical manner, as quasicompactness is also preserved by base change (Exercise 9.4.B(a)). □

10.1.11. Proposition. — *The condition of being separated is local on the target. Precisely, a morphism $\pi : X \to Y$ is separated if and only if for any cover of Y by open subsets U_i, $\pi^{-1}(U_i) \to U_i$ is separated for each i.*

10.1.12. Hence affine morphisms are separated, as every morphism of affine schemes is separated (Exercise 10.1.C). In particular, finite morphisms are separated.

Proof. If $\pi : X \to Y$ is separated, then for any $U_i \hookrightarrow Y$, $\pi^{-1}(U_i) \to U_i$ is separated, as separatedness is preserved by base change (Theorem 10.1.10). Conversely, to check if $\Delta \hookrightarrow X \times_Y X$ is a closed subset, it suffices to check this on an open cover of $X \times_Y X$. Let $\rho : X \times_Y X \to Y$ be the natural map. We will use the open cover $\rho^{-1}(U_i)$, which by construction of the fiber product is $f^{-1}(U_i) \times_{U_i} f^{-1}(U_i)$. As $f^{-1}(U_i) \to U_i$ is separated, $\pi^{-1}(U_i) \to \pi^{-1}(U_i) \times_{U_i} \pi(U_i)$ is a closed embedding by definition of separatedness. □

10.1.I. EXERCISE. Prove that the condition of being quasiseparated is local on the target. (Hint: the condition of being quasicompact is local on the target by Exercise 7.3.C(a); use a similar argument as in Proposition 10.1.11.)

10.1.13. Proposition. —
(a) The condition of being separated is closed under composition. In other words, if $\pi : X \to Y$ is separated and $\rho : Y \to Z$ is separated, then $\rho \circ \pi : X \to Z$ is separated.
(b) The condition of being quasiseparated is closed under composition.

Proof. (a) Let $\tau = \rho \circ \pi$. We are given that $\delta_\pi : X \hookrightarrow X \times_Y X$ and $\delta_\rho : Y \hookrightarrow Y \times_Z Y$ are closed embeddings, and we wish to show that $\delta_\tau : X \to X \times_Z X$ is a closed embedding. Consider the diagram

$$\begin{tikzcd} X \arrow[r, hook, "\delta_\pi"] & X \times_Y X \arrow[r, "\psi"] \arrow[d] & X \times_Z X \arrow[d] \\ & Y \arrow[r, hook, "\delta_\rho"] & Y \times_Z Y. \end{tikzcd}$$

The square is the magic diagram (Exercise 1.3.S). As δ_ρ is a closed embedding, ψ is too (closed embeddings are preserved by base change, §9.2 **(3)**). Thus $\psi \circ \delta_\pi$ is a closed embedding (the composition of two closed embeddings is also a closed embedding, Exercise 8.1.C).

(b) The identical argument (with "closed embedding" replaced by "quasicompact") shows that the condition of being quasiseparated is closed under composition. $\square$

10.1.14. Corollary. — *Every quasiprojective A-scheme is separated over A. In particular, every reduced quasiprojective k-scheme is a k-variety.*

Proof. Suppose $X \to \operatorname{Spec} A$ is a quasiprojective A-scheme. The structure morphism can be factored into an open embedding composed with a closed embedding followed by $\mathbb{P}^n_A \to \operatorname{Spec} A$. Open embeddings and closed embeddings are separated (Exercise 10.1.B), and $\mathbb{P}^n_A \to \operatorname{Spec} A$ is separated (Proposition 10.1.5). Compositions of separated morphisms are separated (Proposition 10.1.13), so we are done. $\square$

10.1.15. Proposition. — *Suppose $\pi : X \to Y$ and $\pi' : X' \to Y'$ are separated (resp. quasiseparated) morphisms of S-schemes (where S is a scheme). Then the product morphism $\pi \times \pi' : X \times_S X' \to Y \times_S Y'$ is separated (resp. quasiseparated).*

Proof. Apply Exercise 9.4.F. $\square$

10.1.16. Applications.

As a first application, we define the graph of a morphism.

10.1.17. *Definition.* Suppose $\pi : X \to Y$ is a morphism of Z-schemes. The morphism $\Gamma_\pi : X \to X \times_Z Y$ given by $\Gamma_\pi = (\mathrm{id}, \pi)$ is called the **graph morphism**. Then π factors as $\mathrm{pr}_2 \circ \Gamma_\pi$, where pr_2 is the second projection (see Figure 10.3). The diagram of Figure 10.3 is often called the **graph of a morphism**. (We will discuss graphs of rational maps in §10.2.4.)

10.1.18. Proposition. — *The graph morphism Γ is always a locally closed embedding. If Y is a separated Z-scheme (i.e., the structure morphism $Y \to Z$ is separated), then Γ is a closed embedding. If Y is a quasiseparated Z-scheme, then Γ is quasicompact.*

This will be generalized in Exercise 10.1.M.

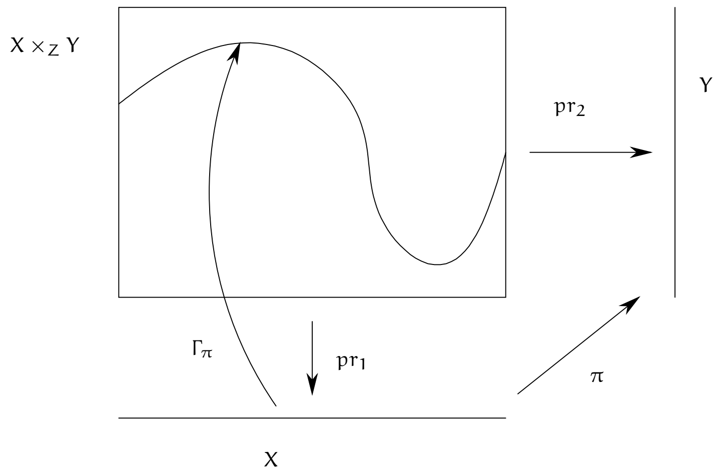


FIGURE 10.3. The graph morphism

Proof by Cartesian diagram. A special case of the magic diagram (Exercise 1.3.S) is:

(10.1.18.1)
$$\begin{tikzcd} X \arrow[r, "\Gamma_\pi"] \arrow[d, "\pi"'] & X \times_Z Y \arrow[d] \\ Y \arrow[r, "\delta"] & Y \times_Z Y. \end{tikzcd}$$

The notions of locally closed embedding and closed embedding are preserved by base change, so if the bottom arrow δ has one of these properties, so does the top. The same argument establishes the last sentence of Proposition 10.1.18. □

We next come to strange-looking, but very useful, result. Like the magic diagram, this result is unexpectedly ubiquitous.

10.1.19. Cancellation Theorem for a Property P of Morphisms. — *Let P be a class of morphisms that is preserved by base change and composition. (Any "reasonable" class of morphisms will satisfy this, see §7.1.1.) Suppose*

$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "\tau"'] & & Y \arrow[dl, "\rho"] \\ & Z & \end{tikzcd}$$

is a commuting diagram of schemes. Suppose that the diagonal morphism $\delta_\rho : Y \to Y \times_Z Y$ is in P and $\tau : X \to Z$ is in P. Then $\pi : X \to Y$ is in P. In particular:

(i) *Suppose that locally closed embeddings are in P. If τ is in P, then π is in P.*
(ii) *Suppose that closed embeddings are in P (e.g. P could be finite morphisms, morphisms of finite type, closed embeddings, affine morphisms). If τ is in P and ρ is separated, then π is in P.*

(iii) *Suppose that quasicompact morphisms are in* P. *If* τ *is in* P *and* ρ *is quasiseparated, then* π *is in* P.

The following diagram summarizes this important theorem:

$$\begin{tikzcd} X \arrow[rr, "\therefore \in P"] \arrow[dr, "\in P"'] & & Y \arrow[dl, "\delta \in P"] \\ & Z & \end{tikzcd}$$

When you plug in different P, you get very different-looking (and nonobvious) consequences, the first of which are given in Exercise 10.1.K. (Here are some facts you can prove easily, but which can be interpreted as applications of the Cancellation Theorem in *Sets*, and which may thus shed light on how the Cancellation Theorem works. If $\pi : X \to Y$ and $\rho : Y \to Z$ are maps of sets, and $\rho \circ \pi$ is injective, then so is π; and if $\rho \circ \pi$ is surjective and ρ is injective, then π is surjective.)

Proof. By the graph Cartesian diagram (10.1.18.1)

$$\begin{tikzcd} X \arrow[r, "\Gamma_\pi"] \arrow[d, "\pi"'] & X \times_Z Y \arrow[d] \\ Y \arrow[r, "\delta_\rho"] & Y \times_Z Y \end{tikzcd}$$

we see that the graph morphism $\Gamma_\pi : X \to X \times_Z Y$ is in P (Definition 10.1.17), as P is closed under base change. By the fibered square

$$\begin{tikzcd} X \times_Z Y \arrow[r, "\tau'"] \arrow[d] & Y \arrow[d, "\rho"] \\ X \arrow[r, "\tau"] & Z \end{tikzcd}$$

the projection $\tau' : X \times_Z Y \to Y$ is in P as well. Thus $\pi = \tau' \circ \Gamma_\pi$ is in P $\square$

Here now are some fun and useful exercises.

10.1.J. LESS IMPORTANT EXERCISE. Show that an A-scheme is separated (over A) if and only if it is separated over $\mathbb{Z}$. In particular, a complex scheme is separated over $\mathbb{C}$ if and only if it is separated over $\mathbb{Z}$, so complex geometers and arithmetic geometers can discuss separated schemes without confusion.

10.1.K. EASY EXERCISE. Suppose we have morphisms $X \xrightarrow{\pi} Y \xrightarrow{\rho} Z$.
(a) Show that if $\rho \circ \pi$ is a locally closed embedding (resp. locally of finite type, separated), then so is π.
(b) If $\rho \circ \pi$ is quasicompact, and Y is Noetherian, show that π is quasicompact. Hint: Exercise 7.3.B(a).
(c) If $\rho \circ \pi$ is quasiseparated, and Y is locally Noetherian, show that π is quasiseparated. Hint: Exercise 7.3.B(b).

10.1.L. EASY EXERCISE. Show that morphisms of k-varieties (i.e., morphisms as k-schemes, see §10.1.7) are finite type and separated.

10.1.M. EXERCISE. Suppose $\mu : Z \to X$ is a morphism, and $\sigma : X \to Z$ is a *section* of μ, i.e., $\mu \circ \sigma$ is the identity on X.

$$\begin{tikzcd} Z \arrow[d, "\mu"', shift right] \\ X \arrow[u, "\sigma"', bend right] \end{tikzcd}$$

Show that σ is a locally closed embedding. Show that if μ is separated, then σ is a closed embedding. (This generalizes Proposition 10.1.18.) Give an example to show that σ need not be a closed embedding if μ is not separated.

10.1.N. LESS IMPORTANT EXERCISE. Suppose P is a class of morphisms such that closed embeddings are in P, and P is closed under fibered product and composition. Show that if $\pi : X \to Y$ is in P then $\pi^{\text{red}} : X^{\text{red}} \to Y^{\text{red}}$ is in P. (Two examples are the classes of separated morphisms and quasiseparated morphisms.) Hint:

$$\begin{tikzcd} X^{\text{red}} \arrow[r] \arrow[dr] & X \times_Y Y^{\text{red}} \arrow[r] \arrow[d] & Y^{\text{red}} \arrow[d] \\ & X \arrow[r, "\pi"] & Y \end{tikzcd}$$

10.1.20. $\star\star$ **Universally injective morphisms and the diagonal.**

10.1.O. EASY EXERCISE. If $\pi : X \to Y$ and $\rho : Y \to Z$ are morphisms, and $\rho \circ \pi$ is universally injective, show that π is universally injective.

10.1.P. EXERCISE.
(a) Show that $\pi : X \to Y$ is universally injective if and only if the diagonal morphism $\delta_\pi : X \to X \times_Y X$ is surjective. (Note that δ_π is always injective, as it is a locally closed embedding, by Proposition 10.1.3.)
(b) Show that universally injective morphisms are separated.
(c) Show that a map between finite type schemes over an algebraically closed field $\overline{k}$ is universally injective if and only if it is injective on closed points.

10.2 Rational maps to separated schemes

When we introduced rational maps in §6.5, we promised that in good circumstances, a rational map has a "largest domain of definition". We are now ready to make precise what "good circumstances" means, in the Reduced-to-Separated Theorem 10.2.2. We first introduce an important result making sense of locus where two morphisms with the same source and target "agree".

10.2.A. USEFUL EXERCISE: THE LOCUS WHERE TWO MORPHISMS AGREE. Suppose $\pi : X \to Y$ and $\pi' : X \to Y$ are two morphisms over some scheme Z. We can now give meaning to the phrase 'the locus where π and π' agree', and that in particular there is a largest locally closed subscheme where they agree — which is closed if Y is separated over Z. Suppose $\mu : W \to X$ is some morphism (not assumed to be a locally closed embedding). We say that π and π' agree on μ if $\pi \circ \mu = \pi' \circ \mu$. Show that there is a locally closed subscheme $i : V \hookrightarrow X$ on which π and π' agree, such that any morphism $\mu : W \to X$ on which π and π' agree factors uniquely through i, i.e., there is a unique $j : W \to V$ such that $\mu = i \circ j$. Show further that if $Y \to Z$ is separated, then $i : V \hookrightarrow X$ is a closed embedding. Hint: define V to be the following fibered product:

$$\begin{tikzcd} V \arrow[r] \arrow[d] & Y \arrow[d, "\delta"] \\ X \arrow[r, "(\pi,\pi')"] & Y \times_Z Y. \end{tikzcd}$$

As δ is a locally closed embedding, $V \to X$ is too. Then if $\mu : W \to X$ is any morphism such that $\pi \circ \mu = \pi' \circ \mu$, then μ factors through V.

The fact that the locus where two maps agree can be nonreduced should not come as a surprise: consider two maps from $\mathbb{A}^1_k$ to itself, $\pi(x) = 0$ and $\pi'(x) = x^2$. They agree when $x = 0$, but the situation is (epsilonically) better than that — they should agree even on $\operatorname{Spec} k[x]/(x^2)$.

10.2.1. *Minor Remarks.*

(i) In the previous exercise, we are describing $V \hookrightarrow X$ by way of a universal property. Taking this as the definition, it is not a priori clear that V is a locally closed subscheme of X, or even that it exists.

(ii) Warning: consider two maps from $\operatorname{Spec} \mathbb{C}$ to itself over $\operatorname{Spec} \mathbb{R}$, the identity and complex conjugation. These are both maps from a point to a point, yet they do not agree despite agreeing as maps of sets. (If you do not find this reasonable, this might help: after base change $\operatorname{Spec} \mathbb{C} \to \operatorname{Spec} \mathbb{R}$, they do not agree even as maps of sets.)

(iii) More generally, the locus where f and g agree can be interpreted as follows: π and π' agree at x if $\pi(x) = \pi'(x)$ and the two maps of residue fields are the same.

10.2.B. EXERCISE: MAPS OF $\overline{k}$-VARIETIES ARE DETERMINED BY THE MAPS ON CLOSED POINTS. Suppose $\pi : X \to Y$ and $\pi' : X \to Y$ are two morphisms of $\overline{k}$-varieties that are the same at the level of closed points (i.e., for each closed point $p \in X$, $\pi(p) = \pi'(p)$). Show that $\pi = \pi'$. (This implies that the functor from the category of "classical varieties over $\overline{k}$", which we won't define here, to the category of $\overline{k}$-schemes, is fully faithful. Can you generalize this appropriately to non-algebraically closed fields?)

10.2.C. LESS IMPORTANT EXERCISE. Show that the line with doubled origin X (Example 4.4.5) is not separated, by finding two morphisms $\pi : W \to X$, $\pi' : W \to X$ whose domain of agreement is not a closed subscheme (cf. Proposition 10.1.3). (Another argument was given above, in Exercise 10.1.D. A fancy argument will be given in Exercise 12.7.C.)

We now come to the central result of this section.

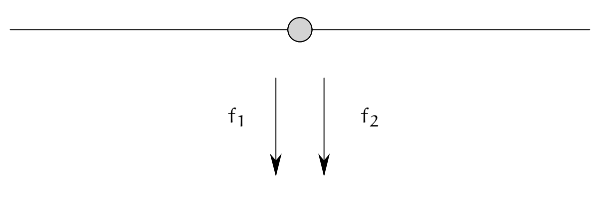
10.2.2. Reduced-to-Separated Theorem. — *Two S-morphisms $\pi : U \to Z$, $\pi' : U \to Z$ from a reduced scheme to a separated S-scheme agreeing on a dense open subset of U are the same.*

Proof. Let V be the locus where π and π' agree. It is a closed subscheme of U (by Exercise 10.2.A), which contains a dense open set. But the only closed subscheme of a reduced scheme U whose underlying set is dense is all of U. □

10.2.3. *Consequence 1.* Hence (as X is reduced and Y is separated) if we have two morphisms from open subsets of X to Y, say $\pi : U \to Y$ and $\pi' : V \to Y$, and they agree on a dense open subset $Z \subset U \cap V$, then they necessarily agree on $U \cap V$.

Consequence 2. A rational map has a largest **domain of definition** on which $\pi : U \dashrightarrow Y$ is a morphism, which is the union of all the domains of definition. In particular, a rational function on a reduced scheme has a largest domain of definition. For example, the domain of definition of $\mathbb{A}^2_k \dashrightarrow \mathbb{P}^1_k$ given by $(x, y) \mapsto [x, y]$ has domain of definition $\mathbb{A}^2_k \setminus \{(0, 0)\}$ (cf. §6.5.3). This partially extends the definition of the domain of a rational function on a locally Noetherian scheme (Definition 5.5.6). The complement of the domain of definition is called the **locus of indeterminacy**, and its points are sometimes called **fundamental points** of the rational map, although we won't use these phrases. (We will see in Exercise 22.4.L that a rational map to a projective scheme can be upgraded to an honest morphism by "blowing up" a scheme-theoretic version of the locus of indeterminacy.)

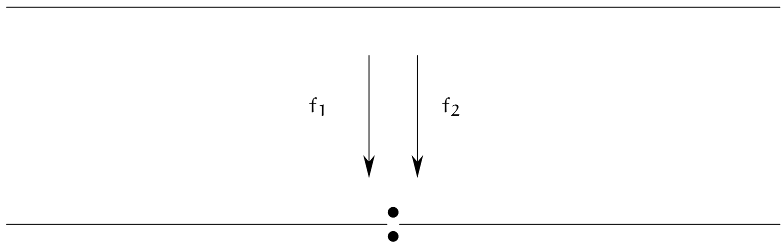
10.2.D. EXERCISE. Show that the Reduced-to-Separated Theorem 10.2.2 is false if we give up reducedness of the source or separatedness of the target. Here are some possibilities. For the first, consider the two maps from $\operatorname{Spec} k[x, y]/(y^2, xy)$ to $\operatorname{Spec} k[t]$, where we take π given by $t \mapsto x$ and π' given by $t \mapsto x + y$; f_1 and f_2 agree on the distinguished open set $D(x)$, see Figure 10.4. For the second, consider the two maps from $\operatorname{Spec} k[t]$ to the line with the doubled origin, one of which maps to the "upper half", and one of which maps to the "lower half". These two morphisms agree on the dense open set $D(t)$, see Figure 10.5.



[change f_1 and f_2 to π and π']

FIGURE 10.4. Two different maps from a nonreduced scheme agreeing on a dense open set

10.2.4. *Graphs of rational maps.* (Graphs of *morphisms* were defined in §10.1.17.) If X is reduced and Y is separated, define the **graph Γ_π of a rational map** $\pi : X \dashrightarrow Y$ as follows. Let (U, π') be any representative of this rational map (so $\pi' : U \to Y$ is a morphism). Let Γ_π be the scheme-theoretic closure of $\Gamma_{\pi'} \hookrightarrow U \times Y \hookrightarrow X \times Y$, where the first map is a closed embedding (Proposition 10.1.18), and the second is an open embedding. The product here should be taken in the category you are



[change f_1 and f_2 to π and π']

FIGURE 10.5. Two different maps to a nonseparated scheme agreeing on a dense open set

working in. For example, if you are working with k-schemes, the fibered product should be taken over k.

10.2.E. EXERCISE. Show that the graph of a rational map π is independent of the choice of representative of π. Hint: Suppose $\xi' : U \to Y$ and $\xi : V \to Y$ are two representatives of π. Reduce to the case where V is the domain of definition of π (§10.2.3), and $\xi' = \xi|_U$. Reduce to the case $V = X$. Show an isomorphism $\Gamma_\pi \cong X$, and $\Gamma_{\xi|_U} \cong U$. Show that the scheme-theoretic closure of U in X is all of X. (Remark: the separatedness of Y is not necessary.)

In analogy with graphs of morphisms, the following diagram of a graph of a rational map π can be useful (cf. Figure 10.3).

$$\begin{tikzcd} \Gamma_\pi \arrow[r, hook, "\text{cl. emb.}"] & X \times Y \arrow[dl] \arrow[dr] & \\ X \arrow[u, dashed] \arrow[rr, dashed, "\pi"] & & Y. \end{tikzcd}$$

10.2.F. EXERCISE (THE BLOW-UP OF THE PLANE AS THE GRAPH OF A RATIONAL MAP). Consider the rational map $\mathbb{A}_k^2 \dashrightarrow \mathbb{P}_k^1$ given by $(x, y) \mapsto [x, y]$. Show that this rational map cannot be extended over the origin. (A similar argument arises in Exercise 6.5.I on the Cremona transformation.) Show that the graph of the rational map is the morphism (the blow-up) described in Exercise 9.3.F. (When we define blow-ups in general, we will see that they are often graphs of rational maps, see Exercise 22.4.M.)

10.2.5. Variations.

Variations of the short proof of Theorem 10.2.2 yield other useful results. Exercise 10.2.B is one example. The next exercise is another.

10.2.G. EXERCISE (MAPS TO A SEPARATED SCHEME CAN BE EXTENDED OVER AN EFFECTIVE CARTIER DIVISOR IN AT MOST ONE WAY). Suppose X is a Z-scheme, and Y is a separated Z-scheme. Suppose further that D is an effective Cartier divisor on X. Show that any Z-morphism $X \setminus D \to Y$ can be extended in at most one way to a Z-morphism $X \to Y$. (Hint: reduce to the case where $X = \operatorname{Spec} A$, and D is the vanishing scheme of $t \in A$. Reduce to showing that the scheme-theoretic image of $D(t)$ in X is all of X. Show this by showing that $A \to A_t$ is an inclusion.)

As noted in §6.5.2, rational maps can be defined from any X that has associated points to any Y. The Reduced-to-Separated Theorem 10.2.2 can be extended to this setting, as follows.

10.2.H. EXERCISE (THE "ASSOCIATED-TO-SEPARATED THEOREM"). Prove that two S-morphisms π and π' from a locally Noetherian scheme U to a separated S-scheme Z, agreeing on a dense open subset containing the associated points of U, are the same.

10.3 Proper morphisms

Recall that a map of topological spaces (also known as a continuous map!) is said to be *proper* if the preimage of any compact set is compact. *Properness* of morphisms is an analogous property. For example, a variety over $\mathbb{C}$ will be proper if it is compact in the classical topology (see [**Se2**, §7]). Alternatively, we will see that projective A-schemes are proper over A — so this is a nice property satisfied by projective schemes, which also is convenient to work with.

Recall (§7.3.8) that a (continuous) map of topological spaces $\pi : X \to Y$ is *closed* if for each closed subset $S \subset X$, $\pi(S)$ is also closed. A morphism of schemes is closed if the underlying continuous map is closed. We say that a morphism of schemes $\pi : X \to Y$ is **universally closed** if for every morphism $Z \to Y$, the induced morphism $Z \times_Y X \to Z$ is closed. In other words, a morphism is universally closed if it remains closed under any base change. (More generally, if P is some property of schemes, then a morphism of schemes is said to be **universally** P if it remains P under any base change.)

To motivate the definition of properness for schemes, we remark that a continuous map $\pi : X \to Y$ of locally compact Hausdorff spaces which have countable bases for their topologies is universally closed if and only if it is proper (i.e., preimages of compact subsets are compact). You are welcome to prove this as an exercise.

10.3.1. Definition. *A morphism* $\pi : X \to Y$ *is* **proper** *if it is separated, finite type, and universally closed.* A scheme X is often said to be proper if some implicit structure morphism is proper. For example, a k-scheme X is often described as proper if $X \to \operatorname{Spec} k$ is proper. (A k-scheme is often said to be **complete** if it is proper. We will not use this terminology.) If A is a ring, one often says that an A-scheme is **proper over** A if it is proper over $\operatorname{Spec} A$.

Let's try this idea out in practice. We expect that $\mathbb{A}^1_{\mathbb{C}} \to \operatorname{Spec} \mathbb{C}$ is not proper, because the complex manifold corresponding to $\mathbb{A}^1_{\mathbb{C}}$ is not compact. However, note that this map is separated (it is a map of affine schemes), finite type, and (trivially) closed. So the "universally" is what matters here.

10.3.A. EASY EXERCISE. Show that $\mathbb{A}^1_{\mathbb{C}} \to \operatorname{Spec} \mathbb{C}$ is not proper, by finding a base change that turns this into a non-closed map. (Possible hint: Consider a well-chosen map $\mathbb{A}^1_{\mathbb{C}} \times_{\mathbb{C}} \mathbb{A}^1_{\mathbb{C}} \to \mathbb{A}^1_{\mathbb{C}}$ or $\mathbb{A}^1_{\mathbb{C}} \times_{\mathbb{C}} \mathbb{P}^1_{\mathbb{C}} \to \mathbb{P}^1_{\mathbb{C}}$.)

10.3.2. *Example.* As a first example: closed embeddings are proper. They are clearly separated, as affine morphisms are separated, §10.1.12. They are finite type. After base change, they remain closed embeddings (§9.2.1), and closed embeddings are always closed. This easily extends further as follows.

10.3.3. Proposition. — *Finite morphisms are proper.*

Proof. Finite morphisms are separated (as they are affine by definition, and affine morphisms are separated, §10.1.12), and finite type (basically because finite modules over a ring are automatically finitely generated). To show that finite morphism are closed after any base change, we note that they remain finite after any base change (finiteness is preserved by base change, Exercise 9.4.B(d)), and finite morphisms are closed (Exercise 7.3.M). □

10.3.4. Proposition. —

(a) *The notion of "proper morphism" is stable under base change.*
(b) *The notion of "proper morphism" is local on the target (i.e., $\pi : X \to Y$ is proper if and only if for any affine open cover $U_i \to Y$, $\pi^{-1}(U_i) \to U_i$ is proper). Note that the "only if" direction follows from (a) — consider base change by $U_i \hookrightarrow Y$.*
(c) *The notion of "proper morphism" is closed under composition.*
(d) *The product of two proper morphisms is proper: if $\pi : X \to Y$ and $\pi' : X' \to Y'$ are proper, where all morphisms are morphisms of Z-schemes, then $\pi \times \pi' : X \times_Z X' \to Y \times_Z Y'$ is proper.*
(e) *Suppose*

(10.3.4.1)
$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "\tau"'] & & Y \arrow[dl, "\rho"] \\ & Z & \end{tikzcd}$$

is a commutative diagram, and τ is proper, and ρ is separated. Then π is proper.

A sample application of (e): a morphism (over $\operatorname{Spec} k$) from a proper k-scheme to a separated k-scheme is always proper.

Proof. (a) The notions of separatedness, finite type, and universal closedness are all preserved by fibered product. (Notice that this is why universal closedness is better than closedness — it is automatically preserved by base change!)

(b) We have already shown that the notions of separatedness and finite type are local on the target. The notion of closedness is local on the target, and hence so is the notion of universal closedness.

(c) The notions of separatedness, finite type, and universal closedness are all preserved by composition.

(d) By (a) and (c), this follows from Exercise 9.4.F.

(e) Closed embeddings are proper (Example 10.3.2), so we invoke the Cancellation Theorem 10.1.19 for proper morphisms. □

10.3.B. UNIMPORTANT EXERCISE ("IMAGE OF PROPER SCHEMES ARE PROPER"). Suppose in the diagram

$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "\tau"'] & & Y \arrow[dl, "\rho"] \\ & Z & \end{tikzcd}$$

that τ is proper and ρ is separated and finite type. Show that the scheme-theoretic image of X under π is a proper Z-scheme. (We won't use this fact, but it reassures us that properness in algebraic geometry behaves like properness in topology.)

We now come to the most important example of proper morphisms.

10.3.5. Theorem. — *Projective A-schemes are proper over A.*

(As finite morphisms to $\operatorname{Spec} A$ are projective A-schemes, Exercise 7.3.J, Theorem 10.3.5 can be used to give a second proof that finite morphisms are proper, Proposition 10.3.3.)

Proof. The structure morphism of a projective A-scheme $X \to \operatorname{Spec} A$ factors as a closed embedding followed by $\mathbb{P}^n_A \to \operatorname{Spec} A$. Closed embeddings are proper (Example 10.3.2), and compositions of proper morphisms are proper (Proposition 10.3.4), so it suffices to show that $\mathbb{P}^n_A \to \operatorname{Spec} A$ is proper. We have already seen that this morphism is finite type (Easy Exercise 5.3.J) and separated (Proposition 10.1.5), so it suffices to show that $\mathbb{P}^n_A \to \operatorname{Spec} A$ is universally closed. As $\mathbb{P}^n_A = \mathbb{P}^n_{\mathbb{Z}} \times_{\mathbb{Z}} \operatorname{Spec} A$, it suffices to show that $\mathbb{P}^n_X := \mathbb{P}^n_{\mathbb{Z}} \times_{\mathbb{Z}} X \to X$ is closed for any scheme X. But the property of being closed is local on the target on X, so by covering X with affine open subsets, it suffices to show that $\mathbb{P}^n_B \to \operatorname{Spec} B$ is closed for all rings B. This is the Fundamental Theorem of Elimination Theory (Theorem 7.4.7). $\square$

10.3.6. *Remark: "Reasonable" proper schemes are projective.* It is not easy to come up with an example of an A-scheme that is proper but not projective! Over a field, all proper curves are projective (see Remark 18.7.2), and all smooth surfaces over a field are projective (see [**Ba**, Thm. 1.28] for a proof of this theorem of Zariski; smoothness of course is not yet defined). We will meet a first example of a proper but not projective variety (a singular threefold) in §16.4.10. We will later see an example of a proper nonprojective surface in §19.11.11, and a simpler one in Exercise 20.2.G. Once we know about flatness, we will see Hironaka's example of a proper nonprojective irreducible smooth threefold over $\mathbb{C}$ (§24.7.6).

10.3.7. Functions on connected reduced proper $\overline{k}$-schemes must be constant. As an enlightening application of these ideas, we show that if X is a connected reduced proper k-scheme where $k = \overline{k}$, then $\Gamma(X, \mathscr{O}_X) = k$. The analogous fact in complex geometry uses the maximum principle. We saw this in the special case $X = \mathbb{P}^n$ in Exercise 4.4.E. This will be vastly generalized by Grothendieck's Coherence Theorem 18.9.1.

Suppose $f \in \Gamma(X, \mathscr{O}_X)$ (f is a function on X). This is the same as a map $\pi : X \to \mathbb{A}^1_k$ (Exercise 6.3.F, discussed further in §6.6.1). Let π' be the composition of π with the open embedding $\mathbb{A}^1 \hookrightarrow \mathbb{P}^1$. By Proposition 10.3.4(e), π' is proper, and in particular closed. As X is irreducible, the image of π' is as well. Thus the image of π' must be either a closed point, or all of $\mathbb{P}^1$. But the image of π' lies in $\mathbb{A}^1$, so it must be a closed point p (which we identify with an element of k).

By Corollary 8.3.5, the support of the scheme-theoretic image of π is the closed point p. By Exercise 8.3.A, the scheme-theoretic image is precisely p (with the reduced structure). Thus π can be interpreted as the structure map to $\operatorname{Spec} k$, followed by a closed embedding to $\mathbb{A}^1$ identifying $\operatorname{Spec} k$ with p. You should be able to verify that this is the map to $\mathbb{A}^1$ corresponding to the constant function $f = p$.

(What are counterexamples if different hypotheses are relaxed?)

10.3.8. Facts (not yet proved) that may help you correctly think about finiteness. The following facts may shed some light on the notion of finiteness. We will prove them later.

A morphism is finite if and only if it is proper and affine, if and only if it is proper and quasifinite (Theorem 29.6.2). We have proved parts of this statement, but we will only finish the proof once we know Zariski's Main Theorem, cf. §7.3.15).

As an application: quasifinite morphisms from proper schemes to separated schemes are finite. Here is why: suppose $\pi : X \to Y$ is a quasifinite morphism over Z, where X is proper over Z. Then by the Cancellation Theorem 10.1.19 for proper morphisms, $X \to Y$ is proper. Hence as π is quasifinite and proper, π is finite.

As an explicit example, consider the map $\pi : \mathbb{P}^1_k \to \mathbb{P}^1_k$ given by $[x, y] \mapsto [f(x, y), g(x, y)]$, where f and g are homogeneous polynomials of the same degree with no common roots in $\mathbb{P}^1$. The fibers are finite, and π is proper (from the Cancellation Theorem 10.1.19 for proper morphisms, as discussed after the statement of Theorem 10.3.4), so π is finite. This could be checked directly as well, but now we can save ourselves the annoyance.

10.3.9. $\star\star$ **Group varieties.** For the rest of this section, *we work in the category of varieties over a field* k. We briefly discuss group varieties, mainly because we can, not because we have anything profound to say. We discuss them right now only because properness gives an unexpected consequence, that proper group varieties are abelian (thanks to the surprisingly simple Rigidity Lemma 10.3.12).

As discussed in §10.1.7, now that we have the category of varieties over k, we immediately have the notion of a **group variety** (thanks to §6.6.4 on group objects in any category). An **algebraic group** (over k) is a smooth group variety. Examples include GL_n (Exercise 6.6.N(b)), which includes $\mathbb{G}_m$ as a special case, and SL_n (Exercise 6.6.Q).

10.3.10. *Side Remarks (that we won't prove or use).* Group varieties are automatically smooth (hence algebraic groups) in characteristic 0. Group varieties are not necessarily smooth in positive characteristic. (Let k be an imperfect field of characteristic p, and consider the "closed sub-group scheme" $x^p = ty^p$ of the "additive group" $\mathbb{G}_a^2 = \operatorname{Spec} k[x, y]$. This group scheme is reduced, but not geometrically reduced; and not smooth, or even regular.) Algebraic groups are automatically quasiprojective. An algebraic group G is affine (as a scheme) if and only if it admits a closed immersion into GL_n for some $n \geq 0$, such that $G \to GL_n$ is a homomorphism of group schemes (Exercise 6.6.N(a)). For further discussion on these and other issues, see [**P2**, §5.2].

10.3.11. *Definition: Abelian varieties.* We can now define one of the most important classes of varieties: abelian varieties. The bad news is that we have no real example yet. We will later see elliptic curves as an example (§19.10), but we will not meet anything more exotic than that. Still, it is pleasant to know that we can make the definition this early.

An **abelian variety** over a field k is an algebraic group that is geometrically integral and projective (over k). It turns out that the analytification of any abelian varieties over $\mathbb{C}$ is a complex torus, i.e., $\mathbb{C}^n/\Lambda$, where Λ is a lattice (of rank $2n$). (The key idea: connected compact complex Lie groups are commutative, see Remark 10.3.14; the universal cover is a simply connected commutative complex Lie group, and thus $\mathbb{C}^n$. See [**BL**, Lem. 1.1.1].)

We now show that abelian varieties are abelian, i.e., abelian varieties are commutative algebraic group varieties. (This is less tautological than it sounds! The adjective "abelian" is used in two different senses; what they have in common is that they are derived from the name of Niels Henrik Abel.) As algebraic groups need not be commutative (witness GL_n), it is somehow surprising that commutativity could be forced by "compactness".

They key result is the following fact, beautiful in its own right.

10.3.12. Rigidity Lemma. — *Let X, Y, and Z be varieties, where X is proper and geometrically integral, with a rational point p, and Y is irreducible. Let $\mathrm{pr} : X \times Y \to Y$ be the projection, and let $\alpha : X \times Y \to Z$ be any morphism. Suppose $q \in Y$ is such that α is constant on $X \times \{q\}$; say $\alpha(X \times \{q\}) = r \in Z$.*
(a) Then there is a morphism $\psi : Y \to Z$ such that $\alpha = \psi \circ \mathrm{pr}$. (In particular, for every $q' \in Y$, α is constant on $X \times \{q'\}$.)
(b) If α is also constant on $\{p\} \times Y$, then α is a constant function on all of $X \times Z$.

$$\begin{tikzcd} & & Z \\ X_q \arrow[r] \arrow[d] & X \times Y \arrow[ur, "\alpha"] \arrow[d, shift right, "\mathrm{pr}"'] & \\ q \arrow[r] & Y \arrow[u, shift right, bend right, "\sigma_p"'] \arrow[uur, bend right, "\psi"'] & \end{tikzcd}$$

The hypotheses can be relaxed in a number of ways, but this version suffices for our purposes. If you have not read §9.5, where geometric integrality is discussed, you need not worry too much; this is automatic if $k = \overline{k}$ and X is integral (Hard Fact 9.5.2). Even if you *have* read §9.5, it is helpful to first read this proof under the assumption that $k = \overline{k}$, to avoid being distracted from the main idea by geometric points. But of course all of §10.3.9 is double-starred, so you shouldn't be reading this anyway.

10.3.C. EXERCISE. Show that the properness hypothesis on X is necessary. (Can you make this accord with your intuition?)

Proof of the Rigidity Lemma 10.3.12. Define $\beta : X \times Y \to Z$ by $\beta(x, y) = \alpha(p, y)$. (More precisely, if σ_p is the section of $Y \to X \times Y$ pulled back from $\operatorname{Spec} k \mapsto p$, then $\beta = \alpha \circ \sigma_p \circ \mathrm{pr}$.) We will show that $\beta(x, y) = \alpha(x, y)$. This will imply (a) (take $\psi = \alpha \circ \sigma_p$), which in turn immediately implies (b).

Proposition 9.5.17 (geometrically reduced times reduced is reduced) implies that $X \times Y$ is reduced. Exercise 9.5.M (geometrically irreducible times irreducible is irreducible) implies that $X \times Y$ is irreducible. As Z is separated, by Exercise 10.2.A, it suffices to show that $\alpha = \beta$ on a nonempty (hence dense) open subset of $X \times Y$.

Let $U \subset Z$ be an affine neighborhood of r. Then $\alpha^{-1}(Z \setminus U)$ is a closed subset of $X \times Y$. As the projection pr is proper (using properness of X, and preservation of properness under base change, Proposition 10.3.4(a)), we have $\mathrm{pr}(\alpha^{-1}(Z \setminus U))$ is closed. Its complement is open, and contains q; let $V \subset Y$ be a neighborhood of q disjoint from $\mathrm{pr}(\alpha^{-1}(Z \setminus U))$.

10.3.D. Exercise. Suppose $\gamma : \operatorname{Spec} \overline{K} \to V$ is a geometric point of V. Show that the fiber $X_{\overline{K}}$ over $\operatorname{Spec} \overline{K}$ (or more precisely, over γ) is mapped to a point in Z by the restriction $\alpha_{\overline{K}}$ of α to $X_{\overline{K}}$. Hint: review §10.3.7.

$$\begin{tikzcd}
 & & Z \\
X_{\overline{K}} \arrow[r] \arrow[urr, shift left, "\alpha_{\overline{K}}"] \arrow[urr, shift right, "\beta_{\overline{K}}"'] \arrow[d] & X \times V \arrow[ur, bend right, "\alpha|_V"'] \arrow[d, "\mathrm{pr}"'] & \\
\operatorname{Spec} \overline{K} \arrow[r, "\gamma"] & V &
\end{tikzcd}$$

By Exercise 10.2.B (maps of varieties over an algebraically closed field are determined by the map of closed points), $\alpha_{\overline{K}} = \beta_{\overline{K}}$.

10.3.E. Exercise. Use Exercise 10.2.A (the fact that the locus where $\alpha = \beta$ is a closed subscheme — and represents a certain functor) to show that $\alpha = \beta$ on $\mathrm{pr}^{-1}(V)$.

$\square$

10.3.13. Corollary. — *Suppose A is an abelian variety, G is a group variety, and $\phi : A \to G$ is any morphism of varieties. Then ϕ is the composition of a translation and a homomorphism.*

Proof. Let e_A and e_G be the identity points of A and G respectively. Composing ϕ with a translation, we may assume that $\phi(e_A) = e_G$. Consider the morphism $\alpha : A \times A \to G$ given by $\alpha(a_1, a_2) = \phi(a_1 a_2)\phi(a_1)^{-1}\phi(a_2)^{-1}$. Then $\alpha(\{e_A\} \times A) = \alpha(A \times \{e_A\}) = \{e_B\}$, so by the Rigidity Lemma 10.3.12, α is a constant, and sends $A \times A$ to e_B. Thus $\phi(a_1 a_2) = \phi(a_1)\phi(a_2)$, so ϕ is a homomorphism. $\square$

10.3.F. Exercise (abelian varieties are abelian). Show that an abelian variety is an abelian group variety. Hint: apply Corollary 10.3.13 to the inversion morphism $i : A \to A$.

10.3.14. *Remark.* A similar idea is used to show that connected compact complex Lie groups are abelian, see for example [**BL**, Lemm. 1.1.1].

Part IV

“Geometric” properties: Dimension and smoothness

CHAPTER 11

Dimension

11.1 Dimension and codimension

Everyone knows what a curve is, until he has studied enough mathematics to become confused ...
— *F. Klein*, [**RS**, p. 90]

At this point, you know a fair bit about schemes, but there are some fundamental notions you cannot yet define. In particular, you cannot use the phrase "smooth surface", as it involves the notion of dimension and of smoothness. You may be surprised that we have gotten so far without using these ideas. You may also be disturbed to find that these notions can be subtle, but you should keep in mind that they are subtle in all parts of mathematics.

In this chapter, we will address the first notion, that of dimension of schemes. This should agree with, and generalize, our geometric intuition. Although we think of dimension as a basic notion in geometry, it is a slippery concept, as it is throughout mathematics. Even in linear algebra, the definition of dimension of a vector space is surprising the first time you see it, even though it quickly becomes second nature. The definition of dimension for manifolds is equally nontrivial. For example, how do we know that there isn't an isomorphism between some 2-dimensional manifold and some 3-dimensional manifold? Your answer will likely use topology, and hence you should not be surprised that the notion of dimension is often quite topological in nature.

A caution for those thinking over the complex numbers: our dimensions will be algebraic, and hence half that of the "real" picture. For example, we will see very shortly that $\mathbb{A}^1_{\mathbb{C}}$, which you may picture as the complex numbers (plus one generic point), has dimension 1.

11.1.1. *Definition(s): dimension.* Surprisingly, the right definition is purely topological — it just depends on the topological space, and not on the structure sheaf. We define the **dimension** of a topological space X (denoted $\dim X$) as the supremum of lengths of chains of closed irreducible sets, starting the indexing of the closed irreducible sets with 0. (The dimension may be infinite.) Scholars of the empty set can take the dimension of the empty set to be $-\infty$. (An analogy from linear algebra: the dimension of a vector space is the supremum of the length of chains of subspaces.) Define the **dimension** of a ring as the supremum of the lengths of the chains of nested prime ideals (where indexing starts at zero). These two definitions of dimension are sometimes called **Krull dimension**. (You might think a

Noetherian ring has finite dimension because all chains of prime ideals are finite, but this isn't necessarily true — see Exercise 11.1.K.)

11.1.A. EASY EXERCISE. Show that $\dim \operatorname{Spec} A = \dim A$. (Hint: Exercise 3.7.E gives a bijection between irreducible closed subsets of $\operatorname{Spec} A$ and prime ideals of A. It is "inclusion-reversing".)

The homeomorphism between $\operatorname{Spec} A$ and $\operatorname{Spec} A/\mathfrak{N}(A)$ (§3.4.5: the Zariski topology disregards nilpotents) implies that $\dim \operatorname{Spec} A = \dim \operatorname{Spec} A/\mathfrak{N}(A)$.

11.1.2. *Examples.* We have identified all the prime ideals of $k[t]$ (they are 0, and $(f(t))$ for irreducible polynomials $f(t)$), $\mathbb{Z}$ ((0) and (p)), k (only (0)), and $k[x]/(x^2)$ (only (x)), so we can quickly check that $\dim \mathbb{A}^1_k = \dim \operatorname{Spec} \mathbb{Z} = 1$, $\dim \operatorname{Spec} k = 0$, $\dim \operatorname{Spec} k[x]/(x^2) = 0$.

11.1.3. We must be careful with the notion of dimension for reducible spaces. If Z is the union of two closed subsets X and Y, then $\dim Z = \max(\dim X, \dim Y)$. Thus dimension is not a "local" characteristic of a space. This sometimes bothers us, so we try to only talk about dimensions of irreducible topological spaces. We say a topological space is **equidimensional** or **pure dimensional** (resp. equidimensional of dimension n or pure dimension n) if each of its irreducible components has the same dimension (resp. they are all of dimension n). An equidimensional dimension 1 (resp. 2, n) topological space is said to be a **curve** (resp. **surface**, **n-fold**).

11.1.B. EXERCISE. Show that a scheme has dimension n if and only if it admits an open cover by affine open subsets of dimension at most n, where equality is achieved for some affine open. Hint: You may find it helpful, here and later, to show the following. For any topological space X and open subset $U \subset X$, there is a bijection between irreducible closed subsets of U, and irreducible closed subsets of X that meet U. (What is this bijection?)

11.1.C. EASY EXERCISE. Show that a Noetherian scheme of dimension 0 has a finite number of points.

11.1.D. EXERCISE (FIBERS OF INTEGRAL MORPHISMS, PROMISED IN §7.3.11). Suppose $\pi : X \to Y$ is an integral morphism. Show that every (nonempty) fiber of π has dimension 0. Hint: As integral morphisms are preserved by base change, we assume that $Y = \operatorname{Spec} k$. Hence we must show that if $\phi : k \to A$ is an integral extension, then $\dim A = 0$. Outline of proof: Suppose $\mathfrak{p} \subset \mathfrak{m}$ are two prime ideals of A. Mod out by $\mathfrak{p}$, so we can assume that A is a domain. I claim that any nonzero element is invertible: Say $x \in A$, and $x \neq 0$. Then the minimal monic polynomial for x has nonzero constant term. But then x is invertible — recall the coefficients are in a field.

11.1.E. IMPORTANT EXERCISE. Show that if $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ corresponds to an integral *extension* of rings, then $\dim \operatorname{Spec} A = \dim \operatorname{Spec} B$. Hint: show that a chain of prime ideals downstairs gives a chain upstairs of the same length, by the Going-up Theorem (Exercise 7.2.F). Conversely, a chain upstairs gives a chain downstairs. Use Exercise 11.1.D to show that no two elements of the chain upstairs go to the same element $[\mathfrak{q}] \in \operatorname{Spec} B$ of the chain downstairs.

11.1.F. EXERCISE. Show that if $\nu : \widetilde{X} \to X$ is the normalization of a scheme (possibly in a finite field extension), then $\dim \widetilde{X} = \dim X$.

11.1.G. EXERCISE. Suppose X is an affine k-scheme, and K/k is an *algebraic* field extension.
(a) Suppose X has pure dimension n. Show that $X_K := X \times_k K$ has pure dimension n. (See Exercise 24.5.F for a generalization, which for example removes the affine hypothesis. Also, see Exercise 11.2.I and Remark 11.2.16 for the fate of possible generalizations to arbitrary field extensions.) Hint: If $X = \operatorname{Spec} A$, reduce to the case where A is an integral domain. An irreducible component of X' corresponds to a minimal prime $\mathfrak{p}$ of $A' := A \otimes_k K$. Suppose $a \in \ker(A \to A'/\mathfrak{p})$. Show that $a = 0$, using the fact that a lies in a minimal prime $\mathfrak{p}$ of A' (and is hence a zerodivisor, by Remark 5.5.12), and A' is a free A-module (so multiplication in A' by $a \in A$ is injective if a is nonzero). Thus $A \to A'/\mathfrak{p}$ is injective. Then use Exercise 11.1.E.
(b) Prove the converse to (a): show that if X_K has pure dimension n, then X has pure dimension n.

11.1.H. EXERCISE. Show that $\dim \mathbb{Z}[x] = 2$. (Hint: The primes of $\mathbb{Z}[x]$ were implicitly determined in Exercise 3.2.Q.)

11.1.4. Codimension. Because dimension behaves oddly for disjoint unions, we need some care when defining codimension, and in using the phrase. For example, if Y is a closed subset of X, we might define the codimension to be $\dim X - \dim Y$, but this behaves badly. For example, if X is the disjoint union of a point Y and a curve Z, then $\dim X - \dim Y = 1$, but this has nothing to do with the local behavior of X near Y.

A better definition is as follows. In order to avoid excessive pathology, we define the codimension of Y in X *only when Y is irreducible*. (Use extreme caution when using this word in any other setting.) Define the **codimension of an irreducible subset** $Y \subset X$ of a topological space as the supremum of lengths of *increasing* chains of irreducible closed subsets starting with $\overline{Y}$ (where indexing starts at 0 — recall that the closure of an irreducible set is irreducible, Exercise 3.6.B(b)). In particular, the **codimension of a point** is the codimension of its closure. The codimension of Y in X is denoted by $\operatorname{codim}_X Y$.

We say that a prime ideal $\mathfrak{p}$ in a ring has **codimension** equal to the supremum of lengths of the chains of decreasing prime ideals starting at $\mathfrak{p}$, with indexing starting at 0. Thus in an integral domain, the ideal (0) has codimension 0; and in $\mathbb{Z}$, the ideal (23) has codimension 1. Note that the codimension of the prime ideal $\mathfrak{p}$ in A is $\dim A_{\mathfrak{p}}$ (see §3.2.8). (This notion is often called **height**.) Thus the codimension of $\mathfrak{p}$ in A is the codimension of $[\mathfrak{p}]$ in $\operatorname{Spec} A$.

(Continuing an analogy with linear algebra: the codimension of a vector subspace $Y \subset X$ is the supremum of lengths of increasing chains of subspaces starting with Y. This is a better definition than $\dim X - \dim Y$, because it works even when $\dim X = \infty$. You might prefer to define $\operatorname{codim}_X Y$ as $\dim(X/Y)$; that is analogous to defining the codimension of $\mathfrak{p}$ in A as the dimension of $A_{\mathfrak{p}}$ — see the previous paragraph.)

11.1.I. EXERCISE. Show that if Y is an irreducible closed subset of a scheme X, and η is the generic point of Y, then the codimension of Y is the dimension of the local ring $\mathscr{O}_{X,\eta}$ (cf. §3.2.8).

Notice that Y is codimension 0 in X if it is an irreducible component of X. Similarly, Y is codimension 1 if it is not an irreducible component, and for every irreducible component Y' it is contained in, there is no irreducible subset strictly between Y and Y'. (See Figure 11.1 for examples.) A closed subset all of whose irreducible components are codimension 1 in some ambient space X is said to be a **hypersurface** in X.

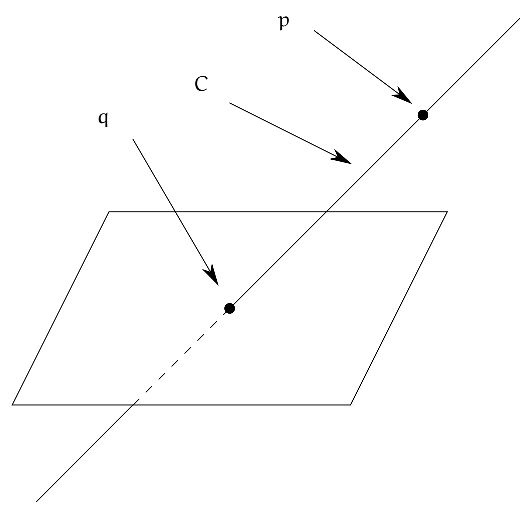


FIGURE 11.1. Behavior of codimension

11.1.J. EASY EXERCISE. If Y is an irreducible closed subset of a scheme X, show that

$$\operatorname{codim}_X Y + \dim Y \leq \dim X. \quad (11.1.4.1)$$

We will soon see that equality always holds if X and Y are varieties (Theorem 11.2.9), but equality doesn't hold in general (§11.3.8).

Warning. The notion of codimension still can behave slightly oddly. For example, consider Figure 11.1. (You should think of this as an intuitive sketch.) Here the total space X has dimension 2, but point p is dimension 0, and codimension 1. We also have an example of a codimension 2 subset q contained in a codimension 0 subset C with no codimension 1 subset "in between".

Worse things can happen; we will soon see an example of a closed point in an *irreducible* surface that is nonetheless codimension 1, not 2, in §11.3.8. However, for irreducible *varieties* this can't happen, and inequality (11.1.4.1) must be an equality (Theorem 11.2.9).

11.1.5. In unique factorization domains, codimension 1 primes are principal. For the sake of further applications, we make a short observation.

11.1.6. Lemma. — *In a unique factorization domain A, all codimension 1 prime ideals are principal.*

This is a first glimpse of the fact that codimension one is rather special — this theme will continue in §11.3. We will see that the converse of Lemma 11.1.6 holds as well (when A is a Noetherian integral domain, Proposition 11.3.5).

Proof. Suppose $\mathfrak{p}$ is a codimension 1 prime. Choose any $f \neq 0$ in $\mathfrak{p}$, and let g be any irreducible/prime factor of f that is in $\mathfrak{p}$ (there is at least one). Then (g) is a nonzero prime ideal contained in $\mathfrak{p}$, so $(0) \subset (g) \subset \mathfrak{p}$. As $\mathfrak{p}$ is codimension 1, we must have $\mathfrak{p} = (g)$, and thus $\mathfrak{p}$ is principal. $\square$

11.1.7. A fun but unimportant counterexample. We end this introductory section with a fun pathology. As a Noetherian ring has no infinite chain of prime ideals, you may think that Noetherian rings must have finite dimension. Nagata, the master of counterexamples, shows you otherwise with the following example.

11.1.K. ⋆⋆ EXERCISE: AN INFINITE-DIMENSIONAL NOETHERIAN RING. Let $A = k[x_1, x_2, \dots]$. Choose an increasing sequence of positive integers $m_1, m_2, \dots$ whose *differences* are also increasing ($m_{i+1} - m_i > m_i - m_{i-1}$). Let $\mathfrak{p}_i = (x_{m_i+1}, \dots, x_{m_{i+1}})$ and $S = A - \cup_i \mathfrak{p}_i$.
(a) Show that S is a multiplicative set.
(b) Show that each $S^{-1}\mathfrak{p}$ in $S^{-1}A$ is the largest prime ideal in a chain of prime ideals of length $m_{i+1} - m_i$. Hence conclude that $\dim S^{-1}A = \infty$.
(c) Suppose B is a ring such that (i) for every maximal ideal $\mathfrak{m}$, $B_{\mathfrak{m}}$ is Noetherian, and (ii) every nonzero $b \in B$ is contained in finitely many maximal ideals. Show that B is Noetherian. (One possible approach: show that for any $x_1, x_2, \dots$, $(x_1, x_2, \dots)$ is finitely generated.)
(d) Use (c) to show that $S^{-1}A$ is Noetherian.

11.1.8. *Remark: Noetherian local rings have finite dimension.* However, we shall see in Exercise 11.3.K(a) that Noetherian *local* rings always have finite dimension. (This requires a surprisingly hard fact, Krull's Height Theorem 11.3.7.) Thus points of locally Noetherian schemes always have finite codimension.

11.2 Dimension, transcendence degree, and Noether normalization

We now give a powerful alternative interpretation for dimension for irreducible varieties, in terms of transcendence degree. The proof will involve a classical construction, *Noether normalization*, which will be useful in other ways as well. In case you haven't seen transcendence theory, here is a lightning introduction.

11.2.A. EXERCISE/DEFINITION. Recall that an *element* of a field extension E/F is *algebraic over* F if it is integral over F. A *field extension* E/F is an *algebraic extension* if it is an integral extension (if all elements of E are algebraic over F). The composition of two algebraic extensions is algebraic, by Exercise 7.2.C. If E/F is a field extension, and F' and F'' are two intermediate field extensions, then we write $F' \sim F''$ if $F'F''$

is algebraic over both F' and F''. Here $F'F''$ is the *compositum* of F' and F'', the smallest field extension in E containing F' and F''. *(a) Show that $\sim$ is an equivalence relation on subextensions of* E/F. A **transcendence basis** of E/F is a set of elements $\{x_i\}$ that are algebraically independent over F (there is no nontrivial polynomial relation among the x_i with coefficients in F) such that $F(\{x_i\}) \sim E$. *(b) Show that if* E/F *has two transcendence bases, and one has cardinality* n, *then both have cardinality* n. (Hint: show that you can substitute elements from the one basis into the other one at a time.) The size of any transcendence basis is called the **transcendence degree** (which may be ∞), and is denoted $\operatorname{tr.deg}$. Any finitely generated field extension necessarily has finite transcendence degree. (Remark: A related result was mentioned in Algebraic Fact 9.5.16.)

11.2.1. Theorem (dimension = transcendence degree). — *Suppose* A *is a* **finitely generated domain over a field** k *(i.e., a finitely generated* k*-algebra that is an integral domain). Then* $\dim \operatorname{Spec} A = \operatorname{tr.deg} K(A)/k$. *Hence if* X *is an irreducible* k*-variety, then* $\dim X = \operatorname{tr.deg} K(X)/k$.

We will prove Theorem 11.2.1 shortly (§11.2.7). We first show that it is useful by giving some immediate consequences. We seem to have immediately $\dim \mathbb{A}^n_k = n$. However, our proof of Theorem 11.2.1 will go *through* this fact, so it isn't really a consequence.

A more substantive consequence is the following. If X is an irreducible k-variety, then $\dim X$ is the transcendence degree of the function field (the stalk at the generic point) $\mathscr{O}_{X,\eta}$ over k. Thus (as the generic point lies in all nonempty open sets) the dimension can be computed in any open set of X. (Warning: this is false without the finite type hypothesis, even in quite reasonable circumstances: let X be the two-point space $\operatorname{Spec} k[x]_{(x)}$, and let U consist of only the generic point, see Exercise 3.4.K.)

Another consequence is a second proof of the Nullstellensatz 3.2.5.

11.2.B. Exercise: the Nullstellensatz from dimension theory. Suppose $A = k[x_1, \dots, x_n]/I$. Show that the residue field of any maximal ideal of A is a finite extension of k. (Hint: the maximal ideals correspond to dimension 0 points, which correspond to transcendence degree 0 extensions of k, i.e., finite extensions of k.)

Yet another consequence is geometrically believable.

11.2.C. Exercise. If $\pi : X \to Y$ is a dominant morphism of irreducible k-varieties, then $\dim X \geq \dim Y$. (This is false more generally: consider the inclusion of the generic point into an irreducible curve.)

11.2.D. Exercise (practice with the concept). Show that the equations $wz - xy = 0$, $wy - x^2 = 0$, $xz - y^2 = 0$ cut out an integral *surface* S in $\mathbb{A}^4_k$. (You may recognize these equations from Exercises 3.6.F and 8.2.A.) You might expect S to be a curve, because it is cut out by three equations in four-space. One of many ways to proceed: cut S into pieces. For example, show that $D(w) \cong \operatorname{Spec} k[x, w]_w$. (You may recognize S as the affine cone over the twisted cubic. The twisted cubic was defined in Exercise 8.2.A.) It turns out that you need three equations to cut out this surface. The first equation cuts out a threefold in $\mathbb{A}^4_k$ (by Krull's Principal Ideal

Theorem 11.3.3, which we will meet soon). The second equation cuts out a surface: our surface, along with another surface. The third equation cuts out our surface, and removes the "extraneous component". One last aside: notice once again that the cone over the quadric surface $k[w,x,y,z]/(wz-xy)$ makes an appearance.)

11.2.2. *Definition: degree of a rational map of irreducible varieties.* If $\pi : X \dashrightarrow Y$ is a dominant rational map of irreducible (hence integral) k-varieties of the same dimension, the degree of the field extension is called the **degree** of the rational map. This readily extends if X is reducible: we add up the degrees on each of the components of X. If π is a rational map of integral affine k-varieties of the same dimension that is *not* dominant we say the degree is 0. We will interpret this degree in terms of counting preimages of general points of Y in §s:intdeggenfin. Note that degree is multiplicative under composition: if $\rho : Y \dashrightarrow Z$ is a rational map of integral k-varieties of the same dimension, then $\deg(\rho\circ\pi) = \deg(\rho)\deg(\pi)$, as degrees of field extensions are multiplicative in towers.

11.2.3. Noether Normalization.

Our proof of Theorem 11.2.1 will use another important classical notion, Noether Normalization.

11.2.4. Noether Normalization Lemma. — *Suppose A is an integral domain, finitely generated over a field k. If $\operatorname{tr.deg}_k K(A) = n$, then there are elements $x_1, \dots, x_n \in A$, algebraically independent over k, such that A is a finite (hence integral by Corollary 7.2.2) extension of $k[x_1, \dots, x_n]$.*

The geometric content behind this result is that given any integral affine k-scheme X, we can find a surjective finite morphism $X \to \mathbb{A}^n_k$, where n is the transcendence degree of the function field of X (over k). Surjectivity follows from the Lying Over Theorem 7.2.5, in particular Exercise 11.1.E. This interpretation is sometimes called *geometric Noether Normalization*.

11.2.5. *Nagata's proof of the Noether Normalization Lemma 11.2.4.* Suppose we can write $A = k[y_1, \dots, y_m]/\mathfrak{p}$, i.e., that A can be chosen to have m generators. Note that $m \geq n$. We show the result by induction on m. The base case $m = n$ is immediate.

Assume now that $m > n$, and that we have proved the result for smaller m. We will find $m-1$ elements $z_1, \dots, z_{m-1}$ of A such that A is finite over $k[z_1, \dots, z_{m-1}]$ (i.e., the subring of A generated by $z_1, \dots, z_{m-1}$). Then by the inductive hypothesis, $k[z_1, \dots, z_{m-1}]$ is finite over some $k[x_1, \dots, x_n]$, and A is finite over $k[z_1, \dots, z_{m-1}]$, so by Exercise 7.3.I, A is finite over $k[x_1, \dots, x_n]$.

$$\begin{tikzcd}
A \arrow[d, dash, "\text{finite}"] \\
k[z_1, \dots, z_{m-1}] \arrow[d, dash, "\text{finite}"] \\
k[x_1, \dots, x_n]
\end{tikzcd}$$

As $y_1, \dots, y_m$ are algebraically dependent in A, there is some nonzero algebraic relation $f(y_1, \dots, y_m) = 0$ among them (where f is a polynomial in m variables over k).

Let $z_1 = y_1 - y_m^{r_1}$, $z_2 = y_2 - y_m^{r_2}$, …, $z_{m-1} = y_{m-1} - y_m^{r_{m-1}}$, where r_1, …, r_{m-1} are positive integers to be chosen shortly. Then

$$f(z_1 + y_m^{r_1}, z_2 + y_m^{r_2}, \dots, z_{m-1} + y_m^{r_{m-1}}, y_m) = 0.$$

Then upon expanding this out, each monomial in f (as a polynomial in m variables) will yield a single term that is a constant times a power of y_m (with no z_i factors). By choosing the r_i so that $0 \ll r_1 \ll r_2 \ll \cdots \ll r_{m-1}$, we can ensure that the powers of y_m appearing are all distinct, and so that in particular there is a leading term y_m^N, and all other terms (including those with factors of z_i) are of smaller degree in y_m. Thus we have described an integral dependence of y_m on z_1, …, z_{m-1} as desired. □

11.2.6. *The geometry behind Nagata's proof.* Here is the geometric intuition behind Nagata's argument. Suppose we have an m-dimensional variety in $\mathbb{A}_k^n$ with $m < n$, for example $xy = 1$ in $\mathbb{A}^2$. One approach is to hope the projection to a hyperplane is a finite morphism. In the case of $xy = 1$, if we projected to the x-axis, it wouldn't be finite, roughly speaking because the asymptote $x = 0$ prevents the map from being closed (cf. Exercise 7.3.L). If we instead projected to a random line, we might hope that we would get rid of this problem, and indeed we usually can: this problem arises for only a finite number of directions. But we might have a problem if the field were finite: perhaps the finite number of directions in which to project each have a problem. (You can show that if k is an infinite field, then the substitution in the above proof $z_i = y_i - y_m^{r_i}$ can be replaced by the linear substitution $z_i = y_i - a_i y_m$ where $a_i \in k$, and that for a nonempty Zariski-open choice of a_i, we indeed obtain a finite morphism.) Nagata's trick in general is to "jiggle" the variables in a nonlinear way, and this jiggling kills the nonfiniteness of the map.

11.2.E. EXERCISE (DIMENSION IS ADDITIVE FOR FIBERED PRODUCTS OF FINITE TYPE k-SCHEMES). If X and Y are irreducible k-varieties, show that $\dim X \times_k Y = \dim X + \dim Y$. (Hint: If we had surjective finite morphisms $X \to \mathbb{A}_k^{\dim X}$ and $Y \to \mathbb{A}_k^{\dim Y}$, we could construct a surjective finite morphism $X \times_k Y \to \mathbb{A}_k^{\dim X + \dim Y}$.)

11.2.7. *Proof of Theorem 11.2.1 on dimension and transcendence degree.* Suppose X is an integral affine k-scheme. We show that $\dim X$ equals the transcendence degree n of its function field, by induction on n. (The idea is that we reduce from X to $\mathbb{A}^n$ to a hypersurface in $\mathbb{A}^n$ to $\mathbb{A}^{n-1}$.) Assume the result is known for all transcendence degrees less than n.

By Noether normalization, there exists a surjective finite morphism $X \to \mathbb{A}_k^n$. By Exercise 11.1.E, $\dim X = \dim \mathbb{A}_k^n$. If $n = 0$, we are done, as $\dim \mathbb{A}_k^0 = 0$.

We now show that $\dim \mathbb{A}_k^n = n$ for $n > 0$, by induction. Clearly $\dim \mathbb{A}_k^n \geq n$, as we can describe a chain of irreducible subsets of length n: if x_1, …, x_n are coordinates on $\mathbb{A}^n$, consider the chain of ideals

$$(0) \subset (x_1) \subset \cdots \subset (x_1, \dots, x_n)$$

in $k[x_1, \dots, x_n]$. Suppose we have a chain of prime ideals of length at least $n+1$:

$$(0) = \mathfrak{p}_0 \subset \cdots \subset \mathfrak{p}_m.$$

Choose any nonzero element g of $\mathfrak{p}_1$, and let f be any irreducible factor of g. Then replace $\mathfrak{p}_1$ by (f). (Of course, $\mathfrak{p}_1$ may have been (f) to begin with...) Then $K(k[x_1, \dots, x_n]/(f(x_1, \dots, x_n)))$ has transcendence degree $n-1$, so by induction,

$$\dim k[x_1, \dots, x_n]/(f) = n.$$

□

11.2.8. Codimension is the difference of dimensions for irreducible varieties. Noether normalization will help us show that codimension is the difference of dimensions for irreducible varieties, i.e., that the inequality (11.1.4.1) is always an equality.

11.2.9. Theorem. — *Suppose X is an irreducible k-variety, Y is an irreducible closed subset, and η is the generic point of Y. Then $\dim Y + \dim \mathscr{O}_{X,\eta} = \dim X$. Hence by Exercise 11.1.I, $\dim Y + \operatorname{codim}_X Y = \dim X$ — inequality (11.1.4.1) is always an equality.*

Proving this will give us an excuse to introduce some useful notions, such as the Going-Down Theorem for finite extensions of integrally closed domains (Theorem 11.2.12). Before we begin the proof, we give an algebraic translation.

11.2.F. EXERCISE. A ring A is called **catenary** if for every nested pair of prime ideals $\mathfrak{p} \subset \mathfrak{q} \subset A$, all maximal chains of prime ideals between $\mathfrak{p}$ and $\mathfrak{q}$ have the same length. (We will not use this term beyond this exercise.) Show that if A is a localization of a finitely generated ring over a field k, then A is catenary.

11.2.10. *Remark.* Most rings arising naturally in algebraic geometry are catenary. Important examples include: localizations of finitely generated $\mathbb{Z}$-algebras; complete Noetherian local rings; Dedekind domains; and Cohen-Macaulay rings (see §26.2.13). It is hard to give an example of a noncatenary ring; see for example [**Stacks**, tag 02JE] or [**He**].

11.2.11. *Proof of Theorem 11.2.9.*

11.2.G. EXERCISE. Reduce the proof of Theorem 11.2.9 to the following problem. If X is an irreducible affine k-variety and Z is a closed irreducible subset maximal among those smaller than X (the only larger closed irreducible subset is X), then $\dim Z = \dim X - 1$.

Let $d = \dim X$ for convenience. By Noether Normalization 11.2.4, we have a finite morphism $\pi : X \to \mathbb{A}^d$ corresponding to a finite extension of rings. Then $\pi(Z)$ is an irreducible closed subset of $\mathbb{A}^d$ (finite morphisms are closed, Exercise 7.3.M).

11.2.H. EXERCISE. Show that it suffices to show that $\pi(Z)$ is a hypersurface. (Hint: the dimension of any hypersurface is $d-1$ by Theorem 11.2.1 on dimension and transcendence degree. Exercise 11.1.E implies that $\dim \pi^{-1}(\pi(Z)) = \dim \pi(Z)$. But be careful: Z is not $\pi^{-1}(\pi(Z))$ in general.)

Now if $\pi(Z)$ is not a hypersurface, then it is properly contained in an irreducible hypersurface H, so by the Going-Down Theorem 11.2.12 for finite extensions of integrally closed domains (which we shall now prove), there is some closed irreducible subset Z' of X properly containing Z, contradicting the maximality of Z. $\square$

11.2.12. Theorem (Going-Down Theorem for finite extensions of integrally closed domains). — *Suppose $\phi : B \hookrightarrow A$ is a finite extension of rings (so A is a finite B-module), B is an integrally closed domain, and A is an integral domain. Then given nested primes $\mathfrak{q} \subset \mathfrak{q}'$ of B, and a prime $\mathfrak{p}'$ of A lying over $\mathfrak{q}'$ (i.e., $\mathfrak{p}' \cap B = \mathfrak{q}'$), there exists a prime $\mathfrak{p}$ of A contained in $\mathfrak{p}'$, lying over $\mathfrak{q}$.*

As usual, you should sketch a geometric picture of the statement. This theorem is usually stated about extending a chain of ideals, in the same way as the Going-Up Theorem (Exercise 7.2.F), and you may want to think this through. (Another Going-Down Theorem, for flat morphisms, will be given in Exercise 24.5.E.)

This theorem is true more generally with "finite" replaced by "integral"; see [**E**, p. 291] ("Completion of the proof of 13.9") for the extension of Theorem 11.2.12, or else see [**AtM**, Thm. 5.16] or [**Mat1**, Thm. 5(v)] for completely different proofs. See [**E**, Fig. 10.4] for an example (in the form of a picture) of why the "integrally closed" hypothesis on B cannot be removed.

In the course of the proof, we will need the following fact. It is not hard, and we could prove it now, but we leave it until Exercise 11.3.D, because the proof uses a trick arising in Exercise 11.3.C.

11.2.13. Proposition (prime avoidance). — *Suppose $I \subset \cup_{i=1}^n \mathfrak{p}_i$, where I is an ideal and the $\mathfrak{p}_i$ are prime ideals of a ring A. (Note: the right side need not be an ideal!) Then $I \subset \mathfrak{p}_i$ for some i.*

(Can you give a geometric interpretation of this result? Can you figure out why it is called "prime avoidance"?)

11.2.14. *Proof of Theorem 11.2.12 (Going-Down Theorem for finite extensions of integrally closed domains).* The proof uses Galois theory. Let L be the normal closure of $K(A)/K(B)$ (the smallest subfield of $\overline{K(B)}$ containing $K(A)$, and that is mapped to itself by any automorphism over $\overline{K(B)}/K(B)$). Let C be the integral closure of B in L (discussed in Exercise 9.7.I). Because $A \hookrightarrow C$ is an integral extension, there is a prime Q of C lying over $\mathfrak{q} \subset B$ (by the Lying Over Theorem 7.2.5), and a prime Q' of C containing Q lying over $\mathfrak{q}'$ (by the Going-Up Theorem, Exercise 7.2.F). Similarly, there is a prime P' of C lying over $\mathfrak{p}' \subset A$ (and thus over $\mathfrak{q}' \subset B$). We would be done if $P' = Q'$ (just take $\mathfrak{p} = Q \cap A$), but this needn't be the case. However, Lemma 11.2.15 below shows there is an automorphism σ of C over B, that sends Q' to P', and then the image of $\sigma(Q)$ in A will do the trick, completing the proof. (The following diagram, in geometric terms, may help.)

$$\begin{tikzcd}
\operatorname{Spec} C/P' \arrow[drr, hook] \arrow[dd, bend right] & & & \\
\operatorname{Spec} C/Q' \arrow[u, "\sigma"'] \arrow[r, hook] \arrow[dd, bend right] & \operatorname{Spec} C/Q \arrow[r, hook] \arrow[dd, bend right] & \operatorname{Spec} C \arrow[d] & L \\
\operatorname{Spec} A/\mathfrak{p}' \arrow[r, hook] \arrow[d] & ? \arrow[r, hook] \arrow[d] & \operatorname{Spec} A \arrow[d] & K(A) \arrow[u, hook] \\
\operatorname{Spec} B/\mathfrak{q}' \arrow[r, hook] & \operatorname{Spec} B/\mathfrak{q} \arrow[r, hook] & \operatorname{Spec} B & K(B) \arrow[u, hook]
\end{tikzcd}$$

11.2.15. Lemma. — *Suppose* B *is an integrally closed domain,* $L/K(B)$ *is a finite normal field extension, and* C *is the integral closure of* B *in* L*. If* $\mathfrak{q}'$ *is a prime ideal of* B*, then automorphisms of* $L/K(B)$ *act transitively on the primes of* C *lying over* $\mathfrak{q}'$*.*

This result is often first seen in number theory, with $B = \mathbb{Z}$ and L a Galois extension of $\mathbb{Q}$.

Proof. Let P and Q_1 be two primes of C lying over $\mathfrak{q}'$, and let $Q_2, \dots, Q_n$ be the primes of C conjugate to Q_1 (the image of Q_1 under $\operatorname{Aut}(L/K(B))$). If P is not one of the Q_i, then P is not contained in any of the Q_i. (Do you see why? Hint: Exercise 11.1.D or 11.1.E.) Hence by prime avoidance (Proposition 11.2.13), P is not contained in their union, so there is some $a \in P$ not contained in any Q_i. Thus no conjugate of a can be contained in Q_1, so the norm $N_{L/K(B)}(a) \in B$ is not contained in $Q_1 \cap B = \mathfrak{q}'$. But since $a \in P$, its norm lies in P, but also in B, and hence in $P \cap B = \mathfrak{q}'$, yielding a contradiction. ∎

We end with two exercises which may give you practice and enlightenment.

11.2.I. Exercise (the dimension of a finite type k-scheme is preserved by any field extension, cf. Exercise 11.1.G(a)). Suppose X is a finite type k-scheme of pure dimension n, and K/k is a field extension (not necessarily algebraic). Show that X_K has pure dimension n. Hint: Reduce to the case where X is affine, so say $X = \operatorname{Spec} A$. Reduce to the case where A is an integral domain. Show (using the axiom of choice) that K/k can be written as an algebraic extension of a pure transcendental extension. Hence by Exercise 11.1.G(a), it suffices to deal with the case where K/k is purely transcendental, say with transcendence basis $\{e_i\}_{i \in I}$ (possibly infinite). Show that $A' := A \otimes_k K$ is an integral domain, by interpreting it as a certain localization of the domain $A[\{e_i\}]$. If $t_1, \dots, t_d$ is a transcendence basis for $K(A)/k$, show that $\{e_i\} \cup \{t_j\}$ is a transcendence basis for $K(A')/k$. Show that $\{t_j\}$ is a transcendence basis for $K(A')/K$.

Exercise 11.2.I is conceptually very useful. For example, if X is described by some equations with $\mathbb{Q}$-coefficients, the dimension of X doesn't depend on whether we consider it as a $\mathbb{Q}$-scheme or as a $\mathbb{C}$-scheme.

11.2.16. *Remark.* Unlike Exercise 11.1.G, Exercise 11.2.I has finite type hypotheses on X. It is not true that if X is an arbitrary k-scheme of pure dimension n, and K/k

is an arbitrary extension, then X_k necessarily has pure dimension n. For example, you can show that $\dim k(x) \otimes_k k(y) = 1$ using the same ideas as in Exercise 9.2.K.

11.2.17. $\star$ **Lines on hypersurfaces, part 1.** Notice: although dimension theory is not central to the following statement, it is essential to the proof.

11.2.J. ENLIGHTENING STRENUOUS EXERCISE: MOST SURFACES IN THREE-SPACE OF DEGREE $d > 3$ HAVE NO LINES. In this exercise, we work over an algebraically closed field $\overline{k}$. For any $d > 3$, show that most degree d surfaces in $\mathbb{P}^3$ contain no lines. Here, "most" means "all closed points of a Zariski-open subset of the parameter space for degree d homogeneous polynomials in 4 variables, up to scalars. As there are $\binom{d+3}{3}$ such monomials, the degree d hypersurfaces are parametrized by $\mathbb{P}^{\binom{d+3}{3}-1}$ (see Remark 4.5.3). Hint: Construct an incidence correspondence

$$X = \{(\ell, H) : [\ell] \in \mathbb{G}(1,3), [H] \in \mathbb{P}^{\binom{d+3}{3}-1}, \ell \subset H\},$$

parametrizing lines in $\mathbb{P}^3$ contained in a hypersurface: define a closed subscheme X of $\mathbb{P}^{\binom{d+3}{3}-1} \times \mathbb{G}(1,3)$ that makes this notion precise. (Recall that $\mathbb{G}(1,3)$ is a Grassmannian.) Show that X is a $\mathbb{P}^{\binom{d+3}{3}-1-(d+1)}$-bundle over $\mathbb{G}(1,3)$. (Possible hint for this: how many degree d hypersurfaces contain the line $x = y = 0$?) Show that $\dim \mathbb{G}(1,3) = 4$ (see §6.7: $\mathbb{G}(1,3)$ has an open cover by $\mathbb{A}^4$'s). Show that $\dim X = \binom{d+3}{3} - 1 - (d+1) + 4$. Show that the image of the projection $X \to \mathbb{P}^{\binom{d+3}{3}-1}$ must lie in a proper closed subset. The following diagram may help.

$$\dim \binom{d+3}{3} - 1 - (d+1) + 4$$

$$\begin{tikzcd}
 & X \arrow[dl] \arrow[dr, "\mathbb{P}^{\binom{d+3}{3}-1-(d+1)}"] & \\
\mathbb{P}^{\binom{d+3}{3}-1} & & \mathbb{G}(1,3)
\end{tikzcd} \qquad \dim 4$$

(The argument readily generalizes to show that if $d > 2n-3$, then "most" degree d hypersurfaces in $\mathbb{P}^n$ have no lines. The case $n = 1$ and $n = 2$ are trivial but worth thinking through.)

11.2.18. *Side Remark.* If you do the previous Exercise, your dimension count will suggest the true facts that degree 1 hypersurfaces — i.e., hyperplanes — have 2-dimensional families of lines, and that most degree 2 hypersurfaces have 1-dimensional families (rulings) of lines, as shown in Exercise 8.2.M. They will also suggest that most degree 3 hypersurfaces contain a finite number of lines, which reflects the celebrated fact that regular cubic surfaces over an algebraically closed field always contain 27 lines (Theorem 27.1.1), and we will use this "incidence correspondence" or "incidence variety" to prove it (§27.4). The statement about quartics generalizes to the Noether-Lefschetz theorem implying that a very general surface of degree d at least 4 contains no curves that are not the intersection of the surface with a hypersurface, see [**Lef**, **GH2**]. **"Very general"** means that in the parameter space (in this case, the projective space parametrizing surfaces of degree d), the statement is true away from a countable union of proper Zariski-closed subsets. Like "general", (which was defined in §9.3.6), "very general" is a weaker version of the phrase "almost every".

11.3 Codimension one miracles: Krull's and Hartogs's Theorems

In this section, we will explore a number of results related to codimension one. We introduce two results that apply in more general situations, and link functions and the codimension one points where they vanish: Krull's Principal Ideal Theorem 11.3.3, and Algebraic Hartogs's Lemma 11.3.10. We will find these two theorems very useful. For example, Krull's Principal Ideal Theorem will help us compute codimensions, and will show us that codimension can behave oddly, and Algebraic Hartogs's Lemma will give us a useful characterization of unique factorization domains (Proposition 11.3.5). The results in this section will require (locally) Noetherian hypotheses. They are harder, in that the proofs are technical, and don't shed much light on the uses of the results. Thus it is more important to understand how to use these results than to be familiar with their proofs.

11.3.1. Krull's Principal Ideal Theorem. In a vector space, a single linear equation always cuts out a subspace of codimension 0 or 1 (and codimension 0 occurs only when the equation is 0). The Principal Ideal Theorem generalizes this linear algebra fact.

11.3.2. Krull's Principal Ideal Theorem (geometric version). — *Suppose X is a locally Noetherian scheme, and f is a function. The irreducible components of $V(f)$ are codimension 0 or 1.*

This is clearly a consequence of the following algebraic statement. You know enough to prove it for varieties (see Exercise 11.3.J), which is where we will use it most often. The full proof is technical, and included in §11.5 (see §11.5.2) only to show you that it isn't excessively long.

11.3.3. Krull's Principal Ideal Theorem (algebraic version). — *Suppose A is a Noetherian ring, and $f \in A$. Then every prime $\mathfrak{p}$ minimal among those containing f has codimension at most 1. If furthermore f is not a zerodivisor, then every minimal prime $\mathfrak{p}$ containing f has codimension precisely 1.*

For example, locally principal closed subschemes have "codimension 0 or 1", and effective Cartier divisors have "pure codimension 1". Here is another example, that you could certainly prove directly, without the Principal Ideal Theorem.

11.3.A. EXERCISE. Show that an irreducible homogeneous polynomial in $n+1$ variables over a field k describes an integral scheme of dimension $n-1$ in $\mathbb{P}^n_k$.

11.3.B. EXERCISE. Suppose $(A, \mathfrak{m})$ is a Noetherian local ring, and $f \in \mathfrak{m}$. Show that $\dim A/(f) \geq \dim A - 1$.

11.3.C. IMPORTANT EXERCISE (TO BE USED REPEATEDLY). This is a cool argument. (a) *(Hypersurfaces meet everything of dimension at least 1 in projective space, unlike in affine space.)* Suppose X is a closed subset of $\mathbb{P}^n_k$ of dimension at least 1, and H is a nonempty hypersurface in $\mathbb{P}^n_k$. Show that H meets X. (Hint: note that the

affine cone over H contains the origin in $\mathbb{A}_k^{n+1}$. Apply Krull's Principal Ideal Theorem 11.3.3 to the cone over X.)
(b) Suppose $X \hookrightarrow \mathbb{P}_k^n$ is a closed subset of dimension r. Show that any codimension r linear space meets X. Hint: Refine your argument in (a). (Exercise 11.3.F generalizes this to show that any two things in projective space that you would expect to meet for dimensional reasons do in fact meet.)
(c) Show further that there is an intersection of $r+1$ nonempty hypersurfaces missing X. (The key step: show that there is a hypersurface of sufficiently high degree that doesn't contain any generic point of X. Show this by induction on the number of generic points. To get from m to $m+1$: take a hypersurface not vanishing on $p_1, \dots, p_m$. If it doesn't vanish on p_{m+1}, we are done. Otherwise, call this hypersurface f_{m+1}. Do something similar with $m+1$ replaced by i for each $1 \leq i \leq m$. Then consider $\sum_i f_1 \cdots \widehat{f_i} \cdots f_{m+1}$.) If k is infinite, show that there is a codimension $r+1$ *linear subspace* missing X. (The key step: show that there is a hyperplane not containing any generic point of a component of X.)
(d) If k is an infinite field, show that there is an intersection of r hyperplanes meeting X in a finite number of points. (We will see in Exercise 12.4.C that if $k = \overline{k}$, for "most" choices of these r hyperplanes, this intersection is reduced, and in Exercise 18.6.N that the number of points is the "degree" of X. But first of course we must define "degree".)

The following exercise has nothing to do with the Principal Ideal Theorem, but its solution is similar to that of Exercise 11.3.C(c) (and you may wish to solve it first).

11.3.D. EXERCISE. Prove Proposition 11.2.13 (prime avoidance). Hint: by induction on n. Don't look in the literature — you might find a much longer argument.

11.3.E. EXERCISE (BOUND ON CODIMENSION OF INTERSECTIONS IN $\mathbb{A}_k^n$). Let k be a field. Suppose X and Y are pure-dimensional subvarieties (possibly singular) of codimension m and n respectively in $\mathbb{A}_k^d$. Show that every component of $X \cap Y$ has codimension at most $m+n$ in $\mathbb{A}_k^d$ as follows. Show that the diagonal $\mathbb{A}_k^d \cong \Delta \subset \mathbb{A}_k^d \times_k \mathbb{A}_k^d$ is a regular embedding of codimension d. (You will quickly guess the d equations for Δ.) Figure out how to identify the intersection of X and Y in $\mathbb{A}^d$ with the intersection of $X \times Y$ with Δ in $\mathbb{A}^d \times_k \mathbb{A}_k^d$. Then show that locally, $X \cap Y$ is cut out in $X \times Y$ by d equations. Use Krull's Principal Ideal Theorem 11.3.3. You will also need Exercise 11.2.E. (See Exercise 12.2.L for a generalization.)

11.3.F. FUN EXERCISE (GENERALIZING EXERCISE 11.3.C(B)). Suppose X and Y are pure-dimensional subvarieties of $\mathbb{P}^n$ of codimensions d and e respectively, and $d + e \leq n$. Show that X and Y intersect. Hint: apply Exercise 11.3.E to the affine cones of X and Y. Recall the argument you used in Exercise 11.3.C(a) or (b).

11.3.G. USEFUL EXERCISE. Suppose f is an element of a Noetherian ring A, contained in no codimension zero or one primes. Show that f is invertible. (Hint: if a function vanishes nowhere, it is invertible, by Exercise 4.3.G(b).)

11.3.4. A useful characterization of unique factorization domains.

We can use Krull's Principal Ideal Theorem to prove one of the four useful criteria for unique factorization domains, promised in §5.4.6.

11.3.5. Proposition. — *Suppose that A is a Noetherian integral domain. Then A is a unique factorization domain if and only if all codimension 1 primes are principal.*

This contains Lemma 11.1.6 and (in some sense) its converse. (We note that the result is true even without Noetherian hypothesis, without too much more work; see [**Ka**, Thm. 5, p. 4].)

Proof. We have already shown in Lemma 11.1.6 that if A is a unique factorization domain, then all codimension 1 primes are principal. Assume conversely that all codimension 1 primes of A are principal. I claim that the generators of these ideals are irreducible, and that we can uniquely factor any element of A into these irreducibles, and invertible. First, suppose (f) is a codimension 1 prime ideal $\mathfrak{p}$. Then if $f = gh$, then either $g \in \mathfrak{p}$ or $h \in \mathfrak{p}$. As $\operatorname{codim} \mathfrak{p} > 0$, $\mathfrak{p} \neq (0)$, so by Nakayama's Lemma 7.2.9 (applied to the local ring $A_\mathfrak{m}$, where $\mathfrak{m}$ is some maximal ideal containing $\mathfrak{p}$, using that $\mathfrak{p}$ is finitely generated), $\mathfrak{p} \neq \mathfrak{p}^2$ (as if we have strict inclusion $\mathfrak{p}^2 \subsetneq \mathfrak{p}$ in $A_\mathfrak{m}$, we must have had strict inclusion $\mathfrak{p}^2 \subsetneq \mathfrak{p}$ in A). Thus g and h cannot both be in $\mathfrak{p}$. Say $g \notin \mathfrak{p}$. Then g is contained in no codimension 1 primes (as f was contained in only one, namely $\mathfrak{p}$), and hence is invertible by Exercise 11.3.G.

11.3.H. Exercise. Show that any nonzero element f of A can be factored into irreducibles. Hint: if f is not irreducible, then factor it into $f_1 f_2$, where neither of f_1 and f_2 are units. If f_1 is not irreducible, then it factors into $f_{11} f_{12}$. Keep on doing this. Show that the Noetherian hypothesis forces the process to terminate.

11.3.I. Exercise. Conclude the proof by showing that this factorization is unique. (Possible hint: the irreducible components of $V(f)$ give you the prime factors, but not the multiplicities.)

11.3.6. Generalizing Krull to more equations. The following generalization of Krull's Principal Ideal Theorem looks like it might follow by induction from Krull, but it is more subtle. A proof is given in §11.5.3.

11.3.7. Krull's Height Theorem. — *Suppose $X = \operatorname{Spec} A$ where A is Noetherian, and Z is an irreducible component of $V(r_1, \dots, r_\ell)$, where $r_1, \dots, r_\ell \in A$. Then the codimension of Z is at most ℓ.*

(The reason for the appearance of the word "Height" in the name of the Theorem is that *height* is another common word for *codimension*, as mentioned in §11.1.4.)

11.3.J. Exercise. Prove Krull's Height Theorem 11.3.7 (and hence Krull's Principal Ideal Theorem 11.3.3) in the special case where X is an irreducible affine variety, i.e., if A is finitely generated domain over some field k. Show that $\dim Z \geq \dim X - \ell$. Hint: Theorem 11.2.9. It can help to localize A so that $Z = V(r_1, \dots, r_\ell)$.

11.3.K. Exercise. Suppose $(A, \mathfrak{m})$ is a Noetherian local ring.
(a) *(Noetherian local rings have finite dimension, promised in Remark 11.1.8)* Use Krull's Height Theorem 11.3.7 to prove that if there are $g_1, \dots, g_\ell$ such that $V(g_1, \dots, g_\ell) =$

$\{[\mathfrak{m}]\}$, then $\dim A \leq \ell$. Hence show that A has finite dimension. (For comparison, Noetherian rings in general may have infinite dimension, see Exercise 11.1.K.)
(b) Let $d = \dim A$. Show that there exist $g_1, \dots, g_d \in A$ such that $V(g_1, \dots, g_d) = \{[\mathfrak{m}]\}$. (Hint: in order to work by induction on d, you need to find a first equation that will knock the dimension down by 1, i.e., $\dim A/(g_d) = \dim A - 1$. Find g_d using prime avoidance, Proposition 11.2.13.) Geometric translation: given a d-dimensional "germ of a reasonable space" around a point p. Then p can be cut out set-theoretically by d equations, and you always need at least d equations. These d elements of A are called a **system of parameters** for the Noetherian local ring A, but we won't use this language except in Exercise 11.4.A.

11.3.8. $\star$ **Pathologies of the notion of "codimension".** We can use Krull's Principal Ideal Theorem to produce the example of pathology in the notion of codimension promised earlier this chapter. Let $A = k[x]_{(x)}[t]$. In other words, elements of A are polynomials in t, whose coefficients are quotients of polynomials in x, where no factors of x appear in the denominator. (Warning: A is not $k[x,t]_{(x)}$.) Clearly, A is an integral domain, so $xt - 1$ is not a zero divisor. You can verify that $A/(xt-1) \cong k[x]_{(x)}[1/x] \cong k(x)$ — "in $k[x]_{(x)}$, we may divide by everything but x, and now we are allowed to divide by x as well" — so $A/(xt-1)$ is a field. Thus $(xt-1)$ is not just prime but also maximal. By Krull's theorem, $(xt-1)$ is codimension 1. Thus $(0) \subset (xt-1)$ is a maximal chain. However, A has dimension at least 2: $(0) \subset (t) \subset (x,t)$ is a chain of primes of length 2. (In fact, A has dimension precisely 2, although we don't need this fact in order to observe the pathology.) Thus we have a codimension 1 prime in a dimension 2 ring that is dimension 0. Here is a picture of this poset of ideals.

$$\begin{tikzcd}
(x,t) \arrow[d, dash] & & \\
(t) \arrow[dr, dash] & & (xt-1) \arrow[dl, dash] \\
& (0) &
\end{tikzcd}$$

This example comes from geometry, and it is enlightening to draw a picture, see Figure 11.2. Spec $k[x]_{(x)}$ corresponds to a "germ" of $\mathbb{A}^1_k$ near the origin, and Spec $k[x]_{(x)}[t]$ corresponds to "this $\times$ the affine line". You may be able to see from the picture some motivation for this pathology — $V(xt-1)$ doesn't meet $V(x)$, so it can't have any specialization on $V(x)$, and there is nowhere else for $V(xt-1)$ to specialize. It is disturbing that this misbehavior turns up even in a relatively benign-looking ring.

11.3.9. Algebraic Hartogs's Lemma for Noetherian normal schemes.

Hartogs's Lemma in several complex variables states (informally) that a holomorphic function defined away from a codimension two set can be extended over that. We now describe an algebraic analog, for Noetherian normal schemes. (It may also be profitably compared to second Riemann extension theorem.) We will use this repeatedly and relentlessly when connecting line bundles and divisors.

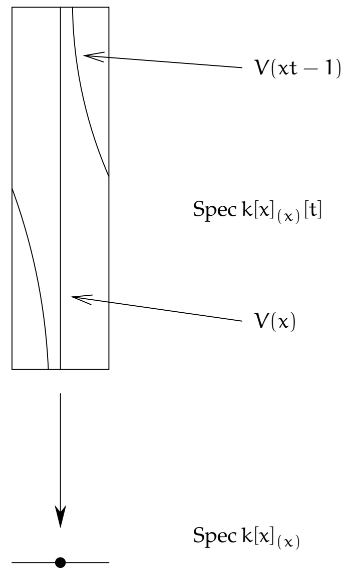


FIGURE 11.2. Dimension and codimension behave oddly on the surface $\operatorname{Spec} k[x]_{(x)}[t]$

11.3.10. Algebraic Hartogs's Lemma. — *Suppose A is a Noetherian normal integral domain. Then*

$$A = \cap_{\mathfrak{p} \text{ codimension } 1} A_{\mathfrak{p}}.$$

The equality takes place in $K(A)$; recall that any localization of an integral domain A is naturally a subset of $K(A)$ (Exercise 1.3.C). Warning: few people call this Algebraic Hartogs's Lemma. I call it this because it parallels the statement in complex geometry.

One might say that if $f \in K(A)$ does not lie in $A_{\mathfrak{p}}$ where $\mathfrak{p}$ has codimension 1, then f has a pole at $[\mathfrak{p}]$, and if $f \in K(A)$ lies in $\mathfrak{p}A_{\mathfrak{p}}$ where $\mathfrak{p}$ has codimension 1, then f has a zero at $[\mathfrak{p}]$. It is worth interpreting Algebraic Hartogs's Lemma as saying that *a rational function on a normal scheme with no poles is in fact regular* (an element of A). Informally: *"Noetherian normal schemes have the Hartogs property."* (We will properly define zeros and poles in §12.5.7, see also Exercise 12.5.H.)

One can state Algebraic Hartogs's Lemma more generally in the case that $\operatorname{Spec} A$ is a Noetherian normal scheme, meaning that A is a product of Noetherian normal integral domains; the reader may wish to do so.

Another generalization (and something closer to the "right" statement) is that if A is a subring of a field K, then the integral closure of A in K is the intersection of all valuation rings of K containing A; see [**AtM**, Cor. 5.22] for explanation and proof.

11.3.11. $\star$ *Proof.* (This proof may be stated completely algebraically, but we state it as geometrically as possible, at the expense of making it longer. See [**Stacks**, tag

031T] for another proof using Serre's criterion for normality, which we prove in §26.3.) The left side is obviously contained in the right, so assume some x lies in every $A_{\mathfrak{p}}$ but not in A. As in the proof of Proposition 5.4.2, we measure the failure of x to be a function (an element of $\operatorname{Spec} A$) with the "ideal of denominators" I of x:

$$I := \{r \in A \ : \ rx \in A\}.$$

(As an important remark not necessary for the proof: it is helpful to interpret the ideal of denominators as scheme-theoretically measuring the failure of x to be regular, or better, giving a scheme-theoretic structure to the locus where x is not regular.) As $1 \notin I$, we have $I \neq A$. Choose a minimal prime $\mathfrak{q}$ containing I.

Our second step in obtaining a contradiction is to focus near the point $[\mathfrak{q}]$, i.e., focus attention on $A_{\mathfrak{q}}$ rather than A, and as a byproduct notice that $\operatorname{codim} \mathfrak{q} > 1$. The construction of the ideal of denominators behaves well with respect to localization — if $\mathfrak{p}$ is any prime, then the ideal of denominators of x in $A_{\mathfrak{p}}$ is $I_{\mathfrak{p}}$, and it again measures "the failure of Algebraic Hartogs's Lemma for x," this time in $A_{\mathfrak{p}}$. But Algebraic Hartogs's Lemma is vacuously true for dimension 1 rings, so no codimension 1 prime contains I. Thus $\mathfrak{q}$ has codimension at least 2. By localizing at $\mathfrak{q}$, we can assume that A is a local ring with maximal ideal $\mathfrak{q}$, and that $\mathfrak{q}$ is the *only* prime containing I.

In the third step, we construct a suitable multiple z of x that is still not a function on $\operatorname{Spec} A$, such that multiplying z by anything vanishing at $[\mathfrak{q}]$ results in a function. (Translation: $z \notin A$, but $z\mathfrak{q} \subset A$.) As $\mathfrak{q}$ is the only prime containing I, $\sqrt{I} = \mathfrak{q}$ (Exercise 3.4.F), so as $\mathfrak{q}$ is finitely generated, there is some n with $I \supset \mathfrak{q}^n$ (do you see why?). Take the minimal such n, so $I \not\supset \mathfrak{q}^{n-1}$, and choose any $y \in \mathfrak{q}^{n-1} - I$. Let $z = yx$. As $y \notin I$, $z \notin A$. On the other hand, $y\mathfrak{q} \subset \mathfrak{q}^n \subset I$, so $z\mathfrak{q} \subset Ix \subset A$, so $z\mathfrak{q}$ is an ideal of A, completing this step.

Finally, we have two cases: either there is function vanishing on $[\mathfrak{q}]$ that, when multiplied by z, doesn't vanish on $[\mathfrak{q}]$; or else every function vanishing on $[\mathfrak{q}]$, multiplied by z, still vanishes on $[\mathfrak{q}]$. Translation: (i) either $z\mathfrak{q}$ is contained in $\mathfrak{q}$, or (ii) it is not.

(i) If $z\mathfrak{q} \subset \mathfrak{q}$, then we would have a finitely generated A-module (namely $\mathfrak{q}$) with a faithful $A[z]$-action, forcing z to be integral over A (and hence in A, as A is integrally closed) by Exercise 7.2.I, yielding a contradiction.

(ii) If $z\mathfrak{q}$ is an ideal of A not contained in the unique maximal ideal $\mathfrak{q}$, then it must be A! Thus $z\mathfrak{q} = A$ from which $\mathfrak{q} = A(1/z)$, from which $\mathfrak{q}$ is principal. But then $\operatorname{codim} \mathfrak{q} = \dim A \leq \dim_{A/\mathfrak{q}} \mathfrak{q}/\mathfrak{q}^2 \leq 1$ (the inequality coming from principality of $\mathfrak{q}$), contradicting $\operatorname{codim} \mathfrak{q} \geq 2$. $\square$

11.3.12. $\star$ **Lines on hypersurfaces, part 2.** (Part 1 was §11.2.17.) We now give a geometric application of Krull's Principal Ideal Theorem 11.3.3, applied through Exercise 11.3.C(a). Throughout, we work over an algebraically closed field $\overline{k}$.

11.3.L. EXERCISE.
(a) Suppose

$$f(x_0, \dots, x_n) = f_d(x_1, \dots, x_n) + x_0 f_{d-1}(x_1, \dots, x_n) + \cdots + x_0^{d-1} f_1(x_1, \dots, x_n)$$

is a homogeneous degree d polynomial (so $\deg f_i = i$) cutting out a hypersurface X in $\mathbb{P}^n$ containing $p := [1, 0, \dots, 0]$. Show that there is a line through p

contained in X if and only if $f_1 = f_2 = \cdots = f_d = 0$ has a common zero in $\mathbb{P}^{n-1} = \operatorname{Proj} \overline{k}[x_1, \ldots, x_n]$. (Hint: given a common zero $[a_1, \ldots, a_n] \in \mathbb{P}^{n-1}$, show that line joining p to $[0, a_1, \ldots, a_n]$ is contained in X.)
(b) If $d \leq n - 1$, show that through any point $p \in X$, there is a line contained in X. Hint: Exercise 11.3.C(a).
(c) If $d \geq n$, show that for "most hypersurfaces" X of degree d in $\mathbb{P}^n$ (for all hypersurfaces whose corresponding point in the parameter space $\mathbb{P}^{\binom{n+d}{d}-1}$ — cf. Remark 4.5.3 and Exercise 8.2.K — lies in some nonempty Zariski-open subset), "most points $p \in X$" (all points in a nonempty dense Zariski-open subset of X) have no lines in X passing through them. (Hint: first show that there is a single p in a single X contained in no line. Chevalley's Theorem 7.4.2 may help.)

11.3.13. *Remark.* A projective (or proper) $\overline{k}$-variety X is **uniruled** if every point $p \in X$ is contained in some $\mathbb{P}^1 \subset X$. (We won't use this word beyond this remark.) Part (b) shows that all hypersurfaces of degree at most $n-1$ are uniruled. One can show (using methods beyond what we know now, see [**Ko1**, Cor. IV.1.11]) that if char $\overline{k} = 0$, then every smooth hypersurface of degree at least $n + 1$ in $\mathbb{P}^n_{\overline{k}}$ is *not* uniruled (thus making the open set in (c) explicit). Furthermore, smooth hypersurfaces of degree n *are* uniruled, but covered by conics rather than lines. Thus there is a strong difference in how hypersurfaces, behave depending on how the degree relates to $n + 1$. This is true in many other ways as well. Smooth hypersurfaces of degree less than $n + 1$ are examples of *Fano* varieties; smooth hypersurfaces of degree $n + 1$ are examples of *Calabi-Yau* varieties (with the possible exception of $n = 1$, depending on the definition); and smooth hypersurfaces of degree greater than $n + 1$ are examples of *general type* varieties. We define these terms in §21.5.5.

11.4 Dimensions of fibers of morphisms of varieties

In this section, we show that the dimensions of fibers of morphisms of varieties behave in a way you might expect from our geometric intuition. The reason we have waited until now to discuss this is because we will use Theorem 11.2.9 (for varieties, codimension is the difference of dimensions). We discuss generalizations in §11.4.5.

Recall that a function f from a topological space X to $\mathbb{R}$ is **upper semicontinuous** if for each $x \in \mathbb{R}$, $f^{-1}((-\infty, x))$ is open. Informally: the function can jump "up" upon taking limits. Our upper semicontinuous functions will map to $\mathbb{Z}$, so informally functions jump "up" on closed subsets. Similarly, f is **lower semicontinuous** if $x \in \mathbb{R}$, $f^{-1}((x, \infty))$ is open.

Before we begin, let's make sure we are on the same page with respect to our intuition. Elimination theory (Theorem 7.4.7) tells us that the projection $\pi : \mathbb{P}^n_A \to \operatorname{Spec} A$ is closed. We can interpret this as follows. A closed subset X of $\mathbb{P}^n_A$ is cut out by a bunch of homogeneous equations in $n + 1$ variables (over A). The image of X is the subset of $\operatorname{Spec} A$ where these equations have a common nontrivial solution. If we try hard enough, we can describe this by saying that the *existence* of a nontrivial solution (or the existence of a preimage of a point under $\pi : X \to \operatorname{Spec} A$) is an "upper semicontinuous" fact. More generally, your intuition might tell you that the locus where a number of homogeneous polynomials in $n + 1$ variables over A have a solution space (in $\mathbb{P}^n_A$) of dimension at least d should be a closed subset of $\operatorname{Spec} A$. (As a special case, consider linear equations. The condition for m linear equations in $n+1$ variables to have a solution space of dimension at least $d+1$ is a closed condition on the coefficients — do you see why, using linear algebra?) This intuition will be correct, and will use properness in a fundamental way (Theorem 11.4.2(b)). We will also make sense of upper semicontinuity in fiber dimension on the *source* (Theorem 11.4.2(a)). A useful example to think through is the map from the xy-plane to the xz-plane ($\operatorname{Spec} k[x,y] \to \operatorname{Spec} k[x,z]$), given by $(x,z) \mapsto (x,xy)$. (This example also came up in §7.4.1.)

We begin our substantive discussion with an inequality that holds more generally in the locally Noetherian setting.

11.4.A. KEY EXERCISE (CODIMENSION BEHAVES AS YOU MIGHT EXPECT FOR A MORPHISM, OR "FIBER DIMENSIONS CAN NEVER BE LOWER THAN EXPECTED"). Suppose $\pi : X \to Y$ is a morphism of locally Noetherian schemes, and $p \in X$ and $q \in Y$ are points such that $q = \pi(p)$. Show that

$$\operatorname{codim}_X p \leq \operatorname{codim}_Y q + \operatorname{codim}_{\pi^{-1}(q)} p$$

(see Figure 11.3). Hint: take a system of parameters for q "in Y", and a system of parameters for p "in $\pi^{-1}(q)$", and use them to find $\operatorname{codim}_Y q + \operatorname{codim}_{\pi^{-1}(q)} p$ elements of $\mathscr{O}_{X,p}$ cutting out $\{[\mathfrak{m}]\}$ in $\operatorname{Spec} \mathscr{O}_{X,p}$. Use Exercise 11.3.K (where "system of parameters" was defined).

[make picture]

FIGURE 11.3. Exercise 11.4.A: the codimension of a point in the total space is bounded by the sum of the codimension of the point in the fiber plus the codimension of the image in the target

Does Exercise 11.4.A agree with your geometric intuition? You should be able to come up with enlightening examples where equality holds, and where equality fails. We will see that equality always holds for sufficiently nice — flat — morphisms, see Proposition 24.5.5.

We now show that the inequality of Exercise 11.4.A is actually an equality over "most of Y" if Y is an irreducible variety.

11.4.1. Proposition. — *Suppose $\pi : X \to Y$ is a (necessarily finite type) morphism of irreducible k-varieties, with $\dim X = m$ and $\dim Y = n$. Then there exists a nonempty open subset $U \subset Y$ such that for all $y \in U$, the fiber over y has pure dimension $m - n$ (or is empty).*

Proof. We begin with three quick reductions. (i) By shrinking Y if necessary, we may assume that Y is affine, say $\operatorname{Spec} B$. (ii) We may also assume that X is affine, say $\operatorname{Spec} A$. (Reason: cover X with a finite number of affine open subsets $X_1, \dots, X_a$, and take the intersection of the U's for each of the $\pi|_{X_i}$.) (iii) If π is not dominant, then we are done, as the image misses a dense open subset U of $\operatorname{Spec} B$. So we assume now that π is dominant.

In order to motivate the rest of the argument, we describe our goal. We will produce a nonempty distinguished open subset U of $\operatorname{Spec} B$ so that $\pi^{-1}(U) \to U$

factors through $\mathbb{A}_U^{m-n} := \mathbb{A}_k^{m-n} \times_k U$ via a finite surjective morphism:

(11.4.1.1)
$$\begin{tikzcd} \operatorname{Spec} A \arrow[d, "\pi"'] & \pi^{-1}(U) \arrow[l, hook', "\text{open emb.}"'] \arrow[d, "\text{finite surj.}"] \\ & \mathbb{A}_U^{m-n} \arrow[d] \\ \operatorname{Spec} B & U \arrow[l, hook', "\text{open emb.}"'] \end{tikzcd}$$

11.4.B. EXERCISE. Show that this suffices to prove the Proposition. (Hint: Use Exercise 11.4.A, and Theorem 11.2.9 that codimension is the difference of dimensions for varieties, to show that each component of the fiber over a point of U has dimension at least $m-n$. Show that any irreducible variety mapping finitely to $\mathbb{A}_k^{m-n}$ has dimension at most $m-n$.)

So we now work to build (11.4.1.1). We begin by noting that we have inclusions of B into both A and $K(B)$, and from both A and $K(B)$ into $K(A)$. The maps from A and $K(B)$ into $K(A)$ both factor through $A \otimes_B K(B)$ (whose Spec is the generic fiber of π), so the maps from both A and $K(B)$ to $A \otimes_B K(B)$ must be inclusions.

(11.4.1.2)
$$\begin{tikzcd} & & K(A) \\ A \arrow[r, hook] \arrow[urr, hook] & A \otimes_B K(B) \arrow[ur] & \\ B \arrow[u, hook] \arrow[r, hook] & K(B) \arrow[u, hook] \arrow[uur, hook] & \end{tikzcd}$$

Clearly $K(A) \otimes_B K(B) = K(A)$ (as $A \otimes_B K(B)$ can be interpreted as taking A and inverting those nonzero elements of B), and $A \otimes_B K(B)$ is a finitely generated ring extension of the field $K(B)$. By transcendence theory (Exercise 11.2.A), $K(A)$ has transcendence degree $m-n$ over $K(B)$ (as $K(A)$ has transcendence degree m over k, and $K(B)$ has transcendence degree n over k). Thus by Noether normalization 11.2.4, we can find elements $t_1, \dots, t_{m-n} \in A \otimes_B K(B)$, algebraically independent over $K(B)$, such that $A \otimes_B K(B)$ is integral over $K(B)[t_1, \dots, t_{m-n}]$.

Now, we can think of the elements $t_i \in A \otimes_B K(B)$ as fractions, with numerators in A and (nonzero) denominators in B. If f is the product of the denominators appearing for each t_i, then by replacing B by B_f (replacing $\operatorname{Spec} B$ by its distinguished open subset $D(f)$), we may assume that the t_i are all in A. Thus (after sloppily renaming B_f as B, and A_f as A) we can trim and extend (11.4.1.2) to the following.

$$\begin{tikzcd} A \arrow[r, hook] & A \otimes_B K(B) \\ B[t_1, \dots, t_{m-n}] \arrow[u, hook] \arrow[r, hook] & K(B)[t_1, \dots, t_{m-n}] \arrow[u, hook, "\text{integral}"'] \\ B \arrow[u, hook] \arrow[r, hook] & K(B) \arrow[u, hook] \end{tikzcd}$$

Now A is finitely generated over B, and hence over $B[t_1, \dots, t_{m-n}]$, say by $u_1, \dots, u_q$. Noether normalization implies that each u_i satisfies some monic equation $f_i(u_i) = 0$, where $f_i \in K(B)[t_1, \dots, t_{m-n}][t]$. The coefficients of f_i are a priori fractions in B, but by multiplying by all those denominators, we can assume each $f_i \in B[t_1, \dots, t_{m-n}][t]$. Let $b \in B$ be the product of the leading coefficients of all the f_i. If $U = D(b)$ (the locus where b is invertible), then over U, the f_i (can be taken to) have leading coefficient 1, so the u_i (in A_b) are integral over $B_b[t_1, \dots, t_n]$. Thus $\operatorname{Spec} A_b \to \operatorname{Spec} B_b[t_1, \dots, t_n]$ is finite and surjective (the latter by the Lying Over Theorem 7.2.5).

We have now constructed (11.4.1.1), as desired. □

There are a couple of things worth pointing out about the proof. First, this result is interesting (and almost exclusively used) for classical varieties over a field k. But the proof of it uses the theory of varieties over *another* field, notably the function field $K(B)$. This is an example of how the introduction of generic points to algebraic geometry is useful even for considering more "classical" questions.

Second, the idea of the main part of the argument is that we have a result over the generic point ($\operatorname{Spec} A \otimes_B K(B)$ finite and surjective over affine space over $K(B)$), and we want to "spread it out" to a neighborhood of the generic point of $\operatorname{Spec} B$. We do this by realizing that "finitely many denominators" appear when correctly describing the problem, and inverting those functions.

11.4.C. EXERCISE (USEFUL CRITERION FOR IRREDUCIBILITY). Suppose $\pi : X \to Y$ is a proper morphism to an irreducible variety, and all the fibers of π are nonempty, and irreducible of the same dimension. Show that X is irreducible.

11.4.2. Theorem (upper semicontinuity of fiber dimension). — *Suppose $\pi : X \to Y$ is a morphism of finite type k-schemes.*
(a) (upper semicontinuity on the source) The dimension of the fiber of π at $p \in X$ (the dimension of the largest component of the fiber containing p) is an upper semicontinuous function of X.
(b) (upper semicontinuity on the target) If furthermore π is closed (e.g. if π is proper), then the dimension of the fiber of π over $q \in Y$ is an upper semicontinuous function in q (i.e., on Y).

You should be able to immediately construct a counterexample to part (b) if the properness hypothesis is dropped. (We also remark that Theorem 11.4.2(b) for projective morphisms is done, in a simple way, in Exercise 18.1.C.)

Proof. (a) Let F_n be the subset of X consisting of points where the fiber dimension is at least n. We wish to show that F_n is a closed subset for all n. We argue by induction on $\dim Y$. The base case $\dim Y = 0$ is trivial. So we fix Y, and assume the result for all smaller-dimensional targets.

11.4.D. EXERCISE. Show that it suffices to prove the result when X and Y are integral, and π is dominant.

Let $r = \dim X - \dim Y$ be the "relative dimension" of π. If $n \leq r$, then $F_n = X$ by Exercise 11.4.A (combined with Theorem 11.2.9, that codimension is the difference of dimensions for varieties).

If $n > r$, then let $U \subset Y$ be the dense open subset of Proposition 11.4.1, where "the fiber dimension is exactly r". Then F_n does not meet the preimage of U. By replacing Y with $Y \setminus U$, we are done by the inductive hypothesis.

11.4.E. EASY EXERCISE. Prove (b) (using (a)).

$\square$

11.4.3. Proposition ("Generically finite implies generally finite"). — *Suppose $\pi : X \to Y$ is a generically finite morphism of irreducible k-varieties of dimension n. Then there is a dense open subset $V \subset Y$ above which π is finite.*

(If you wish, you can later relax the irreducibility hypothesis to simply requiring X and Y to be simply of pure dimension n.)

Proof. As in the proof of Proposition 11.4.1, we may assume that Y is affine, and that π is dominant.

11.4.F. EXERCISE. Prove the result under the additional assumption that X is affine. Hint: follow the appropriate part of the proof of Proposition 11.4.1.

For the general case, suppose that $X = \cup_{i=1}^n U_i$, where the U_i are affine open subschemes of X. By Exercise 11.4.F, there are dense open subsets $V_i \subset Y$ over which $\pi|_{U_i}$ is finite. By replacing Y by an affine open subset of $\cap V_i$, we may assume that $\pi|_{U_i}$ is finite.

11.4.G. EXERCISE. Show that π is closed. Hint: you will just use that $\pi|_{U_i}$ is closed, and that there are a finite number of U_i.

Then $X \setminus U_1$ is a closed subset, so $\pi(X \setminus U_1)$ is closed.

11.4.H. EXERCISE. Show that this closed subset is not all of Y.

Define $V := Y \setminus \pi(X \setminus U_1)$. Then π is finite above V: it is the restriction of the finite morphism $\pi|_{U_1} : U_1 \to Y$ to the open subset V of the target Y. $\square$

11.4.4. Aside: Other semicontinuities.
Semicontinuity is a recurring theme in algebraic geometry. It is worth keeping an eye out for it. Other examples include the following.

(i) fiber dimension (Theorem 11.4.2 above)
(ii) the rank of a matrix of functions (because rank drops on closed subsets, where various discriminants vanish)
(iii) the rank of a finite type quasicoherent sheaf (Exercise 13.7.J)
(iv) degree of a finite morphism, as a function of the target (§13.7.5)

(v) dimension of tangent space at closed points of a variety over an algebraically closed field (Exercise 21.2.J)
(vi) rank of cohomology groups of coherent sheaves, in proper flat families (Theorem 28.1.1)

All but (ii) are upper semicontinuous; (ii) is a lower semicontinuous function.

11.4.5. $\star\star$ **Generalizing results of §11.4 beyond varieties.** The above arguments can be extended to more general situations than varieties. We remain in the locally Noetherian situation for safety, until the last sentence of §11.4.6. One fact used repeatedly was that codimension is the difference of dimensions (Theorem 11.2.9). This holds much more generally; see Remark 11.2.10 on catenary rings. Extensions of Proposition 11.4.1 should require that π be finite type (which was automatic in the statement of Proposition 11.4.1, by the Cancellation Theorem 10.1.19 for finite type morphisms). In the proof of Proposition 11.4.1, we use that the dimension of the the generic fiber of the morphism $\pi : X \to Y$ of irreducible schemes is $\dim X - \dim Y$; this can be proved using Proposition 24.5.5. The remaining results then readily follow without change.

11.4.6. We make one particular generalization explicit because we will mention it later. The ring $\mathbb{Z}$ (along with finitely generated algebras over it, and localizations thereof) is excellent in all possibly ways, including in the ways needed in the previous paragraph, so the argument for upper semicontinuity on the source (Theorem 11.4.2(a)) applies without change if the target is the Spec of a finitely generated $\mathbb{Z}$-algebra. Any finite type morphism of locally Noetherian schemes is locally pulled back from a finite type morphism to the Spec of a finitely-generated $\mathbb{Z}$-algebra (this will essentially be made precise and shown in Exercise 28.2.L), so we then have upper semicontinuity of fiber dimension on the source for all finite type morphisms of locally Noetherian schemes, and better yet, for all locally finitely presented morphisms.

11.5 $\star\star$ Proof of Krull's Principal Ideal and Height Theorems

The details of this proof won't matter to us, so you should probably not read it. It is included so you can glance at it and believe that the proof is fairly short, and that you could read it if you needed to.

If A is a ring, an A-module is **Artinian** if it satisfies the descending chain condition for submodules (any infinite descending sequence of submodules must stabilize, cf. §3.6.14). A **ring is Artinian** if it is Artinian over itself as a module. The notion of Artinian rings is very important, but we will get away without discussing it much.

If $\mathfrak{m}$ is a maximal ideal of A, then any finite-dimensional $(A/\mathfrak{m})$-vector space (interpreted as an A-module) is clearly Artinian, as any descending chain

$$M_1 \supset M_2 \supset \cdots$$

must eventually stabilize (as $\dim_{A/\mathfrak{m}} M_i$ is a non-increasing sequence of non-negative integers).

11.5.A. EXERCISE. Suppose the maximal ideal $\mathfrak{m}$ is finitely generated. Show that for any n, $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ is a finite-dimensional $(A/\mathfrak{m})$-vector space. (Hint: show it for $n = 0$ and $n = 1$. Show surjectivity of $\operatorname{Sym}^n \mathfrak{m}/\mathfrak{m}^2 \to \mathfrak{m}^n/\mathfrak{m}^{n+1}$ to bound the dimension for general n.) Hence $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ is an Artinian A-module.

11.5.B. EXERCISE. Suppose A is a ring with one prime ideal $\mathfrak{m}$. Suppose $\mathfrak{m}$ is finitely generated. Prove that $\mathfrak{m}^n = (0)$ for some n. (Hint: As $\sqrt{0}$ is prime, it must be $\mathfrak{m}$. Suppose $\mathfrak{m}$ can be generated by r elements, each of which has kth power 0, and show that $\mathfrak{m}^{r(k-1)+1} = 0$.)

11.5.C. EXERCISE. Show that if $0 \to M' \to M \to M'' \to 0$ is an exact sequence of modules, then M is Artinian if and only if M' and M'' are Artinian. (Hint: given a descending chain in M, produce descending chains in M' and M''.)

11.5.1. Lemma. — *If A is a Noetherian ring with one prime ideal $\mathfrak{m}$, then A is Artinian, i.e., it satisfies the descending chain condition for ideals.*

Proof. As we have a finite filtration

$$A \supset \mathfrak{m} \supset \cdots \supset \mathfrak{m}^n = (0)$$

all of whose quotients are Artinian, A is Artinian as well. □

11.5.2. *Proof of Krull's Principal Ideal Theorem 11.3.3.* Suppose we are given $f \in A$, with $\mathfrak{p}$ a minimal prime containing f. By localizing at $\mathfrak{p}$, we may assume that A is a local ring, with maximal ideal $\mathfrak{p}$. Suppose $\mathfrak{q}$ is another prime strictly contained in $\mathfrak{p}$.

$$\begin{tikzcd}
(f) \arrow[dr, hook] & & \\
& \mathfrak{p} \arrow[r, hook] & A \\
\mathfrak{q} \arrow[ur, hook] & &
\end{tikzcd}$$

For the first part of the theorem, we must show that $A_{\mathfrak{q}}$ has dimension 0. The second part follows from our earlier work: if any minimal primes are height 0, f is a zerodivisor, by Remark 5.5.12 (or Theorem 5.5.10(c) and §5.5.3).

Now $\mathfrak{p}$ is the only prime ideal containing (f), so $A/(f)$ has one prime ideal. By Lemma 11.5.1, $A/(f)$ is Artinian.

We invoke a useful construction, the n**th symbolic power of a prime ideal**: if A is any ring, and $\mathfrak{q}$ is any prime ideal, then define

$$\mathfrak{q}^{(n)} := \{r \in A \ : \ rs \in \mathfrak{q}^n \text{ for some } s \in A - \mathfrak{q}\}.$$

We return to our particular A and $\mathfrak{q}$. We have a descending chain of ideals in A

$$\mathfrak{q}^{(1)} \supset \mathfrak{q}^{(2)} \supset \cdots,$$

so we have a descending chain of ideals in $A/(f)$

$$\mathfrak{q}^{(1)} + (f) \supset \mathfrak{q}^{(2)} + (f) \supset \cdots$$

which stabilizes, as $A/(f)$ is Artinian. Say $\mathfrak{q}^{(n)} + (f) = \mathfrak{q}^{(n+1)} + (f)$, so

$$\mathfrak{q}^{(n)} \subset \mathfrak{q}^{(n+1)} + (f).$$

Hence for any $g \in \mathfrak{q}^{(n)}$, we can write $g = h + af$ with $h \in \mathfrak{q}^{(n+1)}$. Hence $af \in \mathfrak{q}^{(n)}$. As $\mathfrak{p}$ is minimal over f, $f \notin \mathfrak{q}$, so $a \in \mathfrak{q}^{(n)}$. Thus

$$\mathfrak{q}^{(n)} = \mathfrak{q}^{(n+1)} + (f)\mathfrak{q}^{(n)}.$$

As f is in the maximal ideal $\mathfrak{p}$, the second version of Nakayama's Lemma 7.2.9 gives $\mathfrak{q}^{(n)} = \mathfrak{q}^{(n+1)}$.

We now shift attention to the local ring $A_{\mathfrak{q}}$, which we are hoping has dimension 0. We have $\mathfrak{q}^{(n)}A_{\mathfrak{q}} = \mathfrak{q}^{(n+1)}A_{\mathfrak{q}}$ (the symbolic power construction clearly commutes with localization). For any $r \in \mathfrak{q}^n A_{\mathfrak{q}} \subset \mathfrak{q}^{(n)}A_{\mathfrak{q}}$, there is some $s \in A_{\mathfrak{q}} - \mathfrak{q}A_{\mathfrak{q}}$ such that $rs \in \mathfrak{q}^{n+1}A_{\mathfrak{q}}$. As s is invertible, $r \in \mathfrak{q}^{n+1}A_{\mathfrak{q}}$ as well. Thus $\mathfrak{q}^n A_{\mathfrak{q}} \subset \mathfrak{q}^{n+1}A_{\mathfrak{q}}$, but as $\mathfrak{q}^{n+1}A_{\mathfrak{q}} \subset \mathfrak{q}^n A_{\mathfrak{q}}$, we have $\mathfrak{q}^n A_{\mathfrak{q}} = \mathfrak{q}^{n+1}A_{\mathfrak{q}}$. By Nakayama's Lemma 7.2.9 (version 2),

$$\mathfrak{q}^n A_{\mathfrak{q}} = 0.$$

Finally, any local ring $(R, \mathfrak{m})$ such that $\mathfrak{m}^n = 0$ has dimension 0, as $\operatorname{Spec} R$ consists of only one point: $[\mathfrak{m}] = V(\mathfrak{m}) = V(\mathfrak{m}^n) = V(0) = \operatorname{Spec} R$. □

11.5.3. *Proof of Krull's Height Theorem 11.3.7.* We argue by induction on n. The case $n = 1$ is Krull's Principal Ideal Theorem 11.3.3. Assume $n > 1$. Suppose $\mathfrak{p}$ is a minimal prime containing $r_1, \dots, r_n \in A$. We wish to show that $\operatorname{codim} \mathfrak{p} \leq n$. By localizing at $\mathfrak{p}$, we may assume that $\mathfrak{p}$ is the unique maximal ideal of A. Let $\mathfrak{q} \neq \mathfrak{p}$ be a prime ideal of A with no prime between $\mathfrak{p}$ and $\mathfrak{q}$. We shall show that $\mathfrak{q}$ is minimal over an ideal generated by $c - 1$ elements. Then $\operatorname{codim} \mathfrak{q} \leq c - 1$ by the inductive hypothesis, so we will be done.

Now $\mathfrak{q}$ cannot contain every r_i (as $V(r_1, \dots, r_n) = \{[\mathfrak{p}]\}$), so say $r_1 \notin \mathfrak{q}$. Then $V(\mathfrak{q}, r_1) = \{[\mathfrak{p}]\}$. As each $r_i \in \mathfrak{p}$, there is some N such that $r_i^N \in (\mathfrak{q}, r_1)$ (Exercise 3.4.J), so write $r_i^N = q_i + a_i r_1$ where $q_i \in \mathfrak{q}$ $(2 \leq i \leq n)$ and $a_i \in A$. Note that

$$V(r_1, q_2, \dots, q_n) = V(r_1, r_2^N, \dots, r_n^N) = V(r_1, r_2, \dots, r_n) = \{[\mathfrak{p}]\}. \tag{11.5.3.1}$$

We shall show that $\mathfrak{q}$ is minimal among primes containing $q_2, \dots, q_n$, completing the proof. In the ring $A/(q_2, \dots, q_n)$, $V(r_1) = \{[\mathfrak{p}]\}$ by (11.5.3.1). By Krull's Principal Ideal Theorem 11.3.3, $[\mathfrak{p}]$ is codimension at most 1, so $[\mathfrak{q}]$ must be codimension 0 in $\operatorname{Spec} A/(q_2, \dots, q_n)$, as desired. □

CHAPTER 12

Regularity and smoothness

One natural notion we expect to see for geometric spaces is the notion of when an object is "smooth". In algebraic geometry, this notion, called *regularity*, is easy to define (Definition 12.2.3) but a bit subtle in practice. This will lead us to a different related notion of when *a variety is smooth* (Definition 12.2.6).

This chapter has many moving parts, of which §12.1–12.6 are the important ones. In §12.1, the *Zariski tangent space* is motivated and defined. In §12.2, we define *regularity* and *smoothness over a field*, the central topics of this chapter, and discuss some of their important properties. In §12.3, we give a number of important examples, mostly in the form of exercises. In §12.4, we discuss *Bertini's Theorem*, a fundamental classical result. In §12.5, we give many characterizations of *discrete valuation rings*, which play a central role in algebraic geometry. Having seen clues that "smoothness" is a "relative" notion rather than an "absolute" one, in §12.6 we define *smooth morphisms* (and in particular, *étale morphisms*), and give some of their properties. (We will revisit this definition in §21.3.1, once we know more.)

The remaining sections are less central. In §12.7, we discuss the valuative criteria for separatedness and properness. In §12.8, we mention some more sophisticated facts about regular local rings. In §12.9, we prove the Artin-Rees Lemma, because it was invoked in §12.5 (and will be used later as well).

12.1 The Zariski tangent space

We begin by defining the tangent space of a scheme at a point. It behaves like the tangent space you know and love at "smooth" points, but also makes sense at other points. In other words, geometric intuition at the "smooth" points guides the definition, and then the definition guides the algebra at all points, which in turn lets us refine our geometric intuition.

The definition is short but surprising. The main difficulty is convincing yourself that it deserves to be called the tangent space. This is tricky to explain, because we want to show that it agrees with our intuition, but our intuition is worse than we realize. So you should just accept this definition for now, and later convince yourself that it is reasonable.

12.1.1. *Definition.* The **Zariski cotangent space** of a local ring $(A, \mathfrak{m})$ is defined to be $\mathfrak{m}/\mathfrak{m}^2$; it is a vector space over the residue field $A/\mathfrak{m}$. The dual vector space is the **Zariski tangent space**. If X is a scheme, the **Zariski cotangent space** $T^\vee_{X,p}$ at a point $p \in X$ is defined to be the Zariski cotangent space of the local ring $\mathscr{O}_{X,p}$ (and similarly for the **Zariski tangent space** $T_{X,p}$). Elements of the Zariski cotangent space are called **cotangent vectors** or **differentials**; elements of the tangent space are called **tangent vectors**.

The cotangent space is more algebraically natural than the tangent space, in that the definition is shorter. There is a moral reason for this: the cotangent space is more naturally determined in terms of functions on a space, and we are very much thinking about schemes in terms of "functions on them". This will come up later.

Here are two plausibility arguments that this is a reasonable definition. Hopefully one will catch your fancy.

In differential geometry, the tangent space at a point is sometimes defined as the vector space of derivations at that point. A derivation at a point p of a manifold is an $\mathbb{R}$-linear operation that takes in functions f near p (i.e., elements of $\mathscr{O}_p$), and outputs elements $f'(p)$ of $\mathbb{R}$, and satisfies the Leibniz rule

$$(fg)' = f'g + g'f.$$

(We will later define derivations in a more general setting, §21.2.17.) A derivation is the same as a map $\mathfrak{m} \to \mathbb{R}$, where $\mathfrak{m}$ is the maximal ideal of $\mathscr{O}_p$. (The map $\mathscr{O}_p \to \mathbb{R}$ extends this, via the map $\mathscr{O}_p \to \mathfrak{m}$ given by $f - f(p)$.) But $\mathfrak{m}^2$ maps to 0, as if $f(p) = g(p) = 0$, then

$$(fg)'(p) = f'(p)g(p) + g'(p)f(p) = 0.$$

Thus a derivation induces a map $\mathfrak{m}/\mathfrak{m}^2 \to \mathbb{R}$, i.e., an element of $(\mathfrak{m}/\mathfrak{m}^2)^\vee$.

12.1.A. EXERCISE. Check that this is reversible, i.e., that any map $\mathfrak{m}/\mathfrak{m}^2 \to \mathbb{R}$ gives a derivation. In other words, verify that the Leibniz rule holds.

Here is a second, vaguer, motivation that this definition is plausible for the cotangent space of the origin of $\mathbb{A}^n$. (I prefer this one, as it is more primitive and elementary.) Functions on $\mathbb{A}^n$ should restrict to a linear function on the tangent space. What (linear) function does $x^2 + xy + x + y$ restrict to "near the origin"? You will naturally answer: $x + y$. Thus we "pick off the linear terms". Hence $\mathfrak{m}/\mathfrak{m}^2$ are the linear functionals on the tangent space, so $\mathfrak{m}/\mathfrak{m}^2$ is the cotangent space. In particular, you should picture functions vanishing at a point (i.e., lying in $\mathfrak{m}$) as giving functions on the tangent space in this obvious way.

12.1.2. *Old-fashioned example.* Computing the Zariski-tangent space is actually quite hands-on, because you can compute it just as you did when you learned multivariable calculus. In $\mathbb{A}^3$, we have a curve cut out by $x + y + z^2 + xyz = 0$ and $x - 2y + z + x^2y^2z^3 = 0$. (You can use Krull's Principal Ideal Theorem 11.3.3 to check that this is a curve, but it is not important to do so.) What is the tangent line near the origin? (Is it even smooth there?) Answer: the first surface looks like $x + y = 0$ and the second surface looks like $x - 2y + z = 0$. The curve has tangent line cut out by $x + y = 0$ and $x - 2y + z = 0$. It is smooth (in the traditional sense). In multivariable calculus, the students work hard to get the answer, because we aren't allowed to tell them to just pick out the linear terms.

Let's make explicit the fact that we are using. If A is a ring, $\mathfrak{m}$ is a maximal ideal, and $f \in \mathfrak{m}$ is a function vanishing at the point $[\mathfrak{m}] \in \operatorname{Spec} A$, then the Zariski tangent space of $\operatorname{Spec} A/(f)$ at $\mathfrak{m}$ is cut out in the Zariski tangent space of $\operatorname{Spec} A$ (at $\mathfrak{m}$) by the single linear equation $f \pmod{\mathfrak{m}^2}$. The next exercise will force you to think this through.

12.1.B. IMPORTANT EXERCISE ("KRULL'S PRINCIPAL IDEAL THEOREM FOR TANGENT SPACES" — BUT MUCH EASIER THAN KRULL'S PRINCIPAL IDEAL THEOREM 11.3.3!). Suppose A is a ring, and $\mathfrak{m}$ a maximal ideal. If $f \in \mathfrak{m}$, show that the Zariski tangent space of A/f is cut out in the Zariski tangent space of A by $f \pmod{\mathfrak{m}^2}$. (Note: we can quotient by f and localize at $\mathfrak{m}$ in either order, as quotienting and localizing commute, (4.3.6.1).) Hence the dimension of the Zariski tangent space of $\operatorname{Spec} A/(f)$ at $[\mathfrak{m}]$ is the dimension of the Zariski tangent space of $\operatorname{Spec} A$ at $[\mathfrak{m}]$, or one less. (That last sentence should be suitably interpreted if the dimension is infinite, although it is less interesting in this case.)

Here is another example to see this principle in action, extending Example 12.1.2: $x+y+z^2=0$ and $x+y+x^2+y^4+z^5=0$ cuts out a curve, which obviously passes through the origin. If I asked my multivariable calculus students to calculate the tangent line to the curve at the origin, they would do calculations which would boil down (without them realizing it) to picking off the linear terms. They would end up with the equations $x+y=0$ and $x+y=0$, which cut out a plane, not a line. They would be disturbed, and I would explain that this is because the curve isn't smooth at a point, and their techniques don't work. We on the other hand bravely declare that the cotangent space is cut out by $x+y=0$, and (will soon) use this pathology as *definition* of what makes a point singular (or non-regular). (Intuitively, the curve near the origin is very close to lying in the plane $x+y=0$.) Notice: the cotangent space jumped up in dimension from what it was "supposed to be", not down. We will see that this is not a coincidence soon, in Theorem 12.2.1.

12.1.C. EXERCISE. Suppose Y and Z are closed subschemes of X, both containing the point $p \in X$.
(a) Show that $T_{Z,p}$ is naturally a sub-$\kappa(p)$-vector space of $T_{X,p}$.
(b) Show that $T_{Y\cap Z,p} = T_{Y,p} \cap T_{Z,p}$, where $\cap$ as usual in this context is scheme-theoretic intersection.
(c) Show that $T_{Y\cup Z,p}$ contains the span of $T_{Y,p}$ and $T_{Z,p}$, where $\cup$ as usual in this context is scheme-theoretic union.
(d) Show that $T_{Y\cup Z,p}$ can be strictly larger than the span of $T_{Y,p}$ and $T_{Z,p}$. (Hint: Figure 4.5.)

Here is a pleasant consequence of the notion of Zariski tangent space.

12.1.3. Problem. Consider the ring $A = k[x,y,z]/(xy-z^2)$. Show that (x,z) is not a principal ideal.

As $\dim A = 2$ (why?), and $A/(x,z) \cong k[y]$ has dimension 1, we see that this ideal is codimension 1 (as codimension is the difference of dimensions for irreducible varieties, Theorem 11.2.9). Our geometric picture is that $\operatorname{Spec} A$ is a cone (we can diagonalize the quadric as $xy - z^2 = ((x+y)/2)^2 - ((x-y)/2)^2 - z^2$, at least if $\operatorname{char} k \neq 2$ — see Exercise 5.4.J), and that (x,z) is a line on the cone. (See Figure 12.1 for a sketch.) This suggests that we look at the cone point.

12.1.4. *Solution.* Let $\mathfrak{m} = (x,y,z)$ be the maximal ideal corresponding to the origin. Then $\operatorname{Spec} A$ has Zariski tangent space of dimension 3 at the origin, and $\operatorname{Spec} A/(x,z)$ has Zariski tangent space of dimension 1 at the origin. But $\operatorname{Spec} A/(f)$ must have Zariski tangent space of dimension at least 2 at the origin by Exercise 12.1.B.

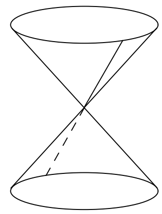


FIGURE 12.1. $V(x,z) \subset \operatorname{Spec} k[x,y,z]/(xy - z^2)$ is a line on a cone

12.1.5. ⋆ *Remark.* Another approach to solving the problem, not requiring the definition of the Zariski tangent space, is to use the fact that the ring is graded (where x, y, and z each have degree 1), and the ideal (x,z) is a graded ideal. (You may enjoy thinking this through.) The advantage of using the tangent space is that it applies to more general situations where there is no grading. For example, (a) (x,z) is not a principal ideal of $k[x,y,z]/(xy - z^2 - z^3)$. As a different example, (b) (x,z) is not a principal ideal of the local ring $(k[x,y,z]/(xy - z^2))_{(x,y,z)}$ (the "germ of the cone"). However, we remark that the graded case is still very useful. The construction of replacing a filtered ring by its "associated graded" ring can turn more general rings into graded rings (and can be used to turn example (a) into the graded case). The construction of completion can turn local rings into graded local rings (and can be used to turn example (b) into, essentially, the graded case). Filtered rings will come up in §12.9, the associated graded construction will implicitly come up in our discussions of the blow-up in §22.3, and many aspects of completions will be described in Chapter 29.

12.1.D. EXERCISE. Show that $(x,z) \subset k[w,x,y,z]/(wz - xy)$ is a codimension 1 ideal that is not principal, using the method of Solution 12.1.4. (See Figure 12.2 for the projectivization of this situation — a line on a smooth quadric surface.) This example was promised just after Exercise 5.4.D. An improvement is given in Exercise 14.2.R.

12.1.E. EXERCISE. Let $A = k[w,x,y,z]/(wz - xy)$. Show that $\operatorname{Spec} A$ is not factorial. (Exercise 5.4.L shows that A is not a unique factorization domain, but this is not enough — why is the localization of A at the prime (w,x,y,z) not factorial? One possibility is to do this "directly", by trying to imitate the solution to Exercise 5.4.L, but this is hard. Instead, use the intermediate result that in a unique factorization domain, any codimension 1 prime is principal, Lemma 11.1.6, and considering Exercise 12.1.D.) As A is integrally closed if $\operatorname{char} k \neq 2$ (Exercise 5.4.I(c)), this yields an example of a scheme that is normal but not factorial, as promised in Exercise 5.4.F. A slight generalization will be given in 22.4.N.

12.1.F. LESS IMPORTANT EXERCISE ("HIGHER-ORDER DATA"). (This exercise is fun, but won't be used.)
(a) In Exercise 3.7.B, you computed the equations cutting out the (union of the) three coordinate axes of $\mathbb{A}^3_k$. (Call this scheme X.) Your ideal should have had three generators. Show that the ideal cannot be generated by fewer than three

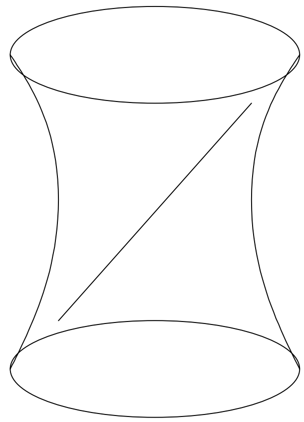


FIGURE 12.2. The line $V(x,z)$ on the smooth quadric surface $V(wz - xy) \subset \mathbb{P}^3$.

elements. (Hint: working modulo $\mathfrak{m} = (x, y, z)$ won't give any useful information, so work modulo a higher power of $\mathfrak{m}$.)
(b) Show that the coordinate axes in $\mathbb{A}^3_k$ are not a regular embedding in $\mathbb{A}^3_k$. (This was promised at the end of §8.4.)

12.1.6. Morphisms and tangent spaces. Suppose $\pi : X \to Y$, and $\pi(p) = q$. Then if we were in the category of manifolds, we would expect a tangent map, from the tangent space of p to the tangent space at q. Indeed that is the case; we have a map of stalks $\mathscr{O}_{Y,q} \to \mathscr{O}_{X,p}$, which sends the maximal ideal of the former $\mathfrak{n}$ to the maximal ideal of the latter $\mathfrak{m}$ (we have checked that this is a "local morphism" when we briefly discussed locally ringed spaces, see §6.3.1). Thus $\mathfrak{n}^2$ maps to $\mathfrak{m}^2$, from which we see that $\mathfrak{n}/\mathfrak{n}^2$ maps to $\mathfrak{m}/\mathfrak{m}^2$. If $(\mathscr{O}_{X,p}, \mathfrak{m})$ and $(\mathscr{O}_{Y,q}, \mathfrak{n})$ have the same residue field κ, so $\mathfrak{n}/\mathfrak{n}^2 \to \mathfrak{m}/\mathfrak{m}^2$ is a linear map of κ-vector spaces, we have a natural map $(\mathfrak{m}/\mathfrak{m}^2)^\vee \to (\mathfrak{n}/\mathfrak{n}^2)^\vee$. This is the map from the tangent space of p to the tangent space at q that we sought. (Aside: note that the *cotangent* map *always* exists, without requiring p and q to have the same residue field — a sign that cotangent spaces are more natural than tangent spaces in algebraic geometry.)

Here are some exercises to give you practice with the Zariski tangent space. If you have some differential geometric background, the first will further convince you that this definition correctly captures the idea of (co)tangent space.

12.1.G. IMPORTANT EXERCISE (THE JACOBIAN COMPUTES THE ZARISKI COTANGENT SPACE). Suppose X is a finite type k-scheme. Then locally it is of the form $\operatorname{Spec} k[x_1, \dots, x_n]/(f_1, \dots, f_r)$. Show that the Zariski cotangent space at a k-valued point (a closed point with residue field k) is given by the cokernel of the Jacobian map $k^r \to k^n$ given by the Jacobian matrix

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_r}{\partial x_1}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n}(p) & \cdots & \frac{\partial f_r}{\partial x_n}(p) \end{pmatrix}. \tag{12.1.6.1}$$

(This makes precise our example of a curve in $\mathbb{A}^3$ cut out by a couple of equations, where we picked off the linear terms, see Example 12.1.2.) You might be alarmed: what does $\frac{\partial f}{\partial x_1}$ mean? Do you need deltas and epsilons? No! Just define derivatives formally, e.g.

$$\frac{\partial}{\partial x_1}(x_1^2 + x_1 x_2 + x_2^2) = 2x_1 + x_2.$$

Hint: Do this first when p is the origin, and consider linear terms, just as in Example 12.1.2 and Exercise 12.1.B. For the general case, "translate p to the origin".

12.1.7. *Remark.* This result can be extended to closed points of X whose residue field is separable over k (and in particular, to *all* closed points if char $k = 0$ or if k is finite), see Remark 21.3.10.

12.1.8. *Warning.* It is more common in mathematics (but not universal) to define the Jacobian matrix as the transpose of this. But it will be more convenient for us to follow this minority convention.

12.1.H. EXERCISE (THE CORANK OF THE JACOBIAN IS INDEPENDENT OF THE PRESENTATION). Suppose A is a finitely-generated k-algebra, generated by $x_1, \dots, x_n$, with ideal of relations I generated by $f_1, \dots, f_r$. Let p be a point of Spec A.
(a) Suppose $g \in I$. Show that appending the column of partials of g to the Jacobian matrix (12.1.6.1) does not change the corank at p. Hence show that the corank of the Jacobian matrix at p does not depend on the choice of generators of I.
(b) Suppose $q(x_1, \dots, x_n) \in k[x_1, \dots, x_n]$. Let h be the polynomial $y - q(x_1, \dots, x_n) \in k[x_1, \dots, x_n, y]$. Show that the Jacobian matrix of $(f_1, \dots, f_r, h)$ with respect to the variables $(x_1, \dots, x_n, y)$ has the same rank at p as the Jacobian matrix of $(f_1, \dots, f_r)$ with respect to $(x_1, \dots, x_n)$. Hence show that the corank of the Jacobian matrix at p is independent of the choice of generators for A.

12.1.I. EXERCISE. Suppose X is a k-scheme. Describe a natural bijection from $\mathrm{Mor}_k(\operatorname{Spec} k[\epsilon]/(\epsilon^2), X)$ to the data of a point p with residue field k (necessarily a closed point) and a tangent vector at p. (This is important, for example in deformation theory.)

12.1.J. EXERCISE. Find the dimension of the Zariski tangent space at the point $[(2, 2i)]$ of $\mathbb{Z}[2i] \cong \mathbb{Z}[x]/(x^2+4)$. Find the dimension of the Zariski tangent space at the point $[(2, x)]$ of $\mathbb{Z}[\sqrt{-2}] \cong \mathbb{Z}[x]/(x^2+2)$. (If you prefer geometric versions of the same examples, replace $\mathbb{Z}$ by $\mathbb{C}$, and 2 by y: consider $\mathbb{C}[x,y]/(x^2+y^2)$ and $\mathbb{C}[x,y]/(x^2+y)$.)

12.2 Regularity, and smoothness over a field

The key idea in the definition of regularity is contained in the following result, that "the dimension of the Zariski tangent space is at least the dimension of the local ring".

12.2.1. Theorem. — *Suppose* $(A, \mathfrak{m}, k)$ *is a Noetherian local ring. Then* $\dim A \leq \dim_k \mathfrak{m}/\mathfrak{m}^2$.

12.2.2. *Proof of Theorem 12.2.1.* Note that $\mathfrak{m}$ is finitely generated (as A is Noetherian), so $\mathfrak{m}/\mathfrak{m}^2$ is a finitely generated $(A/\mathfrak{m} = k)$-module, hence finite-dimensional. Say $\dim_k \mathfrak{m}/\mathfrak{m}^2 = n$. Choose a basis of $\mathfrak{m}/\mathfrak{m}^2$, and lift it to elements $f_1, \dots, f_n$ of $\mathfrak{m}$. Then by Nakayama's Lemma (version 4, Exercise 7.2.H), $(f_1, \dots, f_n) = \mathfrak{m}$.

Then by Exercise 11.3.K (a consequence of Krull's Height Theorem 11.3.7), $\dim A \leq n$. $\square$

12.2.3. *Definition.* If equality holds in Theorem 12.2.1, we say that A is a **regular local ring**. (If a Noetherian ring A is regular at all of its primes, i.e., if $A_{\mathfrak{p}}$ is a regular local ring for all primes $\mathfrak{p}$ of A, then A is said to be a **regular ring**. But we basically won't use this terminology.) A locally Noetherian scheme X is **regular** at a point p if the local ring $\mathscr{O}_{X,p}$ is regular. (The word **nonsingular** is often used as well, notably in the case when X is finite type over a field, but for the sake of consistency we will use "regular" throughout.) It is **singular** at the point otherwise, and we say that the point is a **singularity**. (The word **nonregular** is also used is also used, but for the sake of consistency, we will use "singular", despite the fact that this choice is *inconsistent* with our choice of "regular" over "nonsingular". Possible motivation for the inconsistency: "regular local ring" and "singularity" are both standard terminology, so at some point we are forced to make a choice.) A scheme is **regular** (or **nonsingular**) if it is regular at all points. It is **singular** (or **nonregular**) otherwise (i.e., if it is singular at *at least one* point — if it has a singularity).

12.2.A. EXERCISE. Show that a dimension 0 Noetherian local ring is regular if and only if it is a field.

You will hopefully gradually become convinced that this is the right notion of "smoothness" of schemes. Remarkably, Krull introduced the notion of a regular local ring for purely algebraic reasons, some time before Zariski realized that it was a fundamental notion in geometry in 1947.

12.2.B. EXERCISE (THE SLICING CRITERION FOR REGULARITY). Suppose X is a finite type k-scheme (such as a variety), and D is an effective Cartier divisor on X (Definition 8.4.1), and $p \in X$. Show that if p is a regular point of D then p is a regular point of X. (Hint: Krull's Principal Ideal Theorem for tangent spaces, Exercise 12.1.B.)

12.2.4. The Jacobian criterion.

A finite type k-scheme is locally of the form $\operatorname{Spec} k[x_1, \dots, x_n]/(f_1, \dots, f_r)$. The Jacobian criterion for regularity (Exercise 12.2.C) gives a hands-on method for checking for singularity at closed points, using the equations $f_1, \dots, f_r$, if $k = \overline{k}$.

12.2.C. IMPORTANT EXERCISE (THE JACOBIAN CRITERION — EASY, GIVEN EXERCISE 12.1.G). Suppose $X = \operatorname{Spec} k[x_1, \dots, x_n]/(f_1, \dots, f_r)$ has pure dimension d.

Show that a k-valued point $p \in X$ is regular if and only if the **corank** of the Jacobian matrix (12.1.6.1) (the dimension of the cokernel) at p is d.

12.2.D. EASY EXERCISE. Suppose $k = \overline{k}$. Show that the singular *closed* points of the hypersurface $f(x_1, \dots, x_n) = 0$ in $\mathbb{A}_k^n$ are given by the equations

$$f = \frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0.$$

(Translation: the singular points of $f = 0$ are where the gradient of f vanishes. This is not shocking.)

12.2.5. Smoothness over a field. There seem to be two serious drawbacks with the Jacobian criterion. For finite type schemes over $\overline{k}$, the criterion gives a necessary condition for regularity, but it is not obviously sufficient, as we need to check regularity at non-closed points as well. We can prove sufficiency by working hard to show Fact 12.8.2, which implies that the non-closed points must be regular as well. A second failing is that the criterion requires k to be algebraically closed. These problems suggest that old-fashioned ideas of using derivatives and Jacobians are ill-suited to the fancy modern notion of regularity. But that is wrong — the fault is with the concept of regularity. There is a better notion of *smoothness over a field*. Better yet, this idea generalizes to the notion of a smooth morphism of schemes (to be discussed in §12.6, and again in Chapter 25), which behaves well in all possible ways (including in the sense of §7.1.1). This is another sign that some properties we think of as of objects ("absolute notions") should really be thought of as properties of morphisms ("relative notions"). We know enough to imperfectly (but correctly) define what it means for a scheme to be k**-smooth**, or **smooth over** k.

12.2.6. *Definition.* A k-scheme is k**-smooth of dimension** d, or **smooth of dimension** d **over** k, if it is of pure dimension d, and there exists a cover by affine open sets $\operatorname{Spec} k[x_1, \dots, x_n]/(f_1, \dots, f_r)$ where the Jacobian matrix has corank d at all points. (In particular, it is locally of finite type.) A k-scheme is **smooth** over k if it is smooth of some dimension. The k is often omitted when it is clear from context.

12.2.E. EXERCISE (FIRST EXAMPLES).
(a) Show that $\mathbb{A}_k^n$ is smooth for any n and k. For which characteristics is the curve $y^2 z = x^3 - xz^2$ in $\mathbb{P}_k^2$ smooth (cf. Exercise 12.3.C)?
(b) Suppose $f \in k[x_1, \dots, x_n]$ is a polynomial such that the system of equations

$$f = \frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0$$

has no solutions in $\overline{k}$. Show that the hypersurface $f = 0$ in $\mathbb{A}_k^n$ is smooth. (Compare this to Exercise 12.2.D, which has the additional hypothesis $k = \overline{k}$.)

12.2.F. EXERCISE (SMOOTHNESS IS PRESERVED BY EXTENSION OF BASE FIELD). Suppose X is a finite type k-scheme, and $k \subset \ell$ is a field extension. Show that if X is smooth over k then $X \times_{\operatorname{Spec} k} \operatorname{Spec} \ell$ is smooth over ℓ. (The converse will be proved in Exercise 21.3.C.)

The next exercise shows that we need only check closed points, thereby making a connection to classical geometry.

12.2.G. EXERCISE. Show that if the Jacobian matrix for $X = \operatorname{Spec} k[x_1, \dots, x_n]/(f_1, \dots f_r)$ has corank d at all *closed* points, then it has corank d at *all* points. (Hint: the locus where the Jacobian matrix has corank d can be described in terms of vanishing and nonvanishing of determinants of certain explicit matrices.)

12.2.7. You can check that any open subset of a smooth k-variety is also a smooth k-variety. With what we know now, we could show that this implies that k-smoothness is equivalent to the Jacobian being corank d everywhere for *every* affine open cover (and by *any* choice of generators of the ring corresponding to such an open set). Indeed, you should feel free to do this if you cannot restrain yourself. But the cokernel of the Jacobian matrix is secretly the space of differentials (which might not be surprising if you have experience with differentials in differential geometry), so this will come for free when we give a better version of this definition in Definition 21.3.1. The current imperfect definition will suffice for us to work out examples. And if you don't want to wait until Definition 21.3.1, you can use Exercise 12.2.H below to show that if k algebraically closed, then smoothness can be checked on any open cover.

12.2.8. *In defense of regularity.* Having made a spirited case for smoothness, we should be clear that regularity is still very useful. For example, it is the only concept which makes sense in mixed characteristic. In particular, $\mathbb{Z}$ is regular at its points (it is a regular ring), and more generally, discrete valuation rings are incredibly useful examples of regular rings.

12.2.9. Regularity vs. smoothness.

12.2.H. EXERCISE. Suppose X is a finite type scheme of pure dimension d over an algebraically closed field $k = \overline{k}$. Show that X is regular at its closed points if and only if it is smooth. (We will soon learn that for finite type $\overline{k}$-schemes, regularity at closed points is the same as regularity everywhere, Theorem 12.8.3.) Hint to show regularity implies smoothness: use the Jacobian criterion to show that the corank of the Jacobian is d at the closed points of X. Then use Exercise 12.2.G.

More generally, if k is perfect (e.g. if $\operatorname{char} k = 0$ or k is a finite field), then smoothness is the same as regularity at closed points (see Exercise 21.3.D). More generally still, we will later prove the following fact. We mention it now because it will make a number of statements cleaner long before we finally prove it. (There will be no circularity.)

12.2.10. Smoothness-Regularity Comparison Theorem. —
(a) If k is perfect, every regular finite type k-scheme is smooth over k.
(b) Every smooth k-scheme is regular (with no hypotheses on perfection).

Part (a) will be proved in Exercise 21.3.D. Part (b) will be proved in §25.2.3. The fact that Theorem 12.2.10 will be proved so far in the future does not mean that we truly need to wait that long. We could prove both parts by the end of this chapter, with some work. We instead postpone the proof until we have machinery that will do much of the work for us. But if you wish some insight right away, here is the outline of the argument for (b), which you can even try to implement after reading this chapter. We will soon show (in Exercise 12.3.O) that $\mathbb{A}^n_k$ is regular for all fields k. Once we know the definition of étale morphism, we will realize that every smooth variety locally admits an étale map to $\mathbb{A}^n_k$ (Exercise 12.6.E). You can then show that an if $\pi : X \to Y$ is an étale morphism of locally Noetherian schemes, and $p \in X$, then p is regular if $\pi(p)$ is regular. This will be shown in Exercise 25.2.D, but you could reasonably do this after reading the definition of étaleness.

12.2.11. *Caution: Regularity does not imply smoothness.* If k is not perfect, then regularity does *not* imply smoothness, as demonstrated by the following example. Let $k = \mathbb{F}_p(u)$, and consider the hypersurface $X = \operatorname{Spec} k[x]/(x^p - u)$. Now $k[x]/(x^p - u)$ is a field, hence regular. But if $f(x) = x^p - u$, then $f(u^{1/p}) = \frac{df}{dx}(u^{1/p}) = 0$, so the Jacobian criterion fails — X is not smooth *over* k. (Never forget that smoothness requires a choice of field — it is a "relative" notion, and we will later define smoothness over an arbitrary scheme, in §12.6.) Technically, this argument is not yet complete: as noted in §12.2.7, we have not shown that it suffices to check the Jacobian on *any* affine cover. But as mentioned in §12.2.7, this will be rectified when we give a better definition of smoothness in Definition 21.3.1.

In case the previous example is too "small" to be enlightening (because the scheme in question *is* smooth over a *different* field, namely $k[x]/(x^p - u)$), here is another. Let $k = \mathbb{F}_p(u)$ as before, with $p > 2$, and consider the curve $\operatorname{Spec} k[x,y]/(y^2 - x^p + u)$. Then the closed point $(y, x^p - u)$ is regular but not smooth.

Thus you should not use "regular" and "smooth" interchangeably.

12.2.12. Regular local rings are integral domains.

You might expect from geometric intuition that a scheme is "locally irreducible" at a "smooth" point. Put algebraically:

12.2.13. Theorem. — *Suppose $(A, \mathfrak{m}, k)$ is a regular local ring of dimension n. Then A is an integral domain.*

Before proving it, we give some consequences.

12.2.I. EXERCISE. Suppose p is a regular point of a Noetherian scheme X. Show that only one irreducible component of X passes through p.

12.2.J. EASY EXERCISE. Show that a nonempty regular Noetherian scheme is irreducible if and only if it is connected. (Hint: Exercise 5.3.C.)

12.2.K. IMPORTANT EXERCISE (REGULAR SCHEMES IN REGULAR SCHEMES ARE REGULAR EMBEDDINGS). Suppose $(A, \mathfrak{m}, k)$ is a regular local ring of dimension n, and $I \subset A$ is an ideal of A cutting out a regular local ring of dimension d. Let $r = n - d$. Show that $\operatorname{Spec} A/I$ is a regular embedding in $\operatorname{Spec} A$. Hint: show that there are elements $f_1, \dots, f_r$ of I spanning the k-vector space $I/(I + \mathfrak{m}^2)$. Show that the quotient of A by both $(f_1, \dots, f_r)$ and I yields dimension d regular local rings. Show that a surjection of integral domains of the same dimension must be an isomorphism.

Exercise 12.2.K has the following striking geometric consequence.

12.2.L. EXERCISE (GENERALIZING EXERCISE 11.3.E). Suppose W is a regular variety of pure dimension d, and X and Y are pure-dimensional subvarieties (possibly singular) of codimension m and n respectively. Show that every component of $X \cap Y$ has codimension at most $m + n$ in W as follows. Show that the diagonal

$W \cong \Delta \subset W \times W$ is a regular embedding of codimension d. Then follow the rest of the hint to Exercise 11.3.E.

12.2.14. *Remark.* The following example shows that the regularity hypotheses in Exercise 12.2.L cannot be (completely) dropped. Let $W = \operatorname{Spec} k[w,x,y,z]/(wz - xy)$ be the cone over the smooth quadric surface, which is an integral threefold. Let X be the surface $w = x = 0$ and Y the surface $y = z = 0$; both lie in W. Then $X \cap Y$ is just the origin, so we have two codimension 1 subvarieties meeting in a codimension 3 subvariety. (It is no coincidence that X and Y are the affine cones over two lines in the same ruling, see Exercise 8.2.M. This example will arise again in Exercise 22.4.N.)

12.2.15. *Proof of Theorem 12.2.13 (following* [**Liu**, Prop. 2.11]*).* We prove the result by induction on n. The case $n = 0$ follows from Exercise 12.2.A. We now assume $n > 0$, and that we have proved the result for smaller dimension. Fix any $f \in \mathfrak{m} \backslash \mathfrak{m}^2$. By Exercise 12.1.B, the Zariski tangent space of the local ring $A/(f)$ (at its maximal ideal $\mathfrak{m}$) has dimension $n-1$. By Exercise 11.3.B, $\dim A/(f) \geq n-1$. Thus by Theorem 12.2.1, $\dim A/(f) = n-1$, and $A/(f)$ is a regular local ring. By our inductive hypothesis, $A/(f)$ is an integral domain (of dimension $n-1$).

Now choose any minimal prime ideal $\mathfrak{p}$ of A with $\dim A/\mathfrak{p} = n$. We wish to show that $\mathfrak{p} = (0)$. The Zariski cotangent space of $A/\mathfrak{p}$ is a quotient of that of A, and thus has dimension at most n. But $\dim A/\mathfrak{p} = n$, so by Theorem 12.2.1, $A/\mathfrak{p}$ is a regular local ring of dimension n. By the argument of the previous paragraph with A replaced by $A/\mathfrak{p}$, we see that $A/(\mathfrak{p}+(f))$ is an integral domain of dimension $n-1$. But it is a quotient of $A/(f)$. The only way one integral domain can be the quotient of another of the same dimension is if the quotient is an isomorphism (as discussed in the hint to Exercise 12.2.K).

Thus $\mathfrak{p} + (f) = (f)$, i.e., $\mathfrak{p} \subset fA$. Hence each element u of $\mathfrak{p}$ can be written as fv for some $v \in A$. As $f \notin \mathfrak{p}$ (as $\dim A/(\mathfrak{p}+(f)) = n-1 < n = \dim A/\mathfrak{p}$), we have $v \in \mathfrak{p}$, and so $\mathfrak{p} \subset f\mathfrak{p}$.

Clearly $f\mathfrak{p} \subset \mathfrak{p}$, so $\mathfrak{p} = f\mathfrak{p}$. Then by the second version of Nakayama's Lemma 7.2.9, $\mathfrak{p} = (0)$ as desired. $\square$

12.2.16. $\star\star$ **Checking regularity of k-schemes at closed points by base changing to $\overline{k}$.**

(We revisit these ideas using a different approach in 21.3.G, so you should read this only if you are particularly curious.) The Jacobian criterion is a great criterion for checking regularity of finite type k-schemes at k-valued points. The following result extends its applicability to more general closed points.

Suppose X is a finite type k-scheme of pure dimension n, and $p \in X$ is a closed point with residue field k'. By the Nullstellensatz 3.2.5, k'/k is a finite extension of fields; suppose that it is separable. Define $\pi : X_{\overline{k}} := X \times_k \overline{k} \to X$ by base change from $\operatorname{Spec} \overline{k} \to \operatorname{Spec} k$.

12.2.M. EXERCISE.
(a) Suppose $f(x) \in k[x]$ is a separable polynomial (i.e., f has distinct roots in $\overline{k}$), and irreducible, so $k'' := k[x]/(f(x))$ is a field extension of k. Show that $k'' \otimes_k \overline{k}$ is, as a ring, $\overline{k} \times \cdots \times \overline{k}$, where there are $\deg f = \deg k''/k$ factors.
(b) Show that $\pi^{-1}(p)$ consists of $\deg(k''/k)$ *reduced* points.

12.2.N. EXERCISE. Suppose p is a closed point of X, with residue field k' that is separable over k of degree d. Show that $X_{\overline{k}}$ is regular at all the preimages $p_1, \dots, p_d$ of p if and only if X is regular at p as follows.

(a) Reduce to the case $X = \operatorname{Spec} A$.

(b) Let $\mathfrak{m} \subset A$ be the maximal ideal corresponding to p. By tensoring the exact sequence $0 \to \mathfrak{m} \to A \to k' \to 0$ with $\overline{k}$ (field extensions preserve exactness of sequences of vector spaces), we have the exact sequence

$$0 \to \mathfrak{m} \otimes_k \overline{k} \to A \otimes_k \overline{k} \to k' \otimes_k \overline{k} \to 0.$$

Show that $\mathfrak{m} \otimes_k \overline{k} \subset A \otimes_k \overline{k}$ is the ideal corresponding to the pullback of p to $\operatorname{Spec} A \otimes_k \overline{k}$. Verify that $(\mathfrak{m} \otimes_k \overline{k})^2 = \mathfrak{m}^2 \otimes_k \overline{k}$.

(c) By tensoring the short exact sequence of k-vector spaces $0 \to \mathfrak{m}^2 \to \mathfrak{m} \to \mathfrak{m}/\mathfrak{m}^2 \to 0$ with $\overline{k}$, show that

$$\sum_{i=1}^{d} \dim_{\overline{k}} T_{X_{\overline{k}}, p_i} = d \dim_{k'} T_{X,p}.$$

(d) Use Exercise 11.1.G(a) and the inequalities $\dim_{\overline{k}} T_{X_{\overline{k}}, p_i} \geq \dim X_{\overline{k}}$ and $\dim_{k'} T_{X,p} \geq \dim X$ (Theorem 12.2.1) to conclude.

12.2.17. *Remark.* In fact, regularity at a single p_i is enough to conclude regularity at p. You can show this by following up on Exercise 12.2.N: first deal with the case when k'/k is Galois, and obtain some transitive group action of $\operatorname{Gal}(k'/k)$ on $\{p_1, \dots, p_d\}$. Another approach is given in Exercise 21.3.G.

12.2.18. *Remark.* In Exercise 21.3.D, we will use this to prove that a variety over a perfect field is smooth if and only if it is regular at all closed points (cf. the Smoothness-Regularity Comparison Theorem 12.2.10).

12.3 Examples

We now explore regularity in practice through a series of examples and exercises. Much of this discussion is secretly about smoothness rather than regularity. In particular, in order to use the Jacobian criterion, we will usually work over an algebraically closed field.

12.3.1. Geometric examples.

12.3.A. EASY EXERCISE. Suppose k is a field. Show that $\mathbb{A}^1_k$ and $\mathbb{A}^2_k$ are regular, by directly checking the regularity of all points. Show that $\mathbb{P}^1_k$ and $\mathbb{P}^2_k$ are regular. (The generalization to arbitrary dimension is a bit harder, so we leave it to Exercise 12.3.O.)

12.3.B. EXERCISE (THE EULER OR JACOBIAN TEST FOR PROJECTIVE HYPERSURFACES). Suppose $k = \overline{k}$. Show that the singular *closed* points of the hypersurface $f = 0$ in $\mathbb{P}^n_k$ correspond to the locus

$$f = \frac{\partial f}{\partial x_0} = \cdots = \frac{\partial f}{\partial x_n} = 0.$$

If the degree of the hypersurface is not divisible by char k (e.g. if char $k = 0$), show that it suffices to check $\frac{\partial f}{\partial x_0} = \cdots = \frac{\partial f}{\partial x_n} = 0$. Hint: show that $(\deg f)f = \sum_i x_i \frac{\partial f}{\partial x_i}$. (In fact, this will give the singular points in general, not just the singular closed points, cf. §12.2.5. We won't use this, so we won't prove it.)

12.3.C. EXERCISE. Suppose that $k = \overline{k}$ does not have characteristic 2. Show that $y^2 z = x^3 - xz^2$ in $\mathbb{P}^2_k$ is an irreducible regular curve. (Eisenstein's criterion gives one way of showing irreducibility. Warning: we didn't specify char $k \neq 3$, so be careful when using the Euler test.)

12.3.D. EXERCISE. Suppose $k = \overline{k}$ has characteristic not 2. Show that a quadric hypersurface in $\mathbb{P}^n$ is regular if and only if it is maximal rank. ("Maximal rank" was defined in Exercise 5.4.J.)

12.3.E. EXERCISE. Suppose $k = \overline{k}$ has characteristic 0. Show that there exists a regular (projective) plane curve of degree d. Hint: try a "Fermat curve" $x^d + y^d + z^d = 0$. (Feel free to weaken the hypotheses. Bertini's Theorem 12.4.2 will give another means of showing existence.)

12.3.F. EXERCISE (SEE FIGURE 12.3). Find all the singular closed points of the following plane curves. Here we work over $k = \overline{k}$ of characteristic 0 to avoid distractions.

(a) $y^2 = x^2 + x^3$. This is an example of a *node*.
(b) $y^2 = x^3$. This is called a *cusp*; we met it earlier in Exercise 9.7.F.
(c) $y^2 = x^4$. This is called a *tacnode*; we met it earlier in Exercise 9.7.G.

(A precise definition of a node etc. will be given in Definition 29.3.1.)

[to be made]

FIGURE 12.3. Plane curve singularities

12.3.G. EXERCISE. Suppose $k = \overline{k}$. Use the Jacobian criterion to show that the twisted cubic $\operatorname{Proj} k[w, x, y, z]/(wz - xy, wy - x^2, xz - y^2)$ is regular. (You can do this, without any hypotheses on k, using the fact that it is isomorphic to $\mathbb{P}^1$. But do this with the explicit equations, for the sake of practice. The twisted cubic was defined in Exercise 8.2.A.)

12.3.2. *Tangent planes and tangent lines.*

Suppose a scheme $X \subset \mathbb{A}^n$ is cut out by equations $f_1, \ldots, f_r$, and X is regular of dimension d at the k-valued point $a = (a_1, \ldots, a_n)$. Then the **tangent d-plane to X at p** (sometimes denoted $T_p X$) is given by the r equations

$$\left(\frac{\partial f_i}{\partial x_1}(a)\right)(x_1 - a_1) + \cdots + \left(\frac{\partial f_i}{\partial x_n}(a)\right)(x_n - a_n) = 0,$$

where (as in (12.1.6.1)) $\frac{\partial f_i}{\partial x_1}(a)$ is the evaluation of $\frac{\partial f_i}{\partial x_1}$ at $a = (a_1, \ldots, a_n)$.

12.3.H. EXERCISE. Why is this independent of the *choice* of defining equations f_1, $\ldots$, f_r of X?

The Jacobian criterion (Exercise 12.2.C) ensures that these r equations indeed cut out a d-plane. (If $d = 1$, this is called the **tangent line**.) This can be readily shown to be notion of tangent plane that we see in multivariable calculus, but note that here this is the *definition*, and thus don't have to worry about δ's and ϵ's. Instead we will have to just be careful that it behaves the way we want it to.

12.3.I. EXERCISE. Compute the tangent line to the curve of Exercise 12.3.F(b) at $(1, 1)$.

12.3.J. EXERCISE. Suppose $X \subset \mathbb{P}^n_k$ (k as usual a field) is cut out by homogeneous equations $f_1, \dots, f_r$, and $p \in X$ is a k-valued point that is regular of dimension d. Define the (projective) tangent d-plane to X at p. (Definition 8.2.3 gives the definition of a d-plane in $\mathbb{P}^n_k$, but you shouldn't need to refer there.)

12.3.3. *Side Remark to help you think cleanly.* We would want the definition of tangent k-plane to be natural in the sense that for any automorphism σ of $\mathbb{A}^n_k$ (or, in the case of the previous Exercise, $\mathbb{P}^n_k$), $\sigma(T_p X) = T_{\sigma(p)}\sigma(X)$. You could verify this by hand, but you can also say this in a cleaner way, by interpreting the equations cutting out the tangent space in a coordinate free manner. Informally speaking, we are using the canonical identification of n-space with the tangent space to n-space at p, and using the fact that the Jacobian "linear transformation" cuts out $T_p X$ in $T_p \mathbb{A}^n$ in a way independent of choice of coordinates on $\mathbb{A}^n$ or defining equations of X. Your solution to Exercise 12.3.H will help you start to think in this way.

12.3.K. EXERCISE. Suppose $X \subset \mathbb{P}^n_k$ is a degree d hypersurface cut out by $f = 0$, and L is a line not contained in X. Exercise 8.2.E (a case of Bézout's theorem) showed that X and L meet at d points, counted "with multiplicity". Suppose L meets X "with multiplicity at least 2" at a k-valued point $p \in L \cap X$, and that p is a regular point of X. Show that L is contained in the tangent plane to X at p. (Do you have a picture of this in your mind?)

12.3.4. Arithmetic examples.

12.3.L. EASY EXERCISE. Show that $\operatorname{Spec} \mathbb{Z}$ is a regular curve.

12.3.M. EXERCISE. (This tricky exercise is for those who know about the primes of the Gaussian integers $\mathbb{Z}[i]$.) There are several ways of showing that $\mathbb{Z}[i]$ is dimension 1. (For example: (i) it is a principal ideal domain; (ii) it is the normalization of $\mathbb{Z}$ in the field extension $\mathbb{Q}(i)/\mathbb{Q}$; (iii) using Krull's Principal Ideal Theorem 11.3.3 and the fact that $\dim \mathbb{Z}[x] = 2$ by Exercise 11.1.H.) Show that $\operatorname{Spec} \mathbb{Z}[i]$ is a regular curve. (There are several ways to proceed. You could use Exercise 12.1.B. As an example to work through first, consider the prime $(2, 1+i)$, which is cut out by the equations 2 and $1 + x$ in $\operatorname{Spec} \mathbb{Z}[x]/(x^2 + 1)$.) We will later (§12.5.10) have a simpler approach once we discuss discrete valuation rings.

12.3.N. EXERCISE. Show that $[(5, 5i)]$ is the unique singular point of $\operatorname{Spec} \mathbb{Z}[5i]$. (Hint: $\mathbb{Z}[i]_5 \cong \mathbb{Z}[5i]_5$. Use the previous exercise.)

12.3.5. Back to geometry: $\mathbb{A}^n_k$ is regular.

The key step to showing that $\mathbb{A}^n_k$ is regular (where k is a field) is the following.

12.3.6. Proposition. — *Suppose* $(B, \mathfrak{n}, k)$ *is a regular local ring of dimension* d*. Let* $\phi : B \to B[x]$*. Suppose* $\mathfrak{p}$ *is a prime ideal of* $A := B[x]$ *such that* $\mathfrak{n}B[x] \subset \mathfrak{p}$*. Then* $A_\mathfrak{p}$ *is a regular local ring.*

Proof. Geometrically: we have a morphism $\pi : X = \operatorname{Spec} B[x] \to Y = \operatorname{Spec} B$, and $\pi([\mathfrak{p}]) = [\mathfrak{n}]$. The fiber $\pi^{-1}([\mathfrak{n}]) = \operatorname{Spec}(B[x]/\mathfrak{n}B[x]) = \operatorname{Spec} k[x]$. Thus either (i) $\overline{[\mathfrak{p}]}$ is the fiber $\mathbb{A}^1_k$ above $[\mathfrak{n}]$, or (ii) $\overline{[\mathfrak{p}]}$ is a closed point of the fiber $\mathbb{A}^1_k$.

Before considering these two cases, we make two remarks. As $(B, \mathfrak{n})$ is a regular local ring of dimension d, $\mathfrak{n}$ is generated by d elements of B, say $f_1, \dots, f_d$ (as discussed in the proof of Theorem 12.2.1), and there is a chain of prime ideals

$$\mathfrak{q}_0 \subset \mathfrak{q}_1 \subset \cdots \subset \mathfrak{q}_d = \mathfrak{n}. \tag{12.3.6.1}$$

Case (i): $\overline{[\mathfrak{p}]}$ *is the fiber* $\mathbb{A}^1_k$. In this case, $\mathfrak{p} = \mathfrak{n}B[x]$. Thus Krull's Height Theorem 11.3.7, the height of $\mathfrak{p} = \mathfrak{n}B[x]$ is at most d, because $\mathfrak{p}$ is generated by the d elements $f_1, \dots, f_d$. But $\mathfrak{p}$ has height at least d: given the chain (12.3.6.1) ending with $\mathfrak{n}$, we have a corresponding chain of prime ideals of $B[x]$ ending with $\mathfrak{n}B[x]$. Hence the height of $\mathfrak{p}$ is precisely d, and $\mathfrak{p}A_\mathfrak{p}$ is generated by $f_1, \dots, f_d$, implying that it is a regular local ring.

Case (ii): $\overline{[\mathfrak{p}]}$ *is a closed point of the fiber* $\mathbb{A}^1_k$. The closed point $\mathfrak{p}$ of $k[x]$ corresponds to some monic irreducible polynomials $g(x) \in k[x]$. Arbitrarily lift the coefficients of g to B; we sloppily denote the resulting polynomial in $B[x]$ by $g(x)$ as well. Then $\mathfrak{p} = (f_1, \dots, f_d, g)$, so by Krull's Height Theorem 11.3.7, the height of $\mathfrak{p}$ is at most $d+1$. But $\mathfrak{p}$ has height at least $d+1$: given the chain (12.3.6.1) ending with $\mathfrak{n}$, we have a corresponding chain in $B[x]$ ending with $\mathfrak{n}B[x]$, and we can extend it by appending $\mathfrak{p}$. Hence the height of $\mathfrak{p}$ is precisely $d+1$, and $\mathfrak{p}A_\mathfrak{p}$ is generated by f_1, …, f_d, g, implying that it is a regular local ring. □

12.3.O. EXERCISE. Use Proposition 12.3.6 to show that if X is a regular (locally Noetherian) scheme, then so is $X \times \mathbb{A}^1$. In particular, show that $\mathbb{A}^n_k$ is regular.

12.4 Bertini's Theorem

We now discuss Bertini's Theorem, a fundamental classical result.

12.4.1. *Definition: dual projective space.* The **dual** (or **dual projective space**) to $\mathbb{P}^n_k$ (with coordinates x_0, …, x_n), is informally the space of hyperplanes in $\mathbb{P}^n_k$. Somewhat more precisely, it is a projective space $\mathbb{P}^n_k$ with coordinates $a_0, \dots, a_n$ (which we denote $\mathbb{P}^{n\vee}_k$ with the futile intent of preventing confusion), along with the data of the "incidence variety" or "incidence correspondence" $I \subset \mathbb{P}^n \times \mathbb{P}^{n\vee}$ cut out by the equation $a_0x_0 + \cdots + a_nx_n = 0$. Note that the k-valued points of $\mathbb{P}^{n\vee}$ indeed correspond to hyperplanes in $\mathbb{P}^n$ defined over k, and this is also clearly a duality relation (there is a symmetry in the definition between the x-variables and the a-variables). So this is concrete enough to use in practice, and extends over an arbitrary base (notably $\operatorname{Spec} \mathbb{Z}$). (But if you have a delicate and refined sensibility, you may want to come up with a coordinate-free definition.)

12.4.2. Bertini's Theorem. — *Suppose $k = \overline{k}$, and X is a smooth subvariety of $\mathbb{P}^n_k$. Then there is a nonempty (=dense) open subset U of dual projective space $\mathbb{P}^{n\vee}_k$ such that for any closed point $[H] \in U$, H doesn't contain any component of X, and the scheme $H \cap X$ is k-smooth.*

(Bertini's Theorem is often used just to show that a single H exists.)

We remark that any theorem of this flavor is often called a "Bertini Theorem". One example is the Kleiman-Bertini Theorem 25.3.8, which was *not* proved jointly by Kleiman and Bertini.

As an application of Bertini's Theorem 12.4.2, a general degree $d > 0$ hypersurface in $\mathbb{P}^n_k$ intersects X in a regular subvariety of codimension 1 in X: replace $X \hookrightarrow \mathbb{P}^n$ with the composition

$$X \longrightarrow \mathbb{P}^n \overset{\nu_d}{\hookrightarrow} \mathbb{P}^N$$

where ν_d is the dth Veronese embedding (8.2.8). Here "general" has its usual meaning in algebraic geometry, see §9.3.6, except that we are considering only *closed* points of $U \subset \mathbb{P}^{n\vee}_k$.

Exercise 25.3.D gives a useful improvement of Bertini's Theorem in characteristic 0 (see Exercise 25.3.E).

Proof. In order to keep the proof as clean as possible, we assume X is irreducible, but essentially the same proof applies in general.

The central idea of the proof is quite naive. We will describe the hyperplanes that are "bad", and show that they form a closed subset of dimension at most $n-1$ of $\mathbb{P}^{n\vee}_k$, and hence that the complement is a dense open subset. Somewhat more precisely, we will define a projective variety $Z \subset X \times \mathbb{P}^{n\vee}_k$ that can informally be described as:

$$\begin{aligned} Z \;=\; & \{(p \in X, H \subset \mathbb{P}^n_k) \;:\; p \in H, \\ & \quad \text{and either } p \text{ is a singular point of } H \cap X, \text{ or } X \subset H\} \end{aligned}$$

We will see that the projection $\pi : Z \to X$ has fibers at closed points that are projective spaces of dimension $n - 1 - \dim X$, and use this to show that $\dim Z \leq n - 1$. Thus the image of Z in $\mathbb{P}^{n\vee}_k$ will be a closed subset (Theorem 7.4.7), of dimension of at most $n - 1$, so its complement will be open and non-empty. We now put this strategy into action.

12.4.3. We first define Z more precisely, in terms of equations on $\mathbb{P}^n \times \mathbb{P}^{n\vee}$, where the coordinates on $\mathbb{P}^n$ are $x_0, \dots x_n$, and the dual coordinates on $\mathbb{P}^{n\vee}$ are $a_0, \dots, a_n$. Suppose X is cut out by $f_1, \dots, f_r$. Then we take these equations as the first of the defining equations of Z. (So far we have defined the subscheme $X \times \mathbb{P}^{n\vee}$.) We also add the equation $a_0x_0 + \cdots + a_nx_n = 0$. (So far we have described the subscheme of $\mathbb{P}^n \times \mathbb{P}^{n\vee}$ corresponding to points (p, H) where $p \in X$ and $p \in H$.) Note that the Jacobian matrix (12.1.6.1)

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_r}{\partial x_1}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n}(p) & \cdots & \frac{\partial f_r}{\partial x_n}(p) \end{pmatrix}$$

has corank equal to $\dim X$ at all closed points of X — this is precisely the Jacobian criterion for regularity (Exercise 12.2.C). We then require that the Jacobian matrix with a new column

$$\begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}$$

appended has corank $\geq \dim X$ (hence $= \dim X$). This is cut out by equations (the determinants of certain minors). By the Jacobian description of the Zariski tangent space, this condition encodes the requirement that the Zariski tangent space of $H \cap X$ at p has dimension precisely $\dim X$, which is $\dim H \cap X + 1$ (i.e., $H \cap X$ is singular at p) if H does not contain X, or if H contains X. This is precisely the notion that we wished to capture. (Remark 12.4.4 works through an example, which may help clarify how this works.)

We next show that $\dim Z \leq n-1$. For each closed point $p \in X$, let W_p be the locus of hyperplanes containing p, such that $H \cap X$ is singular at p, or else contains all of X; what is the dimension of W_p? Suppose $\dim X = d$. Then the restrictions on the hyperplanes in definition of W_p correspond to $d+1$ linear conditions. (Do you see why?) This means that W_p is a codimension $d+1$, or dimension $n-d-1$, projective space. Thus the fiber of $\pi : Z \to X$ over each closed point has pure dimension $n-d-1$. By Key Exercise 11.4.A, this implies that $\dim Z \leq n-1$. (If you wish, you can use Exercise 11.4.C to show that $\dim Z = n-1$, and you can later show that Z is a projective bundle over X, once you know what a projective bundle is. But we don't need this for the proof.) $\square$

12.4.4. *Remark.* Here is an example that may help convince you that the algebra of paragraph 12.4.3 is describing the geometry we desire. Consider the plane conic $x_0^2 - x_1^2 - x_2^2 = 0$ over a field of characteristic not 2, which you might picture as the circle $x^2 + y^2 = 1$ from the real picture in the chart U_0. Consider the point $[1,1,0]$, corresponding to $(1,0)$ on the circle. We expect the tangent line in the affine plane to be $x = 1$, which corresponds to $x_0 - x_1 = 0$. Let's see what the algebra gives us. The Jacobian matrix (12.1.6.1) is

$$\begin{pmatrix} 2x_0 \\ -2x_1 \\ -2x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix},$$

which indeed has rank 1 as expected. Our recipe asks that the matrix

$$\begin{pmatrix} 2 & a_0 \\ -2 & a_1 \\ 0 & a_2 \end{pmatrix}$$

have rank 1 (i.e., $a_0 = -a_1$ and $a_2 = 0$), and also that $a_0x_0 + a_1x_1 + a_2x_2 = 0$, which (you should check) is precisely what we wanted.

12.4.A. EXERCISE. Reword the statement of Bertini's Theorem so as to remove the $k = \overline{k}$ hypothesis, so the proof applies without change in any case where the Zariski tangent space at closed points can be computed with the Jacobian criterion. (We will find in Remark 21.3.10 that this is always the case in characteristic 0 or for varieties over finite fields.)

12.4.B. EASY EXERCISE. Prove Bertini's Theorem with the following weaker hypotheses:
(a) if X is singular in dimension 0, and
(b) if $X \to \mathbb{P}^n_k$ is a locally closed embedding.

12.4.C. EXERCISE. Continue to assume $k = \overline{k}$. Show that if X is a projective variety of dimension n and degree d in $\mathbb{P}^m$, then the intersection of X with n general hyperplanes consists of a finite number of reduced points. More precisely: if $\mathbb{P}^{m\vee}$ is the dual projective space, then there is a Zariski-open subset $U \subset (\mathbb{P}^{m\vee})^n$ such that for each closed point $(H_1, \dots, H_n)$ of U, the scheme-theoretic intersection $H_1 \cap \cdots \cap H_n \cap X$ consists of a finite number of reduced points.

12.4.5. Dual varieties. (We continue to assume $k = \overline{k}$ for convenience, although this can be relaxed.) This gives us an excuse to mention a classical construction. The image of Z (see (12.4.2.1)) in $\mathbb{P}^{n\vee}$ is called the **dual variety** of X. As $\dim Z = n-1$, we "expect" the dual of X to be a hypersurface of $\mathbb{P}^{n\vee}$. It is a nonobvious fact that this in fact is a duality: the dual of the dual of X is X itself (see [**H2**, Thm. 15.24]). The following exercise will give you some sense of the dual variety.

12.4.D. EXERCISE. Show that the dual of a hyperplane in $\mathbb{P}^n$ is the corresponding point of the dual space $\mathbb{P}^{n\vee}$. In this way, the duality between $\mathbb{P}^n$ and $\mathbb{P}^{n\vee}$ is a special case of duality between projective varieties.

12.4.E. EXERCISE. Suppose $C \subset \mathbb{P}^2$ is a smooth conic over an algebraically closed field of characteristic not 2. Show that the dual variety to C is also a smooth conic. Thus for example, through a general point in the plane (if $k = \overline{k}$), there are two tangents to C. (The points on a line in the dual plane corresponds to those lines through a point of the original plane.)

12.4.F. $\star$ EXERCISE (THERE IS ONE CONIC TANGENT TO FIVE GENERAL LINES, AND GENERALIZATIONS). Continuing the notation of the previous problem, show that the number of conics C containing i generally chosen points and tangent to $5-i$ generally chosen lines is $1, 2, 4, 4, 2, 1$ respectively for $i = 0, 1, 2, 3, 4, 5$. You might interpret the symmetry of the sequence in terms of the duality between the conic and the dual conic. This fact was likely known in the paleolithic era.

12.5 Discrete valuation rings: Dimension 1 Noetherian regular local rings

The case of (co)dimension 1 is important, because if you understand how primes behave that are separated by dimension 1, then you can use induction to prove facts in arbitrary dimension. This is one reason why Krull's Principal Ideal Theorem 11.3.3 is so useful.

A dimension 1 Noetherian regular local ring can be thought of as a "germ of a smooth curve" (see Figure 12.4). Two examples to keep in mind are $k[x]_{(x)} = \{f(x)/g(x) \ : \ x \not| \ g(x)\}$ and $\mathbb{Z}_{(5)} = \{a/b \ : \ 5 \not| \ b\}$. The first example is "geometric" and the second is "arithmetic", but hopefully it is clear that they have something fundamental in common.

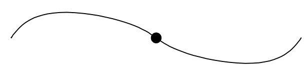


FIGURE 12.4. A germ of a curve

The purpose of this section is to give a long series of equivalent definitions of these rings. Before beginning, we quickly sketch these seven definitions. There are a number of ways a Noetherian local ring can be "nice". It can be regular, or a principal domain, or a unique factorization domain, or normal. In dimension 1, these are the same. Also equivalent are nice properties of ideals: if $\mathfrak{m}$ is principal; or if *all* ideals are either powers of the maximal ideal, or 0. Finally, the ring can have a *discrete valuation*, a measure of "size" of elements that behaves particularly well.

12.5.1. Theorem. — *Suppose $(A, \mathfrak{m})$ is a Noetherian local ring of dimension 1. Then the following are equivalent.*

(a) *$(A, \mathfrak{m})$ is regular.*
(b) *$\mathfrak{m}$ is principal.*

Proof. Here is why (a) implies (b). If A is regular, then $\mathfrak{m}/\mathfrak{m}^2$ is one-dimensional. Choose any element $t \in \mathfrak{m} - \mathfrak{m}^2$. Then t generates $\mathfrak{m}/\mathfrak{m}^2$, so generates $\mathfrak{m}$ by Nakayama's Lemma 7.2.H ($\mathfrak{m}$ is finitely generated by the Noetherian hypothesis). We call such an element a **uniformizer**.

Conversely, if $\mathfrak{m}$ is generated by one element t over A, then $\mathfrak{m}/\mathfrak{m}^2$ is generated by one element t over $A/\mathfrak{m} = k$. Since $\dim_k \mathfrak{m}/\mathfrak{m}^2 \geq 1$ by Theorem 12.2.1, we have $\dim_k \mathfrak{m}/\mathfrak{m}^2 = 1$, so $(A, \mathfrak{m})$ is regular. □

We will soon use a useful fact, which is geometrically motivated, and is a special case of an important result, the Artin-Rees Lemma 12.9.3. We will prove it in §12.9.

12.5.2. Proposition. — *If $(A, \mathfrak{m})$ is a Noetherian local ring, then $\cap_i \mathfrak{m}^i = 0$.*

12.5.3. The geometric intuition for this is that any function that is analytically zero at a point (vanishes to all orders) actually vanishes in a neighborhood of that point. (Exercise 12.9.B will make this precise.) The geometric intuition also suggests an example showing that Noetherianness is necessary: consider the function e^{-1/x^2} in the germs of C^∞-functions on $\mathbb{R}$ at the origin.

It is tempting to argue that

$$\mathfrak{m}(\cap_i \mathfrak{m}^i) = \cap_i \mathfrak{m}^i, \tag{12.5.3.1}$$

and then to use Nakayama's Lemma 7.2.9 to argue that $\cap_i \mathfrak{m}^i = 0$. Unfortunately, it is not obvious that this first equality is true: product does not commute with infinite descending intersections in general. (Aside: product also doesn't commute with finite intersections in general, as for example in $k[x, y, z]/(xz - yz)$, $z((x) \cap (y)) \neq (xz) \cap (yz)$.) We will establish Proposition 12.5.2 in Exercise 12.9.A(b). (We could do it directly right now without too much effort.)

12.5.4. Theorem. — *Suppose $(A, \mathfrak{m})$ is a Noetherian local ring of dimension 1. Then (a) and (b) are equivalent to:*

(c) *all ideals are of the form $\mathfrak{m}^n$ (for $n \geq 0$) or (0).*

Proof. Assume (a): suppose $(A, \mathfrak{m}, k)$ is a Noetherian regular local ring of dimension 1. Then I claim that $\mathfrak{m}^n \neq \mathfrak{m}^{n+1}$ for any n. Otherwise, by Nakayama's Lemma, $\mathfrak{m}^n = 0$, from which $t^n = 0$. But A is an integral domain (by Theorem 12.2.13), so $t = 0$, from which $A = A/\mathfrak{m}$ is a field, which doesn't have dimension 1, contradiction.

I next claim that $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ is dimension 1. Reason: $\mathfrak{m}^n = (t^n)$. So $\mathfrak{m}^n$ is generated as an A-module by one element, and $\mathfrak{m}^n/(\mathfrak{m}\mathfrak{m}^n)$ is generated as a $(A/\mathfrak{m} = k)$-module by 1 element (nonzero by the previous paragraph), so it is a one-dimensional vector space.

So we have a chain of ideals $A \supset \mathfrak{m} \supset \mathfrak{m}^2 \supset \mathfrak{m}^3 \supset \cdots$ with $\cap \mathfrak{m}^i = (0)$ (Proposition 12.5.2). We want to say that there is no room for any ideal besides these, because each pair is "separated by dimension 1", and there is "no room at the end". Proof: suppose $I \subset A$ is an ideal. If $I \neq (0)$, then there is some n such that $I \subset \mathfrak{m}^n$ but $I \not\subset \mathfrak{m}^{n+1}$. Choose some $u \in I - \mathfrak{m}^{n+1}$. Then $(u) \subset I$. But u generates $\mathfrak{m}^n/\mathfrak{m}^{n+1}$, hence by Nakayama it generates $\mathfrak{m}^n$, so we have $\mathfrak{m}^n \subset I \subset \mathfrak{m}^n$, so we are done: (c) holds.

We now show that (c) implies (a). Assume (a) does not hold: suppose we have a dimension 1 Noetherian local integral domain that is not regular, so $\mathfrak{m}/\mathfrak{m}^2$ has dimension at least 2. Choose any $u \in \mathfrak{m} - \mathfrak{m}^2$. Then $(u, \mathfrak{m}^2)$ is an ideal, but $\mathfrak{m}^2 \subsetneq (u, \mathfrak{m}^2) \subsetneq \mathfrak{m}$. ∎

12.5.A. EASY EXERCISE. Suppose $(A, \mathfrak{m})$ is a Noetherian dimension 1 local ring. Show that (a)–(c) above are equivalent to:

(d) A *is a principal ideal domain.*

12.5.5. Discrete valuation rings. We next define the notion of a discrete valuation ring. Suppose K is a field. A **discrete valuation** on K is a **surjective homomorphism** $\nu : K^\times \to \mathbb{Z}$ (in particular, $\nu(xy) = \nu(x) + \nu(y)$) satisfying

$$\nu(x + y) \geq \min(\nu(x), \nu(y))$$

except if $x + y = 0$ (in which case the left side is undefined). (Such a valuation is called *non-archimedean*, although we will not use that term.) It is often convenient to say $\nu(0) = \infty$. More generally, a **valuation** is a surjective homomorphism $\nu : K^\times \to G$ to a totally ordered abelian group G, although this isn't so important to us.

Here are three key examples.

(i) *(the 5-adic valuation)* $K = \mathbb{Q}$, $\nu(r)$ is the "power of 5 appearing in r", e.g. $\nu(35/2) = 1$, $\nu(27/125) = -3$.

(ii) $K = k(x)$, $\nu(f)$ is the "power of x appearing in f."

(iii) $K = k(x)$, $\nu(f)$ is the negative of the degree. This is really the same as (ii), with x replaced by $1/x$.

Then $0 \cup \{x \in K^\times : \nu(x) \geq 0\}$ is a ring, which we denote $\mathscr{O}_\nu$. It is called the **valuation ring** of ν. (Not every valuation is discrete. Consider the ring of *Puisseux series* over a field k, $K = \cup_{n \geq 1} k((x^{1/n}))$, with $\nu : K^\times \to \mathbb{Q}$ given by $\nu(x^q) = q$.)

12.5.B. EXERCISE. Describe the valuation rings in the three examples (i)–(iii) above. (You will notice that they are familiar-looking dimension 1 Noetherian local rings. What a coincidence!)

12.5.C. EXERCISE. Show that $\{0\} \cup \{x \in K^\times : \nu(x) \geq 1\}$ is the unique maximal ideal of the valuation ring. (Hint: show that everything in the complement is invertible.) Thus the valuation ring is a local ring.

An integral domain A is called a **discrete valuation ring** (or **DVR**) if there exists a discrete valuation ν on its fraction field $K = K(A)$ for which $\mathscr{O}_\nu = A$. Similarly, A is a **valuation ring** if there exists a valuation ν on K for which $\mathscr{O}_\nu = A$.

Now if A is a Noetherian regular local ring of dimension 1, and t is a uniformizer (a generator of $\mathfrak{m}$ as an ideal, or equivalently of $\mathfrak{m}/\mathfrak{m}^2$ as a k-vector space) then any nonzero element r of A lies in some $\mathfrak{m}^n - \mathfrak{m}^{n+1}$, so $r = t^n u$ where u is invertible (as t^n generates $\mathfrak{m}^n$ by Nakayama, and so does r), so $K(A) = A_t = A[1/t]$. So *any element of* $K(A)^\times$ *can be written uniquely as* ut^n where u is invertible and $n \in \mathbb{Z}$. Thus we can define a valuation ν by $\nu(ut^n) = n$.

12.5.D. EXERCISE. Show that ν is a discrete valuation.

12.5.E. EXERCISE. Conversely, suppose $(A, \mathfrak{m})$ is a discrete valuation ring. Show that $(A, \mathfrak{m})$ is a Noetherian regular local ring of dimension 1. (Hint: Show that the ideals are all of the form (0) or $I_n = \{r \in A : \nu(r) \geq n\}$, and (0) and I_1 are the only primes. Thus we have Noetherianness, and dimension 1. Show that I_1/I_2 is generated by the image of any element of $I_1 - I_2$.)

Hence we have proved:

12.5.6. Theorem. — *An integral domain* A *is a Noetherian local ring of dimension* 1 *satisfying (a)–(d) if and only if*

(e) A *is a discrete valuation ring.*

12.5.F. EXERCISE. Show that there is only one discrete valuation on a discrete valuation ring.

12.5.7. *Definition.* Thus any Noetherian regular local ring of dimension 1 comes with a unique valuation on its fraction field. If the valuation of an element is $n > 0$, we say that the element has a **zero of order** n. If the valuation is $-n < 0$, we say that the element has a **pole of order** n. We will come back to this shortly, after dealing with (f) and (g).

12.5.8. Theorem. — *Suppose* $(A, \mathfrak{m})$ *is a Noetherian local ring of dimension* 1*. Then (a)–(e) are equivalent to:*

(f) A *is a unique factorization domain, and*
(g) A *is an integral domain, integrally closed in its fraction field* $K = K(A)$.

Proof. (a)–(e) clearly imply (f), because we have the following stupid unique factorization: each nonzero element of A can be written uniquely as ut^n where $n \in \mathbb{Z}^{\geq 0}$ and u is invertible.

Now (f) implies (g), because unique factorization domains are integrally closed in their fraction fields (Exercise 5.4.F).

It remains to check that (g) implies (a)–(e). We will show that (g) implies (b).

Suppose $(A, \mathfrak{m})$ is a Noetherian local integral domain of dimension 1, integrally closed in its fraction field $K = K(A)$. Choose any nonzero $r \in \mathfrak{m}$. Then $S = A/(r)$ is a Noetherian local ring of dimension 0 — its only prime is the image of $\mathfrak{m}$, which we denote $\mathfrak{n}$ to avoid confusion. Then $\mathfrak{n}$ is finitely generated, and each generator is nilpotent (the intersection of all the prime ideals in any ring are the nilpotents, Theorem 3.2.12). Then $\mathfrak{n}^N = 0$, where N is sufficiently large. Hence there is some n such that $\mathfrak{n}^n = 0$ but $\mathfrak{n}^{n-1} \neq 0$.

Now comes the crux of the argument. Thus in A, $\mathfrak{m}^n \subset (r)$ but $\mathfrak{m}^{n-1} \not\subset (r)$. Choose $s \in \mathfrak{m}^{n-1} - (r)$. Consider $s/r \in K(A)$. As $s \notin (r)$, $s/r \notin A$, so as A is integrally closed, s/r is not integral over A.

Now $\frac{s}{r}\mathfrak{m} \not\subset \mathfrak{m}$ (or else $\frac{s}{r}\mathfrak{m} \subset \mathfrak{m}$ would imply that $\mathfrak{m}$ is a faithful $A[\frac{s}{r}]$-module, contradicting Exercise 7.2.I). But $s\mathfrak{m} \subset \mathfrak{m}^n \subset rA$, so $\frac{s}{r}\mathfrak{m} \subset A$. Thus $\frac{s}{r}\mathfrak{m} = A$, from which $\mathfrak{m} = \frac{r}{s}A$, so $\mathfrak{m}$ is principal. $\square$

12.5.9. Geometry of normal Noetherian schemes. We can finally make precise (and generalize) the fact that the function $(x-2)^2x/(x-3)^4$ on $\mathbb{A}^1_{\mathbb{C}}$ has a double zero at $x = 2$ and a quadruple pole at $x = 3$. Furthermore, we can say that 75/34 has a double zero at 5, and a single pole at 2. (What are the zeros and poles of $x^3(x+y)/(x^2+xy)^3$ on $\mathbb{A}^2$?) Suppose X is a locally Noetherian scheme. Then for any regular codimension 1 point p (i.e., any point p where $\mathscr{O}_{X,p}$ is a regular local ring of dimension 1), we have a discrete valuation ν. If f is any nonzero element of the fraction field of $\mathscr{O}_{X,p}$ (e.g. if X is integral, and f is a nonzero element of the function field of X), then if $\nu(f) > 0$, we say that the element has a **zero of order** $\nu(f)$ **at** p, and if $\nu(f) < 0$, we say that the element has a **pole of order** $-\nu(f)$ **at** p. (We are not yet allowed to discuss order of vanishing at a point that is not regular and codimension 1. One can make a definition, but it doesn't behave as well as it does when have you have a discrete valuation.)

12.5.G. EXERCISE (FINITENESS OF ZEROS AND POLES ON NOETHERIAN SCHEMES). Suppose X is an integral Noetherian scheme, and $f \in K(X)^{\times}$ is a nonzero element of its function field. Show that f has a finite number of zeros and poles. (Hint: reduce to $X = \operatorname{Spec} A$. If $f = f_1/f_2$, where $f_i \in A$, prove the result for f_i.)

Suppose A is a Noetherian integrally closed domain. Then it is **regular in codimension** 1 (translation: its points of codimension at most 1 are regular). If A is dimension 1, then obviously A is regular.

12.5.H. EXERCISE. If f is a nonzero rational function on a locally Noetherian normal scheme, and f has no poles, show that f is regular. (Hint: Algebraic Hartogs's Lemma 11.3.10.)

12.5.10. For example (cf. Exercise 12.3.M), $\operatorname{Spec} \mathbb{Z}[i]$ is regular, because it is dimension 1, and $\mathbb{Z}[i]$ is a unique factorization domain. Hence $\mathbb{Z}[i]$ is normal, so all its closed (codimension 1) points are regular. Its generic point is also regular, as $\mathbb{Z}[i]$ is an integral domain.

12.5.11. *Remark.* A (Noetherian) scheme can be singular in codimension 2 and still be normal. For example, you have shown that the cone $x^2 + y^2 = z^2$ in $\mathbb{A}^3$ in characteristic not 2 is normal (Exercise 5.4.I(b)), but it is singular at the origin (the Zariski tangent space is visibly three-dimensional).

But singularities of normal schemes are not so bad in some ways: we have Algebraic Hartogs's Lemma 11.3.10 for Noetherian normal schemes, which states that you can extend functions over codimension 2 sets.

12.5.12. *Remark.* We know that for Noetherian rings we have implications

$$\text{unique factorization domain} \implies \text{integrally closed} \implies \text{regular in codimension 1.}$$

Hence for locally Noetherian schemes, we have similar implications:

$$\text{factorial} \implies \text{normal} \implies \text{regular in codimension 1.}$$

Here are two examples to show you that these inclusions are strict.

12.5.I. EXERCISE (THE **crumpled plane**). Let A be the subring $k[x^3, x^2, xy, y]$ of $k[x, y]$. (Informally, we allow all polynomials that don't include a nonzero multiple of the monomial x.) Show that $\operatorname{Spec} k[x, y] \to \operatorname{Spec} A$ is a normalization. Show that A is not integrally closed. Show that $\operatorname{Spec} A$ is regular in codimension 1. (Hint for the last part: show it is dimension 2, and when you throw out the origin you get something regular, by inverting x^2 and y respectively, and considering A_{x^2} and A_y.)

12.5.13. *Example.* Suppose k is algebraically closed of characteristic not 2. Then $k[w, x, y, z]/(wz - xy)$ is integrally closed, but not a unique factorization domain, see Exercise 5.4.L (and Exercise 12.1.E).

12.5.14. *Aside: Dedekind domains.* A **Dedekind domain** is a Noetherian integral domain of dimension at most one that is normal (integrally closed in its fraction field). The localization of a Dedekind domain at any prime but (0) (i.e., a codimension one prime) is hence a discrete valuation ring. This is an important notion, but we won't use it much. Rings of integers of number fields are examples, see §9.7.1. In particular, if n is a square-free integer congruent to $3 \pmod 4$, then $\mathbb{Z}[\sqrt{n}]$ is a Dedekind domain, by Exercise 5.4.I(a). If you wish you can prove *unique factorization of ideals in a Dedekind domain*: any nonzero ideal in a Dedekind domain can be uniquely factored into prime ideals.

12.5.15. *Final remark: Finitely generated modules over a discrete valuation ring.* We record a useful fact for future reference. Recall that finitely generated modules over a principal ideal domain are finite direct sums of cyclic modules (see for example [**DF**, §12.1, Thm. 5]). Hence any finitely generated module over a discrete valuation ring A with uniformizer t is a finite direct sum of terms A and $A/(t^r)$ (for various r). See Proposition 13.7.3 for an immediate consequence.

12.6 Smooth (and étale) morphisms (first definition)

In §12.2.5, we defined smoothness over a field. It is an imperfect definition (see §12.2.7), and we will improve it when we know more (in §21.3.1).

We know enough to define *smooth morphisms* (and, as a special case, *étale morphisms*) in general. Our definition will be imperfect in a number of ways. For example, it will look a little surprising and unmotivated. As another sign, it will not be obvious that it actually generalizes smoothness over a field. As a third sign, like Definition 12.2.6, because it is of the form "there exist open covers satisfying some property", it is poorly designed to be used to show that some morphism is *not* smooth. But it has a major advantage that we can give the definition right now, and the basic properties of smooth morphisms will be straightforward to show.

12.6.1. *Differential geometric motivation.* The notion of a smooth morphism is motivated by the following idea from differential geometry. For the purposes of this discussion, we say a map $\pi : X \to Y$ of real manifolds (or real analytic spaces, or any variant thereof) is *smooth of relative dimension* n if locally (on X) it looks like $Y' \times \mathbb{R}^n \to Y'$. Translation: it is "locally on the source a smooth fibration". (If you have not heard the word fibration before, don't worry. Informally speaking, a *fibration* in some "category of geometric objects" is a map $\pi : X \to Y$ that locally on the source is a product $V \times W \to V$; in some sort of category of real manifolds, a fibration is a "locally on the source a smooth fibration" if the W in the local description can be taken to be $\mathbb{R}^a$ for some a.)

In particular, "smooth of relative dimension 0" is the same as the important notion of "local isomorphism" ("isomorphism locally on the source" — not quite the same as "covering space"). Carrying this idea too naively into algebraic geometry leads to the "wrong" definition (see Exercise 12.6.G). An explanation of how this differential geometric intuition leads to a good definition of smoothness will be given in §25.1, when we will be ready to *give* a good definition of smoothness. For now we will settle for a *satisfactory* (and correct) definition.

12.6.2. *Definition.* A morphism $\pi : X \to Y$ is **smooth of relative dimension** n if there exist open covers $\{U_i\}$ of X and $\{V_i\}$ of Y, with $\pi(U_i) \subset V_i$, such that for every i we have a commutative diagram

$$\begin{tikzcd} U_i \arrow[r, "\sim"] \arrow[d, "\pi|_{U_i}"'] & W \arrow[d, "\rho|_W"] \\ V_i \arrow[r, "\sim"] & \operatorname{Spec} B \end{tikzcd}$$

where there is a morphism $\rho : \operatorname{Spec} B[x_1, \dots, x_{n+r}]/(f_1, \dots, f_r) \to \operatorname{Spec} B$ (induced by the obvious map of rings in the opposite direction), and W is an open subscheme of $\operatorname{Spec} B[x_1, \dots, x_{n+r}]/(f_1, \dots, f_r)$, such that at every point of W the determinant of the Jacobian matrix of the f_i's with respect to the *first* r x_i's,

$$\det \left(\frac{\partial f_j}{\partial x_i} \right)_{i,j \leq r}, \tag{12.6.2.1}$$

is an invertible (= nowhere zero) function on W. From the definition, smooth morphisms are locally of finite presentation (hence locally of finite type). Also, $\mathbb{A}^n_B \to \operatorname{Spec} B$ is immediately seen to be smooth of relative dimension n.

Étale means **smooth of relative dimension** 0.

As easy examples, open embeddings are étale, and the projection $\mathbb{A}^n \times Y \to Y$ is smooth of relative dimension n. You may already have some sense of what makes this definition imperfect. For example, how do you know a morphism is smooth if it isn't given to you in a form where the "right" variables x_i and "right" equations are clear?

12.6.A. MOTIVATING EXERCISE (FOR THOSE WITH DIFFERENTIAL-GEOMETRIC BACKGROUND). Show how Definition 12.6.2 gives the definition in differential geometry, as described in §12.6.1. (Your argument will use the implicit function theorem, which "doesn't work" in algebraic geometry, see Exercise 12.6.G.)

12.6.B. EXERCISE. Show that the notion of smoothness of relative dimension n is local on both the source and target.

We can thus make sense of the phrase "$\pi : X \to Y$ is smooth of relative dimension n at $p \in X$": it means that there is an open neighborhood U of p such that $\pi|_U$ is smooth of relative dimension n.

The phrase **smooth morphism** (without reference to relative dimension n) often informally means "smooth morphism of some relative dimension n", but sometimes can mean "smooth of some relative dimension in the neighborhood of every point").

12.6.C. EASY EXERCISE. Show that the notion of smoothness of relative dimension n is preserved by base change.

12.6.D. EXERCISE. Suppose $\pi : X \to Y$ is smooth of relative dimension m and $\rho : Y \to Z$ is smooth of relative dimension n. Show that $\rho \circ \pi$ is smooth of relative dimension $m + n$. (In particular, the composition of étale morphisms is étale. Furthermore, this implies that smooth morphisms are closed under product, by Exercise 9.4.F.)

Exercises 12.6.B–12.6.D imply that smooth morphisms (and étale morphisms) form a "reasonable" class in the informal sense of §7.1.1.

12.6.E. EXERCISE. Suppose $\pi : X \to Y$ is smooth of relative dimension n. Show that locally on X, π can be described as an étale cover of $\mathbb{A}^n_Y$. More precisely, show that for every $p \in X$, there is a neighborhood U of p, such that $\pi|_U$ can be factored into

$$U \xrightarrow{\alpha} \mathbb{A}^n_Y \xrightarrow{\beta} Y$$

where α is étale and β is the obvious projection.

12.6.F. EXERCISE. Suppose $\pi : X \to Y$ is a morphism of schemes. Show that the locus on X where π is smooth of relative dimension n is open. (In particular, the locus where π is étale is open.)

12.6.G. EXERCISE. Suppose k is a field of characteristic not 2. Let $Y = \operatorname{Spec} k[t]$, and $X = \operatorname{Spec} k[u, 1/u]$. Show that the morphism $\pi : X \to Y$ induced by $t \mapsto u^2$ is étale. (Sketch π!) Show that there is no nonempty open subset U of X on which π is an isomorphism. (This shows that étale cannot be defined as "local isomorphism", i.e., "isomorphism locally on the source". In particular, the naive extension of the differential intuition of §12.6.1 does not give the right definition of étaleness algebraic geometry.)

We now show that our new Definition 12.6.2 specializes to Definition 12.2.6 when the target is a field. This is the hardest fact in this section. (It is also a consequence of Theorem 25.2.2, so you could skip the proof of Theorem 12.6.3 and take it on faith until Chapter 25. The crux of the proof of Theorem 25.2.2 is the similar the crux of the proof of Theorem 12.6.3.)

12.6.3. Theorem. — *Suppose X is a k-scheme. Then the following are equivalent.*

(i) *X is smooth of relative dimension n over $\operatorname{Spec} k$ (Definition 12.6.2).*
(ii) *X has pure dimension n, and is smooth over k (in the sense of Definition 12.2.6).*

This allows us to use the phrases "smooth over k" and "smooth over $\operatorname{Spec} k$" interchangeably, as we would expect. More generally, **"smooth over a ring A"** means "smooth over $\operatorname{Spec} A$".

$\star$ *Proof. (i) implies (ii).* Suppose X is smooth of relative dimension n over $\operatorname{Spec} k$ (Definition 12.6.2). We will show that X has pure dimension n, and is smooth over k (Definition 12.2.6). Our goal is a Zariski-local statement, so we may assume that X is an open subset

$$W \subset \operatorname{Spec} k[x_1, \dots, x_{n+r}]/(f_1, \dots, f_r)$$

as described in Definition 12.6.2.

The central issue is showing that X has "pure dimension n", as the latter half of the statement would then follow because the corank of the Jacobian matrix is then n.

Let $X_{\overline{k}} := X \times_k \overline{k}$. By Exercise 11.1.G, $X_{\overline{k}}$ is pure dimension n if and only if X has pure dimension n. Let p be a closed point of the (finite type scheme) $X_{\overline{k}}$. Then X has dimension at least n at p, by Krull's Principal Ideal Theorem 11.3.3, combined with Theorem 11.2.9 (codimension is the difference of dimensions for varieties). But the Zariski tangent space at a closed point of a finite type scheme over $\overline{k}$ is cut out by the Jacobian matrix (Exercise 12.1.G), and thus has dimension exactly n (by the corank hypothesis in the definition of k-smooth). As the dimension of the Zariski tangent space bounds the dimension of the scheme (Theorem 12.2.1), $\dim X_{\overline{k}} = n$ as desired.

(ii) implies (i). We will prove the result near any given closed point p of X. (Here we use that any union of neighborhoods of all the closed points is all of X, by Exercise 5.1.E.) We thus may reduce to the following problem. Suppose we are given $X := \operatorname{Spec} k[x_1, \dots, x_{n+r}]/(f_1, \dots, f_N)$, and a point $p \in X$, and an open neighborhood U of p in X such that U is pure dimension n, and the Jacobian matrix of the f_j's with respect to the x_i's has corank n everywhere on U. We wish to show that the structure morphism $\pi : X \to \operatorname{Spec} k$ is smooth of relative dimension n at p (in the sense of Definition 12.6.2). At p, there exists an $r \times r$ submatrix of the Jacobian matrix that is invertible (by the corank condition). Rearrange the x_i's and f_j's so that the Jacobian of the first r of the f_j's with respect to the first r of the x_i's has nonzero determinant. Let $I = (f_1, \dots, f_N)$. We will show that $I = (f_1, \dots, f_r)$ "near p", or more precisely, that there is some $g \in k[x_1, \dots, x_{n+r}]$ with $g(p) \neq 0$ such that $I_g = (f_1, \dots, f_r)_g$ (so $I = (f_1, \dots, f_r)$ "in $D(g)$").

Choose a point $\overline{p} \in X_{\overline{k}}$ above p, i.e., such that if ψ is the morphism $X_{\overline{k}} \to X$ induced by $\operatorname{Spec} \overline{k} \to \operatorname{Spec} k$, then $\psi(\overline{p}) = p$. (Why does such a $\overline{p}$ exist?) Now $X_{\overline{k}}$ is $\overline{k}$-smooth (in the sense of Definition 12.2.6, by Exercise 12.2.F) at $\overline{p}$, hence regular (Exercise 12.2.H). Its Zariski tangent space at $\overline{p}$ is cut out in the tangent space of $\mathbb{A}^N_{\overline{k}}$ by the Jacobian conditions (Exercise 12.1.G), and hence cut out by $f_1, \dots, f_r$.

12.6.H. EXERCISE. Suppose $\overline{\mathfrak{m}}$ is the ideal of $\overline{k}[x_1, \dots, x_{n+r}]$ corresponding to $\overline{p}$. Show that $f_1, \dots, f_r$ generate the ideal $(f_1, \dots, f_N)$ in $\overline{k}[x_1, \dots, x_{n+r}]_{\overline{\mathfrak{m}}}$. Hint: look at your argument for Exercise 12.2.K.

For notational compactness, define $A := k[x_1, \dots, x_{n+r}]/(f_1, \dots, f_r)$, and $\overline{A} := \overline{k}[x_1, \dots, x_{n+r}]/(f_1, \dots, f_r) = A \otimes_k \overline{k}$. Let I (resp. $\overline{I}$) be the ideal of A (resp. $\overline{A}$) generated by $f_{r+1}, \dots, f_N$ (so $\overline{I} = I \otimes_k \overline{k}$). We wish to show that $I = 0$ "near p". Here $p = [\mathfrak{m}]$, where $\mathfrak{m}$ is a maximal ideal (recall p is a closed point).

Exercise 12.6.H means that $\overline{I}_{\overline{\mathfrak{m}}} = 0$, so taking the quotient by the maximal ideal $\overline{\mathfrak{m}}$ of $\overline{A}$ yields

$$I \otimes_A (\overline{A}/\overline{\mathfrak{m}}) = 0$$

which implies

$$(I \otimes_A (A/\mathfrak{m})) \otimes_{A/\mathfrak{m}} (\overline{A}/\overline{\mathfrak{m}}) = 0.$$

But $\overline{k} = \overline{A}/\overline{\mathfrak{m}}$ is a field extension of $\ell := A/\mathfrak{m}$, and an ℓ-vector space V is 0 if (and only if) $V \otimes_\ell \overline{k} = 0$, so $I \otimes_A (A/\mathfrak{m}) = 0$, i.e., $I_\mathfrak{m}/\mathfrak{m}I_\mathfrak{m} = 0$.

12.6.I. EXERCISE. Conclude that there is some $g \in A \setminus \mathfrak{m}$ such that $gI = 0$. Hint: Nakayama, and the finite generation of I. (Be sure you understand why the proof is now complete.)

□

12.7 ⋆ Valuative criteria for separatedness and properness

(The only reason this section is placed here is that we need the theory of discrete valuation rings.)

In reasonable circumstances, it is possible to verify separatedness by checking only maps from spectra of valuations rings. There are four reasons you might like this (even if you never use it). First, it gives useful intuition for what separated morphisms look like. Second, given that we understand schemes by maps to them (the Yoneda philosophy), we might expect to understand morphisms by mapping certain maps of schemes to them, and this is how you can interpret the diagram appearing in the valuative criterion. And the third concrete reason is that one of the two directions in the statement is much easier (a special case of the Reduced-to-Separated Theorem 10.2.2, see Exercise 12.7.A), and this is the direction we will repeatedly use. Finally, the criterion is very useful!

Similarly, there is a valuative criterion for properness.

In this section, we will meet the valuative criteria, but aside from outlining the proof of one result (the DVR version of the valuative criterion of separatedness), we will not give proofs, and satisfy ourselves with references. There are two reasons for this controversial decision. First, the proofs require the development of some commutative algebra involving valuation rings that we will not otherwise need. Second, we will not use these results in any essential way later in this book.

We begin with a valuative criterion for separatedness that applies in a case that will suffice for the interests of most people, that of finite type morphisms of Noetherian schemes. We will then give a more general version for more general readers.

12.7.1. Theorem (Valuative criterion for separatedness, DVR version). — *Suppose $\pi : X \to Y$ is a morphism of finite type of locally Noetherian schemes. Then π is separated if and only if the following condition holds: for any discrete valuation ring A, and any diagram of the form*

(12.7.1.1)
$$\begin{tikzcd} \operatorname{Spec} K(A) \arrow[r] \arrow[d, hook, "\text{open emb.}"'] & X \arrow[d, "\pi"] \\ \operatorname{Spec} A \arrow[r] & Y \end{tikzcd}$$

(where the vertical morphism on the left corresponds to the inclusion $A \hookrightarrow K(A)$), there is at most one morphism $\operatorname{Spec} A \to X$ *such that the diagram*

(12.7.1.2)
$$\begin{tikzcd} \operatorname{Spec} K(A) \arrow[r] \arrow[d, hook, "\text{open emb.}"'] & X \arrow[d, "\pi"] \\ \operatorname{Spec} A \arrow[r] \arrow[ur, dotted, "\leq 1"] & Y \end{tikzcd}$$

commutes.

The idea behind the proof is explained in §12.7.3. We can show one direction right away, in the next exercise.

12.7.A. EXERCISE (THE EASY DIRECTION). Use the Reduced-to-Separated Theorem 10.2.2 to prove one direction of the theorem: that if π is separated, then the valuative criterion holds.

12.7.B. EXERCISE. Suppose X is an irreducible Noetherian separated curve. If $p \in X$ is a regular closed point, then $\mathscr{O}_{X,p}$ is a discrete valuation ring, so each regular point yields a discrete valuation on $K(X)$. Use the previous exercise to show that distinct points yield distinct discrete valuations.

Here is the intuition behind the valuative criterion (see Figure 12.5). We think of Spec of a discrete valuation ring A as a "germ of a curve", and $\operatorname{Spec} K(A)$ as the "germ minus the origin" (even though it is just a point!). Then the valuative criterion says that if we have a map from a germ of a curve to Y, and have a lift of the map away from the origin to X, then there is at most one way to lift the map from the entire germ. In the case where Y is the spectrum of a field, you can think of this as saying that limits of one-parameter families are unique (if they exist).

For example, this captures the idea of what is wrong with the map of the line with the doubled origin over k (Figure 12.6): we take $\operatorname{Spec} A$ to be the germ of the affine line at the origin, and consider the map of the germ minus the origin to the line with doubled origin. Then we have two choices for how the map can extend over the origin.

12.7.C. EXERCISE. Make this precise: show that map of the line with doubled origin over k to $\operatorname{Spec} k$ fails the valuative criterion for separatedness. (Earlier arguments were given in Exercises 10.1.D and 10.2.C.)

12.7.2. $\star\star$ *Remark for experts: moduli spaces and the valuative criterion of separatedness.* If $Y = \operatorname{Spec} k$, and X is a (fine) moduli space (a term we won't define here) of some

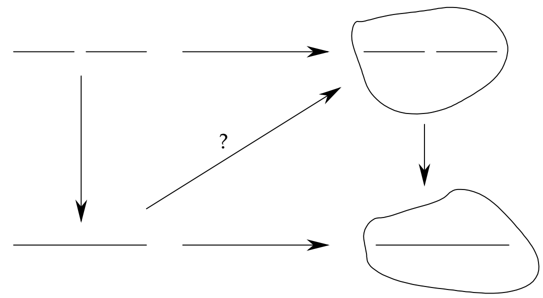


FIGURE 12.5. The valuative criterion for separatedness

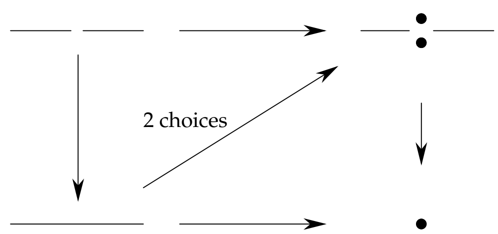


FIGURE 12.6. The line with the doubled origin fails the valuative criterion for separatedness

type of object, then the question of the separatedness of X (over $\operatorname{Spec} k$) has a natural interpretation: given a family of your objects parametrized by a "punctured discrete valuation ring", is there always at most one way of extending it over the closed point?

12.7.3. *Idea behind the proof of Theorem 12.7.1 (the valuative criterion for separatedness, DVR version).* (One direction was done in Exercise 12.7.A.) If π is *not* separated, our goal is to produce a diagram (12.7.1.1) that can be completed to (12.7.1.2) in more than one way. If π is not separated, then $\delta : X \to X \times_Y X$ is a locally closed embedding that is not a closed embedding.

12.7.D. EXERCISE. Show that you can find points p not in the diagonal Δ of $X \times_Y X$ and q in Δ such that $p \in \overline{q}$, and there are no points "between p and q" (no points r distinct from p and q with $p \in \overline{r}$ and $r \in \overline{q}$). (Exercise 7.4.C may shed some light.)

Let Q be the scheme obtained by giving the induced reduced subscheme structure to $\overline{q}$. Let $B = \mathscr{O}_{Q,p}$ be the local ring of Q at p.

12.7.E. EXERCISE. Show that B is a Noetherian local integral domain of dimension 1.

If B were regular, then we would be done: composing the inclusion morphism $Q \to X \times_Y X$ with the two projections induces the same morphism $\mathfrak{q} \to X$ (i.e., $\operatorname{Spec} \kappa(\mathfrak{q}) \to X$) but different extensions to Q precisely because $\mathfrak{p}$ is not in the diagonal. To complete the proof, one shows that the normalization of B is Noetherian; then localizing at any prime above $\mathfrak{p}$ (there is one by the Lying Over Theorem 7.2.5) yields the desired discrete valuation ring A.

For an actual proof, see [**Stacks**, tag 0207] or [**Gr-EGA**, II.72.3].

With a more powerful invocation of commutative algebra, we can prove a valuative criterion with much less restrictive hypotheses.

12.7.4. Theorem (Valuative criterion of separatedness). — *Suppose $\pi : X \to Y$ is a quasiseparated morphism. Then π is separated if and only if for any valuation ring A with function field K, and any diagram of the form (12.7.1.1), there is at most one morphism $\operatorname{Spec} A \to X$ such that the diagram (12.7.1.2) commutes.*

Because I have already failed to completely prove the DVR version, I feel no urge to prove this harder fact. The proof of one direction, that π separated implies that the criterion holds, follows from the identical argument as in Exercise 12.7.A. For a complete proof, see [**Stacks**, tags 01KY and 01KZ] or [**Gr-EGA**, II.7.2.3].

12.7.5. Valuative criteria of (universal closedness and) properness.

There is a valuative criterion for properness too. It is philosophically useful, and sometimes directly useful, although we won't need it. It naturally comes from the valuative criterion for separatedness combined with a valuative criterion for universal closedness.

12.7.6. Theorem (Valuative criterion for universal closedness and properness, DVR version). — *Suppose $\pi : X \to Y$ is a morphism of finite type of locally Noetherian schemes. Then π is universally closed (resp. proper) if and only if for any discrete valuation ring A and any diagram (12.7.1.1), there is at least one (resp. exactly one) morphism $\operatorname{Spec} A \to X$ such that the diagram (12.7.1.2) commutes.*

See [**Gr-EGA**, II.7.3.8], [**Ha1**, Thm. II.4.7], or [**Gr-EGA'**, I.5.5] for proofs. A comparison with Theorem 12.7.1 will convince you that these three criteria naturally form a family.

In the case where Y is a field, you can think of the valuative criterion of properness as saying that limits of one-parameter families in proper varieties always exist, and are unique. This is a useful intuition for the notion of properness.

12.7.F. EASY EXERCISE. Use the valuative criterion of properness to prove that $\mathbb{P}^n_A \to \operatorname{Spec} A$ is proper if A is Noetherian. (Don't be fooled: Because this requires the valuative criterion, this is a difficult way to prove a fact that we already showed in Theorem 10.3.5.)

12.7.G. EXERCISE (CF. EXERCISE 12.7.B). Suppose X is an irreducible regular (Noetherian) curve, proper either over a field k or over $\mathbb{Z}$. Describe a bijection between the discrete valuations on $K(X)$ and the closed points of X.

12.7.7. *Remarks for experts.* There is a moduli-theoretic interpretation similar to that for separatedness (Remark 12.7.2): X is proper if and only if there is always precisely one way of filling in a family over a spectrum of a punctured discrete valuation ring.

12.7.8. Finally, here is a fancier version of the valuative criterion for universal closedness and properness.

12.7.9. Theorem (Valuative criterion of universal closedness and properness). — *Suppose* $\pi : X \to Y$ *is a quasiseparated, finite type (hence quasicompact) morphism. Then* π *is universally closed (resp. proper) if and only if the following condition holds. For any valuation ring* A *and any diagram of the form* (12.7.1.1)*, there is at least one (resp. exactly one) morphism* $\operatorname{Spec} A \to X$ *such that the diagram* (12.7.1.2) *commutes.*

Clearly the valuative criterion of properness is a consequence of the valuative criterion of separateness (Theorem 12.7.4) and the valuative criterion for universal closedness. For proofs, see [**Stacks**, tag 01KF] or [**Gr-EGA**, II.7.3.8].

12.7.10. On the importance of valuation rings in general. Although we have only discussed discrete valuation rings in depth, general valuation rings should not be thought of as an afterthought. Serre makes the case to Grothendieck (in French):

You are very harsh on Valuations! I persist nonetheless in keeping them, for several reasons, of which the first is practical: n *people have sweated over them, there is nothing wrong with the result, and it should not be thrown out without very serious reasons (which you do not have). ... Even an unrepentant Noetherian needs discrete valuations and their extensions; in fact, Tate, Dwork and all the* p*-adic people will tell you that one cannot restrict oneself to the discrete case and the rank* 1 *case is indispensable; Noetherian methods then become a burden, and one understands much better if one considers the general case and not only the rank* 1 *case. ... It is not worth making a mountain out of it, of course, which is why I energetically fought Weil's original plan to make it the central theorem of Commutative Algebra, but on the other hand it must be kept.*
— *Serre, in a letter to Grothendieck* [**GrS**, p. 125]

12.8 ⋆ More sophisticated facts about regular local rings

Regular local rings have essentially every good property you could want, but some of them require hard work. We now discuss a few fancier facts that may help you sleep well at night.

12.8.1. Localizations of regular local rings are regular local rings.

12.8.2. Fact ([E, Cor. 19.14], [Mat2, Thm. 19.3]). — *If* $(A, \mathfrak{m})$ *is a regular local ring, then any localization of* A *at a prime is also a regular local ring.*

(We will not need this, and hence will not prove it.) This major theorem was an open problem in commutative algebra for a long time until settled by Serre and Auslander-Buchsbaum using homological methods.

Hence to check if $\operatorname{Spec} A$ is regular (A Noetherian), it suffices to check at closed points (at maximal ideals). Assuming Fact 12.8.2 (and using Exercise 5.1.E), you can check regularity of a Noetherian scheme by checking at closed points.

We will prove two important cases of Fact 12.8.2. The first you can do right now.

12.8.A. EXERCISE. Suppose X is a Noetherian dimension 1 scheme that is regular at its closed points. Show that X is reduced. Hence show (without invoking Fact 12.8.2) that X is regular.

The second important case will be proved in §21.3.6:

12.8.3. Theorem. — *If X is a finite type scheme over a perfect field k that is regular at its closed points, then X is regular.*

More generally, Exercise 21.3.F will show that Fact 12.8.2 holds if A is the localization of a finite type algebra over a perfect field.

12.8.B. EXERCISE (GENERALIZING EXERCISE 12.3.E). Suppose k is an algebraically closed field of characteristic 0. Assuming Theorem 12.8.3, show that there exists a regular hypersurface of every positive degree d in $\mathbb{P}^n$. (As in Exercise 12.3.E, feel free to weaken the hypotheses.)

12.8.4. Regular local rings are unique factorization domains, integrally closed, and Cohen-Macaulay.

12.8.5. Fact (Auslander-Buchsbaum). — *Regular local rings are unique factorization domains.*

(This is a hard theorem, so we will not prove it, and will therefore not use it. For a proof, see [**E**, Thm. 19.19] or [**Mat2**, Thm. 20.3], or [**Mu7**, §III.7] in the special case of varieties.) Thus regular schemes are factorial, and hence normal by Exercise 5.4.F.

12.8.6. *Remark: Factoriality is weaker than regularity.* The implication "regular implies factorial" is strict. Here is an example showing this. Suppose k is an algebraically closed field of characteristic not 2. Let $A = k[x_1, \dots, x_n]/(x_1^2 + \cdots + x_n^2)$ (cf. Exercise 5.4.I). Note that $\operatorname{Spec} A$ is clearly singular at the origin. In Exercise 14.2.V, we will show that A is a unique factorization domain when $n \geq 5$, so $\operatorname{Spec} A$ is factorial. In particular, $A_{(x_1, \dots, x_n)}$ is a Noetherian local ring that is a unique factorization domain, but not regular. (More generally, it is a consequence of Grothendieck's proof of a conjecture of Samuel that a Noetherian local ring that is a complete intersection — in particular a hypersurface — that is factorial in codimension at most 3 must be factorial, [**SGA2**, Exp. XI, Cor. 3.14].)

12.8.7. *Regular local rings are integrally closed.* Fact 12.8.5 implies that regular local rings are integrally closed (by Exercise 5.4.F). We will prove this (without appealing to Fact 12.8.5) in the most important geometric cases in §26.3.5.

12.8.8. *Regular local rings are Cohen-Macaulay.* In §26.2.5, we will show that regular local rings are Cohen-Macaulay (a notion defined in Chapter 26).

12.9 ⋆ Filtered rings and modules, and the Artin-Rees Lemma

We conclude Chapter 12 by discussing the Artin-Rees Lemma 12.9.3, which was used to prove Proposition 12.5.2. The Artin-Rees Lemma generalizes the intuition behind Proposition 12.5.2, that any function that is analytically zero at a point actually vanishes in a neighborhood of that point (§12.5.3). Because we will use it later (proving the Cohomology and Base Change Theorem 28.1.6), and because it is useful to recognize it in other contexts, we discuss it in some detail.

12.9.1. *Definitions.* Suppose I is an ideal of a ring A. A descending filtration of an A-module M

$$M = M_0 \supset M_1 \supset M_2 \supset \cdots \tag{12.9.1.1}$$

is called an I-**filtration** if $I^d M_n \subset M_{n+d}$ for all $d, n \geq 0$. An example is the I-**adic filtration** where $M_k = I^k M$. We say an I-filtration is I-**stable** if for some s and all $d \geq 0$, $I^d M_s = M_{d+s}$. For example, the I-adic filtration is I-stable.

Let $A_\bullet(I)$ be the graded ring $\oplus_{n \geq 0} I^n$. This is called the **Rees algebra** of the ideal I in A, although we will not need this terminology. Define $M_\bullet(I) := \oplus M_n$. It is naturally a *graded* module over $A_\bullet(I)$.

12.9.2. Proposition. *If A is Noetherian, M is a finitely generated A-module, and (12.9.1.1) is an I-filtration, then $M_\bullet(I)$ is a finitely generated $A_\bullet(I)$-module if and only if the filtration (12.9.1.1) is I-stable.*

Proof. Note that $A_\bullet(I)$ is Noetherian (by Exercise 4.5.D(b), as A is Noetherian, and I is a finitely generated A-module).

Assume first that $M_\bullet(I)$ is finitely generated over the Noetherian ring $A_\bullet(I)$, and hence Noetherian. Consider the increasing chain of $A_\bullet(I)$-submodules whose kth element L_k is

$$M \oplus M_1 \oplus M_2 \oplus \cdots \oplus M_k \oplus IM_k \oplus I^2 M_k \oplus \cdots$$

(which agrees with $M_\bullet(I)$ up until M_k, and then "I-stabilizes"). This chain must stabilize by Noetherianness. But $\cup L_k = M_\bullet(I)$, so for some $s \in \mathbb{Z}$, $L_s = M_\bullet(I)$, so $I^d M_s = M_{s+d}$ for all $d \geq 0$ — (12.9.1.1) is I-stable.

For the other direction, assume that $M_{d+s} = I^d M_s$ for a fixed s and all $d \geq 0$. Then $M_\bullet(I)$ is generated over $A_\bullet(I)$ by $M \oplus M_1 \oplus \cdots \oplus M_s$. But each M_j is finitely generated, so $M_\bullet(I)$ is indeed a finitely generated $A_\bullet(I)$-module. $\square$

12.9.3. Artin-Rees Lemma. — *Suppose A is a Noetherian ring, and (12.9.1.1) is an I-stable filtration of a finitely generated A-module M. Suppose that $L \subset M$ is a submodule, and let $L_n := L \cap M_n$. Then*

$$L = L_0 \supset L_1 \supset L_2 \supset \cdots$$

is an I-stable filtration of L.

Proof. Note that $L_\bullet$ is an I-filtration, as $IL_n \subset IL \cap IM_n \subset L \cap M_{n+1} = L_{n+1}$. Also, $L_\bullet(I)$ is an $A_\bullet(I)$-submodule of the finitely generated $A_\bullet(I)$-module $M_\bullet(I)$, and hence finitely generated by Exercise 3.6.X (as $A_\bullet(I)$ is Noetherian, see the proof of Proposition 12.9.2). $\square$

An important special case is the following.

12.9.4. Corollary. — *Suppose* $I \subset A$ *is an ideal of a Noetherian ring, and* M *is a finitely generated* A*-module, and* L *is a submodule. Then for some integer* s, $I^d(L \cap I^s M) = L \cap I^{d+s}M$ *for all* $d \geq 0$.

Warning: it need not be true that $I^d L = L \cap I^d M$ for all d. (Can you think of a counterexample to this statement?)

Proof. Apply the Artin-Rees Lemma 12.9.3 to the filtration $M_n = I^n M$. □

12.9.A. EXERCISE (KRULL INTERSECTION THEOREM).
(a) Suppose I is an ideal of a Noetherian ring A, and M is a finitely generated A-module. Show that there is some $a \equiv 1 \pmod{I}$ such that $a \cap_{j=1}^{\infty} I^j M = 0$. Hint: Apply the Artin-Rees Lemma 12.9.3 with $L = \cap_{j=1}^{\infty} I^j M$ and $M_n = I^n M$. Show that $L = IL$, and apply the first version of Nakayama (Lemma 7.2.8).
(b) Show that if A is a Noetherian integral domain or a Noetherian local ring, and I is a proper ideal, then $\cap_{j=1}^{\infty} I^j = 0$. In particular, you will have proved Proposition 12.5.2: if $(A, \mathfrak{m})$ is a Noetherian local ring, then $\cap_i \mathfrak{m}^i = 0$.

12.9.B. EXERCISE. Make the following precise, and prove it (thereby justifying the intuition in §12.5.3): if X is a locally Noetherian scheme, and f is a function on X that is analytically zero at a point $p \in X$, then f vanishes in a (Zariski) neighborhood of p.

Part V

Quasicoherent sheaves

CHAPTER 13

Quasicoherent and coherent sheaves

Quasicoherent and coherent sheaves generalize the notion of a vector bundle. To motivate them, we first discuss vector bundles, and their interpretation as locally free sheaves.

Speaking very informally, a vector bundle V on a geometric space X (such as a manifold) is a family of vector spaces continuously parametrized by points of X. In other words, for each point p of X, there is a vector space, and these vector spaces are glued into a space V so that, as p varies, the vector space above p varies continuously. Nontrivial examples to keep in mind are the tangent bundle to a manifold, and the Möbius strip over a circle (interpreted as a line bundle). We will make this somewhat more precise in §13.1, but if you have not seen this idea before, you shouldn't be concerned; the main thing we will use this notion for is motivation.

A **free sheaf** on a ringed space X is an $\mathscr{O}_X$-module isomorphic to $\mathscr{O}_X^{\oplus I}$ where the sum is over some index set I. A **locally free sheaf** on a ringed space X is an $\mathscr{O}_X$-module locally isomorphic to a free sheaf. This corresponds to the notion of a vector bundle (§13.1). Quasicoherent sheaves form a convenient abelian category containing the locally free sheaves that is much smaller than the full category of $\mathscr{O}$-modules. Quasicoherent sheaves generalize free sheaves in much the way that modules generalize free modules. Coherent sheaves are roughly speaking a finite rank version of quasicoherent sheaves, which form a well-behaved abelian category containing finite rank locally free sheaves (or equivalently, finite rank vector bundles). Just as the notion of free modules lead us to the notion of modules in general, and finitely generated modules, the notion of free sheaves will lead us inevitably to the notion of quasicoherent sheaves and coherent sheaves. (There is a slight fib in comparing finitely generated modules to coherent sheaves, as you will find out in §13.6.)

13.1 Vector bundles and locally free sheaves

We recall somewhat more precisely the notion of vector bundles on manifolds. Arithmetically-minded readers shouldn't tune out: for example, fractional ideals of the ring of integers in a number field (defined in §9.7.1) turn out to be an example of a "line bundle on a smooth curve" (Exercise 13.1.M).

A **rank n vector bundle on a manifold** M is a map $\pi : V \to M$ with the structure of an n-dimensional real vector space on $\pi^{-1}(x)$ for each point $x \in M$, such that for every $x \in M$, there is an open neighborhood U and a homeomorphism

$$\phi : U \times \mathbb{R}^n \to \pi^{-1}(U)$$

over U (so that the diagram

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, leftrightarrow, "\cong"] \arrow[dr, "\pi|_{\pi^{-1}(U)}"'] & & U \times \mathbb{R}^n \arrow[dl, "\text{projection to first factor}"] \\ & U & \end{tikzcd} \tag{13.1.0.1}$$

commutes) that is an isomorphism of vector spaces over each $y \in U$. An isomorphism (13.1.0.1) is called a **trivialization over** U.

We call n the **rank** of the vector bundle. A rank 1 vector bundle is called a **line bundle**. (It can also be convenient to be agnostic about the rank of the vector bundle, so it can have different ranks on different connected components. It is also sometimes convenient to consider infinite-rank vector bundles.)

13.1.1. Transition functions. Given trivializations over U_1 and U_2, over their intersection, the two trivializations must be related by an element T_{12} of GL_n with entries consisting of functions on $U_1 \cap U_2$. If $\{U_i\}$ is a cover of M, and we are given trivializations over each U_i, then the $\{T_{ij}\}$ must satisfy the **cocycle condition**:

$$T_{jk}|_{U_i \cap U_j \cap U_k} \circ T_{ij}|_{U_i \cap U_j \cap U_k} = T_{ik}|_{U_i \cap U_j \cap U_K}. \tag{13.1.1.1}$$

(This implies $T_{ij} = T_{ji}^{-1}$.) The data of the T_{ij} are called **transition functions** (or *transition matrices*) for the trivialization.

This is reversible: given the data of a cover $\{U_i\}$ and transition functions T_{ij}, we can recover the vector bundle (up to unique isomorphism) by "gluing together the various $U_i \times \mathbb{R}^n$ along $U_i \cap U_j$ using T_{ij}".

13.1.2. The sheaf of sections. Fix a rank n vector bundle $V \to M$. The sheaf of sections $\mathscr{F}$ of V (Exercise 2.2.G) is an $\mathscr{O}_M$-module — given any open set U, we can multiply a section over U by a function on U and get another section.

Moreover, given a trivialization over U, the sections over U are naturally identified with n-tuples of functions of U.

$$\begin{tikzcd} U \times \mathbb{R}^n \arrow[d, shift right, "\pi"'] \\ U \arrow[u, shift right, bend right, "n\text{-tuple of functions}"'] \end{tikzcd}$$

Thus given a trivialization, over each open set U_i, we have an isomorphism $\mathscr{F}|_{U_i} \cong \mathscr{O}_{U_i}^{\oplus n}$. We say that such an $\mathscr{F}$ is a **locally free sheaf of rank** n. (A sheaf $\mathscr{F}$ is **free of rank** n, or sometimes *trivial of rank* n, if $\mathscr{F} \cong \mathscr{O}^{\oplus n}$.)

13.1.3. Transition functions for the sheaf of sections. Suppose we have a vector bundle on M, along with a trivialization over an open cover U_i. Suppose we have a section of the vector bundle over M. (This discussion will apply with M replaced by any open subset.) Then over each U_i, the section corresponds to an n-tuple functions over U_i, say $\vec{s}^i$.

13.1.A. EXERCISE. Show that over $U_i \cap U_j$, the vector-valued function $\vec{s}^i$ is related to $\vec{s}^j$ by the (same) transition functions: $T_{ij}\vec{s}^i = \vec{s}^j$. (Don't do this too quickly — make sure your i's and j's are on the correct side.)

Given a locally free sheaf $\mathscr{F}$ with rank n, and a trivializing neighborhood of $\mathscr{F}$ (an open cover $\{U_i\}$ such that over each U_i, $\mathscr{F}|_{U_i} \cong \mathscr{O}_{U_i}^{\oplus n}$ as $\mathscr{O}$-modules), we have transition functions $T_{ij} \in GL_n(\mathscr{O}(U_i \cap U_j))$ satisfying the cocycle condition (13.1.1.1). Thus the data of a locally free sheaf of rank n is *equivalent* to the data of a vector bundle of rank n. This change of perspective is useful, and is similar to an earlier change of perspective when we introduced ringed spaces: understanding spaces is the same as understanding (sheaves of) functions on the spaces, and understanding vector bundles (a type of "space over M") is the same as understanding functions.

13.1.4. *Definition.* A rank 1 locally free sheaf is called an **invertible sheaf**. (Unimportant aside: "invertible sheaf" is a heinous term for something that is essentially a line bundle. The motivation is that if X is a locally ringed space, and $\mathscr{F}$ and $\mathscr{G}$ are $\mathscr{O}_X$-modules with $\mathscr{F} \otimes_{\mathscr{O}_X} \mathscr{G} \cong \mathscr{O}_X$, then $\mathscr{F}$ and $\mathscr{G}$ are invertible sheaves [**MO33489**]. Thus in the monoid of $\mathscr{O}_X$-modules under tensor product, invertible sheaves are the invertible elements. We will never use this fact. People often informally use the phrase "line bundle" when they mean "invertible sheaf". The phrase *line sheaf* has been proposed but has not caught on.)

13.1.5. Locally free sheaves on schemes.

We can generalize the notion of locally free sheaves to schemes (or more generally, ringed spaces) without change. A **locally free sheaf of rank n on a scheme** X is defined as an $\mathscr{O}_X$-module $\mathscr{F}$ that is locally a free sheaf of rank n. Precisely, there is an open cover $\{U_i\}$ of X such that for each U_i, $\mathscr{F}|_{U_i} \cong \mathscr{O}_{U_i}^{\oplus n}$. This open cover determines transition functions — the data of a cover $\{U_i\}$ of X, and functions $T_{ij} \in GL_n(\mathscr{O}(U_i \cap U_j))$ satisfying the cocycle condition (13.1.1.1) — which in turn determine the locally free sheaf. As before, given these data, we can find the sections over any open set U. Informally, they are sections of the free sheaves over each $U \cap U_i$ that agree on overlaps. More formally, for each i, they are

$$\vec{s}^i = \begin{pmatrix} s_1^i \\ \vdots \\ s_n^i \end{pmatrix} \in \Gamma(U \cap U_i, \mathscr{O}_X)^n,$$

satisfying $T_{ij}\vec{s}^i = \vec{s}^j$ on $U \cap U_i \cap U_j$.

You should think of these as vector bundles, but just keep in mind that they are not the "same", just equivalent notions. We will later (Definition 17.1.4) define the "total space" of the vector bundle $V \to X$ (a scheme over X) in terms of the sheaf version of Spec (or more precisely, $\mathcal{S}pec$ Sym $V^\bullet$). But the locally free sheaf perspective will prove to be more useful. As one example: the definition of a locally free sheaf is much shorter than that of a vector bundle.

As in our motivating discussion, it is sometimes convenient to let the rank vary among connected components, or to consider infinite rank locally free sheaves.

13.1.6. Useful constructions, in the form of a series of important exercises.

We now give some useful constructions in the form of a series of exercises about locally free sheaves on a scheme. They are useful, important, and surprisingly nontrivial! Two hints: Exercises 13.1.B–13.1.G will apply for ringed spaces in general, so you shouldn't use special properties of schemes. Furthermore, they are all local on X, so you can reduce to the case where the locally free sheaves in question are actually free.

13.1.B. EXERCISE. Suppose $\mathscr{F}$ and $\mathscr{G}$ are locally free sheaves on X of rank m and n respectively. Show that $\mathcal{H}om_{\mathscr{O}_X}(\mathscr{F}, \mathscr{G})$ is a locally free sheaf of rank mn.

13.1.C. EXERCISE. If $\mathscr{E}$ is a locally free sheaf on X of (finite) rank n, Exercise 13.1.B implies that $\mathscr{E}^\vee := \mathcal{H}om(\mathscr{E}, \mathscr{O}_X)$ is also a locally free sheaf of rank n. This is called the **dual** of $\mathscr{E}$ (cf. §2.3.3). Given transition functions for $\mathscr{E}$, describe transition functions for $\mathscr{E}^\vee$. (Note that if $\mathscr{E}$ is rank 1, i.e., invertible, the transition functions of the dual are the inverse of the transition functions of the original.) Show that $\mathscr{E} \cong \mathscr{E}^{\vee\vee}$. (Caution: your argument showing that there is a canonical isomorphism $(\mathscr{F}^\vee)^\vee \cong \mathscr{F}$ better not also show that there is an isomorphism $\mathscr{F}^\vee \cong \mathscr{F}$! We will see an example in §14.1 of a locally free $\mathscr{F}$ that is not isomorphic to its dual: the invertible sheaf $\mathscr{O}(1)$ on $\mathbb{P}^n$.)

13.1.D. EXERCISE. If $\mathscr{F}$ and $\mathscr{G}$ are locally free sheaves, show that $\mathscr{F} \otimes \mathscr{G}$ is a locally free sheaf. (Here $\otimes$ is tensor product as $\mathscr{O}_X$-modules, defined in Exercise 2.5.J.) If $\mathscr{F}$ is an invertible sheaf, show that $\mathscr{F} \otimes \mathscr{F}^\vee \cong \mathscr{O}_X$.

13.1.E. EXERCISE. Recall that tensor products tend to be only right-exact in general. Show that tensoring by a locally free sheaf is exact. More precisely, if $\mathscr{F}$ is a locally free sheaf, and $\mathscr{G}' \to \mathscr{G} \to \mathscr{G}''$ is an exact sequence of $\mathscr{O}_X$-modules, then then so is $\mathscr{G}' \otimes \mathscr{F} \to \mathscr{G} \otimes \mathscr{F} \to \mathscr{G}'' \otimes \mathscr{F}$. (Possible hint: it may help to check exactness by checking exactness at stalks. Recall that the tensor product of stalks can be identified with the stalk of the tensor product, so for example there is a "natural" isomorphism $(\mathscr{G} \otimes_{\mathscr{O}_X} \mathscr{F})_p \cong \mathscr{G}_p \otimes_{\mathscr{O}_{X,p}} \mathscr{F}_p$, Exercise 2.5.J(b).)

13.1.F. EXERCISE. If $\mathscr{E}$ is a locally free sheaf of finite rank, and $\mathscr{F}$ and $\mathscr{G}$ are $\mathscr{O}_X$-modules, show that $\mathcal{H}om(\mathscr{F}, \mathscr{G} \otimes \mathscr{E}) \cong \mathcal{H}om(\mathscr{F} \otimes \mathscr{E}^\vee, \mathscr{G})$. (Possible hint: first consider the case where $\mathscr{E}$ is free.)

13.1.G. EXERCISE AND IMPORTANT DEFINITION. Show that the invertible sheaves on X, up to isomorphism, form an abelian group under tensor product. This is called the **Picard group** of X, and is denoted $\operatorname{Pic} X$.

13.1.H. EXERCISE. If $\pi : X \to Y$ is a morphism of ringed spaces, and $\mathscr{G}$ is a locally free sheaf of rank n on Y, figure out how to define a "pulled back" locally free sheaf of rank n on X, denoted $\pi^*\mathscr{G}$. Then show that Pic is a contravariant functor from the category of ringed spaces to the category *Ab* of abelian groups.

Unlike the previous exercises, the next one is specific to schemes.

13.1.I. EXERCISE. Suppose s is a section of a locally free sheaf $\mathscr{F}$ on a scheme X. Define the notion of the **subscheme cut out by** $s = 0$, denoted (for obvious reasons) $V(s)$. Be sure to check that your definition is independent of choices! Hint: given a trivialization over an open set U, s corresponds to a number of functions $f_1, \dots$ on U; on U, take the scheme cut out by these functions. Alternate hint that avoids coordinates: cleverly apply Exercise 13.1.I. (This exercise gives a new solution to Exercise 4.5.P.)

13.1.7. Random concluding remarks.

We define **rational (and regular) sections of a locally free sheaf** on a scheme X just as we did rational (and regular) functions (see for example §5.5 and §6.5).

13.1.J. EXERCISE. Show that locally free sheaves on locally Noetherian normal schemes satisfy "Hartogs's Lemma": sections defined away from a set of codimension at least 2 extend over that set. (Algebraic Hartogs's Lemma for Noetherian normal schemes is Theorem 11.3.10.)

13.1.K. EASY EXERCISE. Suppose s is a nonzero rational section of an invertible sheaf on a locally Noetherian normal scheme. Show that if s has no poles, then s is regular. (Hint: Exercise 12.5.H.)

13.1.8. *Remark.* Based on your intuition for line bundles on manifolds, you might hope that every point has a "small" open neighborhood on which all invertible sheaves (or locally free sheaves) are trivial. Sadly, this is not the case. We will eventually see (§19.11.1) that for the curve $y^2 - x^3 - x = 0$ in $\mathbb{A}^2_{\mathbb{C}}$, every nonempty open set has nontrivial invertible sheaves. (This will use the fact that it is an open subset of an *elliptic curve*.)

13.1.L. $\star$ EXERCISE (FOR THOSE WITH SUFFICIENT COMPLEX-ANALYTIC BACKGROUND). Recall the analytification functor (Exercises 6.3.K and 10.1.F), that takes a complex finite type reduced scheme and produces a complex analytic space.
(a) If $\mathscr{L}$ is an invertible sheaf on a complex (algebraic) variety X, define (up to unique isomorphism) the corresponding invertible sheaf on the complex variety X_{an}.
(b) Show that the induced map $\operatorname{Pic} X \to \operatorname{Pic} X_{an}$ is a group homomorphism.
(c) Show that this construction is functorial: if $\pi : X \to Y$ is a morphism of complex varieties, the following diagram commutes:

$$\begin{tikzcd} \operatorname{Pic} Y \arrow[r, "\pi^*"] \arrow[d] & \operatorname{Pic} X \arrow[d] \\ \operatorname{Pic} Y_{an} \arrow[r, "\pi^*_{an}"] & \operatorname{Pic} X_{an} \end{tikzcd}$$

where the vertical maps are the ones you have defined.

13.1.M. $\star$ EXERCISE (FOR THOSE WITH SUFFICIENT ARITHMETIC BACKGROUND; SEE ALSO PROPOSITION 14.2.10 AND §14.2.13). Recall the definition of the ring of integers $\mathscr{O}_K$ in a number field K, Remark 9.7.1. A **fractional ideal** $\mathfrak{a}$ of $\mathscr{O}_K$ is a nonzero $\mathscr{O}_K$-submodule of K such that there is a nonzero $a \in \mathscr{O}_K$ such that $a\mathfrak{a} \subset \mathscr{O}_K$. Products of fractional ideals are defined analogously to products of ideals in a ring (defined in Exercise 3.4.C): $\mathfrak{a}\mathfrak{b}$ consists of (finite) $\mathscr{O}_K$-linear combinations of products of elements of $\mathfrak{a}$ and elements of $\mathfrak{b}$. Thus fractional ideals form a semigroup under multiplication, with $\mathscr{O}_K$ as the identity. In fact fractional ideals of $\mathscr{O}_K$ form a group.

(a) Explain how a fractional ideal on a ring of integers in a number field yields an invertible sheaf. (Although we won't need this, it is worth noting that a fractional ideal is the same as an invertible sheaf *with a trivialization at the generic point*.)

(b) A fractional ideal is **principal** if it is of the form $r\mathscr{O}_K$ for some $r \in K^\times$. Show that any two that differ by a principal ideal yield the same invertible sheaf.

(c) Show that two fractional ideals that yield the same invertible sheaf differ by a principal ideal.

(d) The *class group* is defined to be the group of fractional ideals modulo the principal ideals (i.e., modulo $K^\times$). Give an isomorphism of the class group with the Picard group of $\mathscr{O}_K$.

(This discussion applies to any Dedekind domain. See Exercise 14.2.S for a follow-up.)

13.1.9. The problem with locally free sheaves.

Recall that $\mathscr{O}_X$-modules form an abelian category: we can talk about kernels, cokernels, and so forth, and we can do homological algebra. Similarly, vector spaces form an abelian category. But locally free sheaves (i.e., vector bundles), along with reasonably natural maps between them (those that arise as maps of $\mathscr{O}_X$-modules), don't form an abelian category. As a motivating example in the category of differentiable manifolds, consider the map of the trivial line bundle on $\mathbb{R}$ (with coordinate t) to itself, corresponding to multiplying by the coordinate t. Then this map jumps rank, and if you try to define a kernel or cokernel you will get confused.

This problem is resolved by enlarging our notion of nice $\mathscr{O}_X$-modules in a natural way, to quasicoherent sheaves.

$\mathscr{O}_X$-modules	$\supset$	quasicoherent sheaves	$\supset$	locally free sheaves
(abelian category)		(abelian category)		(not an abelian category)

You can turn this into two *definitions* of quasicoherent sheaves, equivalent to those we will give in §13.2. We want a notion that is local on X of course. So we ask for the smallest abelian subcategory of $Mod_{\mathscr{O}_X}$ that is "local" and includes vector bundles. It turns out that the main obstruction to vector bundles to be an abelian category is the failure of cokernels of maps of locally free sheaves — as $\mathscr{O}_X$-modules — to be locally free; we could define quasicoherent sheaves to be those $\mathscr{O}_X$-modules that are locally cokernels, yielding a description that works more generally on ringed spaces, as described in Exercise 13.4.B. You may wish to later check that our future definitions are equivalent to these.

Similarly, in the locally Noetherian setting, finite rank locally free sheaves will sit in a nice smaller abelian category, that of *coherent sheaves*.

quasicoherent sheaves	$\supset$	coherent sheaves	$\supset$	finite rank locally free sheaves
(abelian category)		(abelian category)		(not an abelian category)

13.1.10. *Remark: Quasicoherent and coherent sheaves on ringed spaces.* We will discuss quasicoherent and coherent sheaves on schemes, but they can be defined more generally (see Exercise 13.4.B for quasicoherent sheaves, and [**Se1**, Def. 2] for coherent sheaves). Many of the results we state will hold in greater generality, but because the proofs look slightly different, we restrict ourselves to schemes to avoid distraction.

13.2 Quasicoherent sheaves

We now define the notion of *quasicoherent sheaf*. In the same way that a scheme is defined by "gluing together rings", a quasicoherent sheaf over that scheme is obtained by "gluing together modules over those rings". Given an A-module M, we defined an $\mathscr{O}$-module $\widetilde{M}$ on $\operatorname{Spec} A$ long ago (Exercise 4.1.D) — the sections over $D(f)$ were M_f.

13.2.1. Theorem. — *Let X be a scheme, and $\mathscr{F}$ an $\mathscr{O}_X$-module. Suppose P is the property of affine open subschemes $\operatorname{Spec} A$ of X that $\mathscr{F}|_{\operatorname{Spec} A} \cong \widetilde{M}$ for some A-module M. Then P satisfies the two hypotheses of the Affine Communication Lemma 5.3.2.*

We prove this in a moment.

13.2.2. *Definition.* If X is a scheme, then an $\mathscr{O}_X$-module $\mathscr{F}$ is **quasicoherent** if for every affine open subset $\operatorname{Spec} A \subset X$, $\mathscr{F}|_{\operatorname{Spec} A} \cong \widetilde{M}$ for some A-module M. By Theorem 13.2.1, it suffices to check this for a collection of affine open sets covering X. For example, $\widetilde{M}$ is a quasicoherent sheaf on $\operatorname{Spec} A$, and all locally free sheaves on X are quasicoherent.

13.2.A. UNIMPORTANT EXERCISE (NOT EVERY $\mathscr{O}_X$-MODULE IS A QUASICOHERENT SHEAF).
(a) Suppose $X = \operatorname{Spec} k[t]$. Let $\mathscr{F}$ be the skyscraper sheaf supported at the origin $[(t)]$, with group $k(t)$ and the usual $k[t]$-module structure. Show that this is an $\mathscr{O}_X$-module that is not a quasicoherent sheaf. (More generally, if X is an integral scheme, and $p \in X$ is not the generic point, we could take the skyscraper sheaf at p with group the function field of X. Except in a silly circumstances, this sheaf won't be quasicoherent.) See Exercises 8.1.F and 13.3.I for more (pathological) examples of $\mathscr{O}_X$-modules that are not quasicoherent.
(b) Suppose $X = \operatorname{Spec} k[t]$. Let $\mathscr{F}$ be the skyscraper sheaf supported at the generic point $[(0)]$, with group $k(t)$. Give this the structure of an $\mathscr{O}_X$-module. Show that this *is* a quasicoherent sheaf. Describe the restriction maps in the distinguished topology of X. (Remark: Your argument will apply more generally, for example when X is an integral scheme with generic point η, and $\mathscr{F}$ is the skyscraper sheaf $i_{\eta,*}K(X)$.)

13.2.B. UNIMPORTANT EXERCISE (NOT EVERY QUASICOHERENT SHEAF IS LOCALLY FREE). Use the example of Exercise 13.2.A(b) to show that not every quasicoherent sheaf is locally free.

13.2.C. EXERCISE. Show that every (finite) rank n vector bundle on $\mathbb{A}^1_k$ is trivial of rank n. Hint: finitely generated modules over a principal ideal domain are finite direct sums of cyclic modules, as mentioned in Remark 12.5.15. See the aside in §14.2.8 for the difficult generalization to $\mathbb{A}^n_k$.

13.2.3. *Proof of Theorem 13.2.1.* Clearly if $\operatorname{Spec} A$ has property P, then so does the distinguished open $\operatorname{Spec} A_f$: if M is an A-module, then $\widetilde{M}|_{\operatorname{Spec} A_f} \cong \widetilde{M_f}$ as sheaves of $\mathscr{O}_{\operatorname{Spec} A_f}$-modules (both sides agree on the level of distinguished open sets and their restriction maps).

We next show the second hypothesis of the Affine Communication Lemma 5.3.2. Suppose we have modules $M_1, \dots, M_n$, where M_i is an A_{f_i}-module, along with isomorphisms $\phi_{ij} : (M_i)_{f_j} \to (M_j)_{f_i}$ of $A_{f_i f_j}$-modules, satisfying the cocycle condition (13.1.1.1). We want to construct an M such that $\widetilde{M}$ gives us $\widetilde{M_i}$ on $D(f_i) = \operatorname{Spec} A_{f_i}$, or equivalently, isomorphisms $\rho_i : \Gamma(D(f_i), \widetilde{M}) \to M_i$, so that the bottom triangle of

(13.2.3.1)
$$\begin{tikzcd}
 & & & M \arrow[dl, "\otimes A_{f_i}"'] \arrow[dr, "\otimes A_{f_j}"] & & & \\
 & & M_{f_i} \arrow[dl, "\rho_i"', "\sim"] \arrow[dr, "\otimes A_{f_j}"] & & M_{f_j} \arrow[dl, "\otimes A_{f_i}"'] \arrow[dr, "\rho_j", "\sim"'] & & \\
 & M_i \arrow[dr, "\otimes A_{f_j}"'] & & M_{f_i f_j} & & M_j \arrow[dl, "\otimes A_{f_i}"] & \\
 & & (M_i)_{f_j} \arrow[ur, leftrightarrow, "\sim"'] \arrow[rr, "\phi_{ij}", "\sim"'] & & (M_j)_{f_i} \arrow[ul, leftrightarrow, "\sim"] & &
\end{tikzcd}$$

commutes.

13.2.D. EXERCISE. Why does this suffice to prove the result? In other words, why does this imply that $\mathscr{F}|_{\operatorname{Spec} A} \cong \widetilde{M}$?

We already know that M should be $\Gamma(\mathscr{F}, \operatorname{Spec} A)$, as $\mathscr{F}$ is a sheaf. Consider elements of $M_1 \times \cdots \times M_n$ that "agree on overlaps"; let this set be M. In other words,

(13.2.3.2)
$$0 \longrightarrow M \longrightarrow M_1 \times \cdots \times M_n \xrightarrow{\gamma} M_{12} \times M_{13} \times \cdots \times M_{(n-1)n}$$

is an exact sequence (where $M_{ij} = (M_i)_{f_j} \cong (M_j)_{f_i}$, and the map γ is the "difference" map). So M is a kernel of a morphism of A-modules, hence an A-module. We are left to show that $M_i \cong M_{f_i}$ (and that this isomorphism satisfies (13.2.3.1)). (At this point, we may proceed in a number of ways, and the reader may wish to find their own route rather than reading on.)

For convenience assume $i = 1$. Localization is exact (Exercise 1.6.F(a)), so tensoring (13.2.3.2) by A_{f_1} yields

(13.2.3.3)
$$0 \longrightarrow M_{f_1} \longrightarrow (M_1)_{f_1} \times (M_2)_{f_1} \times \cdots \times (M_n)_{f_1}$$
$$\longrightarrow M_{12} \times \cdots \times M_{1n} \times (M_{23})_{f_1} \times \cdots \times (M_{(n-1)n})_{f_1}$$

is an exact sequence of A_{f_1}-modules.

We now identify many of the modules appearing in (13.2.3.3) in terms of M_1. First of all, f_1 is invertible in A_{f_1}, so $(M_1)_{f_1}$ is canonically M_1. Also, $(M_j)_{f_1} \cong (M_1)_{f_j}$ via ϕ_{ij}. Hence if $i, j \neq 1$, $(M_{ij})_{f_1} \cong (M_1)_{f_i f_j}$ via ϕ_{1i} and ϕ_{1j} (here the cocycle condition is implicitly used). Furthermore, $(M_{1i})_{f_1} \cong (M_1)_{f_i}$ via ϕ_{1i}. Thus we can write (13.2.3.3) as

(13.2.3.4)
$$0 \longrightarrow M_{f_1} \longrightarrow M_1 \times (M_1)_{f_2} \times \cdots \times (M_1)_{f_n}$$
$$\xrightarrow{\alpha} (M_1)_{f_2} \times \cdots \times (M_1)_{f_n} \times (M_1)_{f_2 f_3} \times \cdots \times (M_1)_{f_{n-1} f_n}$$

By assumption, $\mathscr{F}|_{\operatorname{Spec} A_{f_1}} \cong \widetilde{M_1}$ for some M_1, so by considering the cover

$$\operatorname{Spec} A_{f_1} = \operatorname{Spec} A_{f_1} \cup \operatorname{Spec} A_{f_1 f_2} \cup \operatorname{Spec} A_{f_1 f_3} \cup \cdots \cup \operatorname{Spec} A_{f_1 f_n}$$

(notice the "redundant" first term), and identifying sections of $\mathscr{F}$ over $\operatorname{Spec} A_{f_1}$ in terms of sections over the open sets in the cover and their pairwise overlaps, we have an exact sequence of A_{f_1}-modules

$$0 \longrightarrow M_1 \longrightarrow M_1 \times (M_1)_{f_2} \times \cdots \times (M_1)_{f_n}$$

$$\xrightarrow{\beta} (M_1)_{f_2} \times \cdots \times (M_1)_{f_n} \times (M_1)_{f_2 f_3} \times \cdots \times (M_1)_{f_{n-1} f_n}$$

which is very similar to (13.2.3.4). Indeed, the final map β of the above sequence is the same as the map α of (13.2.3.4), so $\ker \alpha = \ker \beta$, i.e., we have an isomorphism $M_1 \cong M_{f_1}$.

Finally, the triangle of (13.2.3.1) is commutative, as each vertex of the triangle can be identified as the sections of $\mathscr{F}$ over $\operatorname{Spec} A_{f_1 f_2}$. □

13.3 Characterizing quasicoherence using the distinguished affine base

Because quasicoherent sheaves are locally of a very special form, in order to "know" a quasicoherent sheaf, we need only know what the sections are over every affine open set, and how to restrict sections from an affine open set U to a *distinguished* affine open subset of U. We make this precise by defining what I will call the *distinguished affine base* of the Zariski topology — not a base in the usual sense. The point of this discussion is to give a useful characterization of quasicoherence, but you may wish to just jump to §13.3.3.

The open sets of the distinguished affine base are the affine open subsets of X. We have already observed that this forms a base. But forget that fact. We like distinguished open sets $\operatorname{Spec} A_f \hookrightarrow \operatorname{Spec} A$, and we don't really understand open embeddings of one random affine open subset in another. So we just remember the "nice" inclusions.

13.3.1. Definition. The **distinguished affine base** of a scheme X is the data of the affine open sets and the distinguished inclusions.

In other words, we remember only some of the open sets (the affine open sets), and *only some of the morphisms between them* (the distinguished morphisms). For experts: if you think of a topology as a category (the category of open sets), we have described a subcategory.

We can define a sheaf on the distinguished affine base in the obvious way: we have a set (or abelian group, or ring) for each affine open set, and we know how to restrict to distinguished open sets.

Given a sheaf $\mathscr{F}$ on X, we get a sheaf on the distinguished affine base. You can guess where we are going: we will show that all the information of the sheaf is contained in the information of the sheaf on the distinguished affine base.

As a warm-up, we can recover stalks as follows. (We will be implicitly using only the following fact. We have a collection of open subsets, and *some* inclusions among these subsets, such that if we have any $x \in U, V$ where U and V are in our collection of open sets, there is some W containing x, and contained in U and V such that $W \hookrightarrow U$ and $W \hookrightarrow V$ are both in our collection of inclusions. In the case we are considering here, this is the key Proposition 5.3.1 that given any two affine open sets $\operatorname{Spec} A$, $\operatorname{Spec} B$ in X, $\operatorname{Spec} A \cap \operatorname{Spec} B$ could be covered by affine open sets that were simultaneously distinguished in $\operatorname{Spec} A$ and $\operatorname{Spec} B$. In fancy language: the category of affine open sets, and distinguished inclusions, forms a filtered set.)

The stalk $\mathscr{F}_x$ is the colimit $\varinjlim(f \in \mathscr{F}(U))$ where the colimit is over all open sets contained in X. We compare this to $\varinjlim(f \in \mathscr{F}(U))$ where the colimit is over all affine open sets, and all distinguished inclusions. You can check that the elements of one correspond to elements of the other. (Think carefully about this!)

13.3.A. EXERCISE. Show that a section of a sheaf on the distinguished affine base is determined by the section's germs.

13.3.2. Theorem. —

(a) *A sheaf on the distinguished affine base $\mathscr{F}^b$ determines a unique sheaf $\mathscr{F}$, which when restricted to the affine base is $\mathscr{F}^b$. (Hence if you start with a sheaf, and take the sheaf on the distinguished affine base, and then take the induced sheaf, you get the sheaf you started with.)*

(b) *A morphism of sheaves on a distinguished affine base uniquely determines a morphism of sheaves.*

(c) *An $\mathscr{O}_X$-module "on the distinguished affine base" yields an $\mathscr{O}_X$-module.*

This proof is identical to our argument of §2.7 showing that sheaves are (essentially) the same as sheaves on a base, using the "sheaf of compatible germs" construction. The main reason for repeating it is to let you see that all that is needed is for the open sets to form a filtered set (or in the current case, that the category of open sets and distinguished inclusions is filtered).

For experts: (a) and (b) are describing an equivalence of categories between sheaves on the Zariski topology of X and sheaves on the distinguished affine base of X.

Proof. (a) Suppose $\mathscr{F}^b$ is a sheaf on the distinguished affine base. Then we can define stalks.

For *any* open set U of X, define the sheaf of compatible germs

$$\begin{aligned} \mathscr{F}(U) \quad := \quad & \{(f_x \in \mathscr{F}^b_x)_{x \in U} : \text{ for all } x \in U, \\ & \text{there exists } U_x \text{ with } x \subset U_x \subset U, F^x \in \mathscr{F}^b(U_x) \\ & \text{such that } F^x_y = f_y \text{ for all } y \in U_x\} \end{aligned}$$

where each U_x is in our base, and F^x_y means "the germ of F^x at y". (As usual, those who want to worry about the empty set are welcome to.)

This really is a sheaf: convince yourself that we have restriction maps, identity, and gluability, really quite easily.

I next claim that if U is in our base, that $\mathscr{F}(U) = \mathscr{F}^b(U)$. We clearly have a map $\mathscr{F}^b(U) \to \mathscr{F}(U)$. This is an isomorphism on stalks, and hence an isomorphism by Exercise 2.4.E.

13.3.B. EXERCISE. Prove (b) (cf. Exercise 2.7.C).

13.3.C. EXERCISE. Prove (c) (cf. Remark 2.7.3)

□

13.3.3. A characterization of quasicoherent sheaves in terms of distinguished inclusions. We use this perspective to give a useful characterization of quasicoherent sheaves among $\mathscr{O}_X$-modules. Suppose $\mathscr{F}$ is an $\mathscr{O}_X$-module, and $\operatorname{Spec} A_f \hookrightarrow \operatorname{Spec} A \subset X$ is a distinguished open subscheme of an affine open subscheme of X. Let $\phi : \Gamma(\operatorname{Spec} A, \mathscr{F}) \to \Gamma(\operatorname{Spec} A_f, \mathscr{F})$ be the restriction map. The source of ϕ is an A-module, and the target is an A_f-module, so by the universal property of localization (Exercise 1.3.D), ϕ naturally factors as:

$$\begin{tikzcd} \Gamma(\operatorname{Spec} A, \mathscr{F}) \arrow[rr, "\phi"] \arrow[dr] & & \Gamma(\operatorname{Spec} A_f, \mathscr{F}) \\ & \Gamma(\operatorname{Spec} A, \mathscr{F})_f \arrow[ur, "\alpha"'] & \end{tikzcd}$$

13.3.D. VERY IMPORTANT EXERCISE. Show that an $\mathscr{O}_X$-module $\mathscr{F}$ is quasicoherent if and only if for each such distinguished $\operatorname{Spec} A_f \hookrightarrow \operatorname{Spec} A$, α is an isomorphism.

Thus a quasicoherent sheaf is (equivalent to) the data of one module for each affine open subset (a module over the corresponding ring), such that the module over a distinguished open set $\operatorname{Spec} A_f$ is given by localizing the module over $\operatorname{Spec} A$. The next exercise shows that this will be an easy criterion to check.

13.3.E. IMPORTANT EXERCISE (CF. THE QCQS LEMMA 7.3.5). Suppose X is a quasicompact and quasiseparated scheme (i.e., covered by a finite number of affine open sets, the pairwise intersection of which is also covered by a finite number of affine open sets). Suppose $\mathscr{F}$ is a quasicoherent sheaf on X, and let $f \in \Gamma(X, \mathscr{O}_X)$ be a function on X. Show that the restriction map $\operatorname{res}_{X_f \subset X} : \Gamma(X, \mathscr{F}) \to \Gamma(X_f, \mathscr{F})$ (here X_f is the open subset of X where f doesn't vanish) is precisely localization. In other words show that there is an isomorphism $\Gamma(X, \mathscr{F})_f \to \Gamma(X_f, \mathscr{F})$ making the following diagram commute.

$$\begin{tikzcd} \Gamma(X, \mathscr{F}) \arrow[rr, "\operatorname{res}_{X_f \subset X}"] \arrow[dr, "\otimes_A A_f"'] & & \Gamma(X_f, \mathscr{F}) \\ & \Gamma(X, \mathscr{F})_f \arrow[ur, "\sim"'] & \end{tikzcd}$$

(Hint: Apply the exact functor $\otimes_A A_f$ to the exact sequence

$$0 \to \Gamma(X, \mathscr{F}) \to \oplus_i \Gamma(U_i, \mathscr{F}) \to \oplus \Gamma(U_{ijk}, \mathscr{F})$$

where the U_i form a finite affine cover of X and U_{ijk} form a finite affine cover of $U_i \cap U_j$.)

13.3.F. IMPORTANT EXERCISE (COROLLARY TO EXERCISE 13.3.E: PUSHFORWARDS OF QUASICOHERENT SHEAVES ARE QUASICOHERENT IN REASONABLE CIRCUMSTANCES). Suppose $\pi : X \to Y$ is a quasicompact quasiseparated morphism, and $\mathscr{F}$ is a quasicoherent sheaf on X. Show that $\pi_*\mathscr{F}$ is a quasicoherent sheaf on Y.

13.3.G. EXERCISE (GOOD PRACTICE: THE SHEAF OF NILPOTENTS). If A is a ring, and $f \in A$, show that $\mathfrak{N}(A_f) \cong \mathfrak{N}(A)_f$. Use this to define/construct the quasicoherent **sheaf of nilpotents** on any scheme X. This is an example of an ideal sheaf (of $\mathscr{O}_X$).

13.3.H. IMPORTANT EXERCISE (TO BE USED REPEATEDLY). Generalize Exercise 13.3.E as follows. Suppose X is a quasicompact quasiseparated scheme, $\mathscr{L}$ is an invertible sheaf on X with section s, and $\mathscr{F}$ is a quasicoherent sheaf on X. As in Exercise 13.3.E, let X_s be the open subset of X where s doesn't vanish. Show that any section of $\mathscr{F}$ over X_s can be interpreted as the quotient of a global section of $\mathscr{F} \otimes_{\mathscr{O}_X} \mathscr{L}^{\otimes n}$ by s^n. In other words, any section of $\mathscr{F}$ over X_s can be extended over all of X, once you multiply it by a large enough power of s. More precisely: note that $\oplus_{n \geq 0}\Gamma(X, \mathscr{L}^{\otimes n})$ is a graded ring, and we interpret s as a degree 1 element of it. Note also that $\oplus_{n\geq 0}\Gamma(X, \mathscr{F} \otimes_{\mathscr{O}_X} \mathscr{L}^{\otimes n})$ is a graded module over this ring. Describe a natural map

$$\left(\left(\oplus_{n\geq 0}\Gamma(X, \mathscr{F} \otimes_{\mathscr{O}_X} \mathscr{L}^{\otimes n})\right)_s\right)_0 \to \Gamma(X_s, \mathscr{F})$$

and show that it is an isomorphism. (Hint: after showing the existence of the natural map, show it is an isomorphism in the affine case.)

13.3.I. LESS IMPORTANT EXERCISE. Give a counterexample to show that Exercise 13.3.E need not hold without the quasicompactness hypothesis. (Possible hint: take an infinite disjoint union of affine schemes. The key idea is that infinite direct products do not commute with localization.)

13.3.4. $\star\star$ **Grothendieck topologies.** The distinguished affine base isn't a topology in the usual sense — the union of two affine sets isn't necessarily affine, for example. It is however a first new example of a generalization of a topology — the notion of a **site** or a **Grothendieck topology**. We give the definition to satisfy the curious, but we certainly won't use this notion. (For a clean statement, see [**Stacks**, tag 00VH]; this is intended only as motivation.) The idea is that we should abstract away only those notions we need to define sheaves. We need the notion of open set, but it turns out that we won't even need an underling set, i.e., we won't even need the notion of points! Let's think through how little we need. For our discussion of sheaves to work, we needed to know what the open sets were, and what the (allowed) inclusions were, and these should "behave well", and in particular the data of the open sets and inclusions should form a category. (For example, the composition of an allowed inclusion with another allowed inclusion should be an allowed inclusion — in the distinguished affine base, a distinguished open set of a distinguished open set is a distinguished open set.) So we just require the data of *this category*. At this point, we can already define presheaf (as just a contravariant functor from this category of "open sets"). We saw this idea earlier in Exercise 2.2.A.

In order to extend this definition to that of a sheaf, we need to know more information. We want two open subsets of an open set to intersect in an open set, so *we want the category to be closed under fiber products* (cf. Exercise 1.3.O). For the identity and gluability axioms, we need to know *when some open sets cover another*, so we also remember this as part of the data of a Grothendieck topology. The data of the coverings satisfy some obvious properties. Every open set covers itself (i.e., *the identity map in the category of open sets is a covering*). Coverings pull back: *if we have a map* $Y \to X$, *then any cover of* X *pulls back to a cover of* Y. Finally, *a cover of a cover should be a cover*. Such data (satisfying these axioms) is called a *Grothendieck topology* or a *site*. (There are useful variants of this definition in the literature. Again, we are following [**Stacks**].) We can define the notion of a sheaf on a Grothendieck topology in the usual way, with no change. A **topos** is a scary name for a category of sheaves of sets on a Grothendieck topology.

Grothendieck topologies are used in a wide variety of contexts in and near algebraic geometry. Étale cohomology (using the étale topology), a generalization of Galois cohomology, is a central tool, as are more general flat topologies, such as the smooth topology. The definition of a Deligne-Mumford or Artin stack uses the étale and smooth topology, respectively. Tate developed a good theory of non-archimedean analytic geometry over totally disconnected ground fields such as $\mathbb{Q}_p$ using a suitable Grothendieck topology. Work in K-theory (related for example to Voevodsky's work) uses exotic topologies.

13.4 Quasicoherent sheaves form an abelian category

Morphisms from one quasicoherent sheaf on a scheme X to another are defined to be just morphisms as $\mathscr{O}_X$-modules. In this way, the quasicoherent sheaves on a scheme X form a category, denoted $QCoh_X$. (By definition it is a full subcategory of $Mod_{\mathscr{O}_X}$.) We now show that quasicoherent sheaves on X form an *abelian* category.

When you show that something is an abelian category, you have to check many things, because the definition has many parts. However, if the objects you are considering lie in some ambient abelian category, then it is much easier. You have seen this idea before: there are several things you have to do to check that something is a group. But if you have a subset of group elements, it is much easier to check that it forms a subgroup.

You can look at back at the definition of an abelian category, and you will see that in order to check that a subcategory is an abelian subcategory, it suffices to check only the following:

(i) 0 is in the subcategory
(ii) the subcategory is closed under finite sums
(iii) the subcategory is closed under kernels and cokernels

In our case of $QCoh_X \subset Mod_{\mathscr{O}_X}$, the first two are cheap: 0 is certainly quasicoherent, and the subcategory is closed under finite sums: if $\mathscr{F}$ and $\mathscr{G}$ are sheaves on X, and over $\operatorname{Spec} A$, $\mathscr{F} \cong \widetilde{M}$ and $\mathscr{G} \cong \widetilde{N}$, then $\mathscr{F} \oplus \mathscr{G} = \widetilde{M \oplus N}$ (do you see why?), so $\mathscr{F} \oplus \mathscr{G}$ is a quasicoherent sheaf.

We now check (iii), using the characterization of Important Exercise 13.3.3. Suppose $\alpha : \mathscr{F} \to \mathscr{G}$ is a morphism of quasicoherent sheaves. Then on any affine open set U, where the morphism is given by $\beta : M \to N$, define $(\ker \alpha)(U) = \ker \beta$ and $(\operatorname{coker} \alpha)(U) = \operatorname{coker} \beta$. Then these behave well under inversion of a single element: if

$$0 \to K \to M \to N \to P \to 0$$

is exact, then so is

$$0 \to K_f \to M_f \to N_f \to P_f \to 0,$$

from which $(\ker \beta)_f \cong \ker(\beta_f)$ and $(\operatorname{coker} \beta)_f \cong \operatorname{coker}(\beta_f)$. Thus both of these define quasicoherent sheaves. Moreover, by checking stalks, they are indeed the kernel and cokernel of α (exactness can be checked stalk-locally). Thus the quasicoherent sheaves indeed form an abelian category.

13.4.A. EXERCISE. Show that a sequence of quasicoherent sheaves $\mathscr{F} \to \mathscr{G} \to \mathscr{H}$ on X is exact if and only if it is exact on every open set in any given affine cover of X. (In particular, taking sections over an affine open $\operatorname{Spec} A$ is an exact functor from the category of quasicoherent sheaves on X to the category of A-modules. Recall that taking sections is only left-exact in general, see §2.5.F.) In particular, we may check injectivity or surjectivity of a morphism of quasicoherent sheaves by checking on an affine cover of our choice.

Caution: If $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is an exact sequence of quasicoherent sheaves, then for any open set

$$0 \to \mathscr{F}(U) \to \mathscr{G}(U) \to \mathscr{H}(U)$$

is exact, and exactness on the right is guaranteed to hold only if U is affine. (To set you up for cohomology: whenever you see left-exactness, you expect to eventually interpret this as a start of a long exact sequence. So we are expecting H^1's on the right, and now we expect that $H^1(\operatorname{Spec} A, \mathscr{F}) = 0$. This will indeed be the case.)

13.4.1. *Warning for those already familiar with vector bundles.* Morphisms of vector bundles are more restrictive than morphisms of quasicoherent sheaves that happen to be locally free sheaves. A locally free subsheaf of a locally free sheaf does not always yield a subvector bundle of a vector bundle. The archetypal example is the exact sequence of quasicoherent sheaves on $\mathbb{A}^1_k = \operatorname{Spec} k[t]$ corresponding to the following exact sequence of $k[t]$-modules:

$$0 \to tk[t] \to k[t] \to k \to 0.$$

The locally free sheaf $(tk[t])^{\sim}$ is a subsheaf of $k[t]^{\sim}$, but it does not correspond to a "subvector bundle"; the cokernel is not a vector bundle.

13.4.B. LESS IMPORTANT EXERCISE (CONNECTION TO ANOTHER DEFINITION, AND QUASICOHERENT SHEAVES ON RINGED SPACES IN GENERAL). Show that an $\mathscr{O}_X$-module $\mathscr{F}$ on a scheme X is quasicoherent if and only if there exists an open cover by U_i such that on each U_i, $\mathscr{F}|_{U_i}$ is isomorphic to the cokernel of a map of two free sheaves:

$$\mathscr{O}_{U_i}^{\oplus I} \to \mathscr{O}_{U_i}^{\oplus J} \to \mathscr{F}|_{U_i} \to 0$$

is exact. We have thus connected our definitions to the definition given at the very start of the chapter. This is the definition of a quasicoherent sheaf on a ringed space in general. It is useful in many circumstances, for example in complex analytic geometry.

13.5 Module-like constructions

In a similar way, basically any nice construction involving modules extends to quasicoherent sheaves. (One exception: the Hom of two A-modules is an A-module, but the $\mathcal{H}om$ of two quasicoherent sheaves is quasicoherent only in "reasonable" circumstances, see Exercise 13.7.A. The failure of "niceness" is the failure of Hom to commute with localization — see the aside in Exercise 13.7.A(a) for an example.)

13.5.1. *Locally free sheaves from free modules.*

13.5.A. EXERCISE (POSSIBLE HELP FOR LATER PROBLEMS).
(a) Suppose

$$0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0 \tag{13.5.1.1}$$

is a short exact sequence of locally free sheaves on X. Suppose $U = \operatorname{Spec} A$ is an affine open set where $\mathscr{F}'$, $\mathscr{F}''$ are free, say $\mathscr{F}'|_{\operatorname{Spec} A} = \widetilde{A}^{\oplus a}$, $\mathscr{F}''|_{\operatorname{Spec} A} = \widetilde{A}^{\oplus b}$. (Here a and b are assumed to be finite for convenience, but this is not necessary, so feel free to generalize to the infinite rank case.) Show that $\mathscr{F}$ is also free on $\operatorname{Spec} A$, and that $0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0$ can be interpreted as coming from the tautological exact sequence $0 \to A^{\oplus a} \to A^{\oplus(a+b)} \to A^{\oplus b} \to 0$. (As a consequence, given an exact sequence of quasicoherent sheaves (13.5.1.1) where $\mathscr{F}'$ and $\mathscr{F}''$ are locally free, $\mathscr{F}$ must also be locally free.)
(b) In the finite rank case, show that given an open covering by trivializing affine open sets (of the form described in (a)), the transition functions (really, matrices) of $\mathscr{F}$ may be interpreted as block upper triangular matrices, where the top left $a \times a$ blocks are transition functions for $\mathscr{F}'$, and the bottom $b \times b$ blocks are transition functions for $\mathscr{F}''$.

13.5.B. EXERCISE. Suppose (13.5.1.1) is an exact sequence of quasicoherent sheaves on X. By Exercise 13.5.A(a), if $\mathscr{F}'$ and $\mathscr{F}$" are locally free, then $\mathscr{F}$ is too.
(a) If $\mathscr{F}$ and $\mathscr{F}$" are locally free *of finite rank*, show that $\mathscr{F}'$ is too. Hint: Reduce to the case $X = \operatorname{Spec} A$ and $\mathscr{F}$ and $\mathscr{F}''$ free. Interpret the map $\phi : \mathscr{F} \to \mathscr{F}''$ as an $n \times m$ matrix M with values in A, with m the rank of $\mathscr{F}$ and n the rank of $\mathscr{F}''$. For each point p of X, show that there exist n columns $\{c_1, \dots, c_n\}$ of M that are linearly independent at p and hence near p (as linear independence is given by nonvanishing of the appropriate $n \times n$ determinant). Thus X can be covered by distinguished open subsets in bijection with the choices of n columns of M. Restricting to one subset and renaming columns, reduce to the case where the determinant of the first n columns of M is invertible. Then change coordinates on $A^{\oplus m} = \mathscr{F}(\operatorname{Spec} A)$ so that M with respect to the new coordinates is the identity matrix in the first n columns, and 0 thereafter. Finally, in this case interpret $\mathscr{F}'$ as $\widetilde{A^{\oplus(m-n)}}$.
(b) If $\mathscr{F}'$ and $\mathscr{F}$ are both locally free, show that $\mathscr{F}''$ need not be. (Hint: over $k[t]$, consider $0 \to tk[t] \to k[t] \to k[t]/(t) \to 0$. We will soon interpret this as the closed subscheme exact sequence (13.5.6.1) for a point on $\mathbb{A}^1$.)

13.5.2. *Tensor products.* Another important example is tensor products.

13.5.C. EXERCISE. If $\mathscr{F}$ and $\mathscr{G}$ are quasicoherent sheaves, show that $\mathscr{F} \otimes \mathscr{G}$ is a quasicoherent sheaf described by the following information: If $\operatorname{Spec} A$ is an affine open, and $\Gamma(\operatorname{Spec} A, \mathscr{F}) = M$ and $\Gamma(\operatorname{Spec} A, \mathscr{G}) = N$, then $\Gamma(\operatorname{Spec} A, \mathscr{F} \otimes \mathscr{G}) = M \otimes_A N$, and the restriction map $\Gamma(\operatorname{Spec} A, \mathscr{F} \otimes \mathscr{G}) \to \Gamma(\operatorname{Spec} A_f, \mathscr{F} \otimes \mathscr{G})$ is precisely the localization map $M \otimes_A N \to (M \otimes_A N)_f \cong M_f \otimes_{A_f} N_f$. (We are using the algebraic fact that $(M \otimes_A N)_f \cong M_f \otimes_{A_f} N_f$. You can prove this by universal property if you want, or by using the explicit construction.)

Note that thanks to the machinery behind the distinguished affine base, sheafification is taken care of. This is a feature we will use often: constructions involving quasicoherent sheaves that involve sheafification for general sheaves don't require sheafification when considered on the distinguished affine base. Along with the fact that injectivity, surjectivity, kernels and so on may be computed on affine opens, this is the reason that it is particularly convenient to think about quasicoherent sheaves in terms of affine open sets.

Given a section s of $\mathscr{F}$ and a section t of $\mathscr{G}$, we have a section $s \otimes t$ of $\mathscr{F} \otimes \mathscr{G}$. If $\mathscr{F}$ is an invertible sheaf, this section is often denoted st.

13.5.3. *Tensor algebra constructions.*

For the next exercises, recall the following. If M is an A-module, then the **tensor algebra** $T^\bullet(M)$ is a noncommutative algebra, graded by $\mathbb{Z}^{\geq 0}$, defined as follows. $T^0(M) = A$, $T^n(M) = M \otimes_A \cdots \otimes_A M$ (where n terms appear in the product), and multiplication is what you expect.

The **symmetric algebra** $\operatorname{Sym}^\bullet M$ is a symmetric algebra, graded by $\mathbb{Z}^{\geq 0}$, defined as the quotient of $T^\bullet(M)$ by the (two-sided) ideal generated by all elements of the form $x \otimes y - y \otimes x$ for all $x, y \in M$. Thus $\operatorname{Sym}^n M$ is the quotient of $M \otimes \cdots \otimes M$ by the relations of the form $m_1 \otimes \cdots \otimes m_n - m'_1 \otimes \cdots \otimes m'_n$ where $(m'_1, \dots, m'_n)$ is a rearrangement of $(m_1, \dots, m_n)$.

The **exterior algebra** $\wedge^\bullet M$ is defined to be the quotient of $T^\bullet M$ by the (two-sided) ideal generated by all elements of the form $x \otimes x$ for all $x \in M$. Expanding $(a+b) \otimes (a+b)$, we see that $a \otimes b = -b \otimes a$ in $\wedge^2 M$. This implies that if 2 is invertible in A (e.g. if A is a field of characteristic not 2), $\wedge^n M$ is the quotient of $M \otimes \cdots \otimes M$ by the relations of the form $m_1 \otimes \cdots \otimes m_n - (-1)^{\operatorname{sgn}(\sigma)} m_{\sigma(1)} \otimes \cdots \otimes m_{\sigma(n)}$ where σ is a permutation of $\{1, \dots, n\}$. The exterior algebra is a "skew-commutative" A-algebra.

Better: both Sym and $\wedge$ can defined by universal properties. For example, any map of A-modules $M^{\otimes n} \to N$ that is symmetric in the n entries factors uniquely through $\operatorname{Sym}^n_A(M)$.

It is most correct to write $T^\bullet_A(M)$, $\operatorname{Sym}^\bullet_A(M)$, and $\wedge^\bullet_A(M)$, but the "base ring" A is usually omitted for convenience.

13.5.D. EXERCISE. Suppose $\mathscr{F}$ is a quasicoherent sheaf. Define the quasicoherent sheaves $T^n \mathscr{F}$, $\operatorname{Sym}^n \mathscr{F}$, and $\wedge^n \mathscr{F}$. (One possibility: describe them on each affine open set, and use the characterization of Important Exercise 13.3.3.) If $\mathscr{F}$ is locally free of rank m, show that $T^n \mathscr{F}$, $\operatorname{Sym}^n \mathscr{F}$, and $\wedge^n \mathscr{F}$ are locally free, and find their ranks. (Remark: These constructions can be defined for $\mathscr{O}$-modules on an arbitrary ringed space.) We note that in this case, $\wedge^{\operatorname{rank} \mathscr{F}} \mathscr{F}$ is denoted $\det \mathscr{F}$, and is called the **determinant (line) bundle** or (both better and worse) the **determinant locally free sheaf**.

You can also define the sheaf of noncommutative algebras $T^\bullet \mathscr{F}$, the sheaf of commutative algebras $\mathrm{Sym}^\bullet \mathscr{F}$, and the sheaf of skew-commutative algebras $\wedge^\bullet \mathscr{F}$.

13.5.E. EXERCISE. Suppose $0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0$ is an exact sequence of locally free sheaves. Show that for any r, there is a filtration of $\mathrm{Sym}^r \mathscr{F}$

$$\mathrm{Sym}^r \mathscr{F} = \mathscr{G}^0 \supset \mathscr{G}^1 \supset \cdots \supset \mathscr{G}^r \supset \mathscr{G}^{r+1} = 0$$

with subquotients

$$\mathscr{G}^p / \mathscr{G}^{p+1} \cong (\mathrm{Sym}^p \mathscr{F}') \otimes (\mathrm{Sym}^{r-p} \mathscr{F}'').$$

(Here are two different possible hints for this and Exercise 13.5.F: (1) Interpret the transition matrices for $\mathscr{F}$ as block upper triangular, with two blocks, where one diagonal block gives the transition matrices for $\mathscr{F}'$, and the other gives the transition matrices for $\mathscr{F}''$ (cf. Exercise 13.5.1.1(b)). Then appropriately interpret the transition matrices for $\mathrm{Sym}^r \mathscr{F}$ as block upper triangular, with $r+1$ blocks. (2) It suffices to consider a small enough affine open set $\operatorname{Spec} A$, where $\mathscr{F}', \mathscr{F}, \mathscr{F}''$ are free, and to show that your construction behaves well with respect to localization at an element $f \in A$. In such an open set, the sequence is $0 \to A^{\oplus p} \to A^{\oplus (p+q)} \to A^{\oplus q} \to 0$ by the Exercise 13.5.A. Let $e_1, \dots, e_p$ be the standard basis of $A^{\oplus p}$, and $f_1, \dots, f_q$ be the the standard basis of $A^{\oplus q}$. Let $e'_1, \dots, e'_p$ be denote the images of $e_1, \dots, e_p$ in $A^{\oplus(p+q)}$. Let $f'_1, \dots, f'_q$ be any lifts of $f_1, \dots, f_q$ to $A^{\oplus(p+q)}$. Note that f'_i is well-defined modulo $e'_1, \dots, e'_p$. Note that

$$\mathrm{Sym}^r \mathscr{F}|_{\operatorname{Spec} A} \cong \oplus_{i=0}^r \mathrm{Sym}^i \mathscr{F}'|_{\operatorname{Spec} A} \otimes_{\mathscr{O}_{\operatorname{Spec} A}} \mathrm{Sym}^{r-i} \mathscr{F}''|_{\operatorname{Spec} A}.$$

Show that $\mathscr{G}^p := \oplus_{i=p}^r \mathrm{Sym}^i \mathscr{F}'|_{\operatorname{Spec} A} \otimes_{\mathscr{O}_{\operatorname{Spec} A}} \mathrm{Sym}^{r-i} \mathscr{F}''|_{\operatorname{Spec} A}$ gives a well-defined (locally free) subsheaf that is independent of the choices made, e.g. of the basis e_1, $\dots, e_p$, $f_1, \dots, f_q$, and the lifts $f'_1, \dots, f'_q$.)

13.5.F. USEFUL EXERCISE. Suppose $0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0$ is an exact sequence of locally free sheaves. Show that for any r, there is a filtration of $\wedge^r \mathscr{F}$:

$$\wedge^r \mathscr{F} = \mathscr{G}^0 \supset \mathscr{G}^1 \supset \cdots \supset \mathscr{G}^r \supset \mathscr{G}^{r+1} = 0$$

with subquotients

$$\mathscr{G}^p / \mathscr{G}^{p+1} \cong (\wedge^p \mathscr{F}') \otimes (\wedge^{r-p} \mathscr{F}'')$$

for each p. In particular, if the sheaves have finite rank, then $\det \mathscr{F} = (\det \mathscr{F}') \otimes (\det \mathscr{F}'')$.

13.5.G. EXERCISE. Suppose $\mathscr{F}$ is locally free of rank n. Describe a map $\wedge^r \mathscr{F} \times \wedge^{n-r} \mathscr{F} \to \wedge^n \mathscr{F}$ that induces an isomorphism $\wedge^r \mathscr{F} \to (\wedge^{n-r} \mathscr{F})^\vee \otimes \wedge^n \mathscr{F}$. This is called a **perfect pairing of vector bundles**. (If you know about perfect pairings of vector spaces, do you see why this is a generalization?) You might use this later in showing duality of Hodge numbers of regular varieties over algebraically closed fields, Exercise 21.5.M.

13.5.H. EXERCISE (DETERMINANT LINE BUNDLES BEHAVE WELL IN EXACT SEQUENCES). Suppose $0 \to \mathscr{F}_1 \to \cdots \to \mathscr{F}_n \to 0$ is an exact sequence of finite rank locally free sheaves on X. Show that "the alternating product of determinant bundles is trivial":

$$\det(\mathscr{F}_1)\otimes\det(\mathscr{F}_2)^\vee\otimes\det(\mathscr{F}_3)\otimes\det(\mathscr{F}_4)^\vee\otimes\cdots\otimes\det(\mathscr{F}_n)^{(-1)^n}\cong\mathscr{O}_X.$$

(Hint: break the exact sequence into short exact sequences. Use Exercise 13.5.B(a) to show that they are short exact sequences of *finite rank locally free sheaves*. Then use Exercise 13.5.F.)

13.5.4. Torsion-free sheaves (a stalk-local condition) and torsion sheaves. An A-module M is said to be **torsion-free** if $am=0$ implies that either a is a zerodivisor in A or $m=0$.

In the case where A is an integral domain, which is basically the only context in which we will use this concept, the definition of torsion-freeness can be restated as $am=0$ only if $a=0$ or $m=0$. In this case, the **torsion submodule** of M, denoted M_{tors}, consists of those elements of M annihilated by some nonzero element of A. (If A is not an integral domain, this construction needn't yield an A-module.) Clearly M is torsion-free if and only if $M_{\text{tors}}=0$. We say a module M over an integral domain A is **torsion** if $M=M_{\text{tors}}$; this is equivalent to $M\otimes_A K(A)=0$.

If X is a scheme, then an $\mathscr{O}_X$-module $\mathscr{F}$ is said to be **torsion-free** if $\mathscr{F}_p$ is a torsion-free $\mathscr{O}_{X,p}$-module for all p. (Caution: [**Gr-EGA**] calls this "strictly torsion-free".)

13.5.I. Exercise. Assume (for convenience, not necessity) that A is an integral domain. Show that if M is a torsion-free A-module, then so is any localization of M. Hence show that $\widetilde{M}$ is a torsion-free sheaf on $\operatorname{Spec} A$.

13.5.J. Unimportant exercise (torsion-freeness is not an affine local condition for stupid reasons). Find an example on a two-point space showing that $M:=A$ might not be a torsion-free A-module even though $\mathscr{O}_{\operatorname{Spec} A}=\widetilde{M}$ is torsion-free.

13.5.5. *Definition: torsion quasicoherent sheaves on reduced schemes.* Motivated by the definition of M_{tors} above, we say that a quasicoherent sheaf on a *reduced* scheme is **torsion** if its stalk at the generic point of every irreducible component is 0. We will mainly use this for coherent sheaves on regular curves, where this notion is very simple indeed (see Exercise 13.7.G(b)), but in the literature it comes up in more general situations.

13.5.6. Important: Quasicoherent sheaves of ideals correspond to closed subschemes. Recall that if $i: X\hookrightarrow Y$ is a closed embedding, then we have a surjection of sheaves on Y: $\mathscr{O}_Y \twoheadrightarrow i_*\mathscr{O}_X$ (§8.1). (The i_* is often omitted, as we are considering the sheaf on X as being a sheaf on Y.) The kernel $\mathscr{I}_{X/Y}$ is a "sheaf of ideals" in Y: for each open subset U of Y, the sections form an ideal in the ring of functions on U.

Compare (hard) Exercise 8.1.H and the characterization of quasicoherent sheaves given in (possibly hard) Exercise 13.3.D. You will see that a sheaf of ideals is quasicoherent if and only if it comes from a closed subscheme. (An example of a non-quasicoherent sheaf of ideals was given in Exercise 8.1.F.) We call

$$0\to\mathscr{I}_{X/Y}\to\mathscr{O}_Y\to i_*\mathscr{O}_X\to 0 \tag{13.5.6.1}$$

the **closed subscheme exact sequence** corresponding to $X \hookrightarrow Y$.

13.6 Finite type and coherent sheaves

Here are three natural finiteness conditions on an A-module M. In the case when A is a Noetherian ring, which is the case that almost all of you will ever care about, they are all the same.

The first is the most naive: a module could be **finitely generated**. In other words, there is a surjection $A^{\oplus p} \to M \to 0$.

The second is reasonable too. It could be finitely presented — it could have a finite number of generators with a finite number of relations: there exists a **finite presentation**, i.e., an exact sequence

$$A^{\oplus q} \to A^{\oplus p} \to M \to 0.$$

13.6.A. EXERCISE ("FINITELY PRESENTED IMPLIES ALWAYS FINITELY PRESENTED"). Suppose M is a finitely presented A-module, and $\phi : A^{\oplus p'} \to M$ is *any surjection*. Show that $\ker \phi$ is finitely generated. Hint: Write M as the kernel of $A^{\oplus p}$ by a finitely generated module K. Figure out how to map the short exact sequence $0 \to K \to A^{\oplus p} \to M \to 0$ to the exact sequence $0 \to \ker \phi \to A^{\oplus p'} \to M \to 0$, and use the Snake Lemma (Example 1.7.5).

13.6.1. The third notion is frankly a bit surprising. We say that an A-module M is **coherent** if (i) it is finitely generated, and (ii) whenever we have a map $A^{\oplus p} \to M$ (not necessarily surjective!), the kernel is finitely generated.

13.6.2. Proposition. — *If A is Noetherian, then these three definitions are the same.*

Proof. Clearly coherent implies finitely presented, which in turn implies finitely generated. So suppose M is finitely generated. Take any $A^{\oplus p} \xrightarrow{\alpha} M$. Then $\ker \alpha$ is a submodule of a finitely generated module over A, and is thus finitely generated by Exercise 3.6.X. Thus M is coherent. $\square$

Hence most people can think of these three notions as the same thing.

13.6.3. Proposition. — *The coherent A-modules form an abelian subcategory of the category of A-modules.*

The proof in general is given in §13.8 in a series of short exercises. You should read this only if you are particularly curious.

Proof if A is Noetherian. Recall from our discussion at the start of §13.4 that we must check three things:

- (i) The 0-module is coherent.
- (ii) The category of coherent modules is closed under finite sums.
- (iii) The category of coherent modules is closed under kernels and cokernels.

The first two are clear. For (iii), suppose that $f : M \to N$ is a map of finitely generated modules. Then $\operatorname{coker} f$ is finitely generated (it is the image of N), and

$\ker f$ is too (it is a submodule of a finitely generated module over a Noetherian ring, Exercise 3.6.X). $\square$

13.6.B. $\star$ EASY EXERCISE (ONLY IMPORTANT FOR NON-NOETHERIAN PEOPLE). Show A is coherent as an A-module if and only if the notion of finitely presented agrees with the notion of coherent.

13.6.C. EXERCISE. If $f \in A$, show that if M is a finitely generated (resp. finitely presented, coherent) A-module, then M_f is a finitely generated (resp. finitely presented, coherent) A_f-module. (The "coherent" case is the tricky one.)

13.6.D. EXERCISE. If $(f_1, \dots, f_n) = A$, and M_{f_i} is a finitely generated (resp. finitely presented, coherent) A_{f_i}-module for all i, then M is a finitely generated (resp. finitely presented, coherent) A-module. Hint for the finitely presented case: Exercise 13.6.A.

13.6.4. Definition. A quasicoherent sheaf $\mathscr{F}$ is **finite type** (resp. **finitely presented**, **coherent**) if for every affine open $\operatorname{Spec} A$, $\Gamma(\operatorname{Spec} A, \mathscr{F})$ is a finitely generated (resp. finitely presented, coherent) A-module. Note that coherent sheaves are always finite type, and that on a locally Noetherian scheme, all three notions are the same (by Proposition 13.6.2). Proposition 13.6.3 implies that the coherent sheaves on X form an abelian category, which we denote Coh_X.

Thanks to the Affine Communication Lemma 5.3.2, and the two previous exercises 13.6.C and 13.6.D, it suffices to check "finite typeness" (resp. finite presentation, coherence) on the open sets in a single affine cover. Notice that finite rank locally free sheaves are always finite type, and if $\mathscr{O}_X$ is coherent, finite rank locally free sheaves on X are coherent. (If $\mathscr{O}_X$ is not coherent, then coherence is a pretty useless notion on X.)

13.6.5. *Associated points of coherent sheaves.* Our discussion of associated points in §5.5 immediately implies a notion of **associated point** for a coherent sheaf on a locally Noetherian scheme, with all the good properties described in §5.5. (The affine case was done there, and the only obstacle to generalizing them to coherent sheaves was that we didn't know what coherent sheaves were.) The phrase **associated point of a locally Noetherian scheme** X (without explicit mention of a coherent sheaf) means "associated point of $\mathscr{O}_X$", and similarly for **embedded points**.

13.6.6. *A few words on the notion of coherence.* Proposition 13.6.3 is a good motivation for the definition of coherence: it gives a small (in a non-technical sense) abelian category in which we can think about vector bundles.

There are two sorts of people who should care about the details of this definition, rather than living in a Noetherian world where coherent means finite type. Complex geometers should care. They consider complex-analytic spaces with the classical topology. One can define the notion of coherent $\mathscr{O}_X$-module in a way analogous to this (see [**Se1**, Def. 2]). Then Oka's Theorem states that the structure sheaf of $\mathbb{C}^n$ (hence of any complex manifold) is coherent, and this is very hard (see [**GR**, §2.5] or [**Rem**, §7.2]).

The second sort of people who should care are the sort of arithmetic people who may need to work with non-Noetherian rings, see §3.6.19, or work in non-archimedean analytic geometry.

Warning: it is not uncommon in the later literature to incorrectly define coherent as finitely generated. Please only use the correct definition, as the wrong definition causes confusion. Besides doing this for the reason of honesty, it will also help you see what hypotheses are actually necessary to prove things. And that always helps you remember what the proofs are — and hence why things are true.

13.7 Pleasant properties of finite type and coherent sheaves

We begin with an exercise that $\mathcal{H}om$ behaves reasonably if the source is coherent.

13.7.A. EXERCISE.
(a) Suppose $\mathscr{F}$ is a coherent sheaf on X, and $\mathscr{G}$ is a quasicoherent sheaf on X. Show that $\mathcal{H}om(\mathscr{F},\mathscr{G})$ is a quasicoherent sheaf. Hint: Describe it on affine open sets, and show that it behaves well with respect to localization with respect to f. To show that $\operatorname{Hom}_A(M,N)_f \cong \operatorname{Hom}_{A_f}(M_f,N_f)$, use Exercise 1.6.G. Up to here, you need only the fact that $\mathscr{F}$ is locally finitely presented. (Aside: For an example of quasicoherent sheaves $\mathscr{F}$ and $\mathscr{G}$ on a scheme X such that $\mathcal{H}om(\mathscr{F},\mathscr{G})$ is not quasicoherent, let A be a discrete valuation ring with uniformizer t, let $X = \operatorname{Spec} A$, let $\mathscr{F} = \widetilde{M}$ and $\mathscr{G} = \widetilde{N}$ with $M = \oplus_{i=1}^{\infty} A$ and $N = A$. Then $M_t = \oplus_{i=1}^{\infty} A_t$, and of course $N = A_t$. Consider the homomorphism $\phi : M_t \to N_t$ sending 1 in the ith factor of M_t to $1/t^i$. Then ϕ is not the localization of any element of $\operatorname{Hom}_A(M,N)$.)
(b) If further $\mathscr{G}$ is coherent and $\mathscr{O}_X$ is coherent, show that $\mathcal{H}om(\mathscr{F},\mathscr{G})$ is also coherent.
(c) Suppose $\mathscr{F}$ is a coherent sheaf on X, and $\mathscr{G}$ is a quasicoherent sheaf on X. Show that $\mathcal{H}om(\mathscr{F},\cdot)$ is a left-exact covariant functor $\mathit{QCoh}_X \to \mathit{QCoh}_X$, and that $\mathcal{H}om(\cdot,\mathscr{G})$ is a left-exact contravariant functor $\mathit{Coh}_X \to \mathit{QCoh}_X$ (cf. Exercise 2.5.H). (In fact left-exactness has nothing to do with coherence or quasicoherence — it is true even for $\mathscr{O}_X$-modules, as remarked in §2.5.4. But the result is easier in the category of quasicoherent sheaves.)

13.7.1. *Duals of coherent sheaves.* From Exercise 13.7.A(b), assuming $\mathscr{O}_X$ is coherent, if $\mathscr{F}$ is coherent, its **dual** $\mathscr{F}^\vee := \mathcal{H}om(\mathscr{F},\mathscr{O}_X)$ is too. This generalizes the notion of duals of vector bundles in Exercise 13.1.C. Your argument there generalizes to show that there is always a natural morphism $\mathscr{F} \to (\mathscr{F}^\vee)^\vee$. Unlike in the vector bundle case, this is not always an isomorphism. (For an example, let $\mathscr{F}$ be the coherent sheaf associated to $k[t]/(t)$ on $\mathbb{A}^1 = \operatorname{Spec} k[t]$, and show that $\mathscr{F}^\vee = 0$.) Coherent sheaves for which the "double dual" map is an isomorphism are called **reflexive sheaves**, but we won't use this notion. The canonical map $\mathscr{F} \otimes \mathscr{F}^\vee \to \mathscr{O}_X$ is called the *trace* map — can you see why?

13.7.B. EXERCISE. Suppose $\mathscr{F}$ is a finite rank locally free sheaf, and $\mathscr{G}$ is a quasicoherent sheaf. Describe an isomorphism $\mathscr{H}om(\mathscr{F},\mathscr{G}) \cong \mathscr{F}^\vee \otimes \mathscr{G}$. (This holds more generally if $\mathscr{G}$ is an $\mathscr{O}$-module, but we won't use that, so you may as well prove the simpler result given in this exercise.)

13.7.C. EXERCISE. Suppose

$$0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0 \tag{13.7.1.1}$$

is an exact sequence of quasicoherent sheaves on a scheme X, where $\mathscr{H}$ is a locally free quasicoherent sheaf, and suppose $\mathscr{E}$ is a quasicoherent sheaf. By left-exactness of $\mathscr{H}om$ (Exercise 2.5.H),

$$0 \to \mathscr{H}om(\mathscr{H},\mathscr{E}) \to \mathscr{H}om(\mathscr{G},\mathscr{E}) \to \mathscr{H}om(\mathscr{F},\mathscr{E}) \to 0$$

is exact except possibly on the right. Show that it is also exact on the right. (Hint: this is local, so you can assume that X is affine, say $\operatorname{Spec} A$, and $\mathscr{H} = \widetilde{A^{\oplus n}}$, so (13.7.1.1) can be written as $0 \to M \to N \to A^{\oplus n} \to 0$. Show that this exact sequence splits, so we can write $N = M \oplus A^{\oplus n}$ in a way that respects the exact sequence.) In particular, if $\mathscr{F}$, $\mathscr{G}$, $\mathscr{H}$, and $\mathscr{O}_X$ are all coherent, and $\mathscr{H}$ is locally free, then we have an exact sequence of coherent sheaves

$$0 \to \mathscr{H}^\vee \to \mathscr{G}^\vee \to \mathscr{F}^\vee \to 0.$$

13.7.D. EXERCISE (THE SUPPORT OF A FINITE TYPE QUASICOHERENT SHEAF IS CLOSED). Suppose $\mathscr{F}$ is a sheaf of abelian groups. Recall Definition 2.4.2 of the *support* of a section s of $\mathscr{F}$, and definition (cf. Exercise 2.6.F(b)) of the *support* of $\mathscr{F}$. (Support is a stalk-local notion, and hence behaves well with respect to restriction to open sets, or to stalks. Warning: Support is where the *germ(s)* are nonzero, not where the *value(s)* are nonzero.) Show that the support of a finite type quasicoherent sheaf on a scheme X is a closed subset. (Hint: Reduce to the case X affine. Choose a finite set of generators of the corresponding module.) Show that the support of a quasicoherent sheaf need not be closed. (Hint: If $A = \mathbb{C}[t]$, then $\mathbb{C}[t]/(t-a)$ is an A-module supported at a. Consider $\oplus_{a\in\mathbb{C}}\mathbb{C}[t]/(t-a)$. Be careful: this example won't work if $\oplus$ is replaced by $\prod$.)

13.7.2. *Remark.* In particular, if X is a locally Noetherian scheme, the sheaf of nilpotents (Exercise 13.3.G) is coherent and hence finite type, and thus has closed support. This makes precise the statement promised in §4.2.1, that in good (Noetherian) situations, the fuzz on a scheme is supported on a closed subset. Also, as promised in Remark 5.2.2, if X is a locally Noetherian scheme, the reduced locus forms an open subset. (We already knew all of this as of Remark 5.5.5, but now we know it twice as well.)

We next come to a geometric interpretation of Nakayama's Lemma, which is why Nakayama's Lemma should be considered a geometric fact (with an algebraic proof).

13.7.E. USEFUL EXERCISE: GEOMETRIC NAKAYAMA (GENERATORS OF A FIBER GENERATE A FINITE TYPE QUASICOHERENT SHEAF NEARBY). Suppose X is a scheme, and $\mathscr{F}$ is a finite type quasicoherent sheaf. Show that if $U \subset X$ is a neighborhood of $p \in X$ and $a_1, \dots, a_n \in \mathscr{F}(U)$ so that the images $\overline{a}_1, \dots, \overline{a}_n \in \mathscr{F}_p$ generate $\mathscr{F}|_p$ (defined as $\mathscr{F}_p \otimes \kappa(p)$, §4.3.7), then there is an affine neighborhood

$\mathfrak{p} \subset \operatorname{Spec} A \subset U$ of $\mathfrak{p}$ such that "$a_1|_{\operatorname{Spec} A}, \dots, a_n|_{\operatorname{Spec} A}$ generate $\mathscr{F}|_{\operatorname{Spec} A}$" in the following senses:

(i) $a_1|_{\operatorname{Spec} A}, \dots, a_n|_{\operatorname{Spec} A}$ generate $\mathscr{F}(\operatorname{Spec} A)$ as an A-module;
(ii) for any $\mathfrak{q} \in \operatorname{Spec} A$, $a_1, \dots, a_n$ generate the stalk $\mathscr{F}_\mathfrak{q}$ as an $\mathscr{O}_{X,\mathfrak{q}}$-module (and hence for any $\mathfrak{q} \in \operatorname{Spec} A$, the fibers $a_1|_\mathfrak{q}, \dots, a_n|_\mathfrak{q}$ generate the fiber $\mathscr{F}|_\mathfrak{q}$ as a $\kappa(\mathfrak{q})$-vector space).

In particular, if $\mathscr{F}_\mathfrak{p} \otimes \kappa(\mathfrak{p}) = 0$, then there exists a neighborhood V of $\mathfrak{p}$ such that $\mathscr{F}|_V = 0$.

13.7.F. USEFUL EXERCISE (LOCAL FREENESS OF A COHERENT SHEAF IS A STALK-LOCAL PROPERTY; AND FREE STALKS IMPLY LOCAL FREENESS NEARBY). Suppose $\mathscr{F}$ is a coherent sheaf on scheme X. Show that if $\mathscr{F}_\mathfrak{p}$ is a free $\mathscr{O}_{X,\mathfrak{p}}$-module for some $\mathfrak{p} \in X$, then $\mathscr{F}$ is locally free in some open neighborhood of $\mathfrak{p}$. Hence $\mathscr{F}$ is locally free if and only if $\mathscr{F}_\mathfrak{p}$ is a free $\mathscr{O}_{X,\mathfrak{p}}$-module for all $\mathfrak{p} \in X$. Hint: Find an open neighborhood U of $\mathfrak{p}$, and n elements of $\mathscr{F}(U)$ that generate $\mathscr{F}|_\mathfrak{p} := \mathscr{F}_\mathfrak{p}/\mathfrak{m}\mathscr{F}_\mathfrak{p} = \mathscr{F}_\mathfrak{p} \otimes_{\mathscr{O}_{X,\mathfrak{p}}} \kappa(\mathfrak{p})$ and hence by Nakayama's Lemma they generate $\mathscr{F}_\mathfrak{p}$. Using Geometric Nakayama, Exercise 13.7.E, show that the sections generate $\mathscr{F}_y$ for all y in some neighborhood Y of $\mathfrak{p}$ in U. Thus you have described a surjection $\mathscr{O}_Y^{\oplus n} \to \mathscr{F}|_Y$. Show that the kernel of this map is finite type, and hence has closed support (say $Z \subset Y$), which does not contain $\mathfrak{p}$. Thus $\mathscr{O}_{Y \setminus Z}^{\oplus n} \to \mathscr{F}|_{Y\setminus Z}$ is an isomorphism.

This is enlightening in a number of ways. It shows that for coherent sheaves, local freeness is a stalk-local condition. Furthermore, on an integral scheme, any coherent sheaf $\mathscr{F}$ is automatically free over the generic point (do you see why?), so every coherent sheaf on an integral scheme is locally free over a dense open subset. And any coherent sheaf that is 0 at the generic point of an irreducible scheme is necessarily 0 on a dense open subset. The last two sentences show the utility of generic points; such statements would have been more mysterious in classical algebraic geometry.

13.7.G. EXERCISE. (Torsion-free and torsion sheaves were defined in §13.5.4.)
(a) Show that torsion-free coherent sheaves on a regular (hence implicitly locally Noetherian) curve are locally free.
(b) Show that torsion coherent sheaves on a regular integral curve are supported at a finite number of closed points.
(c) Suppose $\mathscr{F}$ is a coherent sheaf on a regular curve. Describe a canonical short exact sequence $0 \to \mathscr{F}_{\text{tors}} \to \mathscr{F} \to \mathscr{F}_{\text{lf}} \to 0$, where $\mathscr{F}_{\text{tor}}$ is a torsion sheaf, and $\mathscr{F}_{\text{lf}}$ is locally free.

To answer Exercise 13.7.G, use Useful Exercise 13.7.F (local freeness can be checked at stalks) to reduce to the discrete valuation ring case, and recall Remark 12.5.15, the structure theorem for finitely generated modules over a principal ideal domain A: any such module can be written as the direct sum of principal modules $A/(a)$. For discrete valuation rings, this means that the summands are of the form A or $A/\mathfrak{m}^k$. Hence:

13.7.3. Proposition. — *If M is a finitely generated module over a discrete valuation ring, then M is torsion-free if and only if M is free.*

(Exercise 24.2.B is closely related.)

Proposition 13.7.3 is false without the finite generation hypothesis: consider $M = K(A)$ for a suitably general ring A. It is also false if we give up the "dimension 1" hypothesis: consider $(x, y) \subset \mathbb{C}[x, y]$. And it is false if we give up the "regular" hypothesis: consider $(x, y) \subset \mathbb{C}[x, y]/(xy)$. (These examples require some verification.) Hence Exercise 13.7.G(a) is false if we give up the "dimension 1" or "regular" hypothesis.

13.7.4. Rank of a quasicoherent sheaf at a point.
Suppose $\mathscr{F}$ is a quasicoherent sheaf on a scheme X, and p is a point of X. The vector space $\mathscr{F}|_p := \mathscr{F}_p/\mathfrak{m}\mathscr{F}_p = \mathscr{F}_p \otimes_{\mathscr{O}_{X,p}} \kappa(p)$ can be interpreted as the fiber of the sheaf at the point, where $\mathfrak{m}$ is the maximal ideal of $\mathscr{O}_{X,p}$, and $\kappa(p)$ is as usual the residue field $\mathscr{O}_{X,p}/\mathfrak{m}$ at p. A section of $\mathscr{F}$ over an open set containing p can be said to take on a value at that point, which is an element of this vector space. The **rank** of a quasicoherent sheaf $\mathscr{F}$ at a point p is $\dim_{\kappa(p)} \mathscr{F}_p/\mathfrak{m}\mathscr{F}_p$ (possibly infinite). More explicitly, on any affine set $\operatorname{Spec} A$ where $p = [\mathfrak{p}]$ and $\mathscr{F}(\operatorname{Spec} A) = M$, then the rank is $\dim_{K(A/\mathfrak{p})} M_\mathfrak{p}/\mathfrak{p}M_\mathfrak{p}$. Note that this definition of rank is consistent with the notion of rank of a locally free sheaf. In the locally free case, the rank is a (locally) constant function of the point. The converse is sometimes true, see Exercise 13.7.K below.

If X is irreducible, and $\mathscr{F}$ is a quasicoherent (usually coherent) sheaf on X on X, then $\operatorname{rank} \mathscr{F}$ (with no mention of a point) by convention means at the generic point. (For example, a rank 0 quasicoherent sheaf on an integral scheme is a torsion quasicoherent sheaf, see Definition 13.5.5.)

13.7.H. EXERCISE. Consider the coherent sheaf $\mathscr{F}$ on $\mathbb{A}^1_k = \operatorname{Spec} k[t]$ corresponding to the module $k[t]/(t)$. Find the rank of $\mathscr{F}$ at every point of $\mathbb{A}^1$. Don't forget the generic point!

13.7.I. EXERCISE. Show that at any point, $\operatorname{rank}(\mathscr{F} \oplus \mathscr{G}) = \operatorname{rank}(\mathscr{F}) + \operatorname{rank}(\mathscr{G})$ and $\operatorname{rank}(\mathscr{F} \otimes \mathscr{G}) = \operatorname{rank} \mathscr{F} \operatorname{rank} \mathscr{G}$. (Hint: Show that direct sums and tensor products commute with ring quotients and localizations, i.e., $(M \oplus N) \otimes_R (R/I) \cong M/IM \oplus N/IN$, $(M \otimes_R N) \otimes_R (R/I) \cong (M \otimes_R R/I) \otimes_{R/I} (N \otimes_R R/I) \cong M/IM \otimes_{R/I} N/IM$, etc.)

If $\mathscr{F}$ is finite type, then the rank is finite, and by Nakayama's Lemma, the rank is the minimal number of generators of $M_\mathfrak{p}$ as an $A_\mathfrak{p}$-module.

13.7.J. IMPORTANT EXERCISE. If $\mathscr{F}$ is a finite type quasicoherent sheaf on X, show that $\operatorname{rank}(\mathscr{F})$ is an upper semicontinuous function on X. Hint: generators at a point p are generators nearby by Geometric Nakayama's Lemma, Exercise 13.7.E. (The example in Exercise 13.7.D shows the necessity of the finite type hypothesis.)

13.7.K. IMPORTANT HARD EXERCISE.
(a) If X is reduced, $\mathscr{F}$ is a finite type quasicoherent sheaf on X, and the rank is constant, show that $\mathscr{F}$ is locally free. Then use upper semicontinuity of rank (Exercise 13.7.J) to show that finite type quasicoherent sheaves on an integral scheme are locally free on a dense open set. (By examining your proof, you will see that the integrality hypothesis can be relaxed. In fact, reducedness is all that is necessary.) Hint: Reduce to the case where X is affine. Then show it in a neighborhood of an arbitrary point p as follows. Suppose $n = \operatorname{rank} \mathscr{F}$. Choose n generators of the fiber

$\mathscr{F}|_p$ (a basis as an $\kappa(p)$-vector space). By Geometric Nakayama's Lemma 13.7.E, we can find a smaller neighborhood $p \in \operatorname{Spec} A \subset X$, with $\mathscr{F}|_{\operatorname{Spec} A} = \widetilde{M}$, so that the chosen generators $\mathscr{F}|_p$ lift to generators $m_1, \dots, m_n$ of M. Let $\phi : A^{\oplus n} \to M$ with $(r_1, \dots, r_n) \mapsto \sum r_i m_i$. If $\ker \phi \neq 0$, then suppose $(r_1, \dots, r_n)$ is in the kernel, with $r_1 \neq 0$. As $r_1 \neq 0$, there is some $\mathfrak{p}$ where $r_1 \notin \mathfrak{p}$ — here we use the reduced hypothesis. Then r_1 is invertible in $A_\mathfrak{p}$, so $M_\mathfrak{p}$ has fewer than n generators, contradicting the constancy of rank.
(b) Show that part (a) can be false without the condition of X being reduced. (Hint: $\operatorname{Spec} k[x]/x^2$, $M = k$.)

You can use the notion of rank to help visualize finite type quasicoherent sheaves, or even quasicoherent sheaves. For example, I think of a coherent sheaf as generalizing a finite rank vector bundle as follows: to each point there is an associated vector space, and although the ranks can jump, they fit together in families as well as one might hope. You might try to visualize the example of Example 13.7.H. Nonreducedness can fit into the picture as well — how would you picture the coherent sheaf on $\operatorname{Spec} k[\epsilon]/(\epsilon^2)$ corresponding to $k[\epsilon]/(\epsilon)$? How about $k[\epsilon]/(\epsilon^2) \oplus k[\epsilon]/(\epsilon)$?

13.7.5. *Degree of a finite morphism at a point.* Suppose $\pi : X \to Y$ is a finite morphism. Then $\pi_* \mathscr{O}_X$ is a finite type (quasicoherent) sheaf on Y, and the rank of this sheaf at a point p is called the **degree** of the finite morphism at p. By Exercise 13.7.J, the degree of π is a upper semicontinuous function on Y. The degree can jump: consider the closed embedding of a point into a line corresponding to $k[t] \to k$ given by $t \mapsto 0$. It can also be constant in cases that you might initially find surprising — see Exercise 9.3.3, where the degree is always 2, but the 2 is obtained in a number of different ways.

13.7.L. EXERCISE. Suppose $\pi : X \to Y$ is a finite morphism. By unwinding the definition, verify that the degree of π at p is the dimension of the space of functions of the scheme-theoretic preimage of p, considered as a vector space over the residue field $\kappa(p)$. In particular, the degree is zero if and only if $\pi^{-1}(p)$ is empty.

13.8 $\star\star$ Coherent modules over non-Noetherian rings

This section is intended for people who might work with non-Noetherian rings, or who otherwise might want to understand coherent sheaves in a more general setting. Read this only if you really want to!

Suppose A is a ring. Recall the definition of when an A-module M is finitely generated, finitely presented, and coherent. The reason we like coherence is that coherent modules form an abelian category. Here are some accessible exercises working out why these notions behave well. Some repeat earlier discussion in order to keep this section self-contained.

The notion of coherence of a module is only interesting in the case that a ring is coherent over itself. Similarly, coherent sheaves on a scheme X will be interesting only when $\mathscr{O}_X$ is coherent ("over itself"). In this case, coherence is clearly the same as finite presentation. An example where non-Noetherian coherence comes up is the ring $R\langle x_1, \dots, x_n\rangle$ of "restricted power series" over a valuation ring R of a non-discretely valued K (for example, a completion of the algebraic closure of $\mathbb{Q}_p$). This is relevant to Tate's theory of non-archimedean analytic geometry over K (which you can read about in [**BCDKT**], for example). The importance of the coherence of the structure sheaf underlines the importance of Oka's Theorem in complex geometry (stated in §13.6.6).

13.8.A. EXERCISE. Show that coherent implies finitely presented implies finitely generated. (This was discussed at the start of §13.6.)

13.8.B. EXERCISE. Show that 0 is coherent.

Suppose for problems 13.8.C–13.8.I that

$$0 \to M \to N \to P \to 0 \tag{13.8.0.1}$$

is an exact sequence of A-modules. In this series of problems, we will show that if two of $\{M, N, P\}$ are coherent, the third is as well, which will prove very useful.

13.8.1. Hint †. The following hint applies to several of the problems: try to write

$$\begin{tikzcd}
0 \arrow[r] & A^{\oplus p} \arrow[r] \arrow[d] & A^{\oplus(p+q)} \arrow[r] \arrow[d] & A^{\oplus q} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & M \arrow[r] & N \arrow[r] & P \arrow[r] & 0
\end{tikzcd}$$

and possibly use the Snake Lemma 1.7.5.

13.8.C. EXERCISE. Show that N finitely generated implies P finitely generated. (You will only need right-exactness of (13.8.0.1).)

13.8.D. EXERCISE. Show that M, P finitely generated implies N finitely generated. Possible hint: †. (You will only need right-exactness of (13.8.0.1).)

13.8.E. EXERCISE. Show that N, P finitely generated need not imply M finitely generated. (Hint: if I is an ideal, we have $0 \to I \to A \to A/I \to 0$.)

13.8.F. EXERCISE. Show that N coherent, M finitely generated implies M coherent. (You will only need left-exactness of (13.8.0.1).)

13.8.G. EXERCISE. Show that N, P coherent implies M coherent. Hint for (i) in the definition of coherence (§13.6.1):

$$\begin{tikzcd}
 & A^{\oplus q} \arrow[dd] \arrow[dr] & & & \\
 & & A^{\oplus p} \arrow[d] \arrow[dr] & & \\
0 \arrow[r] & M \arrow[r] \arrow[d] & N \arrow[r] \arrow[d] & P \arrow[r] \arrow[dr] & 0 \\
 & 0 & 0 & & 0
\end{tikzcd}$$

(You will only need left-exactness of (13.8.0.1).)

13.8.H. EXERCISE. Show that M finitely generated and N coherent implies P coherent. Hint for (ii) in the definition of coherence (§13.6.1) : †.

13.8.I. EXERCISE. Show that M, P coherent implies N coherent. (Hint: †.)

13.8.J. EXERCISE. Show that a finite direct sum of coherent modules is coherent.

13.8.K. EXERCISE. Suppose M is finitely generated, N coherent. Then if $\phi : M \to N$ is any map, then show that $\operatorname{Im}\phi$ is coherent.

13.8.L. EXERCISE. Show that the kernel and cokernel of maps of coherent modules are coherent.

At this point, we have verified that coherent A-modules form an abelian subcategory of the category of A-modules. (Things you have to check: 0 should be in this set; it should be closed under finite sums; and it should be closed under taking kernels and cokernels.)

13.8.M. EXERCISE. Suppose M and N are coherent submodules of the coherent module P. Show that $M + N$ and $M \cap N$ are coherent. (Hint: consider the right map $M \oplus N \to P$.)

13.8.N. EXERCISE. Show that if A is coherent (as an A-module) then finitely presented modules are coherent. (Of course, if finitely presented modules are coherent, then A is coherent, as A is finitely presented!)

13.8.O. EXERCISE. If M is finitely presented and N is coherent, show that $\operatorname{Hom}(M, N)$ is coherent. (Hint: Hom is left-exact in its first argument.)

13.8.P. EXERCISE. If M is finitely presented, and N is coherent, show that $M \otimes_A N$ is coherent.

13.8.Q. EXERCISE. If $f \in A$, show that if M is a finitely generated (resp. finitely presented, coherent) A-module, then M_f is a finitely generated (resp. finitely presented, coherent) A_f-module. (Hint: localization is exact, Exercise 1.6.F(a).) This exercise is repeated from Exercise 13.6.C to make this section self-contained.

13.8.R. EXERCISE. Suppose $(f_1, \dots, f_n) = A$. Show that if M_{f_i} is finitely generated for all i, then M is too. (Hint: Say M_{f_i} is generated by $m_{ij} \in M$ as an A_{f_i}-module. Show that the m_{ij} generate M. To check surjectivity $\oplus_{i,j} A \to M$, it suffices to check "on $D(f_i)$" for all i.)

13.8.S. EXERCISE. Suppose $(f_1, \dots, f_n) = A$. Show that if M_{f_i} is coherent for all i, then M is too. (Hint: if $\phi : A^{\oplus p} \to M$, then $(\ker\phi)_{f_i} = \ker(\phi_{f_i})$, which is finitely generated for all i. Then apply the previous exercise.)

CHAPTER 14

Line bundles: Invertible sheaves and divisors

We next describe convenient and powerful ways of working with and classifying line bundles (invertible sheaves). We begin with a fundamental example, the line bundles $\mathscr{O}(n)$ on projective space, §14.1. We then introduce Weil divisors (formal sums of codimension 1 subsets), and use them to determine $\operatorname{Pic} X$ in a number of circumstances, §14.2. We finally discuss sheaves of ideals that happen to be invertible (effective Cartier divisors), §14.3. A central theme is that line bundles are closely related to "codimension 1 information".

14.1 Some line bundles on projective space

We now describe an important family of invertible sheaves on projective space over a field k.

As a warm-up, we begin with the invertible sheaf $\mathscr{O}_{\mathbb{P}^1_k}(1)$ on $\mathbb{P}^1_k = \operatorname{Proj} k[x_0, x_1]$. The subscript $\mathbb{P}^1_k$ refers to the space on which the sheaf lives, and is often omitted when it is clear from the context. We describe the invertible sheaf $\mathscr{O}(1)$ using transition functions. It is trivial on the usual affine open sets $U_0 = D(x_0) = \operatorname{Spec} k[x_{1/0}]$ and $U_1 = D(x_1) = \operatorname{Spec} k[x_{0/1}]$. (We continue to use the convention $x_{i/j}$ for describing coordinates on patches of projective space, see §4.4.9.) Thus the data of a section over U_0 is a polynomial in $x_{1/0}$. The transition function from U_0 to U_1 is multiplication by $x_{0/1} = x_{1/0}^{-1}$. The transition function from U_1 to U_0 is hence multiplication by $x_{1/0} = x_{0/1}^{-1}$.

This information is summarized below:

open cover	$U_0 = \operatorname{Spec} k[x_{1/0}]$	$U_1 = \operatorname{Spec} k[x_{0/1}]$
trivialization and transition functions	$\begin{tikzcd} k[x_{1/0}] \arrow[r, bend left, "\times x_{0/1} = x_{1/0}^{-1}"] & k[x_{0/1}] \arrow[l, bend left, "\times x_{1/0} = x_{0/1}^{-1}"] \end{tikzcd}$	

To test our understanding, let's compute the global sections of $\mathscr{O}(1)$. This will generalize our hands-on calculation that $\Gamma(\mathbb{P}^1_k, \mathscr{O}_{\mathbb{P}^1_k}) \cong k$ (Example 4.4.6). A global section is a polynomial $f(x_{1/0}) \in k[x_{1/0}]$ and a polynomial $g(x_{0/1}) \in k[x_{0/1}]$ such that $f(1/x_{0/1})x_{0/1} = g(x_{0/1})$. A little thought will show that f must be linear: $f(x_{1/0}) = a x_{1/0} + b$, and hence $g(x_{0/1}) = a + b x_{0/1}$. Thus

$$\dim \Gamma(\mathbb{P}^1_k, \mathscr{O}(1)) = 2 \neq 1 = \dim \Gamma(\mathbb{P}^1_k, \mathscr{O}).$$

Thus $\mathscr{O}(1)$ is not isomorphic to $\mathscr{O}$, and we have constructed our first (proved) example of a nontrivial line bundle!

We next define more generally $\mathscr{O}_{\mathbb{P}^1_k}(n)$ on $\mathbb{P}^1_k$. It is defined in the same way, except that the transition functions are the nth powers of those for $\mathscr{O}(1)$.

open cover	$U_0 = \operatorname{Spec} k[x_{1/0}]$	$U_1 = \operatorname{Spec} k[x_{0/1}]$
trivialization and transition functions	$k[x_{1/0}] \underset{\times x_{1/0}^n = x_{0/1}^{-n}}{\overset{\times x_{0/1}^n = x_{1/0}^{-n}}{\rightleftarrows}} k[x_{0/1}]$	

In particular, thanks to the explicit transition functions, we see that $\mathscr{O}(n) = \mathscr{O}(1)^{\otimes n}$ (with the obvious meaning if n is negative: $(\mathscr{O}(1)^{\otimes(-n)})^\vee$). Clearly also $\mathscr{O}(m) \otimes \mathscr{O}(n) = \mathscr{O}(m+n)$.

14.1.A. IMPORTANT EXERCISE. Show that $\dim \Gamma(\mathbb{P}^1, \mathscr{O}(n)) = n+1$ if $n \geq 0$, and 0 otherwise.

14.1.1. *Example.* Long ago (§2.5.J), we warned that sheafification was necessary when tensoring $\mathscr{O}_X$-modules: if $\mathscr{F}$ and $\mathscr{G}$ are two $\mathscr{O}_X$-modules on a ringed space, then it is not necessarily true that $\mathscr{F}(X) \otimes_{\mathscr{O}_X(X)} \mathscr{G}(X) \cong (\mathscr{F} \otimes \mathscr{G})(X)$. We now have an example: let $X = \mathbb{P}^1_k$, $\mathscr{F} = \mathscr{O}(1)$, $\mathscr{G} = \mathscr{O}(-1)$, and use the fact that $\mathscr{O}(-1)$ has no nonzero global sections.

14.1.B. EXERCISE. Show that if $m \neq n$, then $\mathscr{O}(m) \not\cong \mathscr{O}(n)$. Hence conclude that we have an injection of groups $\mathbb{Z} \hookrightarrow \operatorname{Pic} \mathbb{P}^1_k$ given by $n \mapsto \mathscr{O}(n)$.

It is useful to identify the global sections of $\mathscr{O}(n)$ with the homogeneous polynomials of degree n in x_0 and x_1, i.e., with the degree n part of $k[x_0, x_1]$ (cf. §14.1.2 for the generalization to $\mathbb{P}^m$). Can you see this from your solution to Exercise 14.1.A? We will see that this identification is natural in many ways. For example, you can show that the definition of $\mathscr{O}(n)$ doesn't depend on a choice of affine cover, and this polynomial description is also independent of cover. (For this, see Example 4.5.12; you can later compare this to Exercise 28.1.M.) As an immediate check of the usefulness of this point of view, ask yourself: where does the section $x_0^3 - x_0 x_1^2$ of $\mathscr{O}(3)$ vanish? The section $x_0 + x_1$ of $\mathscr{O}(1)$ can be multiplied by the section x_0^2 of $\mathscr{O}(2)$ to get a section of $\mathscr{O}(3)$. Which one? Where does the rational section $x_0^4(x_1 + x_0)/x_1^7$ of $\mathscr{O}(-2)$ have zeros and poles, and to what order? (We saw the notion of zeros and poles in Definition 12.5.7, and will meet them again in §14.2, but you should intuitively answer these questions already.)

We now define the invertible sheaf $\mathscr{O}_{\mathbb{P}^m_k}(n)$ on the projective space $\mathbb{P}^m_k$. On the usual affine open set $U_i = \operatorname{Spec} k[x_{0/i}, \dots, x_{m/i}]/(x_{i/i} - 1) = \operatorname{Spec} A_i$, it is trivial, so sections (as an A_i-module) are isomorphic to A_i. The transition function from

U_i to U_j is multiplication by $x_{i/j}^n = x_{j/i}^{-n}$.

$$U_i = \operatorname{Spec} k[x_{0/i}, \dots, x_{m/i}]/(x_{i/i}-1) \qquad U_j = \operatorname{Spec} k[x_{0/j}, \dots, x_{m/j}]/(x_{j/j}-1)$$

$$k[x_{0/i}, \dots, x_{m/i}]/(x_{i/i}-1) \underset{\times x_{j/i}^n = x_{i/j}^{-n}}{\overset{\times x_{i/j}^n = x_{j/i}^{-n}}{\rightleftarrows}} \operatorname{Spec} k[x_{0/j}, \dots, x_{m/j}]/(x_{j/j}-1)$$

Note that these transition functions clearly satisfy the cocycle condition.

14.1.C. ESSENTIAL EXERCISE (CF. EXERCISE 8.2.K). Show that $\dim_k \Gamma(\mathbb{P}_k^m, \mathscr{O}_{\mathbb{P}_k^m}(n)) = \binom{m+n}{m}$.

14.1.2. As in the case of $\mathbb{P}^1$, sections of $\mathscr{O}(n)$ on $\mathbb{P}_k^m$ are naturally identified with homogeneous degree n polynomials in our $m+1$ variables. (Important question: Do you see why? Can you work out this dictionary?) Thus $x+y+2z$ is a section of $\mathscr{O}(1)$ on $\mathbb{P}^2$. It isn't a function, but we know where this section vanishes — precisely where $x+y+2z=0$.

Also, notice that for fixed m, $\binom{m+n}{m}$ is a polynomial in n of degree m *for* $n \geq 0$ (or better: for $n \geq -m-1$). This should be telling you that this function "wants to be a polynomial," but won't succeed without assistance. We will later define $h^0(\mathbb{P}_k^m, \mathscr{O}(n)) := \dim_k \Gamma(\mathbb{P}_k^m, \mathscr{O}(n))$, and later still we will define higher cohomology groups, and we will define the *Euler characteristic* $\chi(\mathbb{P}_k^m, \mathscr{O}(n)) := \sum_{i=0}^{\infty}(-1)^i h^i(\mathbb{P}_k^m, \mathscr{O}(n))$ (cohomology will vanish in degree higher than m). We will discover the moral that the Euler characteristic is better-behaved than h^0, and so we should now suspect (and later prove, see Theorem 18.1.3) that this polynomial is in fact the Euler characteristic, and the reason that it agrees with h^0 for $n \geq 0$ because all the other cohomology groups should vanish.

We finally note that we can define $\mathscr{O}(n)$ on $\mathbb{P}_A^m$ for any ring A: the above definition applies without change.

14.1.3. These are the only line bundles on $\mathbb{P}_k^m$.

Suppose that k is a field. We will see in §14.2.9 that these $\mathscr{O}(n)$ are the only invertible sheaves on $\mathbb{P}_k^m$. The next Exercise shows this when $m=1$, although this approach will be soon be trumped.

14.1.D. EXERCISE. Show that every invertible sheaf on $\mathbb{P}_k^1$ is of the form $\mathscr{O}(n)$ for some n. Hint: use the classification of finitely generated modules over a principal ideal domain (Remark 12.5.15) to show that all invertible sheaves on $\mathbb{A}_k^1$ are trivial. Reduce to determining possible transition functions between the two open subsets in the standard cover of $\mathbb{P}_k^1$.

Caution: there can exist invertible sheaves on $\mathbb{P}_A^1$ not of the form $\mathscr{O}(n)$. You may later be able to think of examples. (Hints to find an example for when you know more: what if Spec A is disconnected? Or if that leads to too silly an example, what if Spec A has nontrivial invertible sheaves?)

14.2 Line bundles and Weil divisors

The notion of Weil divisors gives a great way of understanding and classifying line bundles, at least on Noetherian normal schemes. Some of what we discuss will apply in more general circumstances, and the expert is invited to consider generalizations by judiciously weakening hypotheses in various statements. Before we get started, you should be warned: this is one of those topics in algebraic geometry that is hard to digest — learning it changes the way in which you think about line bundles. But once you become comfortable with the imperfect dictionary to divisors, it becomes second nature.

For the rest of this section, we consider only *Noetherian schemes*. We do this because we will use finite decomposition into irreducible components (Exercise 5.3.B), and Algebraic Hartogs's Lemma 11.3.10.

Define a **Weil divisor** as a formal $\mathbb{Z}$-linear combination of codimension 1 irreducible closed subsets of X. In other words, a Weil divisor is defined to be an object of the form

$$\sum_{Y \subset X \text{ codimension } 1} n_Y [Y]$$

the n_Y are integers, all but a finite number of which are zero. Weil divisors obviously form an abelian group, denoted $\operatorname{Weil} X$. For example, if X is a curve, the Weil divisors are linear combination of closed points.

We say that $[Y]$ is an **irreducible** (Weil) divisor. A Weil divisor $D = \sum n_Y[Y]$ is said to be **effective** if $n_Y \geq 0$ for all Y. In this case we say $D \geq 0$, and by $D_1 \geq D_2$ we mean $D_1 - D_2 \geq 0$. The **support** of a Weil divisor D is the subset $\cup_{n_Y \neq 0} Y$. If $U \subset X$ is an open set, we define the **restriction map** $\operatorname{Weil} X \to \operatorname{Weil} U$ by $\sum n_Y[Y] \mapsto \sum_{Y \cap U \neq \varnothing} n_Y [Y \cap U]$.

Suppose now that X is *regular in codimension* 1. We add this hypothesis because we will use properties of discrete valuation rings. Assume also that X is *reduced*. (This is only so we can talk about rational functions without worrying about them being defined at embedded points. Feel free to relax this hypothesis.) Suppose that $\mathscr{L}$ is an invertible sheaf, and s a rational section not vanishing everywhere on any irreducible component of X. (Rational sections are given by a section over a dense open subset of X, with the obvious equivalence, §13.1.7.) Then s determines a Weil divisor

$$\operatorname{div}(s) := \sum_Y \operatorname{val}_Y(s)[Y]$$

where the summation runs over all irreducible divisors Y of X. We call $\operatorname{div}(s)$ the **divisor of zeros and poles** of the rational section s (cf. Definition 12.5.7). To determine the valuation $\operatorname{val}_Y(s)$ of s along Y, take any open set U containing the generic point of Y where $\mathscr{L}$ is trivializable, along with any trivialization over U; under this trivialization, s is a nonzero rational function on U, which thus has a valuation. Any two such trivializations differ by an invertible function (transition functions are invertible), so this valuation is well-defined. Note that $\operatorname{val}_Y(s) = 0$ for all but finitely many Y, by Exercise 12.5.G. Now consider the set $\{(\mathscr{L}, s)\}$ of pairs of line bundles $\mathscr{L}$ with nonzero rational sections s of $\mathscr{L}$, *up to isomorphism*. This set (after taking quotient by isomorphism) forms an abelian group under tensor product $\otimes$, with identity $(\mathscr{O}_X, 1)$. It is important to notice that if t is an invertible function on X, then multiplication by t gives an isomorphism $(\mathscr{L}, s) \cong (\mathscr{L}, st)$. The map div yields a group homomorphism

(14.2.0.1) $$\operatorname{div} : \{(\mathscr{L}, s)\}/\text{isomorphism} \to \operatorname{Weil} X.$$

14.2.A. EASIER EXERCISE.
(a) *(divisors of rational functions)* Verify that on $\mathbb{A}^1_k$, $\operatorname{div}(x^3/(x+1)) = 3[(x)]-[(x+1)]$ ("$= 3[0] - [-1]$").
(b) *(divisor of rational sections of a nontrivial invertible sheaf)* On $\mathbb{P}^1_k$, there is a rational section of $\mathscr{O}(1)$ "corresponding to" $x^2/(x+y)$. Figure out what this means, and calculate $\operatorname{div}(x^2/(x+y))$.

The homomorphism (14.2.0.1) will be the key to determining all the line bundles on many X. (Note that any invertible sheaf will have such a rational section. For each irreducible component, take a nonempty open set not meeting any other irreducible component; then shrink it so that $\mathscr{L}$ is trivial; choose a trivialization; then take the union of all these open sets, and choose the section on this union corresponding to 1 under the trivialization.) We will see that in reasonable situations, this map div will be injective, and often an isomorphism. Thus by forgetting the rational section (i.e., taking an appropriate quotient), we will have described the Picard group of all line bundles. Let's put this strategy into action.

14.2.1. Proposition. — *If X is normal and Noetherian then the map* div *is injective.*

Proof. Suppose $\operatorname{div}(\mathscr{L}, s) = 0$. Then s has no poles. By Exercise 13.1.K, s is a regular section. We now show that the morphism $\times s : \mathscr{O}_X \to \mathscr{L}$ is in fact an isomorphism; this will prove the Proposition, as it will give an isomorphism $(\mathscr{O}_X, 1) \cong (\mathscr{L}, s)$.

It suffices to show that $\times s$ is an isomorphism on an open subset U of X where $\mathscr{L}$ is trivial, as X is covered by trivializing neighborhoods of $\mathscr{L}$ (as $\mathscr{L}$ is locally trivial). Choose an isomorphism $i : \mathscr{L}|_U \to \mathscr{O}_U$. Composing $\times s$ with i yields a map $\times s' : \mathscr{O}_U \to \mathscr{O}_U$ that is multiplication by a rational function $s' = i(s)$ that has no zeros and no poles. The rational function s' is regular because it has no poles (Exercise 12.5.H), and $1/s'$ is regular for the same reason. Thus s' is an invertible function on U, so $\times s'$ is an isomorphism. Hence $\times s$ is an isomorphism over U. $\square$

Motivated by this, we try to find an inverse to div, or at least to determine the image of div.

14.2.2. Important Definition. Assume now that X is irreducible (purely to avoid making (14.2.2.1) look uglier — but feel free to relax this, see Exercise 14.2.B). **Assume also that X is normal —** this will be a standing assumption for the rest of this section. Suppose D is a Weil divisor. Define the sheaf $\mathscr{O}_X(D)$ by

$$\Gamma(U, \mathscr{O}_X(D)) := \{t \in K(X)^\times \ : \ \operatorname{div}|_U t + D|_U \geq 0\} \cup \{0\}. \tag{14.2.2.1}$$

Here $\operatorname{div}|_U t$ means take the divisor of t considered as a rational function on U, i.e., consider just the irreducible divisors of U. (The subscript X in $\mathscr{O}_X(D)$ is omitted when it is clear from context.) The sections of $\mathscr{O}_X(D)$ over U are the rational functions on U that have poles and zeros "constrained by D": a positive coefficient in D allows a pole of that order; a negative coefficient demands a zero of that order. Away from the support of D, this is (isomorphic to) the structure sheaf (by Algebraic Hartogs's Lemma 11.3.10).

14.2.3. *Remark.* It will be helpful to note that $\mathscr{O}_X(D)$ comes along with a canonical "rational section" corresponding to $1 \in K(X)^\times$. (It is a rational section in the sense that it is a section over a dense open set, namely the complement of $\operatorname{Supp} D$.)

14.2.B. LESS IMPORTANT EXERCISE. Generalize the definition of $\mathscr{O}_X(D)$ to the case when X is not necessarily irreducible. (This is just a question of language. Once you have done this, feel free to drop this hypothesis in the rest of this section.)

14.2.C. EASY EXERCISE. Verify that $\mathscr{O}_X(D)$ is a quasicoherent sheaf. (Hint: the distinguished affine criterion for quasicoherence of Exercise 13.3.D.)

In good situations, $\mathscr{O}_X(D)$ is an invertible sheaf. For example, let $X = \mathbb{A}^1_k$. Consider

$$\mathscr{O}_X\left(-2[(x)] + [(x-1)] + [(x-2)]\right),$$

often written $\mathscr{O}(-2[0] + [1] + [2])$ for convenience. Then $3x^3/(x-1)$ is a global section; it has the required two zeros at $x = 0$ (and even one to spare), and takes advantage of the allowed pole at $x = 1$, and doesn't have a pole at $x = 2$, even though one is allowed. (Unimportant aside: the statement remains true in characteristic 2, although the explanation requires editing.)

14.2.D. EASY EXERCISE. (This is a consequence of later discussion as well, but you should be able to do this by hand.)
(a) Show that any global section of $\mathscr{O}_{\mathbb{A}^1_k}(-2[(x)] + [(x-1)] + [(x-2)])$ is a $k[x]$-multiple of $x^2/(x-1)(x-2)$.
(b) Extend the argument of (a) to give an isomorphism

$$\mathscr{O}_{\mathbb{A}^1_k}\left(-2[(x)] + [(x-1)] + [(x-2)]\right) \cong \mathscr{O}_{\mathbb{A}^1_k}.$$

As suggested by the previous exercise, in good circumstances, $\mathscr{O}_X(D)$ is an invertible sheaf, as shown in the next several exercises. (In fact the $\mathscr{O}_X(D)$ construction can be useful even if $\mathscr{O}_X(D)$ is *not* an invertible sheaf, but this won't concern us here. An example of an $\mathscr{O}_X(D)$ that is not an invertible sheaf is given in Exercise 14.2.H.)

14.2.E. HARD BUT IMPORTANT EXERCISE. Suppose $\mathscr{L}$ is an invertible sheaf, and s is a nonzero rational section of $\mathscr{L}$.
(a) Describe an isomorphism $\mathscr{O}(\operatorname{div} s) \cong \mathscr{L}$. (You will use the normality hypothesis!) Hint: show that those open subsets U for which $\mathscr{O}(\operatorname{div} s)|_U \cong \mathscr{O}_U$ form a base for the Zariski topology. For each such U, define $\phi_U : \mathscr{O}(\operatorname{div} s)(U) \to \mathscr{L}(U)$ sending a rational function t (with zeros and poles "constrained by $\operatorname{div} s$") to st. Show that ϕ_U is an isomorphism (with the obvious inverse map, division by s). Argue that this map induces an isomorphism of sheaves $\phi : \mathscr{O}(\operatorname{div} s) \to \mathscr{L}$.
(b) Let σ be the map from $K(X)$ to the rational sections of $\mathscr{L}$, where $\sigma(t)$ is the rational section of $\mathscr{O}_X(D) \cong \mathscr{L}$ defined via (14.2.2.1) (as described in Remark 14.2.3). Show that the isomorphism of (a) can be chosen such that $\sigma(1) = s$. (Hint: the map in part (a) sends 1 to s.)

14.2.F. EXERCISE (THE EXAMPLE OF §14.1). Suppose $X = \mathbb{P}^n_k$, $\mathscr{L} = \mathscr{O}(1)$, s is the section of $\mathscr{O}(1)$ corresponding to x_0, and $D = \operatorname{div} s$. Verify that $\mathscr{O}(mD) \cong \mathscr{O}(m)$, and the canonical rational section of $\mathscr{O}(mD)$ is precisely s^m. (Watch out for possible confusion: 1 has no pole along $x_0 = 0$, but $\sigma(1) = s^m$ *does* have a zero if $m > 0$.) For this reason, $\mathscr{O}(1)$ is sometimes called the **hyperplane class** in $\operatorname{Pic} X$. (Of course, x_0 can be replaced by any linear form.)

14.2.4. *Definition.* If D is a Weil divisor on (Noetherian normal irreducible) X such that $D = \operatorname{div} f$ for some rational *function* f, we say that D is **principal**. Principal divisors clearly form a subgroup of $\operatorname{Weil} X$; denote this group of principal divisors $\operatorname{Prin} X$. Note that div induces a group homomorphism $K(X)^\times \to \operatorname{Prin} X$. If X can be covered with open sets U_i such that on U_i, D is principal, we say that D is **locally principal**. Locally principal divisors form a subgroup of $\operatorname{Weil} X$, which we denote $\operatorname{LocPrin} X$. (This notation is not standard.)

14.2.5. *Important observation.* As a consequence of Exercise 14.2.E(a) (taking $\mathscr{L} = \mathscr{O}$), if D is principal (and X is normal, a standing hypothesis), then $\mathscr{O}(D) \cong \mathscr{O}$. (Diagram (14.2.7.1) will imply that the converse holds: if $\mathscr{O}(D) \cong \mathscr{O}$, then D is principal.) Thus if D is *locally* principal, $\mathscr{O}_X(D)$ is *locally* isomorphic to $\mathscr{O}_X$ — $\mathscr{O}_X(D)$ is an invertible sheaf.

14.2.G. IMPORTANT EXERCISE. Suppose $\mathscr{O}_X(D)$ is an invertible sheaf.
(a) Show that $\operatorname{div}(\sigma(1)) = D$, where σ was defined in Exercise 14.2.E(b).
(b) Show the converse to Observation 14.2.5: show that D is locally principal.

14.2.6. *Remark.* In definition (14.2.2.1), it may seem cleaner to consider those s such that $\operatorname{div} s \geq D|_U$. The reason for the convention comes from our desire that $\operatorname{div} \sigma(1) = D$. (Taking the "opposite" convention would yield the dual bundle, in the case where D is locally trivial.)

14.2.H. LESS IMPORTANT EXERCISE: A WEIL DIVISOR THAT IS NOT LOCALLY PRINCIPAL. Let $X = \operatorname{Spec} k[x, y, z]/(xy - z^2)$, a cone, and let D be the line $z = x = 0$.
(a) Show that D is not locally principal. (Hint: consider the stalk at the origin. Use the Zariski tangent space, see Problem 12.1.3.) In particular $\mathscr{O}_X(D)$ is not an invertible sheaf.
(b) Show that $\operatorname{div}(x) = 2D$. This corresponds to the fact that the plane $x = 0$ is tangent to the cone X along D.

14.2.I. IMPORTANT EXERCISE. If X is Noetherian and factorial, show that for any Weil divisor D, $\mathscr{O}(D)$ is an invertible sheaf. (Hint: It suffices to deal with the case where D is irreducible, say $D = [Y]$, and to cover X by open sets so that on each open set U there is a function whose divisor is $[Y \cap U]$. One open set will be $X - Y$. Next, we find an open set U containing an arbitrary $p \in Y$, and a function on U. As $\mathscr{O}_{X,p}$ is a unique factorization domain, the prime corresponding to Y is codimension 1 and hence principal by Lemma 11.1.6. Let f be a generator of this prime ideal, interpreted as an element of $K(X)$. It is regular at p, it has a finite number of zeros and poles, and through p, $[Y]$ is the "only zero" (the only component of the divisor of zeros). Let U be X minus all the other zeros and poles.)

14.2.7. The class group. We can now get a handle on the Picard group of a normal Noetherian scheme. Define the **class group** of X, $\operatorname{Cl} X$, by $\operatorname{Weil} X / \operatorname{Prin} X$. By taking the quotient of the inclusion (14.2.0.1) by $\operatorname{Prin} X$, we have the inclusion

$\operatorname{Pic} X \hookrightarrow \operatorname{Cl} X$. This is summarized in the convenient and enlightening diagram.

(14.2.7.1)
$$\begin{tikzcd} & \{(\mathscr{L}, s)\}/\text{iso.} \arrow[r, "\sim", "\text{div}"'] \arrow[d] & \operatorname{LocPrin} X \arrow[r, hook] \arrow[d, "/\operatorname{Prin} X"] \arrow[l, bend right, "(\mathscr{O}(D), \sigma(1)) \leftarrow\!\shortmid D"'] & \operatorname{Weil} X \arrow[d, "/\operatorname{Prin} X"] \\ \operatorname{Pic} X \arrow[r, equal] & \{\mathscr{L}\}/\text{iso.} \arrow[r, "\sim"] & \operatorname{LocPrin} X/\operatorname{Prin} X \arrow[r, hook] \arrow[l, bend left, "\mathscr{O}(D) \leftarrow\!\shortmid D"] & \operatorname{Cl} X \end{tikzcd}$$

This diagram is very important, and although it is short to state, it takes time to internalize.

In particular, if A is a unique factorization domain, then all Weil divisors on $\operatorname{Spec} A$ are principal by Lemma 11.1.6, so $\operatorname{Cl}\operatorname{Spec} A = 0$, and hence $\operatorname{Pic}\operatorname{Spec} A = 0$.

14.2.8. As $k[x_1, \dots, x_n]$ has unique factorization, $\operatorname{Cl}(\mathbb{A}_k^n) = 0$, so $\boxed{\operatorname{Pic}(\mathbb{A}_k^n) = 0}$. Geometers might find this believable — "$\mathbb{C}^n$ is a contractible manifold, and hence should have no nontrivial line bundles" — even if some caution is in order, as the kinds of line bundles being considered are entirely different: holomorphic vs. topological or C^∞. (Aside: for this reason, you might expect that $\mathbb{A}_k^n$ also has no nontrivial vector bundles. This is the Quillen-Suslin Theorem, formerly known as Serre's conjecture, part of Quillen's work leading to his 1978 Fields Medal. The case $n = 1$ was Exercise 13.2.C. For a short proof by Vaserstein, see [**Lan**, p. 850].)

Removing a closed subset of X of codimension greater than 1 doesn't change the class group, as it doesn't change the Weil divisor group or the principal divisors. (Warning: it *can* affect the Picard group, see Exercise 14.2.Q.)

Removing a subset of codimension 1 changes the Weil divisor group in a controllable way. For example, suppose Z is an *irreducible* codimension 1 subset of X. Then we clearly have an exact sequence:

$$0 \longrightarrow \mathbb{Z} \xrightarrow{1 \mapsto [Z]} \operatorname{Weil} X \longrightarrow \operatorname{Weil}(X - Z) \longrightarrow 0.$$

When we take the quotient by principal divisors, taking into account the fact that we may lose exactness on the left, we get an **excision exact sequence for class groups**:

(14.2.8.1)
$$\mathbb{Z} \xrightarrow{1 \mapsto [Z]} \operatorname{Cl} X \longrightarrow \operatorname{Cl}(X - Z) \longrightarrow 0.$$

(Do you see why?)

For example, if U is an open subscheme of $X = \mathbb{A}^n$, $\operatorname{Pic} U = \{0\}$.

As another application, let $X = \mathbb{P}_k^n$, and Z be the hyperplane $x_0 = 0$. We have

$$\mathbb{Z} \longrightarrow \operatorname{Cl}\mathbb{P}_k^n \longrightarrow \operatorname{Cl}\mathbb{A}_k^n \longrightarrow 0$$

from which $\operatorname{Cl}\mathbb{P}_k^n$ is generated by the class $[Z]$, and $\operatorname{Pic}\mathbb{P}_k^n$ is a subgroup of this.

14.2.9. By Exercise 14.2.F, $[Z] \mapsto \mathscr{O}(1)$, and as $\mathscr{O}(n)$ is nontrivial for $n \neq 0$ (Exercise 14.1.B), $[Z]$ is not torsion in $\operatorname{Cl}\mathbb{P}_k^n$. Hence $\operatorname{Pic}\mathbb{P}_k^n \hookrightarrow \operatorname{Cl}\mathbb{P}_k^n$ is an isomorphism, and $\boxed{\operatorname{Pic}\mathbb{P}_k^n \cong \mathbb{Z}}$, with generator $\mathscr{O}(1)$. The **degree** of an invertible sheaf on $\mathbb{P}^n$ is defined using this: define $\deg \mathscr{O}(d)$ to be d.

We have gotten good mileage from the fact that the Picard group of the spectrum of a unique factorization domain is trivial. More generally, Exercise 14.2.I gives us:

14.2.10. Proposition. — *If* X *is Noetherian and factorial, then for any Weil divisor* D, $\mathscr{O}(D)$ *is invertible, and hence the map* $\operatorname{Pic} X \to \operatorname{Cl} X$ *is an isomorphism.*

This can be used to make the connection to the class group in number theory precise, see Exercise 13.1.M; see also §14.2.13.

14.2.11. *Mild but important generalization: twisting line bundles by divisors.* The above constructions can be extended, with $\mathscr{O}_X$ replaced by an arbitrary invertible sheaf, as follows. Let $\mathscr{L}$ be an invertible sheaf on a normal Noetherian scheme X. Then define $\mathscr{L}(D)$ by $\mathscr{O}_X(D) \otimes \mathscr{L}$.

14.2.J. EASY EXERCISE.
(a) Show that sections of $\mathscr{L}(D)$ can be interpreted as rational sections of $\mathscr{L}$ with zeros and poles constrained by D, just as in (14.2.2.1):

$$\Gamma(U, \mathscr{L}(D)) := \{t \text{ rational section of } \mathscr{L} \ : \ \operatorname{div}|_U t + D|_U \geq 0\} \cup \{0\}.$$

(b) Suppose D_1 and D_2 are locally principal. Show that $(\mathscr{O}(D_1))(D_2) \cong \mathscr{O}(D_1 + D_2)$.

14.2.12. Fun examples.

We can now actually calculate some Picard and class groups. First, a useful observation: notice that you can restrict invertible sheaves on Y to any subscheme X, and this can be a handy way of checking that an invertible sheaf is not trivial. Effective Cartier divisors (§8.4.1) sometimes restrict too: if you have an effective Cartier divisor on Y, then it restricts to a closed subscheme on X, locally cut out by one equation. If you are fortunate and this equation doesn't vanish on any associated point of X (§13.6.5), then you get an effective Cartier divisor on X. You can check that the restriction of effective Cartier divisors corresponds to restriction of invertible sheaves (in the sense of Exercise 13.1.H).

14.2.K. EXERCISE: A TORSION PICARD GROUP. Suppose that Y is an irreducible degree d hypersurface of $\mathbb{P}^n_k$. Show that $\operatorname{Pic}(\mathbb{P}^n_k - Y) \cong \mathbb{Z}/(d)$. (For differential geometers: this is related to the fact that $\pi_1(\mathbb{P}^n_k - Y) \cong \mathbb{Z}/(d)$.) Hint: (14.2.8.1).

The next two exercises explore consequences of Exercise 14.2.K, and provide us with some examples promised in Exercise 5.4.N.

14.2.L. EXERCISE (GENERALIZING EXERCISE 5.4.N). Keeping the same notation, assume $d > 1$ (so $\operatorname{Pic}(\mathbb{P}^n - Y) \neq 0$), and let $H_0, \ldots, H_n$ be the $n+1$ coordinate hyperplanes on $\mathbb{P}^n$. Show that $\mathbb{P}^n - Y$ is affine, and $\mathbb{P}^n - Y - H_i$ is a distinguished open subset of it. Show that the $\mathbb{P}^n - Y - H_i$ form an open cover of $\mathbb{P}^n - Y$. Show that $\operatorname{Pic}(\mathbb{P}^n - Y - H_i) = 0$. Then by Exercise 14.2.T, each $\mathbb{P}^n - Y - H_i$ is the Spec of a unique factorization domain, but $\mathbb{P}^n - Y$ is not. Thus the property of being a unique factorization domain is not an affine-local property — it satisfies only one of the two hypotheses of the Affine Communication Lemma 5.3.2.

14.2.M. EXERCISE. Keeping the same notation as the previous exercise, show that on $\mathbb{P}^n - Y$, H_i (restricted to this open set) is an effective Cartier divisor that is not cut out by a single equation. (Hint: Otherwise it would give a trivial element of the class group.)

14.2.N. EXERCISE. Show that $A := \mathbb{R}[x,y]/(x^2+y^2-1)$ is not a unique factorization domain, but $A \otimes_{\mathbb{R}} \mathbb{C}$ is. Hint: Exercise 14.2.L.

14.2.O. EXERCISE: PICARD GROUP OF $\mathbb{P}^1 \times \mathbb{P}^1$. Consider

$$X = \mathbb{P}^1_k \times_k \mathbb{P}^1_k \cong \operatorname{Proj} k[w,x,y,z]/(wz-xy),$$

a smooth quadric surface (see Figure 8.2, and Example 9.6.2). Show that $\operatorname{Pic} X \cong \mathbb{Z} \oplus \mathbb{Z}$ as follows: Show that if $L = \{\infty\} \times_k \mathbb{P}^1 \subset X$ and $M = \mathbb{P}^1 \times_k \{\infty\} \subset X$, then $X - L - M \cong \mathbb{A}^2$. This will give you a surjection $\mathbb{Z} \oplus \mathbb{Z} \twoheadrightarrow \operatorname{Cl} X$. Show that $\mathscr{O}(L)$ restricts to $\mathscr{O}$ on L and $\mathscr{O}(1)$ on M. Show that $\mathscr{O}(M)$ restricts to $\mathscr{O}$ on M and $\mathscr{O}(1)$ on L. (This exercise takes some time, but is enlightening.)

14.2.P. EXERCISE. Show that irreducible smooth projective surfaces (over k) can be birational but not isomorphic. Hint: show $\mathbb{P}^2$ is not isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$ using the Picard group. (Aside: we will see in Exercise 20.2.D that the Picard group of the "blown up plane" is $\mathbb{Z}^2$, but in Exercise 20.2.E we will see that the blown up plane is not isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$, using a little more information in the Picard group.)

This is unlike the case for curves: birational irreducible smooth projective curves (over k) must be isomorphic, as we will see in Theorem 17.4.3. Nonetheless, any two surfaces are related in a simple way: if X and X' are projective, regular, and birational, then X can be sequentially blown up at judiciously chosen points, and X' can too, such that the two results are isomorphic (see [**Ha1**, Thm. V.5.5]; blowing up will be discussed in Chapter 22).

14.2.Q. EXERCISE: PICARD GROUP OF THE CONE. Let $X = \operatorname{Spec} k[x,y,z]/(xy-z^2)$, a cone, where $\operatorname{char} k \neq 2$. (The characteristic hypothesis is not necessary for the result, but is included so you can use Exercise 5.4.H to show normality of X.) Show that $\operatorname{Pic} X = 0$, and $\operatorname{Cl} X \cong \mathbb{Z}/2$. Hint: show that the class of $Z = \{x = z = 0\}$ (the "affine cone over a line") generates $\operatorname{Cl} X$ by showing that its complement $D(x)$ is isomorphic to an open subset of $\mathbb{A}^2_k$. Show that $2[Z] = \operatorname{div}(x)$ and hence principal, and that Z is not principal, Exercise 14.2.H. (Remark: You know enough to show that $X - \{(0,0,0)\}$ is factorial. So although the class group is insensitive to removing loci of codimension greater than 1, §14.2.7, this is not true of the Picard group.)

A Weil divisor (on a normal scheme) with a nonzero multiple corresponding to a line bundle is called $\mathbb{Q}$-**Cartier**. (We won't use this terminology beyond the next exercise.) Exercise 14.2.Q gives an example of a Weil divisor that does not correspond to a line bundle, but is nonetheless $\mathbb{Q}$-Cartier. We now give an example of a Weil divisor that is *not* $\mathbb{Q}$-Cartier.

14.2.R. EXERCISE (A NON-$\mathbb{Q}$-CARTIER DIVISOR). On the cone over the smooth quadric surface $X = \operatorname{Spec} k[w,x,y,z]/(wz-xy)$, let Z be the Weil divisor cut out by $w = x = 0$. Exercise 12.1.D showed that Z is not cut out scheme-theoretically by a single equation. Show more: that if $n \neq 0$, then $n[Z]$ is not locally principal. Hint: show that the complement of an effective Cartier divisor on an affine scheme is also affine, using Proposition 7.3.4. Then if some multiple of Z were locally principal, then the closed subscheme of the complement of Z cut out by $y = z = 0$ would be affine — any closed subscheme of an affine scheme is affine. But this is the scheme $y = z = 0$ (also known as the wx-plane) minus the point $w = x = 0$, which we have seen is non-affine, §4.4.1.

14.2.S. $\star$ EXERCISE (FOR THOSE WITH SUFFICIENT ARITHMETIC BACKGROUND). Identify the (ideal) class group of the ring of integers $\mathscr{O}_K$ in a number field K, as defined in Exercise 13.1.M, with the class group of $\operatorname{Spec} \mathscr{O}_K$, as defined in this section. In particular, you will recover the common description of the class group as formal sums of primes, modulo an equivalence relation coming from principal fractional ideals.

14.2.13. More on class groups and unique factorization.

As mentioned in §5.4.6, there are few commonly used means of checking that a ring is a unique factorization domain. The next exercise is one of them, and it is useful. For example, it implies the classical fact that for rings of integers in number fields, the class group is the obstruction to unique factorization (see Exercise 13.1.M and Proposition 14.2.10).

14.2.T. EXERCISE. Suppose that A is a Noetherian integral domain. Show that A is a unique factorization domain if and only if A is integrally closed and $\operatorname{Cl}\operatorname{Spec} A = 0$. (One direction is easy: we have already shown that unique factorization domains are integrally closed in their fraction fields. Also, Lemma 11.1.6 shows that all codimension 1 primes of a unique factorization domain are principal, so that implies that $\operatorname{Cl}\operatorname{Spec} A = 0$. It remains to show that if A is integrally closed and $\operatorname{Cl}\operatorname{Spec} A = 0$, then all codimension 1 prime ideals are principal, as this characterizes unique factorization domains (Proposition 11.3.5). Algebraic Hartogs's Lemma 11.3.10, may arise in your argument.) This is the third important characterization of unique factorization domains promised in §5.4.6.

My final favorite method of checking that a ring is a unique factorization domain (§5.4.6) is Nagata's Lemma. It is also the least useful.

14.2.U. $\star\star$ EXERCISE (NAGATA'S LEMMA). Suppose A is a Noetherian domain, $x \in A$ an element such that (x) is prime and $A_x = A[1/x]$ is a unique factorization domain. Then A is a unique factorization domain. (Hint: Exercise 14.2.T. Use the short exact sequence

$$[(x)] \to \operatorname{Cl}\operatorname{Spec} A \to \operatorname{Cl}\operatorname{Spec} A_x \to 0$$

(14.2.8.1) to show that $\operatorname{Cl}\operatorname{Spec} A = 0$. Prove that $A[1/x]$ is integrally closed, then show that A is integrally closed as follows. Suppose $T^n + a_{n-1}T^{n-1} + \cdots + a_0 = 0$, where $a_i \in A$, and $T \in K(A)$. Then by integral closure of A_x, we have that $T = r/x^m$, where if $m > 0$, then $r \notin (x)$. Then we quickly get a contradiction if $m > 0$.)

This leads to a fun algebra fact promised in Remark 12.8.6. Suppose k is an algebraically closed field of characteristic not 2. Let $A = k[x_1, \dots, x_n]/(x_1^2 + \cdots + x_m^2)$ where $m \leq n$. When $m \leq 2$, we get some special behavior. (If $m = 0$, we get affine space; if $m = 1$, we get a nonreduced scheme; if $m = 2$, we get a reducible scheme that is the union of two affine spaces.)

If $m \geq 3$, we have verified that $\operatorname{Spec} A$ is normal, in Exercise 5.4.I(b). In fact, if $m \geq 3$, then A is a unique factorization domain *unless* $m = 3$ *or* $m = 4$ (Exercise 5.4.L; see also Exercise 12.1.E). For the case $m = 3$: $A = k[x, y, z, w_1, \dots, w_{n-3}]/(x^2 + y^2 - z^2)$ is not a unique factorization domain, as it is has nonzero class group 0 (by essentially the same argument as for Exercise 14.2.Q).

The failure at 4 comes from the geometry of the quadric surface: we have checked that in $\operatorname{Spec} k[w, x, y, z]/(wz - xy)$, there is a codimension 1 irreducible subset — the cone over a line in a ruling — that is not principal.

14.2.V. EXERCISE (THE CASE $m \geq 5$). Suppose that k is algebraically closed of characteristic not 2. Show that if $\ell \geq 3$, then $A = k[a, b, x_1, \dots, x_n]/(ab - x_1^2 - \cdots - x_\ell^2)$ is a unique factorization domain, by using Nagata's Lemma with $x = a$.

14.3 ⋆ Effective Cartier divisors "=" invertible ideal sheaves

We now give a different means of describing invertible sheaves on a scheme. One advantage of this over Weil divisors is that it can give line bundles on everywhere nonreduced schemes (such a scheme can't be regular at any codimension 1 prime). But we won't use this, so it is less important.

Suppose $D \hookrightarrow X$ is a closed subscheme such that corresponding ideal sheaf $\mathscr{I}$ is an invertible sheaf. Then $\mathscr{I}$ is locally trivial; suppose U is a trivializing affine open set $\operatorname{Spec} A$. Then the closed subscheme exact sequence (13.5.6.1)

$$0 \to \mathscr{I} \to \mathscr{O}_X \to \mathscr{O}_D \to 0$$

corresponds to

$$0 \to I \to A \to A/I \to 0$$

with $I \cong A$ as A-modules. Thus I is generated by a single element, say a, and this exact sequence starts as

$$0 \longrightarrow A \xrightarrow{\times a} A$$

As multiplication by a is injective, a is not a zerodivisor. We conclude that D is locally cut out by a single equation, that is not a zerodivisor. This was the definition of *effective Cartier divisor* given in §8.4.1. This argument is clearly reversible, so we have a quick new definition of effective Cartier divisor (an ideal sheaf $\mathscr{I}$ that is an invertible sheaf — or equivalently, the corresponding closed subscheme).

14.3.A. EASY EXERCISE. Show that a is unique up to multiplication by an invertible function.

In the case where X is locally Noetherian, we can use the language of associated points (§13.6.5), so we can restate this definition as: D is locally cut out by a single equation, not vanishing at any associated point of X.

We now define an invertible sheaf corresponding to D. The seemingly obvious definition would be to take $\mathscr{I}_D$, but instead we define the invertible sheaf $\mathscr{O}(D)$ corresponding to an effective Cartier divisor to be the *dual*: $\mathscr{I}_D^\vee$. (The reason for the dual is Exercise 14.3.B.) The ideal sheaf $\mathscr{I}_D$ is sometimes denoted $\mathscr{O}(-D)$. We have an exact sequence

$$0 \to \mathscr{O}(-D) \to \mathscr{O} \to \mathscr{O}_D \to 0.$$

The invertible sheaf $\mathscr{O}(D)$ has a canonical section s_D: Tensoring $0 \to \mathscr{I} \to \mathscr{O}$ with $\mathscr{I}^\vee$ gives us $\mathscr{O} \to \mathscr{I}^\vee$. (Easy unimportant fact: instead of tensoring $\mathscr{I} \to \mathscr{O}$ with $\mathscr{I}^\vee$, we could have dualized $\mathscr{I} \to \mathscr{O}$, and we would get the same section.)

14.3.B. IMPORTANT AND SURPRISINGLY TRICKY EXERCISE. Recall that a section of a locally free sheaf on X cuts out a closed subscheme of X (Exercise 13.1.I). Show that this section s_D cuts out D. (Compare this to Remark 14.2.6.)

This construction is reversible:

14.3.C. EXERCISE. Suppose $\mathscr{L}$ is an invertible sheaf, and s is a section that is locally not a zerodivisor. (Make sense of this! In particular, if X is locally Noetherian, this means "s does not vanish at an associated point of X", see §13.6.5.) Show that $s = 0$ cuts out an effective Cartier divisor D, and $\mathscr{O}(D) \cong \mathscr{L}$.

14.3.D. EXERCISE. Suppose $\mathscr{I}$ and $\mathscr{J}$ are invertible ideal sheaves (hence corresponding to effective Cartier divisors, say D and D' respectively). Show that $\mathscr{I}\mathscr{J}$ is an invertible ideal sheaf. (We define the **product of two quasicoherent ideal sheaves** $\mathscr{I}\mathscr{J}$ as you might expect: on each affine, we take the product of the two corresponding ideals. To make sure this is well-defined, we need only check that if A is a ring, and $f \in A$, and $I, J \subset A$ are two ideals, then $(IJ)_f = I_f J_f$ in A_f.) We define the corresponding Cartier divisor to be $D + D'$. Verify that $\mathscr{O}(D + D') \cong \mathscr{O}(D) \otimes \mathscr{O}(D')$.

We thus have an important correspondence between *effective Cartier divisors* (closed subschemes whose ideal sheaves are invertible, or equivalently locally cut out by one non-zerodivisor, or in the locally Noetherian case, locally cut out by one equation not vanishing at an associated point) and *ordered pairs* $(\mathscr{L}, s)$ where $\mathscr{L}$ is an invertible sheaf, and s is a section that is not locally a zerodivisor (or in the locally Noetherian case, not vanishing at an associated point). The effective Cartier divisors form an abelian semigroup. We have a map of semigroups, from effective Cartier divisors to invertible sheaves with sections not locally zerodivisors (and hence also to the Picard group of invertible sheaves).

We get lots of invertible sheaves, by taking differences of two effective Cartier divisors. In fact we "usually get them all" — it is very hard to describe an invertible sheaf on a finite type k-scheme that is not describable in such a way. For example, there are none if the scheme is regular or even factorial (basically by Proposition 14.2.10 for factoriality; and regular schemes are factorial by the Auslander-Buchsbaum Theorem 12.8.5). Exercise 16.6.F will imply that there are none if the scheme is projective. It holds in all other reasonable circumstances, see [**Gr-EGA**, IV$_4$.21.3.4]. However, it does not always hold; the first and best example is due to Kleiman, see [**Kl3**].

CHAPTER 15

Quasicoherent sheaves on projective A-schemes

The first two sections of this chapter are relatively straightforward, and the last two are trickier.

15.1 The quasicoherent sheaf corresponding to a graded module

We now describe quasicoherent sheaves on a projective A-scheme. Recall that a projective A-scheme is produced from the data of $\mathbb{Z}^{\geq 0}$-graded ring $S_\bullet$, with $S_0 = A$, and S_+ is a finitely generated ideal (a "finitely generated graded ring over A", §4.5.6). The resulting scheme is denoted $\operatorname{Proj} S_\bullet$.

Suppose $M_\bullet$ is a graded $S_\bullet$-module, *graded by* $\mathbb{Z}$. (While reading the next section, you may wonder why we don't grade by $\mathbb{Z}^{\geq 0}$. You will see that it doesn't matter. A $\mathbb{Z}$-grading will make things cleaner when we produce an $M_\bullet$ from a quasicoherent sheaf on $\operatorname{Proj} S_\bullet$.) We define the quasicoherent sheaf $\widetilde{M_\bullet}$ as follows. (I will avoid calling it $\widetilde{M}$, as this might cause confusion with the affine case; but $\widetilde{M_\bullet}$ is *not* graded in any way.) For each homogeneous f of *positive degree*, we define a quasicoherent sheaf $\widetilde{M_\bullet(f)}$ on the distinguished open $D(f) = \{\mathfrak{p} : f(\mathfrak{p}) \neq 0\}$ by $\widetilde{((M_\bullet)_f)_0}$ — note that $((M_\bullet)_f)_0$ is an $((S_\bullet)_f)_0$-module, and recall that $D(f)$ is identified with $\operatorname{Spec}((S_\bullet)_f)_0$ (Exercise 4.5.F). As in (4.5.7.1), the subscript 0 means "the 0-graded piece". We have obvious isomorphisms of the restriction of $\widetilde{M_\bullet(f)}$ and $\widetilde{M_\bullet(g)}$ to $D(fg)$, satisfying the cocycle conditions. (Think through this yourself, to be sure you agree with the word "obvious"!) By Exercise 2.7.D, these sheaves glue together to a single sheaf $\widetilde{M_\bullet}$ on $\operatorname{Proj} S_\bullet$. We then discard the temporary notation $\widetilde{M_\bullet(f)}$.

The $\mathscr{O}$-module $\widetilde{M_\bullet}$ is clearly quasicoherent, because it is quasicoherent on each $D(f)$, and quasicoherence is local.

15.1.A. EXERCISE. Give an isomorphism between the stalk of $\widetilde{M_\bullet}$ at a point corresponding to homogeneous prime $\mathfrak{p} \subset S_\bullet$ and $((M_\bullet)_\mathfrak{p})_0$. (Remark: You can use this exercise to give an alternate definition of $\widetilde{M_\bullet}$ in terms of "compatible stalks", cf. Exercise 4.5.M.)

Given a map of graded modules $\phi : M_\bullet \to N_\bullet$, we get an induced map of sheaves $\widetilde{M_\bullet} \to \widetilde{N_\bullet}$. Explicitly, over $D(f)$, the map $M_\bullet \to N_\bullet$ induces $(M_\bullet)_f \to (N_\bullet)_f$, which induces $\phi_f : ((M_\bullet)_f)_0 \to ((N_\bullet)_f)_0$; and this behaves well with respect to restriction to smaller distinguished open sets, i.e., the following diagram commutes.

$$\begin{tikzcd} ((M_\bullet)_f)_0 \arrow[r, "\phi_f"] \arrow[d] & ((N_\bullet)_f)_0 \arrow[d] \\ ((M_\bullet)_{fg})_0 \arrow[r, "\phi_{fg}"] & ((N_\bullet)_{fg})_0. \end{tikzcd}$$

Thus $\sim$ is a functor from the category of graded $S_\bullet$-modules to the category of quasicoherent sheaves on $\operatorname{Proj} S_\bullet$.

15.1.B. EASY EXERCISE. Show that $\sim$ is an exact functor. (Hint: everything in the construction is exact.)

15.1.C. EXERCISE. Show that if $M_\bullet$ and $M'_\bullet$ agree in high enough degrees, then $\widetilde{M_\bullet} \cong \widetilde{M'_\bullet}$. Then show that the map from graded $S_\bullet$-modules (up to isomorphism) to quasicoherent sheaves on $\operatorname{Proj} S_\bullet$ (up to isomorphism) is not a bijection. (Really: show this isn't an equivalence of categories.)

Exercise 15.1.C shows that $\sim$ isn't an isomorphism (or equivalence) of categories, but it is close. The relationship is akin to that between presheaves and sheaves, and the sheafification functor (see §15.4).

15.1.D. EXERCISE. Describe a map of S_0-modules $M_0 \to \Gamma(\operatorname{Proj} S_\bullet, \widetilde{M_\bullet})$. (This foreshadows the "saturation map" of §15.4.5 that takes a graded module to its saturation, see Exercise 15.4.C.)

15.1.1. Example: Graded ideals of $S_\bullet$ give closed subschemes of $\operatorname{Proj} S_\bullet$**.** Recall that a graded ideal $I_\bullet \subset S_\bullet$ yields a closed subscheme $\operatorname{Proj} S_\bullet/I_\bullet \hookrightarrow \operatorname{Proj} S_\bullet$. For example, suppose $S_\bullet = k[w,x,y,z]$, so $\operatorname{Proj} S_\bullet \cong \mathbb{P}^3$. The ideal $I_\bullet = (wz - xy, x^2 - wy, y^2 - xz)$ yields our old friend, the twisted cubic (defined in Exercise 8.2.A)

15.1.E. EXERCISE. Show that if the functor $\sim$ is applied to the exact sequence of graded $S_\bullet$-modules

$$0 \to I_\bullet \to S_\bullet \to S_\bullet/I_\bullet \to 0$$

we obtain the closed subscheme exact sequence (13.5.6.1) for $\operatorname{Proj} S_\bullet/I_\bullet \hookrightarrow \operatorname{Proj} S_\bullet$.

We will soon see (Exercise 15.4.H) that all closed subschemes of $\operatorname{Proj} S_\bullet$ arise in this way; this was also promised in Exercise 8.2.B.

15.1.2. Remark. If $M_\bullet$ is finitely generated (resp. finitely presented, coherent), then so is $\widetilde{M_\bullet}$. We will not need this fact. See [**Ro**] for a proof.

15.2 Invertible sheaves (line bundles) on projective A-schemes

Suppose $M_\bullet$ is a graded $S_\bullet$-module. Define the graded module $M(n)_\bullet$ by $M(n)_m := M_{n+m}$. Thus the quasicoherent sheaf $\widetilde{M(n)_\bullet}$ satisfies

$$\Gamma(D(f), \widetilde{M(n)_\bullet}) = ((M_\bullet)_f)_n$$

where here the subscript means we take the nth graded piece. (These subscripts are admittedly confusing!)

15.2.A. EXERCISE. If $S_\bullet = A[x_0, \dots, x_m]$, so $\operatorname{Proj} S_\bullet = \mathbb{P}^m_A$, show $\widetilde{S(n)_\bullet} \cong \mathscr{O}(n)$ using transition functions (cf. §14.1). (Recall from §14.1.2 that the global sections of $\mathscr{O}(n)$ should be identified with the homogeneous degree n polynomials in x_0, $\dots$, x_m. Can you see that in the context of this exercise?)

15.2.1. *Definition.* Motivated by this, if $S_\bullet$ is a graded ring generated in degree 1, we define $\mathscr{O}_{\operatorname{Proj} S_\bullet}(n)$ (or simply $\mathscr{O}(n)$, where $S_\bullet$ is implicit) by $\widetilde{S(n)_\bullet}$.

15.2.B. IMPORTANT EXERCISE. If $S_\bullet$ is generated in degree 1, show that $\mathscr{O}_{\operatorname{Proj} S_\bullet}(n)$ is an invertible sheaf.

If $S_\bullet$ is generated in degree 1, and $\mathscr{F}$ is a quasicoherent sheaf on $\operatorname{Proj} S_\bullet$, define $\mathscr{F}(n) := \mathscr{F} \otimes \mathscr{O}(n)$. This is often called **twisting $\mathscr{F}$ by $\mathscr{O}(n)$**, or **twisting $\mathscr{F}$ by n**. More generally, if $\mathscr{L}$ is an invertible sheaf, then $\mathscr{F} \otimes \mathscr{L}$ is often called **twisting $\mathscr{F}$ by $\mathscr{L}$**.

15.2.C. EXERCISE. If $S_\bullet$ is generated in degree 1, show that $\widetilde{M_\bullet}(n) \cong \widetilde{M(n)_\bullet}$. (Hereafter, we can be cavalier with the placement of the "dot" in such situations.)

15.2.D. EXERCISE. If $S_\bullet$ is generated in degree 1, show that $\mathscr{O}(m+n) \cong \mathscr{O}(m) \otimes \mathscr{O}(n)$.

15.2.2. *Unimportant remark.* Even if $S_\bullet$ is not generated in degree 1, then by Exercise 6.4.G, $S_{d\bullet}$ is generated in degree 1 for some d. In this case, we may define the invertible sheaves $\mathscr{O}(dn)$ for $n \in \mathbb{Z}$. This does *not* mean that we *can't* define $\mathscr{O}(1)$; this depends on $S_\bullet$. For example, if $S_\bullet$ is the polynomial ring $k[x, y]$ with the usual grading, except without linear terms (so $S_\bullet = k[x^2, xy, y^2, x^3, x^2y, xy^2, y^3]$), then $S_{2\bullet}$ and $S_{3\bullet}$ are both generated in degree 1, meaning that we may define $\mathscr{O}(2)$ and $\mathscr{O}(3)$. There is good reason to call their "difference" $\mathscr{O}(1)$.

15.3 Globally generated and base-point-free line bundles

We now come to a topic that is harder, but that will be important. Throughout this section, $S_\bullet$ will be a finitely generated graded ring over A, generated in degree 1. We will prove the following result.

15.3.1. Theorem. — *Any finite type sheaf $\mathscr{F}$ on* $\operatorname{Proj} S_\bullet$ *can be presented in the form*

$$\oplus_{\mathit{finite}} \mathscr{O}(-n) \to \mathscr{F} \to 0.$$

Because we can work with the line bundles $\mathscr{O}(-n)$ in a hands-on way, this result will give us great control over *all* coherent sheaves (and in particular, vector bundles) on $\operatorname{Proj} S_\bullet$. As just a first example, it will allow us to show that every coherent sheaf on a projective k-scheme has a finite-dimensional space of global sections (Corollary 18.1.5). (This fact will grow up to be the fact that the higher pushforward of coherent sheaves under proper morphisms are also coherent, see Theorem 18.8.1(d) and Grothendieck's Coherence Theorem 18.9.1.)

Rather than proceeding directly to a proof, we use this as an excuse to introduce notions that are useful in wider circumstances (*global generation*, *base-point-freeness*, *ampleness*), and their interrelationships. But first we use it as an excuse to mention an important classical result.

15.3.2. *The Hilbert Syzygy Theorem.*

Given any coherent sheaf $\mathscr{F}$ on $\mathbb{P}^n_k$, Theorem 15.3.1 gives the existence of a surjection $\phi : \oplus_{\text{finite}} \mathscr{O}(-m) \to \mathscr{F} \to 0$. The kernel of the surjection is also coherent, so iterating this construction, we can construct an infinite resolution of $\mathscr{F}$ by a direct sum of line bundles:

$$\cdots \oplus_{\text{finite}} \mathscr{O}(m_{2,j}) \to \oplus_{\text{finite}} \mathscr{O}(m_{1,j}) \to \oplus_{\text{finite}} \mathscr{O}(m_{0,j}) \to \mathscr{F} \to 0.$$

The Hilbert Syzygy Theorem states that there is in fact a *finite* resolution, of length at most n. (The Hilbert Syzygy Theorem in fact states more.) Because we won't use this, we don't give a proof, but [**E**] (especially [**E**, Thm. 1.13] and the links thereafter) has an excellent discussion. See the comments after Theorem 3.6.17 for the original history of this result.

15.3.3. Globally generated sheaves. Suppose X is a scheme, and $\mathscr{F}$ is an $\mathscr{O}$-module. The most important definition of this section is the following: $\mathscr{F}$ is **globally generated** (or **generated by global sections**) if it admits a surjection from a free sheaf on X:

$$\mathscr{O}^{\oplus I} \twoheadrightarrow \mathscr{F} .$$

Here I is some index set. The global sections in question are the images of the $|I|$ sections corresponding to 1 in the various summands of $\mathscr{O}_X^{\oplus I}$; those images generate the stalks of $\mathscr{F}$. We say $\mathscr{F}$ is **finitely globally generated** (or **generated by a finite number of global sections**) if the index set I can be taken to be finite.

More definitions in more detail: we say that $\mathscr{F}$ is **globally generated at a point** p (or sometimes **generated by global sections at** p) if we can find $\phi : \mathscr{O}^{\oplus I} \to \mathscr{F}$ that is surjective on stalks at p:

$$\mathscr{O}_p^{\oplus I} \xrightarrow{\phi_p} \mathscr{F}_p.$$

(It would be more precise to say that the stalk of $\mathscr{F}$ at p is generated by global sections of $\mathscr{F}$.) The key insight is that global generation at p means that every germ at p is a linear combination (over the local ring $\mathscr{O}_{X,p}$) of germs of global sections.

Note that $\mathscr{F}$ is *globally generated* if it is globally generated at all points p. (Reason: Exercise 2.4.E showed that isomorphisms can be checked on the level of stalks. An easier version of the same argument shows that surjectivity can also be checked on the level of stalks.) Notice that we can take a single index set for all of X, by taking the union of all the index sets for each p.

15.3.A. Easy Exercise (reality check). Show that every quasicoherent sheaf on every *affine* scheme is globally generated. Show that every finite type quasicoherent sheaf on every affine scheme is generated by a finite number of global sections. (Hint for both: for any A-module M, there is a surjection onto M from a free A-module.)

15.3.B. EASY EXERCISE. Show that if quasicoherent sheaves $\mathscr{F}$ and $\mathscr{G}$ are globally generated at a point p, then so is $\mathscr{F} \otimes \mathscr{G}$.

15.3.C. EASY BUT IMPORTANT EXERCISE. Suppose $\mathscr{F}$ is a finite type quasicoherent sheaf on X.
(a) Show that $\mathscr{F}$ is globally generated at p if and only if "the fiber of $\mathscr{F}$ is generated by global sections at p", i.e., the map from global sections to the fiber $\mathscr{F}_p/\mathfrak{m}\mathscr{F}_p$ is surjective, where $\mathfrak{m}$ is the maximal ideal of $\mathscr{O}_{X,p}$. (Hint: Geometric Nakayama, Exercise 13.7.E.)
(b) Show that if $\mathscr{F}$ is globally generated at p, then "$\mathscr{F}$ is globally generated near p": there is an open neighborhood U of p such that $\mathscr{F}$ is globally generated at every point of U.
(c) Suppose further that X is a quasicompact scheme. Show that if $\mathscr{F}$ is globally generated at all closed points of X, then $\mathscr{F}$ is globally generated at all points of X. (Note that nonempty quasicompact schemes *have* closed points, Exercise 5.1.E.)

15.3.D. EASY EXERCISE. If $\mathscr{F}$ is a finite type quasicoherent sheaf on X, and X is quasicompact, show that $\mathscr{F}$ is globally generated if and only if it is generated by a *finite number* of global sections.

15.3.E. EASY EXERCISE. An invertible sheaf $\mathscr{L}$ on X is globally generated if and only if for any point $x \in X$, there is a global section of $\mathscr{L}$ not vanishing at x. (See Theorem 16.4.1 for why we care.)

15.3.4. *Definitions.* If $\mathscr{L}$ is an invertible sheaf on X, then those points where all sections of $\mathscr{L}$ vanish are called the **base points** of $\mathscr{L}$, and the set of base points is called the **base locus** of $\mathscr{L}$; it is a closed subset of X. (We can refine this to a closed subscheme: by taking the scheme-theoretic intersection of the vanishing loci of the sections of $\mathscr{L}$, we obtain the **scheme-theoretic base locus**.) The complement of the **base locus** is the **base-point-free locus**. If $\mathscr{L}$ has no base-points, it is **base-point-free**. By the previous discussion, (i) the base-point-free locus is an open subset of X, and (ii) $\mathscr{L}$ is generated by global sections if and only if it is base-point free. By Exercise 15.3.B, the tensor of two base-point-free line bundles is base-point-free.

(Remark: We will see in Exercise 18.2.I that if X is a k-scheme, and $\mathscr{L}$ is an invertible sheaf on X, and K/k is any field extension, then $\mathscr{L}$ is base-point-free if and only if it is "base-point-free after base change to K". You could reasonably prove this now.)

15.3.5. *Base-point-free line bundles and maps to projective space.* The main reason we care about the definitions above is the following. Recall Exercise 6.3.N(a), which shows that $n+1$ functions on a scheme X with no common zeros yield a map to $\mathbb{P}^n$. This notion generalizes.

15.3.F. EASY EXERCISE (A VITALLY IMPORTANT CONSTRUCTION). Suppose s_0, $\dots$, s_n are $n+1$ global sections of an invertible sheaf $\mathscr{L}$ on a scheme X, with no common zero. Define a corresponding map to $\mathbb{P}^n$:

$$X \xrightarrow{[s_0,\dots,s_n]} \mathbb{P}^n$$

Hint: If U is an open subset on which $\mathscr{L}$ is trivial, choose a trivialization, then translate the s_i into functions using this trivialization, and use Exercise 6.3.N(a) to obtain a morphism $U \to \mathbb{P}^n$. Then show that all of these maps (for different U and different trivializations) "agree", using Exercise 6.3.N(b).

(In Theorem 16.4.1, we will see that this yields *all* maps to projective space.) Note that this exercise as written "works over $\mathbb{Z}$" (as *all* morphisms are "over" the final object in the category of schemes), although many readers will just work over a particular base such as a given field k. Here is some convenient classical language which is used in this case.

15.3.6. *Definitions.* A **linear series** on a k-scheme X is a k-vector space V (usually finite-dimensional), an invertible sheaf $\mathscr{L}$, and a linear map $\lambda : V \to \Gamma(X, \mathscr{L})$. Such a linear series is often called "V", with the rest of the data left implicit. If the map λ is an isomorphism, it is called a **complete linear series**, and is often written $|\mathscr{L}|$. The language of base-points (Definition 15.3.4) readily translates to this situation. For example, given a linear series, any point $x \in X$ on which all elements of the linear series V vanish, we say that x is a **base-point** of V. If V has no base-points, we say that it is **base-point-free**. The union of base-points is called the **base locus** of the linear series. One can similarly define the **scheme-theoretic base locus** (or **base scheme**, although this phrase can have another meaning) of the linear series.

As a reality check, you should understand why an $n+1$-dimensional linear series on a k-scheme X with base-point-free locus U defines a morphism $U \to \mathbb{P}^n_k$.

15.3.7. *Serre's Theorem A.* We are now able to state a celebrated result of Serre.

15.3.8. Serre's Theorem A. — *Suppose $S_\bullet$ is generated in degree 1, and finitely generated over $A = S_0$. Let $\mathscr{F}$ be any finite type quasicoherent sheaf on $\operatorname{Proj} S_\bullet$. Then there exists some n_0 such that for all $n \geq n_0$, $\mathscr{F}(n)$ is finitely globally generated.*

We could now prove Serre's Theorem A directly, but will continue to use this as an excuse to introduce more ideas; it will be a consequence of Theorem 16.6.2.

15.3.9. *Proof of Theorem 15.3.1 assuming Serre's Theorem A (Theorem 15.3.8).* Suppose we have m global sections $s_1, \ldots, s_m$ of $\mathscr{F}(n)$ that generate $\mathscr{F}(n)$. This gives a map

$$\oplus_m \mathscr{O} \longrightarrow \mathscr{F}(n)$$

given by $(f_1, \ldots, f_m) \mapsto f_1 s_1 + \cdots + f_m s_m$ on any open set. Because these global sections generate $\mathscr{F}(n)$, this is a surjection. Tensoring with $\mathscr{O}(-n)$ (which is exact, as tensoring with any locally free sheaf is exact, Exercise 13.1.E) gives the desired result. $\square$

15.4 Quasicoherent sheaves and graded modules

(This section answers some fundamental questions, but it is surprisingly tricky. You may wish to skip this section, or at least the proofs, on first reading, unless you have a particular need for them.)

Throughout this section, $S_\bullet$ is a finitely generated graded algebra *generated in degree* 1, so in particular we have the invertible sheaf $\mathscr{O}(n)$ for all n by Exercise 15.2.B. Also, throughout, $M_\bullet$ is a graded $S_\bullet$-module, and $\mathscr{F}$ is a quasicoherent sheaf on $\operatorname{Proj} S_\bullet$.

We know how to get quasicoherent sheaves on $\operatorname{Proj} S_\bullet$ from graded $S_\bullet$-modules. We will now see that we can get them all in this way. We will define a functor $\Gamma_\bullet$ from (the category of) quasicoherent sheaves on $\operatorname{Proj} S_\bullet$ to (the category of) graded $S_\bullet$-modules that will attempt to reverse the $\sim$ construction. They are not quite inverses, as $\sim$ can turn two different graded modules into the same quasicoherent sheaf (see for example Exercise 15.1.C). But we will see a natural isomorphism $\widetilde{\Gamma_\bullet(\mathscr{F})} \cong \mathscr{F}$. In fact $\Gamma_\bullet(\widetilde{M_\bullet})$ is a better ("saturated") version of $M_\bullet$, and there is a saturation functor $M_\bullet \to \Gamma_\bullet(\widetilde{M_\bullet})$ that is akin to groupification and sheafification — it is adjoint to the forgetful functor from saturated graded modules to graded modules. And thus we come to the fundamental relationship between $\sim$ and $\Gamma_\bullet$: they are an adjoint pair.

$$
\begin{tikzcd}
 & \text{graded } S_\bullet\text{-modules} \arrow[dl, "\sim"'] \arrow[dd, bend right, "\text{saturate}"'] \\
QCoh_{\operatorname{Proj} S_\bullet} \arrow[dr, "\text{equivalence}", "\Gamma_\bullet"'] & \\
 & \text{saturated graded } S_\bullet\text{-modules} \arrow[uu, bend right, "\text{forget}"']
\end{tikzcd}
$$

We now make some of this precise, but as little as possible to move forward. In particular, we will show that every quasicoherent sheaf on a projective A-scheme arises from a graded module (Corollary 15.4.3), and that every closed subscheme of $\operatorname{Proj} S_\bullet$ arises from a graded ideal $I_\bullet \subset S_\bullet$ (Exercise 15.4.H).

15.4.1. Definition of $\Gamma_\bullet$. When you do Essential Exercise 14.1.C (on global sections of $\mathscr{O}_{\mathbb{P}^m_k}(n)$), you will suspect that in good situations,

$$M_n \cong \Gamma(\operatorname{Proj} S_\bullet, \widetilde{M}(n)).$$

Motivated by this, we define

$$\Gamma_n(\mathscr{F}) := \Gamma(\operatorname{Proj} S_\bullet, \mathscr{F}(n)).$$

15.4.A. EXERCISE. Describe a morphism of S_0-modules $M_n \to \Gamma(\operatorname{Proj} S_\bullet, \widetilde{M(n)_\bullet})$, extending the $n = 0$ case of Exercise 15.1.D.

15.4.B. EXERCISE. Show that $\Gamma_\bullet(\mathscr{F})$ is a graded $S_\bullet$-module. (Hint: consider $S_n \to \Gamma(\operatorname{Proj} S_\bullet, \mathscr{O}(n))$.)

15.4.C. EXERCISE. Show that the map $M_\bullet \to \Gamma_\bullet(\widetilde{M_\bullet})$ arising from the previous two exercises is a map of $S_\bullet$-modules. We call this the **saturation map**.

15.4.D. EXERCISE.
(a) Show that the saturation map need not be injective, nor need it be surjective. (Hint: $S_\bullet = k[x]$, $M_\bullet = k[x]/x^2$ or $M_\bullet = xk[x]$.)

(b) On the other hand, show that if $S_\bullet$ is a finitely generated graded ring over a field k, and $M_\bullet$ is finitely generated, then the saturation map is an isomorphism in large degree. In other words, show that there exists an n_0 such that $M_n \to \Gamma(\operatorname{Proj} S_\bullet, \widetilde{M(n)_\bullet})$, is an isomorphism for $n \geq n_0$.

15.4.E. EXERCISE. Show that $\Gamma_\bullet$ is a functor from $QCoh_{\operatorname{Proj} S_\bullet}$ to the category of graded $S_\bullet$-modules. In other words, if $\mathscr{F} \to \mathscr{G}$ is a morphism of quasicoherent sheaves on $\operatorname{Proj} S_\bullet$, describe the natural map $\Gamma_\bullet \mathscr{F} \to \Gamma_\bullet \mathscr{G}$, and show that such maps respect the identity and composition.

15.4.2. $\star$ **The reverse map.** Now that we have defined the saturation map $M_\bullet \to \Gamma_\bullet \widetilde{M_\bullet}$, we will describe a map $\widetilde{\Gamma_\bullet \mathscr{F}} \to \mathscr{F}$. While subtler to define, it will have the advantage of being an isomorphism.

15.4.F. EXERCISE. Define the natural map $\widetilde{\Gamma_\bullet \mathscr{F}} \to \mathscr{F}$ as follows. First describe the map on sections over $D(f)$. Note that sections of the left side are of the form m/f^n where $m \in \Gamma_{n \deg f}(\mathscr{F})$, and $m/f^n = m'/f^{n'}$ if there is some N with $f^N(f^{n'}m - f^n m') = 0$. Sections on the right are implicitly described in Exercise 13.3.H. Show that your map behaves well on overlaps $D(f) \cap D(g) = D(fg)$.

15.4.G. TRICKY EXERCISE. Show that the natural map $\widetilde{\Gamma_\bullet \mathscr{F}} \to \mathscr{F}$ is an isomorphism, by showing that it is an isomorphism of sections over $D(f)$ for any f. First show surjectivity, using Exercise 13.3.H to show that any section of $\mathscr{F}$ over $D(f)$ is of the form m/f^n where $m \in \Gamma_{n \deg f}(\mathscr{F})$. Then verify that it is injective.

15.4.3. Corollary. — *Every quasicoherent sheaf on a projective A-scheme arises from the $\sim$ construction.*

15.4.H. EXERCISE. Show that each closed subscheme of $\operatorname{Proj} S_\bullet$ arises from a graded ideal $I_\bullet \subset S_\bullet$. (Hint: Suppose Z is a closed subscheme of $\operatorname{Proj} S_\bullet$. Consider the exact sequence $0 \to \mathscr{I}_Z \to \mathscr{O}_{\operatorname{Proj} S_\bullet} \to \mathscr{O}_Z \to 0$. Apply $\Gamma_\bullet$, and then $\sim$. Be careful: $\Gamma_\bullet$ is left-exact, but not necessarily exact.) This fulfills promises made in Exercises 8.2.B and 15.1.E.

For the first time, we see that every closed subscheme of a projective scheme is cut out by homogeneous equations. This is the analog of the fact that every closed subscheme of an affine scheme is cut out by equations. It is disturbing that it is so hard to prove this fact. (For comparison, this was easy on the level of the Zariski topology — see Exercise 4.5.H(c).)

15.4.I. $\star\star$ EXERCISE ($\Gamma_\bullet$ AND $\sim$ ARE ADJOINT FUNCTORS). Describe a natural bijection $\operatorname{Hom}(M_\bullet, \Gamma_\bullet \mathscr{F}) \cong \operatorname{Hom}(\widetilde{M_\bullet}, \mathscr{F})$, as follows.

(a) Show that maps $M_\bullet \to \Gamma_\bullet \mathscr{F}$ are the "same" as maps $((M_\bullet)_f)_0 \to ((\Gamma_\bullet \mathscr{F})_f)_0$ as f varies through S_+, that are "compatible" as f varies, i.e., if $D(g) \subset$

D(f), there is a commutative diagram

$$\begin{tikzcd} ((M_\bullet)_f)_0 \arrow[r] \arrow[d] & ((\Gamma_\bullet \mathscr{F})_f)_0 \arrow[d] \\ ((M_\bullet)_g)_0 \arrow[r] & ((\Gamma_\bullet \mathscr{F})_g)_0 \end{tikzcd}$$

More precisely, give a bijection between $\mathrm{Hom}(M_\bullet, \Gamma_\bullet \mathscr{F})$ and the set of compatible maps

$$\left(\mathrm{Hom}((M_\bullet)_f)_0 \to ((\Gamma_\bullet \mathscr{F})_f)_0 \right)_{f \in S_+}.$$

(b) Describe a bijection between the set of compatible maps $(\mathrm{Hom}((M_\bullet)_f)_0 \to ((\Gamma_\bullet \mathscr{F})_f)_0)_{f \in S_+}$ and the set of compatible maps $\Gamma(D(f), \widetilde{M_\bullet}) \to \Gamma(D(f), \mathscr{F})$.

15.4.4. *The special case* $M_\bullet = S_\bullet$. We have a saturation map $S_\bullet \to \Gamma_\bullet \widetilde{S_\bullet}$, which is a map of $S_\bullet$-modules. But $\Gamma_\bullet \widetilde{S_\bullet}$ has the structure of a graded ring (basically because we can multiply sections of $\mathscr{O}(m)$ by sections of $\mathscr{O}(n)$ to get sections of $\mathscr{O}(m+n)$, see Exercise 15.2.D).

15.4.J. EXERCISE. Show that the map of graded rings $S_\bullet \to \Gamma_\bullet \widetilde{S_\bullet}$ induces (via the construction of Essential Exercise 6.4.A) an isomorphism $\mathrm{Proj}\, \Gamma_\bullet \widetilde{S_\bullet} \to \mathrm{Proj}\, S_\bullet$, and under this isomorphism, the respective $\mathscr{O}(1)$'s are identified.

This addresses the following question: to what extent can we recover $S_\bullet$ from $(\mathrm{Proj}\, S_\bullet, \mathscr{O}(1))$? The answer is: we cannot recover $S_\bullet$, but we can recover its "saturation". And better yet: given a projective A-scheme $\pi : X \to \mathrm{Spec}\, A$, along with $\mathscr{O}(1)$, we obtain it as a Proj of a graded algebra in a canonical way, via

$$X \cong \mathrm{Proj}\left(\oplus_{n \geq 0} \Gamma(X, \mathscr{O}(n))\right).$$

There is one last worry you might have, which is assuaged by the following exercise.

15.4.K. EXERCISE. Suppose $S_\bullet$ is a finitely generated graded ring over A, so $X = \mathrm{Proj}\, S_\bullet$ is a projective A-scheme. Show that $(\oplus_{n \geq 0} \Gamma(X, \mathscr{O}(n)))$ is a finitely generated A-algebra. (Hint: $S_\bullet$ and $(\oplus_{n \geq 0} \Gamma(X, \mathscr{O}(n)))$ agree in sufficiently high degrees, by Exercise 15.4.D.)

15.4.5. ⋆⋆ **Saturated $S_\bullet$-modules.** We end with a remark: different graded $S_\bullet$-modules give the same quasicoherent sheaf on $\mathrm{Proj}\, S_\bullet$, but the results of this section show that there is a "best" (saturated) graded module for each quasicoherent sheaf, and there is a map from each graded module to its "best" version, $M_\bullet \to \Gamma_\bullet \widetilde{M_\bullet}$. A module for which this is an isomorphism (a "best" module) is called *saturated*. We won't use this term later.

This "saturation" map $M_\bullet \to \Gamma_\bullet \widetilde{M_\bullet}$ is analogous to the sheafification map, taking presheaves to sheaves. For example, the saturation of the saturation equals the saturation.

There is a bijection between saturated quasicoherent sheaves of ideals on $\mathrm{Proj}\, S_\bullet$ and closed subschemes of $\mathrm{Proj}\, S_\bullet$.

CHAPTER 16

Pushforwards and pullbacks of quasicoherent sheaves

16.1 Introduction

This chapter is devoted to pushforward and pullbacks of quasicoherent sheaves, their properties, and some applications.

Suppose $B \to A$ is a morphism of rings. Recall (from Exercise 1.5.E) that $(\cdot \otimes_B A, \cdot_B)$ is an adjoint pair between the categories of A-modules and B-modules: we have a bijection

$$\operatorname{Hom}_A(N \otimes_B A, M) \cong \operatorname{Hom}_B(N, M_B)$$

functorial in both arguments. These constructions behave well with respect to localization (in an appropriate sense), and hence work (often) in the category of quasicoherent sheaves on schemes (and indeed always in the category of $\mathscr{O}$-modules on ringed spaces, see Definition 16.3.5, although we won't particularly care). The easier construction ($M \mapsto M_B$) will turn into our old friend pushforward. The other ($N \mapsto A \otimes_B N$) will be a relative of pullback, whom I'm reluctant to call an "old friend".

16.2 Pushforwards of quasicoherent sheaves

We begin with the pushforwards, for which we have already done much of the work.

The main moral of this section is that in "reasonable" situations, the pushforward of a quasicoherent sheaf is quasicoherent, and that this can be understood in terms of one of the module constructions defined above. We begin with a motivating example:

16.2.A. EXERCISE. Let $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ be a morphism of affine schemes, and suppose M is an A-module, so $\widetilde{M}$ is a quasicoherent sheaf on $\operatorname{Spec} A$. Give an isomorphism $\pi_* \widetilde{M} \to \widetilde{M_B}$. (Hint: There is only one reasonable way to proceed: look at distinguished open sets.)

In particular, $\pi_* \widetilde{M}$ is quasicoherent. Perhaps more important, this implies that the pushforward of a quasicoherent sheaf under an affine morphism is also quasicoherent.

16.2.B. EXERCISE. If $\pi : X \to Y$ is an affine morphism, show that π_* is an exact functor $QCoh_X \to QCoh_Y$.

The following result, proved earlier, generalizes the fact that the pushforward of a quasicoherent sheaf under an affine morphism is also quasicoherent.

16.2.1. Theorem (Exercise 13.3.F). — *Suppose $\pi : X \to Y$ is a quasicompact quasiseparated morphism, and $\mathscr{F}$ is a quasicoherent sheaf on X. Then $\pi_*\mathscr{F}$ is a quasicoherent sheaf on Y.*

Coherent sheaves do not always push forward to coherent sheaves. For example, consider the structure morphism $\pi : \mathbb{A}^1_k \to \operatorname{Spec} k$, corresponding to $k \to k[t]$. Then $\pi_*\mathscr{O}_{\mathbb{A}^1_k}$ is the quasicoherent sheaf corresponding to $k[t]$, which is not a finitely generated k-module. But in good situations, coherent sheaves do push forward. For example:

16.2.C. EXERCISE. Suppose $\pi : X \to Y$ is a finite morphism of Noetherian schemes. If $\mathscr{F}$ is a coherent sheaf on X, show that $\pi_*\mathscr{F}$ is a coherent sheaf. Hint: Show first that $\pi_*\mathscr{O}_X$ is finite type. (Noetherian hypotheses are stronger than necessary, see Remark 18.1.7, but this suffices for most purposes.)

Once we define cohomology of quasicoherent sheaves, we will quickly prove that if $\mathscr{F}$ is a coherent sheaf on $\mathbb{P}^n_k$, then $\Gamma(\mathbb{P}^n_k, \mathscr{F})$ is a finite-dimensional k-module, and more generally if $\mathscr{F}$ is a coherent sheaf on $\operatorname{Proj} S_\bullet$, then $\Gamma(\operatorname{Proj} S_\bullet, \mathscr{F})$ is a coherent A-module (where $S_0 = A$). This is a special case of the fact that the "pushforwards of coherent sheaves by projective morphisms are also coherent sheaves". (The notion of projective morphism, a relative version of $\operatorname{Proj} S_\bullet \to \operatorname{Spec} A$, will be defined in §17.3.)

More generally, given Noetherian hypotheses, pushforwards of coherent sheaves by proper morphisms are also coherent sheaves (Theorem 18.9.1).

16.3 Pullbacks of quasicoherent sheaves

We next discuss the pullback of a quasicoherent sheaf: if $\pi : X \to Y$ is a morphism of schemes, π^* is a covariant functor $QCoh_Y \to QCoh_X$. The notion of the pullback of a quasicoherent sheaf can be confusing on first (and second) glance. (For example, it is *not* the inverse image sheaf, although we will see that it is related.)

Here are three contexts in which you have seen the pullback, or can understand it quickly. It may be helpful to keep these in mind, to keep you anchored in the long discussion that follows. Suppose $\mathscr{G}$ is a quasicoherent sheaf on a scheme Y.

(i) *(restriction to open subsets)* If $i : U \hookrightarrow Y$ is an open immersion, then $i^*\mathscr{G}$ is $\mathscr{G}|_U$, the restriction of $\mathscr{G}$ to U (Example 2.2.8).

(ii) *(restriction to points)* If $i : p \hookrightarrow Y$ is the "inclusion" of a point p in Y (for example, the closed embedding of a closed point; see Exercise 6.3.J(b)), then $i^*\mathscr{G}$ is $\mathscr{G}|_p$, the fiber of $\mathscr{G}$ at p (Definition 4.3.7).

The similarity of the notation $\mathscr{G}|_U$ and $\mathscr{G}|_p$ is precisely because both are pullbacks. Pullbacks (especially to locally closed subschemes or generic points) are often called *restriction*. For this reason, if $\pi : X \to Y$ is some sort of inclusion (such as a locally closed embedding, or an "inclusion of a generic point") then $\pi^*\mathscr{G}$ is often written as $\mathscr{G}|_X$ and called the **restriction of $\mathscr{G}$ to** X, when π can be interpreted as some type of "inclusion".

(iii) *(pulling back vector bundles)* Suppose $\mathscr{G}$ is a locally free sheaf on Y, $\pi : X \to Y$ is any morphism. If $\{U_i\}$ are trivializing neighborhoods for $\mathscr{G}$, and $T_{ij} \in GL_r(\mathscr{O}_X(U_i \cap U_j))$ are transition matrices for $\mathscr{G}$ between U_i and U_j, then $\{\pi^{-1}U_i\}$ are trivializing neighborhoods for $\pi^*\mathscr{G}$, and π^*T_{ij} are transition matrices for $\pi^*\mathscr{G}$. (This will be established in Theorem 16.3.7(3).)

16.3.1. *Strategy.* We will see three different ways of thinking about the pullback. Each has significant disadvantages, but together they give a good understanding.

(a) Because we are understanding quasicoherent sheaves in terms of affine open sets, and modules over the corresponding rings, we begin with an interpretation in this vein. This will be very useful for proving and understanding facts. The disadvantage is that it is annoying to make a definition out of this (when the target is not affine), because gluing arguments can be tedious.

(b) As we saw with fibered product, gluing arguments can be made simpler using universal properties, so our second "definition" will be by universal property. This is elegant, but has the disadvantage that it still needs a construction, and because it works in the larger category of $\mathscr{O}$-modules, it isn't clear from the universal property that it takes quasicoherent sheaves to quasicoherent sheaves. But if the target is affine, our construction of (a) is easily seen to satisfy universal property. Furthermore, the universal property is "local on the target": if $\pi : X \to Y$ is any morphism, $i : U \hookrightarrow Y$ is an open immersion, and $\mathscr{G}$ is a quasicoherent sheaf on Y, then if $\pi^*\mathscr{G}$ exists, then its restriction to $\pi^{-1}(U)$ is (canonically identified with) $(\pi|_U)^*(\mathscr{G}|_U)$. Thus if the pullback exists in general (even as an $\mathscr{O}$-module), affine-locally on Y it looks like the construction of (a) (and thus is quasicoherent).

(c) The third definition is one that works on ringed spaces in general. It is short, and is easily seen to satisfy the universal property. It doesn't obviously take quasicoherent sheaves to quasicoherent sheaves (at least in the way that we have defined quasicoherent sheaves) — a priori it takes quasicoherent sheaves to $\mathscr{O}$-modules. But thanks to the discussion at the end of (b) above, which used (a), this shows that the pullback of a quasicoherent sheaf is indeed quasicoherent.

16.3.2. First attempt at describing the pullback, using affines. Suppose $\pi : X \to Y$ is a morphism of schemes, and $\mathscr{G}$ is a quasicoherent sheaf on Y. We want to define the pullback quasicoherent sheaf $\pi^*\mathscr{G}$ on X in terms of affine open sets on X and Y. Suppose $\operatorname{Spec} A \subset X$, $\operatorname{Spec} B \subset Y$ are affine open sets, with $\pi(\operatorname{Spec} A) \subset \operatorname{Spec} B$. Suppose $\mathscr{G}|_{\operatorname{Spec} B} \cong \widetilde{N}$. Perhaps motivated by the fact that pullback should relate to tensor product, we want

$$\Gamma(\operatorname{Spec} A, \pi^*\mathscr{G}) = N \otimes_B A. \tag{16.3.2.1}$$

More precisely, we would like $\Gamma(\operatorname{Spec} A, \pi^*\mathscr{G})$ and $N \otimes_B A$ to be identified. This could mean that we use this to construct a definition of $\pi^*\mathscr{G}$, by "gluing all this information together" (and showing it is well-defined). Or it could mean that we define $\pi^*\mathscr{G}$ in some other way, and then find a natural identification (16.3.2.1). The first approach can be made to work (and §16.3.3 is the first step), but we will follow the second.

[Figure to be made: X maps to Y; inside X is Spec A, inside Y is Spec B. Over Spec A is $\widetilde{N \otimes_B A}$, and inside Spec B is $\widetilde{N}$]

FIGURE 16.1. The pullback of a quasicoherent sheaf

16.3.3. We begin this project by *fixing* an affine open subset $\operatorname{Spec} B \subset Y$. To avoid confusion, let $\phi = \pi|_{\pi^{-1}(\operatorname{Spec} B)} : \pi^{-1}(\operatorname{Spec} B) \to \operatorname{Spec} B$. We will define a quasicoherent sheaf on $\pi^{-1}(\operatorname{Spec} B)$ that will turn out to be $\phi^*(\mathscr{G}|_{\operatorname{Spec} B})$ (and will also be the restriction of $\pi^*\mathscr{G}$ to $\pi^{-1}(\operatorname{Spec} B)$).

If $\operatorname{Spec} A_f \subset \operatorname{Spec} A$ is a distinguished open set, then

$$\Gamma(\operatorname{Spec} A_f, \phi^*\mathscr{G}) = N \otimes_B A_f = (N \otimes_B A)_f = \Gamma(\operatorname{Spec} A, \phi^*\mathscr{G})_f$$

where "=" means "canonically isomorphic". Define the restriction map $\Gamma(\operatorname{Spec} A, \phi^*\mathscr{G}) \to \Gamma(\operatorname{Spec} A_f, \phi^*\mathscr{G})$,

$$\Gamma(\operatorname{Spec} A, \phi^*\mathscr{G},) \to \Gamma(\operatorname{Spec} A, \phi^*\mathscr{G}) \otimes_A A_f, \tag{16.3.3.1}$$

by $\alpha \mapsto \alpha \otimes 1$ (of course). Thus $\phi^*\mathscr{G}$ is (or: extends to) a quasicoherent sheaf on $\pi^{-1}(\operatorname{Spec} B)$ (by Exercise 13.3.D).

We have now defined a quasicoherent sheaf on $\pi^{-1}(\operatorname{Spec} B)$, for every affine open subset $\operatorname{Spec} B \subset Y$. We want to show that this construction, as $\operatorname{Spec} B$ varies, glues into a single quasicoherent sheaf on X.

You are welcome to do this gluing appropriately, for example using the distinguished affine base of Y (§13.3.1). This works, but can be confusing, so we take another approach.

16.3.4. Universal property definition of pullback. Suppose $\pi : X \to Y$ is a morphism of ringed spaces, and $\mathscr{G}$ is an $\mathscr{O}_Y$-module. (We are of course interested in the case where π is a morphism of schemes, and $\mathscr{G}$ is quasicoherent. Even once we specialize our discussion to schemes, much of our discussion will extend without change to this more general situation.) We "define" the pullback $\pi^*\mathscr{G}$ as an $\mathscr{O}_X$-module, using the following adjointness universal property: for any $\mathscr{O}_X$-module $\mathscr{F}$, there is a bijection $\operatorname{Hom}_{\mathscr{O}_X}(\pi^*\mathscr{G}, \mathscr{F}) \leftrightarrow \operatorname{Hom}_{\mathscr{O}_Y}(\mathscr{G}, \pi_*\mathscr{F})$, and these bijections are functorial in $\mathscr{F}$. By universal property nonsense, this determines $\pi^*\mathscr{G}$ up to unique isomorphism; we just need to make sure that it exists (which is why the word "define" is in quotes). Notice that we avoid worrying about when the pushforward of a quasicoherent sheaf $\mathscr{F}$ is quasicoherent by working in the larger category of $\mathscr{O}$-modules.

16.3.A. IMPORTANT EXERCISE. If Y is affine, say $Y = \operatorname{Spec} B$, show that the construction of the quasicoherent sheaf in §16.3.3 satisfies this universal property of pullback of $\mathscr{G}$. Thus calling this sheaf $\pi^*\mathscr{G}$ is justified. (Hint: Interpret both sides of the alleged bijection explicitly. The adjointness in the ring/module case should turn up.)

16.3.B. IMPORTANT EXERCISE. Suppose $i : U \hookrightarrow X$ is an open embedding of ringed spaces, and $\mathscr{F}$ is an $\mathscr{O}_X$-module. Show that $\mathscr{F}|_U$ satisfies the universal property of $i^*\mathscr{F}$ (and thus deserves to be called $i^*\mathscr{F}$). In other words, for each

$\mathscr{O}_U$-module $\mathscr{E}$, describe a bijection

$$\operatorname{Hom}_{\mathscr{O}_U}(\mathscr{F}|_U, \mathscr{E}) \longleftrightarrow \operatorname{Hom}_{\mathscr{O}_X}(\mathscr{F}, i_*\mathscr{E}),$$

functorial in $\mathscr{E}$.

We next show that if $\pi^*\mathscr{G}$ satisfies the universal property (for the morphism $\pi : X \to Y$), then if $j : V \hookrightarrow Y$ is any open subset, and $i : U = \pi^{-1}(V) \hookrightarrow X$ (see (16.3.4.1)), then $(\pi^*\mathscr{G})|_U$ satisfies the universal property for $\pi|_U : U \to V$. Thus $(\pi^*\mathscr{G})|_U$ deserves to be called $\pi|_U^*(\mathscr{G}|_V)$. You will notice that we really need to work with $\mathscr{O}$-modules, not just with quasicoherent sheaves.

(16.3.4.1)
$$\begin{tikzcd} \pi^{-1}(V) \arrow[r, equal] & U \arrow[r, hook, "i"] \arrow[d, "\pi|_U"'] & X \arrow[d, "\pi"] \\ & V \arrow[r, hook, "j"] & Y \end{tikzcd}$$

If $\mathscr{F}'$ is an $\mathscr{O}_U$-module, we have a series of bijections (using Important Exercise 16.3.B and adjointness of pullback and pushforward):

$$\begin{aligned} \operatorname{Hom}_{\mathscr{O}_U}((\pi^*\mathscr{G})|_U, \mathscr{F}') &\cong \operatorname{Hom}_{\mathscr{O}_U}(i^*(\pi^*\mathscr{G}), \mathscr{F}') \\ &\cong \operatorname{Hom}_{\mathscr{O}_X}(\pi^*\mathscr{G}, i_*\mathscr{F}') \\ &\cong \operatorname{Hom}_{\mathscr{O}_Y}(\mathscr{G}, \pi_* i_*\mathscr{F}') \\ &\cong \operatorname{Hom}_{\mathscr{O}_Y}(\mathscr{G}, j_*(\pi|_U)_*\mathscr{F}') \\ &\cong \operatorname{Hom}_{\mathscr{O}_V}(j^*\mathscr{G}, (\pi|_U)_*\mathscr{F}') \\ &\cong \operatorname{Hom}_{\mathscr{O}_V}(\mathscr{G}|_V, (\pi|_U)_*\mathscr{F}'). \end{aligned}$$

We have thus described a bijection

$$\operatorname{Hom}_{\mathscr{O}_U}((\pi^*\mathscr{G})|_U, \mathscr{F}') \leftrightarrow \operatorname{Hom}_{\mathscr{O}_V}(\mathscr{G}|_V, (\pi|_U)_*\mathscr{F}'),$$

which is clearly (by construction) functorial in $\mathscr{F}'$. Hence the discussion in the first paragraph of §16.3.4 is justified. For example, thanks to Important Exercise 16.3.A, the pullback π^* exists if Y is an open subset of an affine scheme.

At this point, we could show that the pullback exists, following the idea behind the construction of the fibered product: we would start with the definition when Y is affine, and "glue". We will instead take another route.

16.3.5. Third definition: pullback of $\mathscr{O}$-modules via explicit construction. Suppose $\pi : X \to Y$ is a morphism of ringed spaces, and $\mathscr{G}$ is an $\mathscr{O}_Y$-module. Of course, our case of interest is if π is a morphism of schemes, and $\mathscr{G}$ is quasicoherent. Now $\pi^{-1}\mathscr{G}$ is a $\pi^{-1}\mathscr{O}_Y$-module. (Notice that we are using the ringed space $(X, \pi^{-1}\mathscr{O}_Y)$, not $(X, \mathscr{O}_X)$. The inverse image construction π^{-1} was discussed in §2.6.) Furthermore, $\mathscr{O}_X$ is also a $\pi^{-1}\mathscr{O}_Y$-module, via the map $\pi^{-1}\mathscr{O}_Y \to \mathscr{O}_X$ that is part of the data of the morphism π. Define the **pullback** of $\mathscr{G}$ by π as the $\mathscr{O}_X$-module

(16.3.5.1)
$$\pi^*\mathscr{G} := \pi^{-1}\mathscr{G} \otimes_{\pi^{-1}\mathscr{O}_Y} \mathscr{O}_X.$$

It is immediate that pullback is a covariant functor $\pi^* : Mod_{\mathscr{O}_Y} \to Mod_{\mathscr{O}_X}$.

16.3.C. IMPORTANT EXERCISE. Show that this definition (16.3.5.1) of pullback satisfies the universal property. Thus the pullback exists, at least as a functor $Mod_{\mathscr{O}_Y} \to Mod_{\mathscr{O}_X}$.

16.3.D. IMPORTANT EXERCISE. Show that if $\pi : X \to Y$ is a morphism of schemes, then π^* gives a covariant functor $QCoh_Y \to QCoh_X$. (You will use §16.3.3, Exercise 16.3.B, and Exercise 16.3.A.)

The following is then immediate from the universal property.

16.3.6. Proposition. — *Suppose $\pi : X \to Y$ is a quasicompact, quasiseparated morphism. Then $(\pi^* : QCoh_Y \to QCoh_X, \pi_* : QCoh_X \to QCoh_Y)$ are an adjoint pair: there is an isomorphism*

$$\operatorname{Hom}_{\mathscr{O}_X}(\pi^*\mathscr{G}, \mathscr{F}) \cong \operatorname{Hom}_{\mathscr{O}_Y}(\mathscr{G}, \pi_*\mathscr{F}), \tag{16.3.6.1}$$

functorial in both $\mathscr{F} \in QCoh_X$ and $\mathscr{G} \in QCoh_Y$.

The "quasicompact and quasiseparated" hypotheses are solely to ensure that π_* indeed sends $QCoh_X$ to $QCoh_Y$ (Theorem 13.3.F).

We are now ready to show that pullback has all sorts of desired properties.

16.3.7. Theorem. — *Suppose $\pi : X \to Y$ is a morphism of schemes, and $\mathscr{G}$ is a quasicoherent sheaf on Y.*

(1) *(pullback preserves the structure sheaf) There is a canonical isomorphism $\pi^*\mathscr{O}_Y \cong \mathscr{O}_X$.*
(2) *(pullback preserves finite type quasicoherent sheaves) If $\mathscr{G}$ is a finite type quasicoherent sheaf, so is $\pi^*\mathscr{G}$. Hence if X is locally Noetherian, and $\mathscr{G}$ is coherent, then so is $\pi^*\mathscr{G}$. (It is not always true that the pullback of a coherent sheaf is coherent, and the interested reader can think of a counterexample.)*
(3) *(pullback preserves vector bundles, and their transition functions) If $\mathscr{G}$ is locally free sheaf of rank r, then so is $\pi^*\mathscr{G}$. (In particular, the pullback of an invertible sheaf is invertible.) Furthermore, if $\{U_i\}$ are trivializing neighborhoods for $\mathscr{G}$, and $T_{ij} \in GL_r(\mathscr{O}_X(U_i \cap U_j))$ are transition matrices for $\mathscr{G}$ between U_i and U_j, then $\{\pi^{-1}U_i\}$ are trivializing neighborhoods for $\pi^*\mathscr{G}$, and π^*T_{ij} are transition matrices for $\pi^*\mathscr{G}$.*
(4) *(functoriality in the morphism) If $\xi : W \to X$ is a morphism of schemes, then there is a canonical isomorphism $\xi^*\pi^*\mathscr{G} \cong (\pi \circ \xi)^*\mathscr{G}$.*
(5) *(functoriality in the quasicoherent sheaf) π^* is a functor $QCoh_Y \to QCoh_X$.*
(6) *(pulling back a section) Note that a section of $\mathscr{G}$ is the data of a map $\mathscr{O}_Y \to \mathscr{G}$. By (1) and (5), if $s : \mathscr{O}_Y \to \mathscr{G}$ is a section of $\mathscr{G}$ then there is a natural section $\pi^*s : \mathscr{O}_X \to \pi^*\mathscr{G}$ of $\pi^*\mathscr{G}$. The pullback of the locus where s vanishes is the locus where the pulled-back section π^*s vanishes.*
(7) *(pullback on stalks) If $\pi : X \to Y$, $\pi(p) = q$, then pullback induces an isomorphism*

$$(\pi^*\mathscr{G})_p \xrightarrow{\sim} \mathscr{G}_q \otimes_{\mathscr{O}_{Y,q}} \mathscr{O}_{X,p}\ . \tag{16.3.7.1}$$

(8) *(pullback on fibers of the quasicoherent sheaves) Pullback of fibers are given as follows: if $\pi : X \to Y$, where $\pi(p) = q$, then the map*

$$(\pi^*\mathscr{G})|_p \xrightarrow{\sim} \mathscr{G}|_q \otimes_{\kappa(q)} \kappa(p)$$

induced by (16.3.7.1) is an isomorphism.
(9) *(pullback preserves tensor product) $\pi^*(\mathscr{G} \otimes_{\mathscr{O}_Y} \mathscr{G}') = \pi^*\mathscr{G} \otimes_{\mathscr{O}_X} \pi^*\mathscr{G}'$. (Here $\mathscr{G}'$ is also a quasicoherent sheaf on Y.)*
(10) *Pullback is a right-exact functor.*

All of the above are interconnected in obvious ways that you should be able to prove by hand. (As just one example: the germ of a pulled back section, (6), is the expected element of the pulled back stalk, (7).) In fact much more is true, that you should be able to prove on a moment's notice, such as for example that the pullback of the symmetric power of a locally free sheaf is naturally isomorphic to the symmetric power of the pullback, and similarly for wedge powers and tensor powers.

16.3.E. IMPORTANT EXERCISE. Prove Theorem 16.3.7. Possible hints: You may find it convenient to do right-exactness (10) early; it is related to right-exactness of $\otimes$. For the tensor product fact (9), show that $(M \otimes_B A) \otimes (N \otimes_B A) \cong (M \otimes_B N) \otimes_B A$, and that this behaves well with respect to localization. The proof of the fiber fact (8) is as follows. Given a ring map $B \to A$ with $[\mathfrak{m}] \mapsto [\mathfrak{n}]$, where $\mathfrak{m} \subset A$ and $\mathfrak{n} \subset B$ are maximal ideals, show that $(N \otimes_B A) \otimes_A (A/\mathfrak{m}) \cong (N \otimes_B (B/\mathfrak{n})) \otimes_{B/\mathfrak{n}} (A/\mathfrak{m})$ by showing both sides are isomorphic to $N \otimes_B (A/\mathfrak{m})$.

16.3.F. UNIMPORTANT EXERCISE. Verify that the following is an example showing that pullback is not left-exact: consider the exact sequence of sheaves on $\mathbb{A}^1_k$, where $i : p \hookrightarrow \mathbb{A}^1_k$ is the origin:

$$0 \to \mathscr{O}_{\mathbb{A}^1_k}(-p) \to \mathscr{O}_{\mathbb{A}^1_k} \to i_* \mathscr{O}|_p \to 0.$$

(This is the closed subscheme exact sequence for $p \in \mathbb{A}^1_k$, and corresponds to the exact sequence of $k[t]$-modules $0 \to tk[t] \to k[t] \to k \to 0$. Warning: here $\mathscr{O}|_p$ is not the stalk $\mathscr{O}_p$; it is the structure sheaf of the scheme p.) Restrict to p.

16.3.G. EXERCISE (THE PUSH-PULL FORMULA, CF. EXERCISE 18.8.B). Suppose $\psi : Z \to Y$ is any morphism, and $\pi : X \to Y$ is quasicompact and quasiseparated (so pushforwards send quasicoherent sheaves to quasicoherent sheaves). Suppose $\mathscr{F}$ is a quasicoherent sheaf on X. Suppose

(16.3.7.2)
$$\begin{tikzcd} W \arrow[r, "\psi'"] \arrow[d, "\pi'"'] & X \arrow[d, "\pi"] \\ Z \arrow[r, "\psi"] & Y \end{tikzcd}$$

is a commutative diagram. Describe a natural morphism $\psi^* \pi_* \mathscr{F} \to \pi'_* (\psi')^* \mathscr{F}$ of quasicoherent sheaves on Z. (Possible hint: first do the special case where (16.3.7.2) is a fiber diagram.)

By applying the above exercise in the special case where Z is a point q of Y, we see that there is a natural map from the fiber of the pushforward to the sections over the fiber:

(16.3.7.3)
$$\pi_* \mathscr{F} \otimes \kappa(q) \to \Gamma(\pi^{-1}(q), \mathscr{F}|_{\pi^{-1}(q)}).$$

One might hope that (16.3.7.3) is an isomorphism, i.e., that $\pi_* \mathscr{F}$ "glues together" the fibers $\Gamma(\pi^{-1}(q), \mathscr{F}|_{\pi^{-1}(q)})$. This is too much to ask, but at least (16.3.7.3) gives a map. (In fact, under just the right circumstances, (16.3.7.3) is an isomorphism, see §28.1.)

16.3.H. EXERCISE (PROJECTION FORMULA, TO BE GENERALIZED IN EXERCISE 18.8.E). Suppose $\pi : X \to Y$ is quasicompact and quasiseparated, and $\mathscr{F}$ and $\mathscr{G}$ are quasicoherent sheaves on X and Y respectively.
(a) Describe a natural morphism $(\pi_* \mathscr{F}) \otimes \mathscr{G} \to \pi_*(\mathscr{F} \otimes \pi^* \mathscr{G})$.
(b) If $\mathscr{G}$ is locally free, show that this natural morphism is an isomorphism. (Hint: what if $\mathscr{G}$ is free?)

16.3.8. *Remark: Flatness.* Given $\pi : X \to Y$, if the functor π^* from quasicoherent sheaves on Y to quasicoherent sheaves on X is exact (not just right-exact as given by Theorem 16.3.7(10)), we will say that π is a **flat morphism**. This is an incredibly important notion. We will return to it (and define it in a different-looking but equivalent way) in Chapter 24; see Exercise 24.2.L.

16.3.9. *Remark: Pulling back ideal sheaves.* There is one subtlety in pulling back quasicoherent *ideal* sheaves. Suppose $i : X \hookrightarrow Y$ is a closed embedding, and $\mu : Y' \to Y$ is an arbitrary morphism. Let $X' := X \times_Y Y'$. As "closed embeddings pull back" (§9.2.1), the pulled back map $i' : X' \to Y'$ is a closed embedding.

$$\begin{tikzcd} X' \arrow[r] \arrow[d, hook, "i'"'] & X \arrow[d, hook, "i"] \\ Y' \arrow[r, "\mu"] & Y \end{tikzcd}$$

Now μ^* induces canonical isomorphisms $\mu^* \mathscr{O}_Y \cong \mathscr{O}_{Y'}$ and $\mu^*(i_* \mathscr{O}_X) \cong (i'_* \mathscr{O}_{X'})$, but it is *not* necessarily true that $\mu^* \mathscr{I}_{X/Y} = \mathscr{I}_{X'/Y'}$. (Exercise 16.3.F yields an example.) This is because the application of μ^* to the closed subscheme exact sequence

$$0 \to \mathscr{I}_{X/Y} \to \mathscr{O}_Y \to i_* \mathscr{O}_X \to 0$$

yields something that is a priori only right-exact:

$$\mu^* \mathscr{I}_{X/Y} \to \mathscr{O}_{Y'} \to i'_* \mathscr{O}_{X'} \to 0.$$

Thus, as $\mathscr{I}_{X'/Y'}$ is the kernel of $\mathscr{O}_{Y'} \to i'_* \mathscr{O}_{X'}$, we see that $\mathscr{I}_{X'/Y'}$ is the image of $\mu^* \mathscr{I}_{X/Y}$ in $\mathscr{O}_{Y'}$. We can also see this explicitly from Exercise 9.2.B: affine-locally, the ideal of the pullback is generated by the pullback of the ideal.

Note also that if μ is flat (Remark 16.3.8), then $\mu^* \mathscr{I}_{X/Y} \to \mathscr{I}_{X'/Y'}$ *is* an isomorphism.

16.4 Line bundles and maps to projective schemes

Theorem 16.4.1, the converse or completion to Exercise 15.3.F, will give one reason why line bundles are crucially important: they tell us about maps to projective space, and more generally, to quasiprojective A-schemes. Given that we have had a hard time naming any non-quasiprojective schemes, they tell us about maps to essentially all schemes that are interesting to us.

16.4.1. Important Theorem. — *For a fixed scheme X, morphisms $X \to \mathbb{P}^n$ are in bijection with the data $(\mathscr{L}, s_0, \dots, s_n)$, where $\mathscr{L}$ is an invertible sheaf and $s_0, \dots, s_n$ are sections of $\mathscr{L}$ with no common zeros, up to isomorphism of these data.*

(This works over $\mathbb{Z}$, or indeed any base.) Informally: morphisms to $\mathbb{P}^n$ correspond to $n+1$ sections of a line bundle, not all vanishing at any point, modulo global sections of $\mathscr{O}_X^*$, as multiplication by an invertible function gives an automorphism of $\mathscr{L}$. This is one of those important theorems in algebraic geometry that is easy to prove, but quite subtle in its effect on how one should think. It takes some time to properly digest. A "coordinate-free" version is given in Exercise 16.4.J.

16.4.2. Theorem 16.4.1 describes all morphisms to projective space, and hence by the Yoneda philosophy, this can be taken as the *definition* of projective space: it defines projective space up to unique isomorphism. *Projective space $\mathbb{P}^n$ (over $\mathbb{Z}$) is the moduli space of line bundles $\mathscr{L}$ along with $n+1$ sections of $\mathscr{L}$ with no common zeros.* (Can you give an analogous definition of projective space over X, denoted $\mathbb{P}^n_X$?)

Every time you see a map to projective space, you should immediately simultaneously keep in mind the invertible sheaf and sections.

Maps to projective schemes can be described similarly. For example, if $Y \hookrightarrow \mathbb{P}^2_k$ is the curve $x_2^2 x_0 = x_1^3 - x_1 x_0^2$, then maps from a scheme X to Y are given by an invertible sheaf on X along with three sections s_0, s_1, s_2, with no common zeros, satisfying $s_2^2 s_0 - s_1^3 + s_1 s_0^2 = 0$.

Here more precisely is the correspondence of Theorem 16.4.1. Any $n+1$ sections of $\mathscr{L}$ with no common zeros determine a morphism to $\mathbb{P}^n$, by Exercise 15.3.F. Conversely, if you have a map to projective space $\pi : X \to \mathbb{P}^n$, then we have $n+1$ sections of $\mathscr{O}_{\mathbb{P}^n}(1)$, corresponding to the hyperplane sections, $x_0, \ldots, x_{n+1}$. Then $\pi^* x_0, \ldots, \pi^* x_{n+1}$ are sections of $\pi^* \mathscr{O}_{\mathbb{P}^n}(1)$, and they have no common zero. (Reminder: it is helpful to think of pulling back invertible sheaves in terms of pulling back transition functions, see Theorem 16.3.7(3).)

So to prove Theorem 16.4.1, we just need to show that these two constructions compose to give the identity in either direction.

Proof of Important Theorem 16.4.1. Suppose we are given $n+1$ sections $s_0, \ldots, s_n$ of an invertible sheaf $\mathscr{L}$, with no common zeros, which (via Exercise 15.3.F) induce a morphism $\pi : X \to \mathbb{P}^n$. For each s_i, we get a trivialization on $\mathscr{L}$ on the open set X_{s_i} where s_i doesn't vanish. (More precisely, we have an isomorphism $(\mathscr{L}, s_i) \cong (\mathscr{O}, 1)$, cf. Important Exercise 14.2.E(a).) The transition functions for $\mathscr{L}$ are precisely s_i/s_j on $X_{s_i} \cap X_{s_j}$. As $\mathscr{O}(1)$ is trivial on the standard affine open sets $D(x_i)$ of $\mathbb{P}^n$, $\pi^*(\mathscr{O}(1))$ is trivial on $X_{s_i} = \pi^{-1}(D(x_i))$. Moreover, $s_i/s_j = \pi^*(x_i/x_j)$ (directly from the construction of π in Exercise 15.3.F). This gives an isomorphism $\mathscr{L} \cong \pi^* \mathscr{O}(1)$ — the two invertible sheaves have the same transition functions.

16.4.A. EXERCISE. Show that this isomorphism can be chosen so that $(\mathscr{L}, s_0, \ldots, s_n) \cong (\pi^* \mathscr{O}(1), \pi^* x_0, \ldots, \pi^* x_n)$, thereby completing one of the two implications of the theorem.

For the other direction, suppose we are given a map $\pi : X \to \mathbb{P}^n$. Let $s_i = \pi^* x_i \in \Gamma(X, \pi^*(\mathscr{O}(1)))$. As the x_i's have no common zeros on $\mathbb{P}^n$, the s_i's have no common zeros on X. The map $[s_0, \ldots, s_n]$ is the same as the map π. We see this as follows. The preimage of $D(x_i)$ (by the morphism $[s_0, \ldots, s_n]$) is $D(s_i) = D(\pi^* x_i) = \pi^{-1} D(x_i)$, so "the right open sets go to the right open sets". To show the two morphisms $D(s_i) \to D(x_i)$ (induced from $(s_1, \ldots, s_n)$ and π) are the same, we use the fact that maps to an affine scheme $D(x_i)$ are determined by their maps of global sections in the opposite direction (Essential Exercise 6.3.F). Both morphisms $D(s_i) \to D(x_i)$ corresponds to the ring map $\pi^\sharp : x_{j/i} = x_j/x_i \mapsto s_j/s_i$. $\square$

16.4.3. *Remark: Extending Theorem 16.4.1 to rational maps.* Suppose $s_0, \ldots, s_n$ are sections of an invertible sheaf $\mathscr{L}$ on a scheme X. Then Theorem 16.4.1 yields a morphism $X - V(s_0, \ldots, s_n) \to \mathbb{P}^n$. In particular, if X is integral, and the s_i are not all 0, these data yield a rational map $X \dashrightarrow \mathbb{P}^n$.

16.4.4. Examples and applications.

16.4.B. EXERCISE (AUTOMORPHISMS OF PROJECTIVE SPACE). Show that all the automorphisms of projective space $\mathbb{P}^n_k$ (fixing k) correspond to $(n+1) \times (n+1)$ invertible matrices over k, modulo scalars (also known as $PGL_{n+1}(k)$). (Hint: Suppose $\pi : \mathbb{P}^n_k \to \mathbb{P}^n_k$ is an automorphism. Show that $\pi^*\mathscr{O}(1) \cong \mathscr{O}(1)$. Show that $\pi^* : \Gamma(\mathbb{P}^n, \mathscr{O}(1)) \to \Gamma(\mathbb{P}^n, \mathscr{O}(1))$ is an isomorphism.)

Exercise 16.4.B will be useful later, especially for the case $n = 1$. In this case, these automorphisms are called *fractional linear transformations*. (For experts: why was Exercise 16.4.B not stated over an arbitrary base ring A? Where does the argument go wrong in that case?)

16.4.C. EXERCISE. Show that $\mathrm{Aut}(\mathbb{P}^1_k)$ is strictly three-transitive on k-valued points, i.e., given two triplets (p_1, p_2, p_3) and (q_1, q_2, q_3) each of distinct k-valued points of $\mathbb{P}^1$, there is precisely one automorphism of $\mathbb{P}^1$ sending p_i to q_i ($i = 1, 2, 3$).

Here are more examples of these ideas in action.

16.4.5. *Example: the Veronese embedding is $|\mathscr{O}_{\mathbb{P}^n_k}(d)|$.* Consider the line bundle $\mathscr{O}_{\mathbb{P}^n_k}(d)$ on $\mathbb{P}^n_k$. We have checked that the sections of this line bundle form a vector space of dimension $\binom{n+d}{d}$, with a basis corresponding to homogeneous degree d polynomials in the projective coordinates for $\mathbb{P}^n_k$. Also, they have no common zeros (as for example the subset of sections $x_0^d, x_1^d, \ldots, x_n^d$ have no common zeros). Thus the complete linear series is base-point-free, and determines a morphism $\mathbb{P}^n \to \mathbb{P}^{\binom{n+d}{d}-1}$. This is the Veronese embedding (Definition 8.2.8). For example, if $n = 2$ and $d = 2$, we get a map $\mathbb{P}^2_k \to \mathbb{P}^5_k$.

In §8.2.8, we saw that this is a closed embedding. The following is a more general method of checking that maps to projective space are closed embeddings.

16.4.D. LESS IMPORTANT EXERCISE. Suppose $\pi : X \to \mathbb{P}^n_A$ corresponds to an invertible sheaf $\mathscr{L}$ on X, and sections $s_0, \ldots, s_n$. Show that π is a closed embedding if and only if

(i) each open set X_{s_i} is affine, and
(ii) for each i, the map of rings $A[y_0, \ldots, y_n] \to \Gamma(X_{s_i}, \mathscr{O})$ given by $y_j \mapsto s_j/s_i$ is surjective.

16.4.6. *Special case of Example 16.4.5.* Recall that the image of the Veronese embedding when $n = 1$ is called a *rational normal curve of degree* d (Exercise 8.2.J). Our map is $\mathbb{P}^1_k \to \mathbb{P}^d_k$ given by $[x, y] \mapsto [x^d, x^{d-1}y, \ldots, xy^{d-1}, y^d]$.

16.4.E. EXERCISE. (For this exercise, we work over a field k.) If the scheme-theoretic image of X in $\mathbb{P}^n$ lies in a hyperplane, we say that the linear series (or X itself) is **degenerate** (and otherwise, **nondegenerate**). Show that a base-point-free linear series V with invertible sheaf $\mathscr{L}$ is nondegenerate if and only if the map $V \to \Gamma(X, \mathscr{L})$ is an inclusion. Hence a complete linear series is always nondegenerate.

16.4.F. EXERCISE. Suppose we are given a map $\pi : \mathbb{P}^1_k \to \mathbb{P}^n_k$ where the corresponding invertible sheaf on $\mathbb{P}^1_k$ is $\mathscr{O}(d)$. (This can reasonably be called a *degree* d *map*, cf. Exercises 17.4.F and 18.6.I.) Show that if $d < n$, then the image is degenerate. Show that if $d = n$ and the image is nondegenerate, then the image is isomorphic (via an automorphism of projective space, Exercise 16.4.B) to a rational normal curve.

16.4.G. EXERCISE. Define the graded rings $R_\bullet = k[u, v, w]/(uw - v^2)$ and $S_\bullet = k[x, y]$ (with all variables having degree 1). By Exercise 8.2.I, we have an isomorphism $\operatorname{Proj} R_\bullet \cong \operatorname{Proj} S_\bullet$ (via the Veronese embedding ν_2). Show that this isomorphism is not induced by a map of graded rings $S_\bullet \to R_\bullet$.

16.4.7. *Remark.* You may be able to show that after "regrading", the isomorphism $\operatorname{Proj} R_\bullet \cong \operatorname{Proj} S_\bullet$ *does* arise from a map of graded rings ($S_{2\bullet} \to R_\bullet$). Exercise 19.11.B gives an example where it is not possible to patch the lack of maps of graded rings by just regrading.

16.4.H. EXERCISE: AN EARLY LOOK AT INTERSECTION THEORY, RELATED TO BÉZOUT'S THEOREM. A classical definition of the degree of a curve in projective space is as follows: intersect it with a "general" hyperplane, and count the number of points of intersection, with appropriate multiplicity. We interpret this in the case of $\pi : \mathbb{P}^1_k \to \mathbb{P}^n_k$. Show that there is a hyperplane H of $\mathbb{P}^n_k$ not containing $\pi(\mathbb{P}^1_k)$. Equivalently, $\pi^*H \in \Gamma(\mathbb{P}^1, \mathscr{O}_{\mathbb{P}^1}(d))$ is not 0. Show that the number of zeros of π^*H is precisely d. (You will have to define "appropriate multiplicity". What does it mean geometrically if π is a closed embedding, and π^*H has a double zero? Aside: Can you make sense of this even if π is not a closed embedding?) Thus this classical notion of degree agrees with the notion of degree in Exercise 16.4.F. (See Exercise 8.2.E for another case of Bézout's theorem. Here we intersect a degree d curve with a degree 1 hyperplane; there we intersect a degree 1 curve with a degree d hypersurface. Exercise 18.6.K will give a common generalization.)

16.4.8. *Example: The Segre embedding revisited.* The Segre embedding can also be interpreted in this way. This is a useful excuse to define some notation. Suppose $\mathscr{F}$ is a quasicoherent sheaf on a Z-scheme X, and $\mathscr{G}$ is a quasicoherent sheaf on a Z-scheme Y. Let π_X, π_Y be the projections from $X \times_Z Y$ to X and Y respectively. Then $\mathscr{F} \boxtimes \mathscr{G}$ (pronounced "$\mathscr{F}$ box-times $\mathscr{G}$") is defined to be $\pi_X^*\mathscr{F} \otimes \pi_Y^*\mathscr{G}$. In particular, $\mathscr{O}_{\mathbb{P}^m \times \mathbb{P}^n}(a, b)$ is defined to be $\mathscr{O}_{\mathbb{P}^m}(a) \boxtimes \mathscr{O}_{\mathbb{P}^n}(b)$ (over any base Z). The Segre embedding $\mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^{mn+m+n}$ corresponds to the complete linear series for the invertible sheaf $\mathscr{O}(1, 1)$.

When we first saw the Segre embedding in §9.6, we saw (in different language) that this complete linear series is base-point-free. We also checked by hand (§9.6.1) that it is a closed embedding, essentially by Exercise 16.4.D.

Recall that if $\mathscr{L}$ and $\mathscr{M}$ are both base-point-free invertible sheaves on a scheme X, then $\mathscr{L} \otimes \mathscr{M}$ is also base-point-free (Exercise 15.3.B, see also Definition 15.3.4).

We may interpret this fact using the Segre embedding (under reasonable hypotheses on X). If $\phi_{\mathscr{L}} : X \to \mathbb{P}^M$ is a morphism corresponding to a (base-point-free) linear series based on $\mathscr{L}$, and $\phi_{\mathscr{M}} : X \to \mathbb{P}^N$ is a morphism corresponding to a linear series on $\mathscr{M}$, then the Segre embedding yields a morphism $X \to \mathbb{P}^M \times \mathbb{P}^N \to \mathbb{P}^{(M+1)(N+1)-1}$, which corresponds to a base-point-free series of sections of $\mathscr{L} \otimes \mathscr{M}$.

16.4.I. FUN EXERCISE. Suppose X is a quasiprojective k-scheme, and $\pi : \mathbb{P}^n_k \to X$ is any morphism (over k). Show that either the image of π has dimension n, or π contracts $\mathbb{P}^n_k$ to a point. In particular, there are no nonconstant maps from projective space to a smaller-dimensional variety. Hint: if $X \subset \mathbb{P}^N$, define d by $\pi^* \mathscr{O}_{\mathbb{P}^N}(1) \cong \mathscr{O}_{\mathbb{P}^n_k}(d)$. Try to show that $d = 0$. To do that, show that if $m < n$ then m nonempty hypersurfaces in $\mathbb{P}^n$ have nonempty intersection. For this, use the fact that any nonempty hypersurface in $\mathbb{P}^n_k$ has nonempty intersection with any subscheme of dimension at least 1 (Exercise 11.3.C(a)).

16.4.J. EXERCISE. Show that a (finite-dimensional) base-point-free linear series V on X corresponding to $\mathscr{L}$ induces a morphism to projective space

$$\phi_V : X \longrightarrow \mathbb{P}V^\vee .$$

(This should be seen as a coordinate-free version of Theorem 16.4.1.)

16.4.K. EXERCISE. Explain why GL_n acts (nontrivially!) on $\mathbb{P}^{n-1}$ (over $\mathbb{Z}$, or over a field of your choice). (The group scheme GL_n was defined in Exercise 6.6.N. The *action* of a group scheme appeared earlier in Exercise 6.6.S(a).) Hint: this is much more easily done with the language of functors, §6.6, using our functorial description of projective space (§16.4.2), than with our old description of projective space in terms of patches. (A generalization to the Grassmannian will be given in Exercise 16.7.L.)

16.4.9. *Remark.* Over an algebraically closed field $\overline{k}$, GL_n acts transitively on the closed points of $\mathbb{P}^n_{\overline{k}}$, and the stabilizer of the point $[1, 0, \cdots, 0]$ consists of the subgroup P of matrices with 0's in the first column below the first row. (Side Remark: the point is better written as a column vector, so the GL_n-action can be interpreted as matrix multiplication in the usual way.) This suggests that $\mathbb{P}^{n-1}_{\overline{k}}$ is the quotient GL_n/P. This is largely true; but we first would have to make sense of the notion of group quotient.

16.4.10. $\star\star$ **A proper nonprojective k-scheme — and gluing schemes along closed subschemes.**

We conclude by using what we have developed to describe an example of a scheme that is proper but not projective (promised in Remark 10.3.6). We use a construction that looks so fundamental that you may be surprised to find that we won't use it in any meaningful way later.

Fix an algebraically closed field k. For $i = 1, 2$, let $X_i \cong \mathbb{P}^3_k$, Z_i be a line in X_i, and Z_i' be a regular conic in X_i disjoint from X_i (both Z_i and Z_i' isomorphic to $\mathbb{P}^1_k$). The construction of §16.4.11 will allow us to glue X_1 to X_2 so that Z_1 is identified with Z_2' and Z_1' is identified with Z_2, see Figure 16.2. (You will be able to make this precise after reading §16.4.11.) The result, call it X, is proper, by Exercise 16.4.O.

[To be made]

FIGURE 16.2. Building a proper nonprojective variety

Then X is not projective. For if it were, then it would be embedded in projective space by some invertible sheaf $\mathscr{L}$. If X is embedded, then X_1 is too, so $\mathscr{L}$ must restrict to an invertible sheaf on X_1 of the form $\mathscr{O}_{X_1}(n_1)$, where $n_1 > 0$. You can check that the restriction of $\mathscr{L}$ to Z_1 is $\mathscr{O}_{Z_1}(n_1)$, and the restriction of $\mathscr{L}$ to Z_1' is $\mathscr{O}_{Z_1'}(2n_1)$. Symmetrically, the restriction of $\mathscr{L}$ to Z_2 is $\mathscr{O}_{Z_2}(n_2)$ for some $n_2 > 0$, and the restriction of $\mathscr{L}$ to Z_2' is $\mathscr{O}_{Z_2'}(2n_2)$. But after gluing, $Z_1 = Z_2'$, and $Z_1' = Z_2$, so we have $n_1 = 2n_2$ and $2n_1 = n_2$, which is impossible.

16.4.11. *Gluing two schemes together along isomorphic closed subschemes.*
It is straightforward to show that you can glue two schemes along isomorphic open subschemes. (More precisely, if X_1 and X_2 are schemes, with open subschemes U_1 and U_2 respectively, and an isomorphism $U_1 \cong U_2$, you can make sense of gluing X_1 and X_2 along $U_1 \cong U_2$. You should think this through.) You can similarly glue two schemes along isomorphic *closed* subschemes. We now make this precise. Suppose $Z_1 \hookrightarrow X_1$ and $Z_2 \hookrightarrow X_2$ are closed embeddings, and $\phi : Z_1 \xrightarrow{\sim} Z_2$ is an isomorphism. We will explain how to glue X_1 to X_2 along ϕ. The result will be called $X_1 \coprod_\phi X_2$.

16.4.12. *Motivating example.* Our motivating example is if $X_i = \operatorname{Spec} A_i$ and $Z_i = \operatorname{Spec} A_i/I_i$, and ϕ corresponds to $\phi^\sharp : A_2/I_2 \xrightarrow{\sim} A_1/I_1$. Then the result will be $\operatorname{Spec} R$, where R is the ring of consisting of ordered pairs $(a_1, a_2) \in A_1 \times A_2$ that "agree via ϕ". More precisely, this is a fibered product of rings:

$$R := A_1 \times_{\phi^\sharp : A_1/I_1 \to A_2/I_2} A_2.$$

16.4.13. *The general construction, as a locally ringed space.* In our general situation, we might wish to cover X_1 and X_2 by open charts of this form. We would then have to worry about gluing and choices, so to avoid this, we instead first construct $X_1 \coprod_\phi X_2$ as a locally ringed space. As a topological space, the definition is clear: we glue the underlying sets together along the underlying sets of $Z_1 \cong Z_2$, and topologize it so that a subset of $X_1 \coprod_\phi X_2$ is open if and only if its restrictions to X_1 and X_2 are both open. For convenience, let Z be the image of Z_1 (or equivalently Z_2) in $X_1 \coprod_\phi X_2$. We next define the stalk of the structure sheaf at any point $p \in X_1 \coprod_\phi X_2$. If $p \in X_i \setminus Z = (X_1 \coprod_\phi X_2) \setminus X_{3-i}$ (hopefully the meaning of this is clear), we define the stalk as $\mathscr{O}_{X_i,p}$. If $p \in X_1 \cap X_2$, we define the stalk to consist of elements (s_1, s_2) $\mathscr{O}_{X_1,p} \times \mathscr{O}_{X_2,p}$ such that agree in $\mathscr{O}_{Z_1,p} \cong \mathscr{O}_{Z_2,p}$. The meaning of everything in this paragraph will be clear to you if you can do the following.

16.4.L. EXERCISE. Define the structure sheaf of $\mathscr{O}_{X_1 \coprod_\phi X_2}$ in terms of compatible germs. (What should it mean for germs to be compatible? Hint: for $z \in Z$, suppose we have open subsets U_1 of X_1 and U_2 of X_2, with $U_1 \cap Z = U_2 \cap Z$, so U_1 and U_2 glue together to give an open subset U of $X_1 \coprod_\phi X_2$. Suppose we also have functions f_1 on X_1 and f_2 on U_2 that "agree on $U \cap Z$" — what does that mean? Then we declare that the germs of the "function on U obtained by gluing together

f_1 and f_2" are compatible.) Show that the resulting ringed space is a locally ringed space.

We next want to show that the locally ringed space $X_1 \coprod_\phi X_2$ is a scheme. Clearly it is a scheme away from Z. We first verify a special case.

16.4.M. EXERCISE. Show that in Example 16.4.12 the construction of §16.4.13 indeed yields $\operatorname{Spec}(A_1 \times_{\phi^\sharp} A_2)$.

16.4.N. EXERCISE. In the general case, suppose $x \in Z$. Show that there is an affine open subset $\operatorname{Spec} A_i \subset X_i$ such that $Z \cap \operatorname{Spec} A_1 = Z \cap \operatorname{Spec} A_2$. Then use Exercise 16.4.L to show that $X_1 \coprod_\phi X_2$ is a scheme in a neighborhood of x, and thus a scheme.

16.4.14. *Remarks.*
(a) As the notation suggests, this is a fibered coproduct in the category of schemes, and indeed in the category of locally ringed spaces. We won't need this fact, but you can prove it if you wish; it isn't hard. Unlike the situation for products, fibered coproducts don't exist in general in the category of schemes. Miraculously (and for reasons that are specific to schemes), the resulting cofibered diagram is *also* a fibered diagram. This has pleasant ramifications. For example, this construction "behaves well with respect to" (or "commutes with") base change; this can help with Exercise 16.4.O(a), but if you use it, you have to prove it.
(b) Here are some interesting questions to think through: Can we recover the gluing locus from the "glued scheme" $X_1 \coprod_\phi X_2$ and the two closed subschemes X_1 and X_2? (Yes.) When is a scheme the gluing of two closed subschemes along their scheme-theoretic intersection? (When their scheme-theoretic union is the entire scheme.)
(c) You might hope that if you have a single scheme X with two disjoint closed subschemes W' and W'', and an isomorphism $W' \to W''$, then you should be able to glue X to itself along $W' \to W''$. This construction doesn't work, and indeed it may not be possible. You can still make sense of the quotient as an *algebraic space*, which I will not define here.

16.4.O. EXERCISE. We continue to use the notation X_i, ϕ, etc. Suppose we are working in the category of A-schemes.
(a) If X_1 and X_2 are universally closed, show that $X_1 \coprod_\phi X_2$ is as well.
(b) If X_1 and X_2 are separated, show that $X_1 \coprod_\phi X_2$ is as well.
(c) If X_1 and X_2 are finite type over a *Noetherian* ring A, show that $X_1 \coprod_\phi X_2$ is as well. (Hint: Reduce to the "affine" case of the Motivating Example 16.4.12. Choose generators $x_1, \dots, x_n$ of A_1, and $y_1, \dots, y_n$, such that x_i modulo I_1 agrees with y_i modulo I_2 via ϕ. Choose generators $g_1, \dots, g_m$ of I_2 — here use Noetherianness of A. Show that (x_i, y_i) and $(0, g_i)$ generate $R \subset A_1 \times A_2$, as follows. Suppose $(a_1, a_2) \in R$. Then there is some polynomial m such that $a_1 = m(x_1, \dots, x_n)$. Hence $(a_1, a_2) - m((x_1, y_1), \dots, (x_n, y_n)) = (0, a_2')$ for some $a_2' \in I_2$. Then a_2' can be written as $\sum_{i=1}^m \ell_i(y_1, \dots, y_n) g_i$. But then $(0, a_2') = \sum_{i=1}^m \ell_i((x_1, y_1), \dots, (x_n, y_n))(0, g_i)$.)

Thus if X_1 and X_2 are proper, so is $X_1 \coprod_\phi X_2$.

16.5 The Curve-to-Projective Extension Theorem

We now use the main theorem of the previous section, Theorem 16.4.1, to prove something useful and concrete.

16.5.1. The Curve-to-Projective Extension Theorem. — *Suppose C is a pure dimension 1 Noetherian scheme over an affine base S, and $p \in C$ is a regular closed point of it. Suppose Y is a projective S-scheme. Then any morphism $C \setminus \{p\} \to Y$ (of S-schemes) extends to all of C.*

In practice, we will use this theorem when $S = \operatorname{Spec} k$, and C is a k-variety. The only reason we assume S is affine is because we won't know the meaning of "projective S-scheme" until we know what a projective morphism is (§17.3). But the proof below extends immediately to general S once we know the meaning of the statement.

Note that if such an extension exists, then it is unique: the nonreduced locus of C is a closed subset (Exercise 5.5.E). Hence by replacing C by an open neighborhood of p that is reduced, we can use the Reduced-to-Separated Theorem 10.2.2 that maps from reduced schemes to separated schemes are determined by their behavior on a dense open set. Alternatively, maps to a separated scheme can be extended over an effective Cartier divisor in at most one way (Exercise 10.2.G).

The following exercise shows that the hypotheses are necessary.

16.5.A. EXERCISE. In each of the following cases, prove that the morphism $C \setminus \{p\} \to Y$ cannot be extended to a morphism $C \to Y$.

(a) *Projectivity of Y is necessary.* Suppose $C = \mathbb{A}^1_k$, $p = 0$, $Y = \mathbb{A}^1_k$, and $C \setminus \{p\} \to Y$ is given by "$t \mapsto 1/t$".
(b) *One-dimensionality of C is necessary.* Suppose $C = \mathbb{A}^2_k$, $p = (0,0)$, $Y = \mathbb{P}^1_k$, and $C \setminus \{p\} \to Y$ is given by $(x,y) \mapsto [x,y]$.
(c) *Non-singularity of C is necessary.* Suppose $C = \operatorname{Spec} k[x,y]/(y^2 - x^3)$, $p = 0$, $Y = \mathbb{P}^1_k$, and $C \setminus \{p\} \to Y$ is given by $(x,y) \mapsto [x,y]$.

We remark that by combining this (easy) theorem with the (hard) valuative criterion of properness (Theorem 12.7.6), one obtains a proof of the properness of projective space bypassing the (tricky) Fundamental Theorem of Elimination Theory 7.4.7 (see Exercise 12.7.F). Fancier remark: the valuative criterion of properness can be used to show that Theorem 16.5.1 remains true if Y is only required to be proper, but it requires some thought.

16.5.2. *Central idea of proof.* The central idea of the proof may be summarized as "clear denominators", and is illustrated by the following motivating example. Suppose you have a morphism from $\mathbb{A}^1 - \{0\}$ to projective space, and you wanted to extend it to $\mathbb{A}^1$. Suppose the map was given by $t \mapsto [t^4 + t^{-3}, t^{-2} + 4t]$. Then of course you would "clear the denominators", and replace the map by $t \mapsto [t^7 + 1, t + 4t^4]$. Similarly, if the map was given by $t \mapsto [t^2 + t^3, t^2 + t^4]$, you would divide by t^2, to obtain the map $t \mapsto [1 + t, 1 + t^2]$.

Proof. Our plan is to maneuver ourselves into the situation where we can apply the idea of §16.5.2. We begin with some quick reductions. Say $S = \operatorname{Spec} A$. The nonreduced locus of C is closed and doesn't contain p (Exercise 5.5.E), so by replacing C by an appropriate neighborhood of p, we may assume that C is reduced and affine.

We next reduce to the case where $Y = \mathbb{P}^n_A$. Choose a closed embedding $Y \to \mathbb{P}^n_A$. If the result holds for $\mathbb{P}^n$, and we have a morphism $C \to \mathbb{P}^n$ with $C \setminus \{p\}$ mapping to Y, then C must map to Y as well. Reason: we can reduce to the case where the source is an affine open subset, and the target is $\mathbb{A}^n_A \subset \mathbb{P}^n_A$ (and hence affine). Then the functions vanishing on $Y \cap \mathbb{A}^n_A$ pull back to functions that vanish at the generic points of the irreducible components of C and hence vanish everywhere on C (using reducedness of C), i.e., C maps to Y.

Choose a uniformizer $t \in \mathfrak{m} - \mathfrak{m}^2$ in the local ring of C at p. This is an element of $K(C)^\times$, with a finite number of poles (from Exercise 12.5.G on finiteness of number of zeros and poles). The complement of these finite number of points is an open neighborhood of p, so by replacing C by a smaller open affine neighborhood of p, we may assume that t is a function on C. Then $V(t)$ is also a finite number of points (including p), again from Exercise 12.5.G) so by replacing C by an open affine neighborhood of p in $C \setminus V(t) \cup p$, we may assume that p is only zero of the function t (and of course t vanishes to multiplicity 1 at p).

We have a map $C \setminus \{p\} \to \mathbb{P}^n_A$, which by Theorem 16.4.1 corresponds to a line bundle $\mathscr{L}$ on $C \setminus \{p\}$ and $n+1$ sections of it with no common zeros in $C \setminus \{p\}$. Let U be a nonempty open set of $C \setminus \{p\}$ on which $\mathscr{L} \cong \mathscr{O}$. Then by replacing C by $U \cup p$, we interpret the map to $\mathbb{P}^n$ as $n+1$ rational *functions* $f_0, \dots, f_n$, defined away away from p, with no common zeros away from p. Let $N = \min_i(\operatorname{val}_p f_i)$. Then $t^{-N} f_0, \dots, t^{-N} f_n$ are $n+1$ functions with no common zeros. Thus they determine a morphism $C \to \mathbb{P}^n_A$ extending $C \setminus \{p\} \to \mathbb{P}^n_A$ as desired. $\square$

16.5.B. Exercise (useful practice). Suppose X is a Noetherian k-scheme, and Z is an irreducible codimension 1 subvariety whose generic point is a regular point of X (so the local ring $\mathscr{O}_{X,Z}$ is a discrete valuation ring). Suppose $\pi : X \dashrightarrow Y$ is a rational map to a projective k-scheme. Show that the domain of definition of the rational map includes a dense open subset of Z. In other words, rational maps from Noetherian k-schemes to projective k-schemes can be extended over regular codimension 1 sets. (We have seen this principle in action, see Exercise 6.5.I on the Cremona transformation.)

16.6 Ample and very ample line bundles

Suppose $\pi : X \to \operatorname{Spec} A$ is a proper morphism, and $\mathscr{L}$ is an invertible sheaf on X. (The case when A is a field is the one of most immediate interest.) We say that $\mathscr{L}$ is **very ample over** A or π**-very ample**, or **relatively very ample** if $X \cong \operatorname{Proj} S_\bullet$ where $S_\bullet$ is a finitely generated graded ring over A generated in degree 1 (Definition 4.5.6), and $\mathscr{L} \cong \mathscr{O}_{\operatorname{Proj} S_\bullet}(1)$. One often just says **very ample** if the structure morphism is clear from the context. Note that the existence of a very ample line bundle implies that X is projective.

16.6.A. Easy but Important Exercise (equivalent definition of very ample over A). Suppose $\pi : X \to \operatorname{Spec} A$ is proper, and $\mathscr{L}$ is an invertible sheaf on

X. Show that $\mathscr{L}$ is very ample if and only if the sections of $\mathscr{L}$ (the complete linear series $|\mathscr{L}|$) give a closed embedding of X into $\mathbb{P}^n_A$ (as A-schemes) for some n.

16.6.B. EASY EXERCISE (VERY AMPLE IMPLIES BASE-POINT-FREE). Show that a very ample invertible sheaf $\mathscr{L}$ on a proper A-scheme must be base-point-free.

16.6.C. EXERCISE (VERY AMPLE $\otimes$ BASE-POINT-FREE IS VERY AMPLE, HENCE VERY AMPLE $\otimes$ VERY AMPLE IS VERY AMPLE). Suppose $\mathscr{L}$ and $\mathscr{M}$ are invertible sheaves on a proper A-scheme X, and $\mathscr{L}$ is very ample over A and $\mathscr{M}$ is base-point-free, then $\mathscr{L} \otimes \mathscr{M}$ is very ample. (Hint: $\mathscr{L}$ gives a closed embedding $X \hookrightarrow \mathbb{P}^m$, and $\mathscr{M}$ gives a morphism $X \to \mathbb{P}^n$. Show that the product map $X \to \mathbb{P}^m \times \mathbb{P}^n$ is a closed embedding, using the Cancellation Theorem 10.1.19 for closed embeddings on $X \to \mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^m$. Finally, consider the composition $X \hookrightarrow \mathbb{P}^m \times \mathbb{P}^n \hookrightarrow \mathbb{P}^{mn+m+n}$, where the last closed embedding is the Segre embedding.)

16.6.D. EXERCISE (VERY AMPLE $\boxtimes$ VERY AMPLE IS VERY AMPLE, CF. EXAMPLE 16.4.8). Suppose X and Y are proper A-schemes, and $\mathscr{L}$ (resp. $\mathscr{M}$) is a very ample invertible sheaf on X (resp. Y). If $\pi_X : X \times_A Y \to X$ and $\pi_Y : X \times_A Y \to Y$ are the usual projections, show that $\pi_X^* \mathscr{L} \otimes \pi_Y^* \mathscr{M}$ (also known as $\mathscr{L} \boxtimes \mathscr{M}$, see §16.4.8) is very ample on $X \times_A Y$.

16.6.1. *Definition.* We say an invertible sheaf $\mathscr{L}$ on a proper A-scheme X is **ample over** A or π**-ample** (where $\pi : X \to \operatorname{Spec} A$ is the structure morphism), or **relatively ample** if one of the following equivalent conditions holds.

16.6.2. Theorem. — *Suppose $\pi : X \to \operatorname{Spec} A$ is proper, and $\mathscr{L}$ is an invertible sheaf on X. The following are equivalent.*

(a) *For some $N > 0$, $\mathscr{L}^{\otimes N}$ is very ample over A.*
(a') *For all $n \gg 0$, $\mathscr{L}^{\otimes n}$ is very ample over A.*
(b) *For all finite type quasicoherent sheaves $\mathscr{F}$, there is an n_0 such that for $n \geq n_0$, $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated.*
(c) *As f runs over all the global sections of $\mathscr{L}^{\otimes n}$ (over all $n > 0$), the open subsets $X_f = \{p \in X : f(p) \neq 0\}$ form a base for the topology of X.*
(c') *As f runs over the global sections of $\mathscr{L}^{\otimes n}$ ($n > 0$), those open subsets X_f which are affine form a base for the topology of X.*

(Variants of this Theorem 16.6.2 in the "absolute" and "relative" settings will be given in Theorems 16.6.6 and 17.3.9 respectively.)

Properties (a) and (a') relate to projective geometry, and property (b) relates to global generation (stalks). Properties (c) and (c') are somehow more topological, and while they may seem odd, they will provide the connection between (a)/(a') and (b). Note that (c) and (c') make no reference to the structure morphism π. In Theorem 18.7.1, we will meet a cohomological criterion (due, unsurprisingly, to Serre) later. Kodaira also gives a criterion for ampleness in the complex category: if X is a complex projective variety, then an invertible sheaf $\mathscr{L}$ on X is ample if and only if it admits a Hermitian metric with curvature positive everywhere.

The different flavor of these conditions gives some indication that ampleness is better-behaved than very ampleness in a number of ways. We mention without proof another property: if $\pi : X \to T$ is a finitely presented proper morphism, and $\mathscr{L}$ is an invertible sheaf on X, then those points on T where $\mathscr{L}$ is ample on the fiber is ample forms an open subset of T. Furthermore, on this open subset, $\mathscr{L}$ is relatively ample over the base. We won't use these facts (proved in [**Gr-EGA**, IV$_3$.9.6.4]), but they are good to know.

Before getting to the proof of Theorem 16.6.2, we give some sample applications. We begin by noting that the fact that (a) implies (b) gives Serre's Theorem A (Theorem 15.3.8).

16.6.E. IMPORTANT EXERCISE. Suppose $\mathscr{L}$ and $\mathscr{M}$ are invertible sheaves on a proper A-scheme X, and $\mathscr{L}$ is ample. Show that $\mathscr{L}^{\otimes n} \otimes \mathscr{M}$ is very ample for $n \gg 0$. (Hint: use both (a) and (b) of Theorem 16.6.2, and Exercise 16.6.C.)

16.6.F. IMPORTANT EXERCISE. Show that every line bundle on a projective A-scheme X is the difference of two very ample line bundles. More precisely, for any invertible sheaf $\mathscr{L}$ on X, we can find two very ample invertible sheaves $\mathscr{M}$ and $\mathscr{N}$ such that $\mathscr{L} \cong \mathscr{M} \otimes \mathscr{N}^\vee$. (Hint: use the previous Exercise.)

16.6.G. IMPORTANT EXERCISE (USED REPEATEDLY). Suppose $\pi : X \to Y$ is a finite morphism of proper A-schemes, and $\mathscr{L}$ is an ample line bundle on Y. Show that $\pi^*\mathscr{L}$ is ample on X. Hint: use the criterion of Theorem 16.6.2(b). Suppose $\mathscr{F}$ is a finite type quasicoherent sheaf on X. We wish to show that $\mathscr{F} \otimes (\pi^*\mathscr{L})^{\otimes n}$ is globally generated for $n \gg 0$. Note that $(\pi_*\mathscr{F}) \otimes \mathscr{L}^{\otimes n}$ is globally generated for $n \gg 0$ by ampleness of $\mathscr{L}$ on Y, i.e., there exists a surjection

$$\mathscr{O}_Y^{\oplus I} \twoheadrightarrow (\pi_*\mathscr{F}) \otimes \mathscr{L}^{\otimes n},$$

where I is some index set. Show that

$$\mathscr{O}_X^{\oplus I} \cong \pi^*(\mathscr{O}_Y^{\oplus I}) \longrightarrow \pi^*(\pi_*\mathscr{F} \otimes \mathscr{L}^{\otimes n})$$

is surjective. Pullback π^* preserves tensor products (Theorem 16.3.7(9)), so we have an isomorphism $\pi^*(\pi_*\mathscr{F} \otimes \mathscr{L}^{\otimes n}) \cong \pi^*(\pi_*\mathscr{F}) \otimes (\pi^*\mathscr{L})^{\otimes n}$. Show (using only affineness of π) that $\pi^*\pi_*\mathscr{F} \to \mathscr{F}$ is surjective. Connect these pieces together to describe a surjection

$$\mathscr{O}_X^{\oplus I} \twoheadrightarrow \mathscr{F} \otimes (\pi^*\mathscr{L})^{\otimes n}.$$

(Remark for those who have read about ampleness in the absolute setting in §16.6.5: the argument applies in that situation, i.e., with "proper A-schemes" changed to "schemes", without change. The only additional thing to note is that ampleness of $\mathscr{L}$ on Y implies that Y is quasicompact from the definition, and separated from Theorem 16.6.6(d). A relative version of this result appears in §17.3.8. It can be generalized even further, with "π finite" replaced by "π quasiaffine" — to be defined in §17.3.11 — see [**Gr-EGA**, II.5.1.12].)

16.6.H. EXERCISE (AMPLE $\otimes$ AMPLE IS AMPLE, AMPLE $\otimes$ BASE-POINT-FREE IS AMPLE. Suppose $\mathscr{L}$ and $\mathscr{M}$ are invertible sheaves on a proper A-scheme X, and $\mathscr{L}$ is ample. Show that if $\mathscr{M}$ is ample or base-point-free, then $\mathscr{L} \otimes \mathscr{M}$ is ample.

16.6.I. LESS IMPORTANT EXERCISE (AMPLE $\boxtimes$ AMPLE IS AMPLE). Solve Exercise 16.6.D with "very ample" replaced by "ample".

16.6.3. *Proof of Theorem 16.6.2 in the case X is Noetherian.* Noetherian hypotheses are used at only one point in the proof, and we explain how to remove them, and give a reference for the details.

Obviously, *(a′) implies (a)*.

Clearly *(c′) implies (c)*. We now show that *(c) implies (c′)*. Suppose we have a point p in an open subset U of X. We seek an *affine* X_f containing p and contained in U. By shrinking U, we may assume that U is affine. From (c), U contains some X_f containing p. But this X_f is affine, as it is the complement of the vanishing locus of a section of a line bundle on an affine scheme (Exercise 7.3.F), so (c′) holds. Note for future reference that the equivalence of (c) and (c′) did not require the hypothesis of properness.

We next show that *(a) implies (c)*. We embed X in projective space by some power of $\mathscr{L}$. Given a closed subset $Z \subset X$, and a point p of the complement $X \setminus Z$, we seek a section of some $\mathscr{L}^{\otimes N}$ that vanishes on Z and not on p. The existence of such a section follows from the fact that $V(I(Z)) = Z$ (Exercise 4.5.H(c)): there is some element of $I(Z)$ that does not vanish on p.

We next show that *(b) implies (c)*. Suppose we have a point p in an open subset U of X. We seek a section of $\mathscr{L}^{\otimes N}$ that doesn't vanish at p, but vanishes on $X \setminus U$. Let $\mathscr{I}$ be the sheaf of ideals of functions vanishing on $X \setminus U$ (the quasicoherent sheaf of ideals cutting out $X \setminus U$, with reduced structure). As X is Noetherian, $\mathscr{I}$ is finite type, so by (b), $\mathscr{I} \otimes \mathscr{L}^{\otimes N}$ is generated by global sections for some N, so there is some section of it not vanishing at p. (*Noetherian note:* This is the only part of the argument where we use Noetherian hypotheses. They can be removed as follows. Show that for a quasicompact quasiseparated scheme, every ideal sheaf is generated by its finite type subideal sheaves. Indeed, any quasicoherent sheaf on a quasicompact quasiseparated scheme is the union of its finite type quasicoherent subsheaves, see [**Gr-EGA′**, (6.9.9)] or [**GW**, Cor. 10.50]. One of these finite type ideal sheaves doesn't vanish at p; use this as $\mathscr{I}$ instead.)

We now have to start working harder.

We next show that *(c′) implies (b)*. We wish to show that $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated for $n \gg 0$.

We first show that (c′) implies that for some $N > 0$, $\mathscr{L}^{\otimes N}$ is globally generated, as follows. For each closed point $p \in X$, there is some $f \in \Gamma(X, \mathscr{L}^{\otimes N(p)})$ not vanishing at p, so $p \in X_f$. (Don't forget that quasicompact schemes have closed points, Exercise 5.1.E!) As p varies, these X_f cover all of X. Use quasicompactness of X to select a finite number of these X_f that cover X. To set notation, say these are $X_{f_1}, \dots, X_{f_n}$, where $f_i \in \Gamma(X, \mathscr{L}^{\otimes N_i})$. By replacing f_i with $f_i^{\otimes (\prod_j N_j)/N_i}$, we may assume that they are all sections of the *same* power $\mathscr{L}^{\otimes N}$ of $\mathscr{L}$ ($N = \prod_j N_j$). Then $\mathscr{L}^{\otimes N}$ is generated by these global sections.

We next show that it suffices to show that for all finite type quasicoherent sheaves $\mathscr{F}$, $\mathscr{F} \otimes \mathscr{L}^{\otimes mN}$ is globally generated for $m \gg 0$. For if we knew this, we could apply it to $\mathscr{F}, \mathscr{F} \otimes \mathscr{L}, \dots, \mathscr{F} \otimes \mathscr{L}^{\otimes (N-1)}$ (a finite number of times), and the result would follow. For this reason, we can replace $\mathscr{L}$ by $\mathscr{L}^{\otimes N}$. In other words, to show that (c′) implies (b), we may also assume the additional hypothesis that $\mathscr{L}$ is globally generated.

For each closed point p, choose an affine neighborhood of the form X_f, using (c′). Then $\mathscr{F}|_{X_f}$ is generated by a finite number of global sections (Easy Exercise 15.3.A). By Exercise 13.3.H, each of these generators can be expressed as a quotient of a section (over X) of $\mathscr{F} \otimes \mathscr{L}^{\otimes M(p)}$ by $f^{M(p)}$. (Note: we can take a single $M(p)$ for each p.) Then $\mathscr{F} \otimes \mathscr{L}^{\otimes M(p)}$ is globally generated at p by a finite number of global sections. By Exercise 15.3.C(b), $\mathscr{F} \otimes \mathscr{L}^{\otimes M(p)}$ is globally generated at all points in some neighborhood U_p of p. As $\mathscr{L}$ is also globally generated, this implies that $\mathscr{F} \otimes \mathscr{L}^{\otimes M'}$ is globally generated at all points of U_p for $M' \geq M(p)$ (cf. Easy Exercise 15.3.B). From quasicompactness of X, a finite number of these U_p cover X, so we are done (by taking the maximum of these $M(p)$).

Our penultimate step is to show that *(c′) implies (a)*. Our goal is to assume (c′), and to find sections of some $\mathscr{L}^{\otimes N}$ that embeds X into projective space. Choose a cover of (quasicompact) X by n affine open subsets $X_{a_1}, \dots, X_{a_n}$, where $a_1, \dots, a_n$ are all sections of powers of $\mathscr{L}$. By replacing each section with a suitable power, we may assume that they are all sections of the *same* power of $\mathscr{L}$, say $\mathscr{L}^{\otimes N}$. Say $X_{a_i} = \operatorname{Spec} A_i$, where (using that π is finite type) $A_i = \operatorname{Spec} A[a_{i1}, \dots, a_{ij_i}]/I_i$. By Exercise 13.3.H, each a_{ij} is of the form $s_{ij}/a_i^{m_{ij}}$, where $s_{ij} \in \Gamma(X, \mathscr{L}^{\otimes m_{ij}})$ (for some m_{ij}). Let $m = \max_{i,j} m_{ij}$. Then for each i, j, $a_{ij} = (s_{ij} a_i^{m-m_{ij}})/a_i^m$. For convenience, let $b_i = a_i^m$, and $b_{ij} = s_{ij} a_i^{m-m_{ij}}$; these are all global sections of $\mathscr{L}^{\otimes mN}$. Now consider the linear series generated by the b_i and b_{ij}. As the $D(b_i) = X_{a_i}$ cover X, this linear series is base-point-free, and hence (by Exercise 15.3.F) gives a morphism to $\mathbb{P}^Q$ (where $Q = \#b_i + \#b_{ij} - 1$). Let $x_1, \dots, x_n, \dots, x_{ij}, \dots$ be the projective coordinates on $\mathbb{P}^Q$, so $f^* x_i = b_i$, and $f^* x_{ij} = b_{ij}$. Then the morphism of affine schemes $X_{a_i} \to D(x_i)$ is a closed embedding, as the associated maps of rings is a surjection (the generator a_{ij} of A_i is the image of x_{ij}/x_i).

At this point, we note for future reference that we have shown the following. If $X \to \operatorname{Spec} A$ is finite type, and $\mathscr{L}$ satisfies (c)=(c′), then X is an open embedding into a projective A-scheme. (We did not use separatedness.) We conclude our proof that (c′) implies (a) by using properness to show that the image of this open embedding into a projective A-scheme is in fact closed, so X is a projective A-scheme.

Finally, we note that *(a) and (b) together imply (a′)*: if $\mathscr{L}^{\otimes N}$ is very ample (from (a)), and $\mathscr{L}^{\otimes n}$ is base-point-free for $n \geq n_0$ (from (b)), then $\mathscr{L}^{\otimes n}$ is very ample for $n \geq n_0 + N$ by Exercise 16.6.C. □

16.6.4. ⋆⋆ *Semiample line bundles.* Just as an invertible sheaf is ample if some tensor power of it is very ample, an invertible sheaf $\mathscr{L}$ is said to be **semiample** if some tensor power of it is base-point-free. (Translation: $\mathscr{L}$ is ample if some power gives a closed embedding into projective space, and $\mathscr{L}$ is semiample if some power gives just a *morphism* to projective space.) We won't use this notion.

16.6.5. ⋆ **Ampleness in the absolute setting.** (We will not use this section in any serious way later.) Note that global generation is already an absolute notion, i.e., is defined for a quasicoherent sheaf on a scheme, with no reference to any morphism. An examination of the proof of Theorem 16.6.2 shows that ampleness may similarly be interpreted in an absolute setting. We make this precise. Suppose $\mathscr{L}$ is an invertible sheaf on a *quasicompact* scheme X. We say that $\mathscr{L}$ is **ample** if as f runs over the sections of $\mathscr{L}^{\otimes n}$ ($n > 0$), the open subsets $X_f = \{p \in X : f(p) \neq 0\}$ form a base for the topology of X. (We emphasize that quasicompactness in X is part of the condition of ampleness of $\mathscr{L}$.) For example, (i) if X is an affine scheme, every invertible sheaf is ample, and (ii) if X is a projective A-scheme, $\mathscr{O}(1)$ is ample.

16.6.J. EASY EXERCISE (PROPERTIES OF ABSOLUTE AMPLENESS).
(a) Fix a positive integer n. Show that $\mathscr{L}$ is ample if and only if $\mathscr{L}^{\otimes n}$ is ample.
(b) Show that if $Z \hookrightarrow X$ is a closed embedding, and $\mathscr{L}$ is ample on X, then $\mathscr{L}|_Z$ is ample on Z.

The following result will give you some sense of how ampleness behaves. We will not use it, and hence omit the proof (which is given in [**Stacks**, tag 01Q3]). However, many parts of the proof are identical to (or generalize) the corresponding arguments in Theorem 16.6.2. The labeling of the statements parallels the labelling of the statements in Theorem 16.6.2.

16.6.6. Theorem (cf. Theorem 16.6.2). — *Suppose $\mathscr{L}$ is an invertible sheaf on a quasicompact scheme X. The following are equivalent.*

(b) *X is quasiseparated, and for every finite type quasicoherent sheaf $\mathscr{F}$, there is an n_0 such that for $n \geq n_0$, $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated.*
(c) *As f runs over the global sections of $\mathscr{L}^{\otimes n}$ ($n > 0$), the open subsets $X_f = \{p \in X : f(p) \neq 0\}$ form a base for the topology of X (i.e., $\mathscr{L}$ is ample).*
(c′) *As f runs over the global sections of $\mathscr{L}^{\otimes n}$ ($n > 0$), those open subsets X_f which are affine form a base for the topology of X.*
(d) *Let $S_\bullet$ be the graded ring $\oplus_{n \geq 0} \Gamma(X, \mathscr{L}^{\otimes n})$. Then the open sets X_s with $s \in S_+$ cover X, and the associated map $X \to \operatorname{Proj} S$ is an open embedding. (Warning: $S_\bullet$ need not be finitely generated, and $\operatorname{Proj} S$ is not necessarily finite type.)*

Part (d) implies that X is separated (and thus quasiseparated).

16.6.7. $\star$ **Transporting global generation, base-point-freeness, and ampleness to the relative situation.**
These notions can be "relativized". We could do this right now, but we wait until §17.3.7, when we will have defined the notion of a projective morphism, and thus a "relatively very ample" line bundle.

16.7 $\star$ The Grassmannian as a moduli space

We first defined projective space inelegantly in §4.4.9, and in §16.4.2 we gave a clean (if perhaps surprising) functorial definition. Similarly, in §6.7, we gave a preliminary description of the Grassmannian. We are now in a position to give a better definition.

We describe the "Grassmannian functor" (which we also denote $G(k, n)$), then show that it is representable (§6.6.2). The construction works over an arbitrary base scheme, so we work over the final object $\operatorname{Spec} \mathbb{Z}$. (You should think through what to change if you wish to work with, for example, complex schemes.) The functor is defined as follows. To a scheme B, we associate the set of *locally free rank* k *quotients of the rank* n *free sheaf,*

$$\mathscr{O}_B^{\oplus n} \twoheadrightarrow \mathscr{Q} \tag{16.7.0.1}$$

up to isomorphism. An isomorphism of two such quotients $\phi : \mathscr{O}_B^{\oplus n} \to \mathscr{Q} \to 0$ and $\phi' : \mathscr{O}_B^{\oplus n} \to \mathscr{Q}' \to 0$ is an isomorphism $\sigma : \mathscr{Q} \to \mathscr{Q}'$ such that the diagram

$$\begin{tikzcd} \mathscr{O}^{\oplus n} \arrow[r, "\phi"] \arrow[dr, "\phi'"'] & \mathscr{Q} \arrow[d, "\sigma"] \\ & \mathscr{Q}' \end{tikzcd}$$

commutes. By Exercise 13.5.B(a), $\ker \phi$ is locally free of rank $n-k$. (Thus if you prefer, you can extend (16.7.0.1), and instead consider the functor to take B to short exact sequences

$$0 \to \mathscr{S} \to \mathscr{O}^{\oplus n} \to \mathscr{Q} \to 0 \tag{16.7.0.2}$$

of locally free sheaves over B, of ranks $n-k$, n, and k respectively.)

It may surprise you that we are considering rank k *quotients* of a rank n sheaf, not rank k *subobjects*, given that the Grassmannian should parametrize k-dimensional subspace of an n-dimensional space. This is done for several reasons. One is that the kernel of a surjective map of locally free sheaves must be locally free, while the cokernel of an injective map of locally free sheaves need not be locally free (Exercise 13.5.B(a) and (b) respectively). Another reason: we will see in §28.3.3 that the geometric incarnation of this problem indeed translates to this. We can already see a key example here: if $k=1$, our definition yields one-dimensional quotients $\mathscr{O}^{\oplus n} \to \mathscr{L} \to 0$. But this is precisely the data of n sections of $\mathscr{L}$, with no common zeros, which by Theorem 16.4.1 (the functorial description of projective space) corresponds precisely to maps to $\mathbb{P}^n$, so the $k=1$ case parametrizes what we want.

We now show that the Grassmannian functor is representable for given n and k.

16.7.A. EXERCISE. Show that the Grassmannian functor is a Zariski sheaf (§9.1.7).

Hence by Key Exercise 9.1.I, to show that the Grassmannian functor is representable, we need only cover it with open subfunctors that are representable.

Throughout the rest of this section, a k-*subset* is a subset of $\{1, \dots, n\}$ of size k.

16.7.B. EXERCISE.
(a) Suppose I is a k-subset. Make the following statement precise: there is an open subfunctor $G(k,n)_I$ of $G(k,n)$ where the k sections of $\mathscr{Q}$ corresponding to I (of the n sections of $\mathscr{Q}$ coming from the surjection $\phi : \mathscr{O}^{\oplus n} \to \mathscr{Q}$) are linearly independent. Hint: in a trivializing neighborhood of $\mathscr{Q}$, where we can choose an isomorphism $\mathscr{Q} \xrightarrow{\sim} \mathscr{O}^{\oplus k}$, ϕ can be interpreted as a $k \times n$ matrix M, and this locus is where the determinant of the $k \times k$ matrix consisting of the I columns of M is nonzero. Show that this locus behaves well under transitions between trivializations.
(b) Show that these open subfunctors $G(k,n)_I$ cover the functor $G(k,n)$ (as I runs through the k-subsets).

Hence by Exercise 9.1.I, to show $G(k,n)$ is representable, we need only show that $G(k,n)_I$ is representable for arbitrary I. After renaming the summands of $\mathscr{O}^{\oplus n}$, without loss of generality we may assume $I = \{1, \dots, k\}$.

16.7.C. EXERCISE. Show that $G(k,n)_{\{1\dots,k\}}$ is represented by $\mathbb{A}^{k(n-k)}$ as follows. (You will have to make this precise.) Given a surjection $\phi : \mathscr{O}^{\oplus n} \to \mathscr{Q}$, let $\phi_i : \mathscr{O} \to \mathscr{Q}$ be the map from the ith summand of $\mathscr{O}^{\oplus n}$. (Really, ϕ_i is just a section of $\mathscr{Q}$.) For the open subfunctor $G(k,n)_I$, show that

$$\phi_1 \oplus \cdots \oplus \phi_k : \mathscr{O}^{\oplus k} \to \mathscr{Q}$$

is an isomorphism. For a scheme B, the bijection $G(k,n)_I(B) \leftrightarrow \operatorname{Hom}(B, \mathbb{A}^{nk})$ is given as follows. Given an element $\phi \in G(k,n)_I(B)$, for $j \in \{k+1,\dots,n\}$, $\phi_j = a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{kj}\phi_k$, where a_{ij} are functions on B. But $k(n-k)$ functions on B is the same as a map to $\mathbb{A}^{k(n-k)}$ (Exercise 6.6.E). Conversely, given $k(n-k)$ functions a_{ij} $(1 \leq i \leq k < j \leq n)$, define a surjection $\phi : \mathscr{O}^{\oplus n} \to \mathscr{O}^{\oplus k}$ as follows: $(\phi_1 \dots, \phi_k)$ is the identity, and $\phi_j = a_{1j}\phi_1 + a_{2j}\phi_2 + \cdots + a_{kj}\phi_k$ for $j > k$.

You have now shown that $G(k,n)$ is representable, by covering it with $\binom{n}{k}$ copies of $\mathbb{A}^{k(n-k)}$. (You might wish to relate this to the description you gave in §6.7.) In particular, the Grassmannian over a field is smooth, and irreducible of dimension $k(n-k)$. (The Grassmannian over any base is smooth over that base, because $\mathbb{A}_B^{k(n-k)} \to B$ is smooth, see §12.6.2.)

16.7.1. *The universal exact sequence over the Grassmannian.* Note that we have a tautological exact sequence

$$0 \to \mathscr{S} \to \mathscr{O}^{\oplus n} \to \mathscr{Q} \to 0.$$

16.7.2. The Plücker embedding.

By applying $\wedge^k$ to a surjection $\phi : \mathscr{O}^{\oplus n} \to \mathscr{Q}$ (over an arbitrary base B), we get a surjection $\wedge^k \phi : \mathscr{O}^{\oplus\binom{n}{k}} \to \det \mathscr{Q}$ (Exercise 13.5.F). But a surjection from a rank N free sheaf to a line bundle is the same as a map to $\mathbb{P}^{N-1}$ (Theorem 16.4.1).

16.7.D. EXERCISE. Use this to describe a map $P : G(k,n) \to \mathbb{P}^{\binom{n}{k}-1}$. (This is just a tautology: a natural transformation of functors induces a map of the representing schemes. This is Yoneda's Lemma, although if you didn't do Exercise 1.3.Z, you may wish to do this exercise by hand. But once you do, you may as well go back to prove Yoneda's Lemma and do Exercise 1.3.Z, because the argument is just the same!)

16.7.E. EXERCISE. The projective coordinate x_I on $\mathbb{P}^{\binom{n}{k}-1}$ corresponding to the Ith factor of $\mathscr{O}^{\oplus\binom{n}{k}}$ may be interpreted as the determinant of the map $\phi_I : \mathscr{O}^{\oplus k} \to \mathscr{Q}$, where the $\mathscr{O}^{\oplus k}$ consists of the summands of $\mathscr{O}^{\oplus n}$ corresponding to I. Make this precise.

16.7.F. EXERCISE. Show that the standard open set U_I of $\mathbb{P}^{\binom{n}{k}-1}$ corresponding to k-subset I (i.e., where the corresponding coordinate x_I doesn't vanish) pulls back to the open subscheme $G(k,n)_I \subset G(k,n)$. Denote this map $P_I : G(k,n)_I \to U_I$.

16.7.G. EXERCISE. Show that P_I is a closed embedding as follows. We may deal with the case $I = \{1,\dots,k\}$. Note that $G(k,n)_I$ is affine — you described it $\operatorname{Spec} \mathbb{Z}[a_{ij}]_{1 \leq i \leq k < j \leq n}$ in Exercise 16.7.C. Also, U_I is affine, with coordinates $x_{I'/I}$, as I' varies over the other k-subsets. You want to show that the map

$$P_I^\sharp : \mathbb{Z}[x_{I'/I}]_{I'\subset\{1,\dots,n\},|I'|=k\}}/(x_{I/I}-1) \to \mathbb{Z}[a_{ij}]_{1\leq i\leq k<j\leq n}$$

is a surjection. By interpreting the map $\phi : \mathscr{O}^{\oplus n} \to \mathscr{O}^{\oplus k}$ as a $k \times n$ matrix M whose left k columns are the identity matrix and whose remaining entries are a_{ij} ($1 \leq i \leq k < j \leq n$), interpret $P_I^\sharp$ as taking $x_{I'/I}$ to the determinant of the $k \times k$ submatrix corresponding to the columns in I'. For each (i,j) (with $1 \leq i \leq k < j \leq n$), find some I' so that $x_{I'/I} \mapsto \pm a_{ij}$. (Let $I' = \{1, \dots, i-1, i+1, \dots, k, j\}$.)

Hence $G(k,n) \to \mathbb{P}^{\binom{n}{k}-1}$ is a closed embedding, so $G(k,n)$ is projective over $\mathbb{Z}$.

16.7.H. UNIMPORTANT EXERCISE. As an entertaining geometric consequence: if V is a vector space over a field, show that the "pure tensors in $\wedge^k V$ are pure in exactly one way": if $v_1 \wedge \cdots \wedge v_k = w_1 \wedge \cdots \wedge w_k \neq 0$ in $\wedge^k V$, show that there is a $k \times k$ matrix of determinant 1 relating the v_i to the w_i.

16.7.3. *The Plücker equations.* The equations of $G(k,n) \to \mathbb{P}^{\binom{n}{k}-1}$ are particularly nice. There are quadratic relations among the $k \times k$ minors of a $k \times (n-k)$ matrix, called the Plücker relations. By our construction, they are equations satisfied by $G(k,n)$. It turns out that these equations cut out $G(k,n)$, and in fact generate the homogeneous ideal of $G(k,n)$, but this takes more work (see [**MS**, §14.2]). We explore this in one example.

16.7.I. EASY EXERCISE. Suppose v_1, v_2, v_3, and v_4 are four vectors in a two-dimensional vector space V over some field. Show that

$$(v_1 \wedge v_2)(v_3 \wedge v_4) - (v_1 \wedge v_3)(v_2 \wedge v_4) + (v_1 \wedge v_4)(v_2 \wedge v_3) = 0.$$

16.7.J. EXERCISE. Note that the Plücker embedding embeds the Grassmannian $G(2,4)$ into $\mathbb{P}^5$.
(a) Show that $G(2,4)$ is cut out by the quadratic equation $x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23} = 0$. (Hint: Use Exercise 16.7.I to show that the quadratic vanishes on $G(2,4)$. But that isn't enough.)
(b) Show that every smooth quadric hypersurface in $\mathbb{P}^5$ over an algebraically closed field $\overline{k}$ is isomorphic to the Grassmannian (over $\overline{k}$). (For comparison, every smooth quadric hypersurface in $\mathbb{P}^1_{\overline{k}}$ is two points; every smooth quadric hypersurface in $\mathbb{P}^2_{\overline{k}}$ is isomorphic to $\mathbb{P}^1_{\overline{k}}$, §6.5.9; and every smooth quadric hypersurface in $\mathbb{P}^3_{\overline{k}}$ is isomorphic to $\mathbb{P}^1_{\overline{k}} \times_{\overline{k}} \mathbb{P}^1_{\overline{k}}$, Example 9.6.2.)

16.7.K. EXERCISE (RULINGS ON QUADRICS IN $\mathbb{P}^5$). Suppose $k = \overline{k}$. From Remark 8.2.10, we expect two three-dimensional families of planes in $G(2,4)$ (interpreted as a hypersurface in $\mathbb{P}^5$ via the Plücker embedding, see Exercise 16.7.I). One of them may be described as follows: for each point $p \in \mathbb{P}^3_k$, we have a two-dimensional family of lines through p; this is a plane in $G(2,4)$. There is a three-dimensional family of planes corresponding to the choice of p. This is one of the two rulings. What is the other one? Prove (as rigorously as you can manage, given what you know) that these are both rulings.

16.7.4. *Further discussion.*

16.7.L. EXERCISE. Show that the group scheme GL_n acts on the Grassmannian $G(k, n)$. (The *action* of a group scheme appeared earlier in Exercise 6.6.S(a).) Hint: this is much more easily done with the language of functors, §6.6, than with the description of the Grassmannian in terms of patches, §6.7. (Exercise 16.4.K was the special case of projective space.)

16.7.M. ⋆⋆ EXERCISE (GRASSMANNIAN BUNDLES). Suppose $\mathscr{F}$ is a rank n locally free sheaf on a scheme X. Define the Grassmannian bundle $G(k, \mathscr{F})$ over X. Intuitively, if $\mathscr{F}$ is a varying family of n-dimensional vector spaces over X, $G(k, \mathscr{F})$ should parametrize k-dimensional quotients of the fibers. You may want to define the functor first, and then show that it is representable. Your construction will behave well under base change.

16.7.5. *(Partial) flag varieties.* The discussion here extends without change to partial flag varieties (§6.7.1), and the interested reader should think this through.

CHAPTER 17

Relative versions of Spec and Proj, and projective morphisms

In this chapter, we will use universal properties to define two useful constructions, $\mathcal{S}pec$ of a sheaf of algebras $\mathscr{A}$, and $\mathcal{P}roj$ of a sheaf of graded algebras $\mathscr{A}_\bullet$ on a scheme X. These will both generalize (globalize) our constructions of Spec of A-algebras and Proj of graded A-algebras. We will see that affine morphisms are precisely those of the form $\mathcal{S}pec\,\mathscr{A} \to X$, and so we will *define* projective morphisms to be those of the form $\mathcal{P}roj\,\mathscr{A}_\bullet \to X$.

In both cases, our plan is to make a notion we know well over a ring work more generally over a scheme. The main issue is how to glue the constructions over each affine open subset together. The slick way we will proceed is to give a universal property, then show that the affine construction satisfies this universal property, then that the universal property behaves well with respect to open subsets, then to use the idea that let us glue together the fibered product (or normalization) together to do all the hard gluing work. The most annoying part of this plan is finding the right universal property, especially in the $\mathcal{P}roj$ case.

17.1 Relative Spec of a (quasicoherent) sheaf of algebras

Given an A-algebra, B, we can take its Spec to get an affine scheme over Spec A: $\operatorname{Spec} B \to \operatorname{Spec} A$. We will now see a universal property description of a globalization of that notation. Consider an arbitrary scheme X, and a quasicoherent sheaf of algebras $\mathscr{B}$ on it. We will define how to take Spec of this sheaf of algebras, and we will get a scheme $\mathcal{S}pec\,\mathscr{B} \to X$ that is "affine over X", i.e., the structure morphism is an affine morphism. You can think of this in two ways.

17.1.1. First, and most concretely, for any affine open set $\operatorname{Spec} A \subset X$, $\Gamma(\operatorname{Spec} A, \mathscr{B})$ is some A-algebra; call it B. Then above $\operatorname{Spec} A$, $\mathcal{S}pec\,\mathscr{B}$ will be $\operatorname{Spec} B$.

17.1.2. Second, it will satisfy a universal property. We could define the A-scheme $\operatorname{Spec} B$ by the fact that morphisms to $\operatorname{Spec} B$ (from an A-scheme W, over $\operatorname{Spec} A$) correspond to maps of A-algebras $B \to \Gamma(W, \mathscr{O}_W)$ (this is our old friend Exercise 6.3.F). The universal property for $\beta : \mathcal{S}pec\,\mathscr{B} \to X$ generalizes this. Given a morphism $\mu : W \to X$, the X-morphisms $W \to \mathcal{S}pec\,\mathscr{B}$ are in functorial (in W)

bijection with morphisms α making

$$\begin{tikzcd} & \mathscr{O}_X \arrow[dl] \arrow[dr] & \\ \mathscr{B} \arrow[rr, "\alpha"] & & \mu_* \mathscr{O}_W \end{tikzcd}$$

commute. Here the map $\mathscr{O}_X \to \mu_* \mathscr{O}_W$ is that coming from the map of ringed spaces, and the map $\mathscr{O}_X \to \mathscr{B}$ comes from the $\mathscr{O}_X$-algebra structure on $\mathscr{B}$. (For experts: it needn't be true that $\mu_* \mathscr{O}_W$ is quasicoherent, but that doesn't matter. Non-experts should completely ignore this parenthetical comment.)

By universal property nonsense, these data determine $\beta : \mathcal{S}pec\, \mathscr{B} \to X$ up to unique isomorphism, assuming that it exists.

Fancy translation: in the category of X-schemes, $\beta : \mathcal{S}pec\, \mathscr{B} \to X$ represents the functor

$$(\mu : W \to X) \longmapsto \{(\alpha : \mathscr{B} \to \mu_* \mathscr{O}_W)\}.$$

17.1.A. EXERCISE. Show that if X is affine, say $\operatorname{Spec} A$, and $\mathscr{B} = \widetilde{B}$, where B is an A-algebra, then $\operatorname{Spec} B \to \operatorname{Spec} A$ satisfies this universal property. (Hint: Exercise 6.3.F.)

17.1.3. Proposition. — *Suppose $\beta : \mathcal{S}pec\, \mathscr{B} \to X$ satisfies the universal property for $(X, \mathscr{B})$, and $U \hookrightarrow X$ is an open subset. Then $\beta|_U : \mathcal{S}pec\, \mathscr{B} \times_X U = (\mathcal{S}pec\, \mathscr{B})|_U \to U$ satisfies the universal property for $(U, \mathscr{B}|_U)$.*

Proof. For convenience, let $V = \mathcal{S}pec\, \mathscr{B} \times_X U$. A U-morphism $W \to V$ is the same as an X-morphism $W \to \mathcal{S}pec\, \mathscr{B}$ (where by assumption $\mu : W \to X$ factors through U). By the universal property of $\mathcal{S}pec\, \mathscr{B}$, this is the same information as a map $\mathscr{B} \to \mu_* \mathscr{O}_W$, which by the universal property definition of pullback (§16.3.4) is the same as $\mu^* \mathscr{B} \to \mathscr{O}_W$, which is the same information as $(\mu|_U)^* \mathscr{B} \to \mathscr{O}_W$. By adjointness again this is the same as $\mathscr{B}|_U \to (\mu|_U)_* \mathscr{O}_W$. $\square$

Combining the above Exercise and Proposition, we have shown the existence of $\mathcal{S}pec\, \mathscr{B}$ in the case that X is an open subscheme of an affine scheme.

17.1.B. EXERCISE. Show the existence of $\mathcal{S}pec\, \mathscr{B}$ in general, following the philosophy of our construction of the fibered product, normalization, and so forth.

We make some quick observations. First $\mathcal{S}pec\, \mathscr{B}$ can be "computed affine-locally on X". We also have an isomorphism $\phi : \mathscr{B} \to \beta_* \mathscr{O}_{\mathcal{S}pec\, \mathscr{B}}$.

17.1.C. EXERCISE. Given an X-morphism

$$\begin{tikzcd} W \arrow[rr, "\gamma"] \arrow[dr, "\mu"'] & & \mathcal{S}pec\, \mathscr{B} \arrow[dl, "\beta"] \\ & X & \end{tikzcd}$$

show that α is the composition

$$\mathscr{B} \xrightarrow{\phi} \beta_* \mathscr{O}_{\mathcal{S}pec\, \mathscr{B}} \longrightarrow \beta_* \gamma_* \mathscr{O}_W = \mu_* \mathscr{O}_W.$$

The $\mathcal{S}pec$ construction gives an important way to understand affine morphisms. Note that $\mathcal{S}pec\,\mathscr{B} \to X$ is an affine morphism. The "converse" is also true:

17.1.D. EXERCISE. Show that if $\mu : Z \to X$ is an affine morphism, then we have a natural isomorphism $Z \cong \mathcal{S}pec\,\mu_* \mathscr{O}_Z$ of X-schemes.

Hence we can recover any affine morphism using the $\mathcal{S}pec$ construction. More precisely, a morphism is affine if and only if it is of the form $\mathcal{S}pec\,\mathscr{B} \to X$.

17.1.E. EXERCISE. Suppose $\mu : \mathcal{S}pec\,\mathscr{B} \to X$ is a morphism. Show that the category of quasicoherent sheaves on $\mathcal{S}pec\,\mathscr{B}$ is equivalent to the category of quasicoherent sheaves on X with the structure of $\mathscr{B}$-modules (quasicoherent $\mathscr{B}$-modules on X).

This is useful if X is quite simple but $\mathcal{S}pec\,\mathscr{B}$ is complicated. We will use this before long when $X \cong \mathbb{P}^1$, and $\mathcal{S}pec\,\mathscr{B}$ is a more complicated curve.

17.1.F. EXERCISE ($\mathcal{S}pec$ BEHAVES WELL WITH RESPECT TO BASE CHANGE). Suppose $\mu : Z \to X$ is any morphism, and $\mathscr{B}$ is a quasicoherent sheaf of algebras on X. Show that there is a natural isomorphism $Z \times_X \mathcal{S}pec\,\mathscr{B} \cong \mathcal{S}pec\,\mu^*\mathscr{B}$.

17.1.4. *Definition.* An important example of the $\mathcal{S}pec$ construction is the **total space of a finite rank locally free sheaf** $\mathscr{F}$, which we define to be $\mathcal{S}pec(\operatorname{Sym}^\bullet \mathscr{F}^\vee)$.

17.1.G. EXERCISE. Suppose $\mathscr{F}$ is a locally free sheaf of rank n. Show that the total space of $\mathscr{F}$ is a rank n *vector bundle*, i.e., that given any point $p \in X$, there is a neighborhood $p \in U \subset X$ such that

$$\mathcal{S}pec\left(\operatorname{Sym}^\bullet \mathscr{F}^\vee|_U\right) \cong \mathbb{A}^n_U.$$

Show that $\mathscr{F}$ is isomorphic to the sheaf of sections of the total space $\mathcal{S}pec(\operatorname{Sym}^\bullet \mathscr{F}^\vee)$. (Possible hint: use transition functions.) For this reason, the total space is also called the **vector bundle associated to a locally free sheaf** $\mathscr{F}$. (Caution: some authors, e.g. [**Stacks**, tag 01M2], call $\mathcal{S}pec(\operatorname{Sym}^\bullet \mathscr{F})$, the *dual* of this vector bundle, the vector bundle associated to $\mathscr{F}$.)

In particular, if $\mathscr{F} = \mathscr{O}_X^{\oplus n}$, then $\mathcal{S}pec(\operatorname{Sym}^\bullet \mathscr{F}^\vee)$ is called $\mathbb{A}^n_X$, generalizing our earlier notions of $\mathbb{A}^n_A$. As the notion of free sheaf behaves well with respect to base change, so does the notion of $\mathbb{A}^n_X$, i.e., given $X \to Y$, $\mathbb{A}^n_Y \times_Y X \cong \mathbb{A}^n_X$. (Aside: you may notice that the construction $\mathcal{S}pec\,\operatorname{Sym}^\bullet$ can be applied to any coherent sheaf $\mathscr{F}$ (without dualizing, i.e., $\mathcal{S}pec(\operatorname{Sym}^\bullet \mathscr{F})$). This is sometimes called the *abelian cone* associated to $\mathscr{F}$. This concept can be useful, but we won't need it.)

17.1.H. EXERCISE (THE TAUTOLOGICAL BUNDLE ON $\mathbb{P}^n$ IS $\mathscr{O}(-1)$). Suppose k is a field. Define the subset $X \subset \mathbb{A}^{n+1}_k \times \mathbb{P}^n_k$ corresponding to "points of $\mathbb{A}^{n+1}_k$ on the corresponding line of $\mathbb{P}^n_k$", so that the fiber of the map $\pi : X \to \mathbb{P}^n$ corresponding to a point $l = [x_0, \cdots, x_n]$ is the line in $\mathbb{A}^{n+1}_k$ corresponding to l, i.e., the scalar multiples of $(x_0, \dots, x_n)$. Show that $\pi : X \to \mathbb{P}^n_k$ is (the line bundle corresponding to) the invertible sheaf $\mathscr{O}(-1)$. (Possible hint: work first over the usual affine open sets of $\mathbb{P}^n_k$, and figure out transition functions.) For this reason, $\mathscr{O}(-1)$ is often called the **tautological bundle** of $\mathbb{P}^n_k$ (even over an arbitrary base, not just a field). (Side Remark: The projection $X \to \mathbb{A}^{n+1}_k$ is the blow-up of $\mathbb{A}^{n+1}_k$ at the "origin", see Exercise 9.3.F.)

17.2 Relative Proj of a sheaf of graded algebras

In parallel with the relative version $\mathcal{S}pec$ of Spec, we define a relative version of Proj, denoted $\mathcal{P}roj$ (called "relative Proj" or "sheaf Proj"), of a quasicoherent graded sheaf of algebras (satisfying some hypotheses) on a scheme X. We have already done the case where the base X is affine, in §4.5.7, using the regular Proj construction over a ring A. The elegant way to proceed would be to state the right universal property, and then use this cleverly to glue together the constructions over each affine, just as we did in the constructions of fibered product, normalization, and $\mathcal{S}pec$. But because graded rings and graded modules make everything confusing, we do not do this. Instead we guiltily take a more pedestrian approach. (But the universal property can be made to work, see [**Stacks**, tag 01O0].)

17.2.A. EXERCISE (Proj COMMUTES WITH AFFINE BASE CHANGE). Suppose $A \to B$ is map of rings, and $S_\bullet$ is a $\mathbb{Z}^{\geq 0}$-graded ring.
(a) Give a canonical isomorphism

$$\alpha : \operatorname{Proj}_B(S_\bullet \otimes_A B) \xrightarrow{\sim} (\operatorname{Proj}_A S_\bullet) \times_{\operatorname{Spec} A} \operatorname{Spec} B \tag{17.2.0.1}$$

(b) (easy) Suppose X is a projective A-scheme (§4.5.9). Show that $X \times_{\operatorname{Spec} A} \operatorname{Spec} B$ is a projective B-scheme.
(c) Suppose $S_\bullet$ is generated in degree 1, so $\mathscr{O}_{\operatorname{Proj}_A S_\bullet}(1)$ is an invertible sheaf (§15.2). Clearly $S_\bullet \otimes_A B$ is generated in degree 1 as a B-algebra. Describe an isomorphism

$$\mathscr{O}_{\operatorname{Proj}_B(S_\bullet \otimes_A B)}(1) \cong \alpha^* \gamma^* \mathscr{O}_{\operatorname{Proj}_A S_\bullet}(1),$$

where γ is the top morphism in the pullback diagram

$$\begin{tikzcd} (\operatorname{Proj}_A S_\bullet) \times_{\operatorname{Spec} A} \operatorname{Spec} B \arrow[r, "\gamma"] \arrow[d] & \operatorname{Proj}_A S_\bullet \arrow[d] \\ \operatorname{Spec} B \arrow[r] & \operatorname{Spec} A \end{tikzcd}$$

Possible hint: transition functions.

We now give a general means of constructing schemes over X (from [**Stacks**, tag 01LH]), if we know what they should be over any affine open set, and how these behave under open embeddings of one affine open set into another.

17.2.B. EXERCISE. Suppose we are given a scheme X, and the following data:

(i) For each affine open subset $U \subset X$, we are given some morphism $\pi_U : Z_U \to U$ (a "scheme over U").
(ii) For each (open) inclusion of affine open subsets $V \subset U \subset X$, we are given an open embedding $\rho^U_V : Z_V \hookrightarrow Z_U$.

Assume this data satisfies:

(a) for each $V \subset U \subset X$, ρ^U_V induces an isomorphism $Z_V \to \pi_U^{-1}(V)$ of schemes over V, and
(b) whenever $W \subset V \subset U \subset X$ are three nested affine open subsets, $\rho^U_W = \rho^U_V \circ \rho^V_W$.

Show that there exists an X-scheme $\pi : Z \to X$, and isomorphisms $i_U : \pi^{-1}(U) \to Z_U$ for each affine open set U, such that for nested affine open sets $V \subset U$, ρ^U_V agrees with the composition

$$Z_V \xrightarrow{i_V^{-1}} \pi^{-1}(V) \hookrightarrow \pi^{-1}(U) \xrightarrow{i_U} Z_U$$

Hint (cf. Exercise 4.4.A): construct Z first as a set, then as a topological space, then as a scheme. (Your construction will be independent of choices. Your solution will work in more general situations, for example when the category of schemes is replaced by ringed spaces, and when the affine open subsets are replaced by any base of the topology.)

17.2.C. IMPORTANT EXERCISE AND DEFINITION (RELATIVE $\mathcal{P}roj$). Suppose $\mathscr{S}_\bullet = \oplus_{n \geq 0} \mathscr{S}_n$ is a quasicoherent sheaf of $\mathbb{Z}^{\geq 0}$-graded algebras on a scheme X. Over each affine open subset $\operatorname{Spec} A \cong U \subset X$, we have an U-scheme $\operatorname{Proj}_A \mathscr{S}_\bullet(U) \to U$. Show that these can be glued together to form an X-scheme, which we call $\mathcal{P}roj_X \mathscr{S}_\bullet$; we have a "structure morphism" $\beta : \mathcal{P}roj_X \mathscr{S}_\bullet \to X$.

By the construction of Exercise 17.2.B, the preimage over any affine open set can be computed using the original Proj construction. (You may enjoy going back and giving constructions of X^{red}, the normalization of X, and $\mathcal{S}pec$ of a quasicoherent sheaf of $\mathscr{O}$-algebras using this idea. But there is a moral price to be paid by giving up the universal property.)

17.2.1. Ongoing (reasonable) hypotheses on $\mathscr{S}_\bullet$: "finite generation in degree 1". The Proj construction is most useful when applied to an A-algebra $S_\bullet$ satisfying some reasonable hypotheses (§4.5.6), notably when $S_\bullet$ is a finitely generated $\mathbb{Z}^{\geq 0}$-graded A-algebra, and ideally if it is generated in degree 1. For this reason, in the rest of the book, we will enforce these assumptions on $\mathscr{S}_\bullet$, once we make sense of them for quasicoherent sheaves of algebras. (If you later need to relax these hypotheses — for example, to keep the finite generation hypothesis but remove the "generation in degree 1" hypothesis — it will not be too difficult.) Precisely, **we now always require that (i) $\mathscr{S}_\bullet$ is "generated in degree 1", and (ii) $\mathscr{S}_1$ is finite type.** The cleanest way to make condition (i) precise is to require the natural map

$$\operatorname{Sym}^\bullet_{\mathscr{O}_X} \mathscr{S}_1 \to \mathscr{S}_\bullet$$

to be surjective. Because the $\operatorname{Sym}^\bullet$ construction may be computed affine-locally (§13.5.3), we can check generation in degree 1 on any affine cover.

17.2.D. IMPORTANT EXERCISE: $\mathscr{O}(1)$ ON $\mathcal{P}roj\, \mathscr{S}_\bullet$. If $\mathscr{S}_\bullet$ is finitely generated in degree 1 (Hypotheses 17.2.1), construct an invertible sheaf $\mathscr{O}_{\mathcal{P}roj\, \mathscr{S}_\bullet}(1)$ on $\mathcal{P}roj\, \mathscr{S}_\bullet$ that "restricts to $\mathscr{O}_{\operatorname{Proj}_A \mathscr{S}_\bullet(\operatorname{Spec} A)}(1)$ over each affine open subset $\operatorname{Spec} A \subset X$".

17.2.E. EXERCISE ("$\mathcal{P}roj$ COMMUTES WITH BASE CHANGE"). Suppose $\mathscr{S}_\bullet$ is a quasicoherent sheaf of $\mathbb{Z}^{\geq 0}$-graded algebras on X. Let $\rho : Z \to X$ be any morphism. Give a natural isomorphism

$$(\mathcal{P}roj\, \rho^*\mathscr{S}_\bullet, \mathscr{O}_{\mathcal{P}roj\, \rho^*\mathscr{S}_\bullet}(1)) \cong (Z \times_X \mathcal{P}roj\, \mathscr{S}_\bullet, \psi^*\mathscr{O}_{\mathcal{P}roj\, \mathscr{S}_\bullet}(1))$$

where ψ is the "top" morphism in the base change diagram

$$\begin{tikzcd} Z \times_X \mathcal{P}roj\, \mathcal{S}_\bullet \arrow[r, "\psi"] \arrow[d] & \mathcal{P}roj\, \mathcal{S}_\bullet \arrow[d, "\beta"] \\ Z \arrow[r, "\rho"] & X. \end{tikzcd}$$

17.2.2. *Definition: π-very ample.* Suppose $\pi : X \to Y$ is proper. If $\mathscr{L}$ is an invertible sheaf on X, then we say that $\mathscr{L}$ is **very ample (with respect to π)**, or (awkwardly) **π-very ample** if we can write $X = \mathcal{P}roj_Y\, \mathcal{S}_\bullet$ with $\mathscr{L} \cong \mathscr{O}(1)$, where $\mathcal{S}_\bullet$ is a quasicoherent sheaf of algebras on Y satisfying Hypotheses 17.2.1 ("finite generation in degree 1"). (The notion of very ampleness can be extended to more general situations, see for example [**Stacks**, tag 01VM]. But this is of interest only to people with esoteric tastes.)

17.2.F. EXERCISE. Suppose $\mathcal{S}_\bullet$ is finitely generated in degree 1 (Hypotheses 17.2.1). Describe a map of graded quasicoherent sheaves $\phi : \mathcal{S}_\bullet \to \oplus_n \beta_* \mathscr{O}(n)$, which is locally an isomorphism in high degrees (given any point of X, there is a neighborhood of the point and an n_0, so that ϕ_n is an isomorphism for $n \geq n_0$). Hint: Exercise 15.4.C.

17.2.G. EXERCISE. Suppose $\mathscr{L}$ is an invertible sheaf on X, and $\mathcal{S}_\bullet$ is a quasicoherent sheaf of graded algebras on X generated in degree 1 (Hypotheses 17.2.1). Define $\mathcal{S}'_\bullet = \oplus_{n=0} (\mathcal{S}_n \otimes \mathscr{L}^{\otimes n})$. Then $\mathcal{S}'_\bullet$ has a natural algebra structure inherited from $\mathcal{S}_\bullet$; describe it. Give a natural isomorphism of "X-schemes with line bundles"

$$(\mathcal{P}roj\, \mathcal{S}'_\bullet, \mathscr{O}_{\mathcal{P}roj\, \mathcal{S}'_\bullet}(1)) \cong (\mathcal{P}roj\, \mathcal{S}_\bullet, \mathscr{O}_{\mathcal{P}roj\, \mathcal{S}_\bullet}(1) \otimes \beta^* \mathscr{L}),$$

where $\beta : \mathcal{P}roj\, \mathcal{S}_\bullet \to X$ is the structure morphism. In other words, informally speaking, the $\mathcal{P}roj$ is the same, but the $\mathscr{O}(1)$ is twisted by $\mathscr{L}$.

17.2.3. *Definition.* If $\mathscr{F}$ is a finite type quasicoherent sheaf on X, then $\mathcal{P}roj(\mathrm{Sym}^\bullet \mathscr{F})$ is called its **projectivization**, and is denoted $\mathbb{P}\mathscr{F}$. You can check that this construction behaves well with respect to base change. Define $\mathbb{P}^n_X := \mathbb{P}(\mathscr{O}_X^{\oplus(n+1)})$. (Then $\mathbb{P}^n_{\operatorname{Spec} A}$ agrees with our earlier definition of $\mathbb{P}^n_A$, cf. Exercise 4.5.N, and $\mathbb{P}^n_X$ agrees with our earlier usage, see for example the proof of Theorem 10.3.5.) More generally, if $\mathscr{F}$ is locally of free of rank $n+1$, then $\mathbb{P}\mathscr{F}$ is a **projective bundle** or **$\mathbb{P}^n$-bundle** over X. By Exercise 17.2.G, if $\mathscr{V}$ is a finite rank locally free sheaf on X, there is a canonical isomorphism $\mathbb{P}\mathscr{V} \cong \mathbb{P}(\mathscr{L} \otimes \mathscr{V})$.

17.2.4. *Example: ruled surfaces.* If C is a regular curve and $\mathscr{F}$ is locally free of rank 2, then $\mathbb{P}\mathscr{F}$ is called a **ruled surface** over C. If C is further isomorphic to $\mathbb{P}^1$, $\mathbb{P}\mathscr{F}$ is called a **Hirzebruch surface**. All vector bundles on $\mathbb{P}^1$ split as a direct sum of line bundles (see §18.5.5 for a proof), so each Hirzebruch surface is of the form $\mathbb{P}(\mathscr{O}(n_1) \oplus \mathscr{O}(n_2))$. By Exercise 17.2.G, this depends only on $n_2 - n_1$. The Hirzebruch surface $\mathbb{P}(\mathscr{O} \oplus \mathscr{O}(n))$ ($n \geq 0$) is often denoted $\mathbb{F}_n$. We will discuss the Hirzebruch surfaces in greater length in §20.2.8. We will see that the $\mathbb{F}_n$ are all distinct in Exercise 20.2.Q.

17.2.H. EXERCISE. If $\mathscr{S}_\bullet$ is finitely generated in degree 1 (Hypotheses 17.2.1), describe a canonical closed embedding

$$\begin{tikzcd} \mathcal{P}roj\, \mathscr{S}_\bullet \arrow[rr, hook, "i"] \arrow[dr, "\beta"'] & & \mathbb{P}\mathscr{S}_1 \arrow[dl] \\ & X & \end{tikzcd}$$

and an isomorphism $\mathscr{O}_{\mathcal{P}roj\, \mathscr{S}_\bullet}(1) \cong i^* \mathscr{O}_{\mathbb{P}\mathscr{S}_1}(1)$ arising from the surjection $\mathrm{Sym}^\bullet \mathscr{S}_1 \to \mathscr{S}_\bullet$.

17.2.I. EXERCISE. Suppose $\mathscr{F}$ is a locally free sheaf of rank $n+1$ on X. Exhibit a bijection between the set of sections $s : X \to \mathbb{P}\mathscr{F}$ of $\mathbb{P}\mathscr{F} \to X$ and the set of surjective homomorphisms $\mathscr{F} \to \mathscr{L} \to 0$ of $\mathscr{F}$ onto invertible sheaves on X.

17.2.5. *Remark (the relative version of the projective and affine cone).* There is a natural morphism from $\mathcal{S}pec\, \mathscr{S}_\bullet$ minus the zero-section to $\mathcal{P}roj\, \mathscr{S}_\bullet$ (cf. Exercise 8.2.P). Just as $\operatorname{Proj} S_\bullet[T]$ contains a closed subscheme identified with $\operatorname{Proj} S_\bullet$ whose complement can be identified with $\operatorname{Spec} S_\bullet$ (Exercise 8.2.Q), $\mathcal{P}roj\, \mathscr{S}_\bullet[T]$ contains a closed subscheme identified with $\mathcal{P}roj\, \mathscr{S}_\bullet$ whose complement can be identified with $\mathcal{S}pec\, \mathscr{S}_\bullet$. You are welcome to think this through.

17.2.6. *Remark.* If you wish, you can describe (with proof) a universal property of $\mathcal{P}roj\, \mathscr{S}_\bullet$. (You may want to describe a universal property of Proj first.) I recommend against it — a universal property should make your life easier, not harder. One possible universal property is given in [**Stacks**, tag 01NS].

17.3 Projective morphisms

In §17.1, we reinterpreted affine morphisms: $X \to Y$ is an affine morphism if there is an isomorphism $X \cong \mathcal{S}pec\, \mathscr{B}$ of Y-schemes for some quasicoherent sheaf of algebras $\mathscr{B}$ on Y. We will *define* the notion of a projective morphism similarly.

You might think that because projectivity is such a classical notion, there should be some obvious definition, that is reasonably behaved. But this is not the case, and there are many possible variant definitions of projective (see [**Stacks**, tag 01W8]). All are imperfect, including the accepted definition we give here. Although projective morphisms are preserved by base change, we will manage to show that they are preserved by composition only when the target is quasicompact (Exercise 17.3.B), and we will only show that the notion is local on the base when we add the data of a line bundle, and even then only under locally Noetherian hypotheses (§17.3.4).

17.3.1. *Definition.* A morphism $\pi : X \to Y$ is **projective** if there is an isomorphism

$$\begin{tikzcd} X \arrow[rr, "\sim"] \arrow[dr, "\pi"'] & & \mathcal{P}roj\, \mathscr{S}_\bullet \arrow[dl] \\ & Y & \end{tikzcd}$$

for a quasicoherent sheaf of algebras $\mathscr{S}_\bullet$ on Y (satisfying "finite generation in degree 1", Hypotheses 17.2.1). We say X is a **projective Y-scheme**, or X is **projective over** Y. This generalizes the notion of a projective A-scheme.

17.3.2. *Warnings.* First, notice that $\mathscr{O}(1)$, an important part of the definition of $\mathcal{P}roj$, is not mentioned. (I would prefer that it be part of the definition, but this isn't accepted practice.) As a result, the notion of affine morphism is affine-local on the target, but the notion of projectivity or a morphism is not clearly affine-local on the target. (In Noetherian circumstances, with the additional data of the invertible sheaf $\mathscr{O}(1)$, it is, as we will see in §17.3.4. We will also later see an example showing that the property of being projective is *not* local, §24.7.7.)

Second, [**Ha1**, p. 103] gives a different definition of projective morphism; we follow the more general definition of Grothendieck. These definitions turn out to be the same in nice circumstances. (But finite morphisms are not always projective in the sense of [**Ha1**], while they *are* projective in our sense.)

17.3.A. EXERCISE.
(a) *(a useful characterization of projective morphisms)* Suppose $\mathscr{L}$ is an invertible sheaf on X, and $\pi : X \to Y$ is a morphism. Show that π is projective, with $\mathscr{O}(1) \cong \mathscr{L}$, if and only if there exist a finite type quasicoherent sheaf $\mathscr{S}_1$ on Y, a closed embedding $i : X \hookrightarrow \mathbb{P}\mathscr{S}_1$ (over Y, i.e., commuting with the maps to Y), and an isomorphism $i^*\mathscr{O}_{\mathbb{P}\mathscr{S}_1}(1) \cong \mathscr{L}$. Hint: Exercise 17.2.H.
(b) If furthermore Y admits an ample line bundle $\mathscr{M}$, show that π is projective if and only if there exists a closed embedding $i : X \hookrightarrow \mathbb{P}^n_Y$ (over Y) for some n. (If you wish, assume Y is proper over $\operatorname{Spec} A$, so you can avoid the starred section §16.6.5.) Hint: the harder direction is the forward implication. Use the finite type quasicoherent sheaf $\mathscr{S}_1$ from (a). Tensor $\mathscr{S}_1$ with a high enough power of $\mathscr{M}$ so that it is finitely globally generated (Theorem 16.6.6, or Theorem 16.6.2 in the proper setting), to obtain a surjection $\mathscr{O}_Y^{\oplus(n+1)} \twoheadrightarrow \mathscr{S}_1 \otimes \mathscr{M}^{\otimes N}$. Then use Exercise 17.2.G.

17.3.3. *Definition: Quasiprojective morphisms.* In analogy with projective and quasiprojective A-schemes (§4.5.9), one may define quasiprojective morphisms. *If Y is quasicompact*, we say that $\pi : X \to Y$ is **quasiprojective** if π can be expressed as a quasicompact open embedding into a scheme projective over Y. This is not a great notion, and we will not use it. (The general definition of quasiprojective is slightly delicate — see [**Gr-EGA**, II.5.3] — but we won't need it.)

17.3.4. Properties of projective morphisms.

We start to establish a number of properties of projective morphisms. First, the property of a morphism being projective is clearly preserved by base change, as the $\mathcal{P}roj$ construction behaves well with respect to base change (Exercise 17.2.E). Also, projective morphisms are proper: properness is local on the target (Theorem 10.3.4(b)), and we saw earlier that projective A-schemes are proper over A (Theorem 10.3.5). In particular (by definition of properness), projective morphisms are separated, finite type, and universally closed.

Exercise 17.3.G (in a future optional section) implies that if $\pi : X \to Y$ is a proper morphism of locally Noetherian schemes, and $\mathscr{L}$ is an invertible sheaf on

X, the question of whether π is a projective morphism with $\mathscr{L}$ as $\mathscr{O}(1)$ is local on Y.

17.3.B. IMPORTANT CHALLENGING EXERCISE (THE COMPOSITION OF PROJECTIVE MORPHISMS IS PROJECTIVE, IF THE FINAL TARGET IS QUASICOMPACT). Suppose $\pi : X \to Y$ and $\rho : Y \to Z$ are projective morphisms, and Z is quasicompact. Show that $\pi \circ \rho$ is projective. Hint: the criterion for projectivity given in Exercise 17.3.A(a) will be useful. (i) Deal first with the case where Z is affine. Build the following commutative diagram, thereby finding a closed embedding $X \hookrightarrow \mathbb{P}\mathscr{F}^{\oplus n}$ over Z. In this diagram, all inclusions are closed embeddings, and all script fonts refer to finite type quasicoherent sheaves.

$$\begin{tikzcd}
X \arrow[r, hook] \arrow[drr, "\pi"'] & \mathbb{P}\mathscr{E} \arrow[r, hook, "(\dagger)"] \arrow[dr] & \mathbb{P}^{n-1}_Z \times_Z Y \arrow[r, hook] \arrow[d] & \mathbb{P}^{n-1}_Z \times_Z \mathbb{P}\mathscr{F} \arrow[r, hook, "\text{Segre}", "\text{cf. Ex. 9.6.D}"'] \arrow[d] & \mathbb{P}(\mathscr{F}^{\oplus n}) \arrow[ddll, bend left] \\
& & Y \arrow[r, hook] \arrow[d, "\rho"] & \mathbb{P}\mathscr{F} \arrow[dl] & \\
& & Z & &
\end{tikzcd}$$

Construct the closed embedding $(\dagger)$ as follows. Suppose $\mathscr{M}$ is the very ample line bundle on Y over Z. Then $\mathscr{M}$ is ample, and so by Theorem 16.6.2, for $m \gg 0$, $\mathscr{E} \otimes \mathscr{M}^{\otimes m}$ is generated by a finite number of global sections. Suppose $\mathscr{O}_Y^{\oplus n} \twoheadrightarrow \mathscr{E} \otimes \mathscr{M}^{\otimes m}$ is the corresponding surjection. This induces a closed embedding $\mathbb{P}(\mathscr{E} \otimes \mathscr{M}^{\otimes m}) \hookrightarrow \mathbb{P}^{n-1}_Y$. But $\mathbb{P}(\mathscr{E} \otimes \mathscr{M}^{\otimes m}) \cong \mathbb{P}\mathscr{E}$ (Exercise 17.2.G), and $\mathbb{P}^{n-1}_Y = \mathbb{P}^{n-1}_Z \times_Z Y$. (ii) Unwind this diagram to show that (for Z affine) if $\mathscr{L}$ is π-very ample and $\mathscr{M}$ is ρ-very ample, then for $m \gg 0$, $\mathscr{L} \otimes \mathscr{M}^{\otimes m}$ is $(\rho \circ \pi)$-very ample. Then deal with the general case by covering Z with a finite number of affines.

17.3.5. *Caution: Consequences of projectivity not being "reasonable" in the sense of §7.1.1.* Because the property of being projective is preserved by base change (§17.3.4), and composition *to quasicompact targets* (Exercise 17.3.B), the property of being projective is "usually" preserved by products (Exercise 9.4.F): if $\pi : X \to Y$ and $\pi' : X' \to Y$ are projective, then so is $\pi \times \pi' : X \times X' \to Y \times Y'$, so long as $Y \times Y'$ is quasicompact. Also, if you follow through the proof of the Cancellation Theorem 10.1.19 for properties of morphisms, you will see that if $\pi : X \to Y$ is a morphism, $\rho : Y \to Z$ is a separated morphism (so the diagonal δ_ρ is a closed embedding and hence projective), and $\rho \circ \pi$ is projective, *and* Y *is quasicompact*, then π is projective.

17.3.C. EXERCISE. Show that a morphism (over $\operatorname{Spec} k$) from a projective k-scheme to a quasicompact separated k-scheme is always projective. (Hint: the Cancellation Theorem 10.1.19 for projective morphisms; see also Caution 17.3.5.)

17.3.6. Finite morphisms are projective.

17.3.D. IMPORTANT EXERCISE: FINITE MORPHISMS ARE PROJECTIVE (CF. EXERCISE 7.3.J). Show that finite morphisms are projective as follows. Suppose $Z \to X$ is finite, so that $Z \cong \mathcal{S}pec\, \mathscr{B}$ where $\mathscr{B}$ is a finite type quasicoherent sheaf on X.

Describe a sheaf of graded algebras $\mathscr{S}_\bullet$ where $\mathscr{S}_0 \cong \mathscr{O}_X$ and $\mathscr{S}_n \cong \mathscr{B}$ for $n > 0$. Describe an X-isomorphism $Z \cong \mathcal{P}roj\, \mathscr{S}_\bullet$.

In particular, closed embeddings are projective. We have the sequence of implications for morphisms

$$\text{closed embedding} \Longrightarrow \text{finite} \Longrightarrow \text{projective} \Longrightarrow \text{proper}.$$

We know that finite morphisms are projective (Exercise 17.3.D), and have finite fibers (Exercise 7.3.K). We will show the converse in Theorem 18.1.9, and state the extension to proper morphisms immediately after.

17.3.7. $\star\star$ **Global generation and (very) ampleness in the relative setting.**
We extend the discussion of §15.3 to the relative setting, in order to give ourselves the language of relatively base-point-freeness. We won't use this discussion, so on a first reading you should jump directly to §17.4. But these ideas come up repeatedly in the research literature.

Suppose $\pi : X \to Y$ is a quasicompact quasiseparated morphism. In $\mathscr{F}$ is a quasicoherent sheaf on X, we say that $\mathscr{F}$ is **relatively globally generated** or **globally generated with respect to** π if the natural map of quasicoherent sheaves $\pi^*\pi_*\mathscr{F} \to \mathscr{F}$ is surjective. (Quasicompactness and quasiseparatedness are needed ensure that $\pi_*\mathscr{F}$ is a quasicoherent sheaf, Exercise 13.3.F). But these hypotheses are not very restrictive. Global generation is most useful only in the quasicompact setting, and most people won't be bothered by quasiseparated hypotheses. Unimportant aside: these hypotheses can be relaxed considerably. If $\pi : X \to Y$ is a morphism of *locally ringed spaces* — not necessarily schemes — with no other hypotheses, and $\mathscr{F}$ is a quasicoherent sheaf on X, then we say that $\mathscr{F}$ is **relatively globally generated** or **globally generated with respect to** π if the natural map $\pi^*\pi_*\mathscr{F} \to \mathscr{F}$ of $\mathscr{O}_X$-modules is surjective.)

Thanks to our hypotheses, as the natural map $\pi^*\pi_*\mathscr{F} \to \mathscr{F}$ is a morphism of quasicoherent sheaves, the condition of being relatively globally generated is affine-local on Y.

Suppose now that $\mathscr{L}$ is a locally free sheaf on X, and $\pi : X \to Y$ is a morphism. We say that $\mathscr{L}$ is **relatively base-point-free** or **base-point-free with respect to** π if it is relatively globally generated.

17.3.E. EXERCISE. Suppose $\mathscr{L}$ is a finite rank locally free sheaf on $\mathscr{L}$, $\pi : X \to Y$ is a quasicompact separated morphism, and $\pi_*\mathscr{L}$ is finite type on Y. (We will later show in Theorem 18.9.1 that this latter statement is true if π is proper and Y is Noetherian. This is much easier if π is projective, see Theorem 18.8.1. We could work hard and prove it now, but it isn't worth the trouble.) Describe a canonical morphism $\psi : X \to \mathbb{P}\mathscr{L}$. (Possible hint: this generalizes the fact that base-point-free line bundles give maps to projective space, so generalize that argument, see §15.3.5.)

We say that $\mathscr{L}$ is **relatively ample** or **π-ample** or **relatively ample with respect to** π if for every affine open subset $\operatorname{Spec} B$ of Y, $\mathscr{L}|_{\pi^{-1}(\operatorname{Spec} B)}$ is ample on $\pi^{-1}(\operatorname{Spec} B)$ over B, or equivalently (by §16.6.5). $\mathscr{L}|_{\pi^{-1}(\operatorname{Spec} B)}$ is (absolutely) ample on $\pi^{-1}(\operatorname{Spec} B)$. By the discussion in §16.6.5, if $\mathscr{L}$ is ample then π is necessarily quasicompact, and (by Theorem 16.6.6) separated; if π is affine, then all invertible sheaves are ample; and if π is projective, then the corresponding $\mathscr{O}(1)$ is ample.

By Exercise 16.6.J, $\mathscr{L}$ is π-ample if and only if $\mathscr{L}^{\otimes n}$ is π-ample, and if $Z \hookrightarrow X$ is a closed embedding, then $\mathscr{L}|_Z$ is ample over Y.

From Theorem 16.6.6(d) implies that we have a natural open embedding $X \to \mathcal{P}roj_Y \oplus f_*\mathscr{L}^{\otimes d}$. (Do you see what this map is? Also, be careful: $\oplus f_*\mathscr{L}^{\otimes d}$ need not be a finitely generated graded sheaf of algebras, so we are using the $\mathcal{P}roj$ construction where one of the usual hypotheses doesn't hold.)

The notions of relative global generation and relative ampleness are most useful in the proper setting, because of Theorem 16.6.2.

17.3.8. Many statements of §15.3 carry over without change. For example, we have the following. Suppose $\pi : X \to Y$ is proper, $\mathscr{F}$ and $\mathscr{G}$ are quasicoherent sheaves on X, and $\mathscr{L}$ and $\mathscr{M}$ are invertible sheaves on X. If π is affine, then $\mathscr{F}$ is relatively globally generated (from Easy Exercise 15.3.A). If $\mathscr{F}$ and $\mathscr{G}$ are relatively globally generated, so is $\mathscr{F} \otimes \mathscr{G}$ (Easy Exercise 15.3.B). If $\mathscr{L}$ is π-very ample (Definition 17.2.2), then it is π-base-point-free (Easy Exercise 16.6.B). If $\mathscr{L}$ is π-very ample, and $\mathscr{M}$ is π-base-point-free (if for example it is π-very ample), then $\mathscr{L} \otimes \mathscr{M}$ is π-very ample (Exercise 16.6.C). Exercise 16.6.G extends immediately to show that if

$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "\tau"'] & & Y \arrow[dl, "\rho"] \\ & S & \end{tikzcd}$$

is a finite morphism of S-schemes, and if $\mathscr{L}$ is a ρ-ample invertible sheaf on Y, then $\pi^*\mathscr{L}$ is τ-ample.

By the nature of the statements, some of the statements of §15.3 require quasicompactness hypotheses on Y, or other patches. For example:

17.3.9. Theorem. — *Suppose $\pi : X \to Y$ is proper, $\mathscr{L}$ is an invertible sheaf on X, and Y is quasicompact. The following are equivalent.*

(a) *For some $N > 0$, $\mathscr{L}^{\otimes N}$ is π-very ample.*
(a') *For all $n \gg 0$, $\mathscr{L}^{\otimes n}$ is π-very ample.*
(b) *For all finite type quasicoherent sheaves $\mathscr{F}$, there is an n_0 such that for $n \geq n_0$, $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is relatively globally generated.*
(c) *The invertible sheaf $\mathscr{L}$ is π-ample.*

17.3.F. EXERCISE. Prove Theorem 17.3.9 using Theorem 16.6.2. (Unimportant remark: The proof given of Theorem 16.6.2 used Noetherian hypotheses, but as stated there, they can be removed.)

After doing the above Exercise, it will be clear how to adjust the statement of Theorem 17.3.9 if you need to remove the quasicompactness assumption on Y.

17.3.G. EXERCISE (A USEFUL EQUIVALENT DEFINITION OF VERY AMPLENESS UNDER NOETHERIAN HYPOTHESES). Suppose $\pi : X \to Y$ is a proper morphism, Y is locally Noetherian (hence X is too, as π is finite type), and $\mathscr{L}$ is an invertible sheaf on X. Suppose that you know that in this situation $\pi_*\mathscr{L}$ is finite type. (We will later show this, as described in Exercise 17.3.E.) Show that $\mathscr{L}$ is very ample if and only if (i) $\mathscr{L}$ is relatively base-point-free, and (ii) the canonical Y-morphism

$i : X \to \mathbb{P}(\pi_* \mathscr{L})$ of Exercise 17.3.E is a closed embedding. Conclude that the notion of relative very ampleness is affine-local on Y (it may be checked on *any* affine cover Y), if Y is locally Noetherian and π is proper.

As a consequence, Theorem 17.3.9 implies the notion of relative ampleness is affine-local on Y (if π is proper and Y is locally Noetherian).

17.3.10. $\star\star$ *Ample vector bundles.* The notion of an **ample vector bundle** is useful in some parts of the literature, so we define it, although we won't use the notion. A locally free sheaf $\mathscr{E}$ on a proper A-scheme X is **ample** if $\mathscr{O}_{\mathbb{P}\mathscr{E}/X}(1)$ is an ample invertible sheaf. In particular, using Exercise 17.2.G, you can verify that an invertible sheaf is ample as a locally free sheaf (this definition) if and only if it is ample as an invertible sheaf (Definition 16.6.1), preventing a notational crisis. (The proper hypotheses can be relaxed; it is included only because Definition 16.6.1 of ampleness is only for proper schemes.)

17.3.11. $\star\star$ **Quasiaffine morphisms.**

Because we have introduced quasiprojective morphisms (Definition 17.3.3), we briefly introduce quasiaffine morphisms (and quasiaffine schemes), as some readers may have cause to use them. Many of these ideas could have been introduced long before, but because we will never use them, we deal with them all at once.

A scheme X is **quasiaffine** if it admits a quasicompact open embedding into an affine scheme. This implies that X is quasicompact and separated. Note that if X is Noetherian (the most relevant case for most people), then any open embedding is of course automatically quasicompact.

17.3.H. EXERCISE. Show that X is quasiaffine if and only if the canonical map $X \to \operatorname{Spec} \Gamma(X, \mathscr{O}_X)$ (defined in Exercise 6.3.F and the paragraph following it) is a quasicompact open embedding. Thus a quasiaffine scheme comes with a *canonical* quasicompact open embedding into an affine scheme. Hint: Let $A = \Gamma(X, \mathscr{O}_X)$ for convenience. Suppose $X \to \operatorname{Spec} R$ is a quasicompact open embedding. We wish to show that $X \to \operatorname{Spec} A$ is a quasicompact open embedding. Factor $X \to \operatorname{Spec} R$ through $X \to \operatorname{Spec} A \to \operatorname{Spec} R$. Show that $X \to \operatorname{Spec} A$ is an open embedding in a neighborhood of any chosen point $p \in X$, as follows. Choose $r \in R$ such that $p \subset D(r) \subset X$. Notice that if $X_r = \{q \in X : r(q) \neq 0\}$, then $\Gamma(X_r, \mathscr{O}_X) = \Gamma(X, \mathscr{O}_X)_r$ by Exercise 13.3.H, using the fact that X is quasicompact and quasiseparated. Use this to show that the map $X_r \to \operatorname{Spec} A_r$ is an isomorphism.

It is not hard to show that X is quasiaffine if and only if $\mathscr{O}_X$ is ample, but we won't use this fact.

A morphism $\pi : X \to Y$ is **quasiaffine** if the inverse image of every affine open subset of Y is a quasiaffine scheme. By Exercise 17.3.H, this is equivalent to π being quasicompact and separated, and the natural map $X \to \mathcal{S}pec\, \pi_* \mathscr{O}_X$ being a quasicompact open embedding. This implies that the notion of quasiaffineness is local on the target (may be checked on an open cover), and also affine-local on a target (one may choose an affine cover, and check that the preimages of these open sets are quasiaffine). Quasiaffine morphisms are preserved by base change: if a morphism $X \hookrightarrow Z$ over Y is a quasicompact open embedding into an affine Y-scheme, then for any $W \to Y$, $X \times_Y W \hookrightarrow Z \times_Y W$ is a quasicompact open embedding into an affine W-scheme. (Interestingly, Exercise 17.3.H is *not* the right tool to use to show this base change property.)

One may readily check that quasiaffine morphisms are preserved by composition, [**Stacks**, tag 01SN]. Thus quasicompact locally closed embeddings are quasiaffine. If X is affine, then $X \to Y$ is quasiaffine if and only if it is quasicompact (as the preimage of any affine open subset of Y is an open subset of an affine scheme, namely X). In particular, from the Cancellation Theorem 10.1.19 for quasicompact morphisms, any morphism from an affine scheme to a quasiseparated scheme is quasiaffine.

17.4 Applications to curves

We now apply what we have learned to curves.

17.4.1. Theorem (every integral curve has a birational model that is regular and projective). — *If C is an integral curve of finite type over a field k, then there exists a regular projective k-curve C' birational to C.*

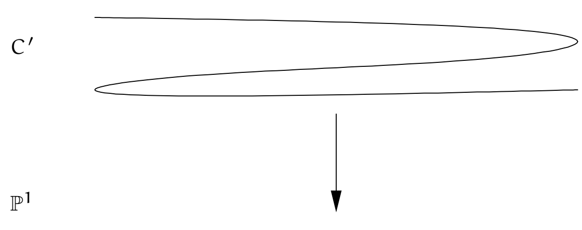


FIGURE 17.1. Constructing a projective regular model of a curve C over k via a finite cover of $\mathbb{P}^1$

Proof. We can assume C is affine. By the Noether Normalization Lemma 11.2.4, we can find some $x \in K(C) \setminus k$ with $K(C)/k(x)$ a finite field extension. By identifying a standard open of $\mathbb{P}^1_k$ with $\operatorname{Spec} k[x]$, and taking the normalization of $\mathbb{P}^1$ in the function field of $K(C)$ (Definition 9.7.I), we obtain a finite morphisms $C' \to \mathbb{P}^1$, where C' is a curve ($\dim C' = \dim \mathbb{P}^1$ by Exercise 11.1.F), and regular (it is reduced hence regular at the generic point, and regular at the closed points by the main theorem on discrete valuation rings in §12.5). Also, C' is birational to C as they have isomorphic function fields (Exercise 6.5.D).

Finally, $C' \to \mathbb{P}^1_k$ is finite (Exercise 9.7.M) hence projective (Exercise 17.3.D), and $\mathbb{P}^1_k \to \operatorname{Spec} k$ is projective, so as composition of projective morphisms (to a quasicompact target) are projective (Exercise 17.3.B), $C' \to \operatorname{Spec} k$ is projective. $\square$

17.4.2. Theorem. — *If C is an irreducible regular curve, finite type over a field k, then there is an open embedding $C \hookrightarrow C'$ into some projective regular curve C' (over k).*

Proof. We first prove the result in the case where C is affine. Then we have a closed embedding $C \hookrightarrow \mathbb{A}^n$, and we consider $\mathbb{A}^n$ as a standard open subset of $\mathbb{P}^n$. Taking the scheme-theoretic closure of C in $\mathbb{P}^n$, we obtain a projective integral curve $\overline{C}$, containing C as an open subset. The normalization $\widetilde{\overline{C}}$ of $\overline{C}$ is a finite morphism (finiteness of integral closure, Theorem 9.7.3(b)), so $\widetilde{\overline{C}}$ is Noetherian, and regular (as normal Noetherian dimension 1 rings are discrete valuation rings, §12.5). Moreover, by the universal property of normalization, normalization of $\overline{C}$ doesn't affect the normal open set C, so we have an open embedding $C \hookrightarrow \widetilde{\overline{C}}$. Finally, $\widetilde{\overline{C}} \to \overline{C}$ is finite hence projective, and $\overline{C} \to \operatorname{Spec} k$ is projective, so (by Exercise 17.3.B) $\widetilde{\overline{C}}$ is projective.

We next consider the case of general C. Let C_1 by any nonempty affine open subset of C. By the discussion in the previous paragraph, we have a regular projective compactification $\widetilde{\overline{C_1}}$. The Curve-to-Projective Extension Theorem 16.5.1 (applied successively to the finite number of points $C \setminus C_1$) implies that the morphism $C_1 \hookrightarrow \widetilde{\overline{C_1}}$ extends to a birational morphism $C \to \widetilde{\overline{C_1}}$. Because points of a regular curve are determined by their valuation (Exercise 12.7.B), this is an inclusion of sets. Because the topology on curves is stupid (cofinite), it expresses C as an open subset of $\widetilde{\overline{C}}$. But why is it an open embedding of schemes?

We show it is an open embedding near a point $p \in C$ as follows. Let C_2 be an affine neighborhood of p in C. We repeat the construction we used on C_1, to obtain the following diagram, with open embeddings marked.

$$\begin{tikzcd}
C_1 \arrow[dr, hook] \arrow[dd, hook] & & C_2 \arrow[dl, hook'] \arrow[dd, hook'] \\
& C \arrow[dl] \arrow[dr] & \\
\widetilde{\overline{C_1}} & & \widetilde{\overline{C_2}}
\end{tikzcd}$$

By the Curve-to-Projective Extension theorem 16.5.1, the map $C_1 \to \widetilde{\overline{C_2}}$ extends to $\pi_{12} : \widetilde{\overline{C_1}} \to \widetilde{\overline{C_2}}$, and we similarly have a morphism $\pi_{21} : \widetilde{\overline{C_2}} \to \widetilde{\overline{C_1}}$, extending $C_2 \to \widetilde{\overline{C_1}}$. The composition $\pi_{21} \circ \pi_{12}$ is the identity morphism (as it is the identity rational map, see Theorem 10.2.2). The same is true for $\pi_{12} \circ \pi_{21}$, so π_{12} and π_{21} are isomorphisms. The enhanced diagram

$$\begin{tikzcd}
C_1 \arrow[dr, hook] \arrow[dd, hook] & & C_2 \arrow[dl, hook'] \arrow[dd, hook'] \\
& C \arrow[dl] \arrow[dr] & \\
\widetilde{\overline{C_1}} \arrow[rr, leftrightarrow] & & \widetilde{\overline{C_2}}
\end{tikzcd}$$

commutes (by Theorem 10.2.2 again, implying that morphisms of reduced separated schemes are determined by their behavior on dense open sets). But $C_2 \to \widetilde{\overline{C_1}}$

is an open embedding (in particular, at p), so $C \to \widetilde{\widetilde{C_1}}$ is an open embedding there as well. $\square$

17.4.A. EXERCISE. Show that all regular proper curves over k are projective.

17.4.3. Theorem (various categories of curves are the same). — *The following categories are equivalent.*

(i) *irreducible regular projective curves over k, and surjective k-morphisms.*
(ii) *irreducible regular projective curves over k, and dominant k-morphisms.*
(iii) *irreducible regular projective curves over k, and dominant rational maps over k.*
(iv) *integral curves of finite type over k, and dominant rational maps over k.*
(v) *the opposite category of finitely generated fields of transcendence degree 1 over k, and k-homomorphisms.*

All morphisms and maps in the following discussion are assumed to be defined over k.

(Aside: The interested reader can tweak the proof below to show the following variation of the theorem: in (i)–(iv), consider only geometrically irreducible curves, and in (v), consider only fields K such that $K \cap k^s = k$ in $\overline{K}$. This variation allows us to exclude "weird" curves we may not want to consider. For example, if $k = \mathbb{R}$, then we are allowing curves such as $\mathbb{P}^1_{\mathbb{C}}$ which are not geometrically irreducible, as $\mathbb{P}^1_{\mathbb{C}} \times_{\mathbb{R}} \operatorname{Spec} \mathbb{C} \cong \mathbb{P}^1_{\mathbb{C}} \coprod \mathbb{P}^1_{\mathbb{C}}$.)

Proof. Any surjective morphism is a dominant morphism, and any dominant morphism is a dominant rational map, and each regular projective curve is a quasiprojective curve, so we have shown (informally speaking) how to get from (i) to (ii) to (iii) to (iv). To get from (iv) to (i), suppose we have a dominant rational map $C_1 \dashrightarrow C_2$ of integral curves. Replace C_1 by a dense open set so the rational map is a morphism $C_1 \to C_2$. This induces a map of normalizations $\widetilde{C_1} \to \widetilde{C_2}$ of regular irreducible curves. Let $\widetilde{\widetilde{C_i}}$ be a regular projective compactification of $\widetilde{C_i}$ (for $i = 1, 2$), as in Theorem 17.4.2. Then the morphism $\widetilde{C_1} \to \widetilde{\widetilde{C_2}}$ extends to a morphism $\widetilde{\widetilde{C_1}} \to \widetilde{\widetilde{C_2}}$ by the Curve-to-Projective Extension Theorem 16.5.1. This morphism is surjective (do you see why?), so we have produced a morphism in category (i).

17.4.B. EXERCISE. Put the above pieces together to describe equivalences of categories (i) through (iv).

It remains to connect (v). This is essentially the content of Exercise 6.5.D; details are left to the reader. $\square$

Theorem 17.4.3 has a number of implications. For example, each quasiprojective reduced curve is birational to precisely one projective regular curve. Here is another interesting consequence, promised in Side Remark 3.1.3.

17.4.4. Degree of a projective morphism from a curve to a regular curve.

You might already have a reasonable sense that a map of compact Riemann surfaces has a well-behaved degree, that the number of preimages of a point of C' is constant, so long as the preimages are counted with appropriate multiplicity. For example, if f locally looks like $z \mapsto z^m = y$, then near $y = 0$ and $z = 0$ (but not *at* $z = 0$), each point has precisely m preimages, but as y goes to 0, the m preimages coalesce. Enlightening Example 9.3.3 showed this phenomenon in a more complicated context.

We now show the algebraic version of this fact. Suppose $\pi : C \to C'$ is a surjective (or equivalently, dominant) map of regular projective curves. We will show that π has a well-behaved degree, in a sense that we will now make precise.

First we show that π is finite. Theorem 18.1.9 (finite = projective + finite fibers) implies this, but we haven't proved it yet. So instead we show the finiteness of π as follows. Let C'' be the normalization of C' in the function field of C. Then we have an isomorphism $K(C) \cong K(C'')$ (really, equality) which leads to birational maps $C \dashleftarrow\dashrightarrow C''$ which extend to morphisms as both C and C'' are regular and projective (by the Curve-to-Projective Extension Theorem 16.5.1). Thus this yields an isomorphism of C and C''. But $C'' \to C$ is a finite morphism by the finiteness of integral closure (Theorem 9.7.3).

17.4.5. Proposition. — *Suppose that $\pi : C \to C'$ is a finite morphism, where C is a (pure dimension 1) curve with no embedded points (the most important case: C is reduced), and C' is a regular curve. Then $\pi_* \mathscr{O}_C$ is locally free of finite rank.*

The "no embedded points" hypothesis is the same as requiring that every associated point of C maps to a generic point of (some component of) C'.

We will prove Proposition 17.4.5 in §17.4.9, after showing how useful it is. The regularity hypothesis on C' is necessary: the normalization of a nodal curve (Figure 7.4) is an example where most points have one preimage, and one point (the "node") has two. (We will later see, in Exercise 24.4.G and §24.4.11, that what matters in the hypotheses of Proposition 17.4.5 is that the morphism is finite and *flat*.)

17.4.6. *Definition.* If C' is irreducible, the rank of this locally free sheaf is the **degree** of π.

17.4.C. EXERCISE. Recall that the degree of a rational map from one irreducible curve to another is defined as the degree of the function field extension (Definition 11.2.2). Show that (with the notation of Proposition 17.4.5) if C and C' are irreducible, the degree of π as a rational map is the same as the rank of $\pi_* \mathscr{O}_C$.

17.4.7. *Remark for those with complex-analytic background (algebraic degree = analytic degree).* If $C \to C'$ is a finite map of regular complex algebraic curves, Proposition 17.4.5 establishes that algebraic degree as defined above is the same as analytic degree (counting preimages, with multiplicity).

17.4.D. EXERCISE. We continue the notation and hypotheses of Proposition 17.4.5. Suppose p is a point of C'. The scheme-theoretic preimage $\pi^*(p)$ of p is a dimension 0 scheme over k.

(a) Suppose C' is finite type over a field k, and n is the dimension of the structure sheaf of $\pi^*(p)$ as a k-vector space. Show that $n = (\deg \pi)(\deg p)$. (The degree of a point was defined in §5.3.8.)

(b) Suppose that C is regular, and $\pi^{-1}(p) = \{p_1, \dots, p_m\}$. Suppose t is a uniformizer of the discrete valuation ring $\mathscr{O}_{C',p}$. Show that

$$\deg \pi = \sum_{i=1}^{m} (\mathrm{val}_{p_i} \pi^* t) \deg(\kappa(p_i)/\kappa(p)),$$

where $\deg(\kappa(p_i)/\kappa(p))$ denotes the degree of the field extension of the residue fields. (Can you extend (a) to remove the hypotheses of working over a field? If you are a number theorist, can you recognize (b) in terms of splitting primes in extensions of rings of integers in number fields?)

17.4.E. EXERCISE. Suppose that C is an irreducible regular curve, and s is a nonzero rational function on C. Show that the number of zeros of s (counted with appropriate multiplicity) equals the number of poles. Hint: recognize this as the degree of a morphism $s : C \to \mathbb{P}^1$. (In the complex category, this is an important consequence of the Residue Theorem. Another approach is given in Exercise 18.4.D.)

17.4.F. EXERCISE. Suppose s_1 and s_2 are two sections of a degree d line bundle $\mathscr{L}$ on an irreducible regular curve C, with no common zeros. Then s_1 and s_2 determine a morphism $\pi : C \to \mathbb{P}^1_k$. Show that the degree of π is d. (Translation: a two-dimensional base-point-free degree d linear system on C defines a degree d cover of $\mathbb{P}^1$.)

17.4.8. *Revisiting Example 9.3.3.* Proposition 17.4.5 and Exercise 17.4.D make precise what general behavior we observed in Example 9.3.3. Suppose C' is irreducible, and that d is the rank of this allegedly locally free sheaf. Then the fiber over any point of C with residue field K is the Spec of an algebra of dimension d over K. This means that the number of points in the fiber, counted with appropriate multiplicity, is always d.

As a motivating example, we revisit Example 9.3.3, the map $\mathbb{Q}[y] \to \mathbb{Q}[x]$ given by $x \mapsto y^2$, the projection of the parabola $x = y^2$ to the x-axis. We observed the following.

(i) The fiber over $x = 1$ is $\mathbb{Q}[y]/(y^2 - 1)$, so we get 2 points.
(ii) The fiber over $x = 0$ is $\mathbb{Q}[y]/(y^2)$ — we get one point, with multiplicity 2, arising because of the nonreducedness.
(iii) The fiber over $x = -1$ is $\mathbb{Q}[y]/(y^2 + 1) \cong \mathbb{Q}(i)$ — we get one point, with multiplicity 2, arising because of the field extension.
(iv) Finally, the fiber over the generic point $\operatorname{Spec} \mathbb{Q}(x)$ is $\operatorname{Spec} \mathbb{Q}(y)$, which is one point, with multiplicity 2, arising again because of the field extension (as $\mathbb{Q}(y)/\mathbb{Q}(x)$ is a degree 2 extension).

We thus see three sorts of behaviors ((iii) and (iv) are really the same). Note that even if you only work with algebraically closed fields, you will still be forced to this third type of behavior, because residue fields at generic points are usually not algebraically closed (witness case (iv) above).

17.4.9. *Proof of Proposition 17.4.5.* The key idea, useful in other circumstances, is to reduce to a fact about discrete valuation rings.

The question is local on the target, so we may assume that C' is affine. By Exercise 5.4.B, we may also assume C' is integral.

By Important Exercise 13.7.K, if the rank of the finite type quasicoherent sheaf $\pi_* \mathscr{O}_C$ is constant, then (as C' is reduced) $\pi_* \mathscr{O}_C$ is locally free. We will show this by showing the rank at any closed point p of C' is the same as the rank at the generic point.

Suppose $C' = \operatorname{Spec} A'$, where A' is an integral domain, and $p = [\mathfrak{m}]$. As π is affine, C is affine as well; say $C = \operatorname{Spec} A$.

We wish to show that (i) $\dim_{A'/\mathfrak{m}}(A/\mathfrak{m})$ (the rank of $\pi_* \mathscr{O}_C$ at p) equals (ii) $\dim_{K(A')}(A'^{\times})^{-1}A$ (the rank of $\pi_* \mathscr{O}_C$ at the generic point). In other words, we take A (considered as an A'-module), and (i) quotient by $\mathfrak{m}$, and (ii) invert all nonzero elements of A', and in each case compute the result's dimension over the appropriate field.

Both (i) and (ii) factor through localizing at $\mathfrak{m}$, so it suffices to show that $A_{\mathfrak{m}}$ is a finite rank free $A'_{\mathfrak{m}}$-module, of rank d, say, as the answers to both (i) and (ii) will then be d.

Now $A'_{\mathfrak{m}}$ is a discrete valuation ring; let t be its uniformizer. We can assume that $t \in A'$ (as otherwise, we replace A' by A'_t). Then $A_{\mathfrak{m}}$ is a finitely generated $A'_{\mathfrak{m}}$-module, and hence by Remark 12.5.15 is a finite sum of principal modules, of the form $A'_{\mathfrak{m}}$ or $A'_{\mathfrak{m}}/(t^n)$ (for various n). We wish to show that there are no summands of the latter type. But if there were, then t (interpreted as an element of $A_{\mathfrak{m}}$) would be a zerodivisor of $A_{\mathfrak{m}}$, and thus (interpreted as an element of A) a zerodivisor of A. But then by §5.5 **(C)**, there is an associated point of C in $\pi^{-1}(p)$, contradicting the hypotheses that C has no embedded points. $\square$

CHAPTER 18

Čech cohomology of quasicoherent sheaves

This topic is surprisingly simple and elegant. You may think cohomology must be complicated, and that this is why it appears so late in the book. But you will see that we need very little background. After defining schemes, we could have immediately defined quasicoherent sheaves, and then defined cohomology, and verified that it had many useful properties.

18.1 (Desired) properties of cohomology

Rather than immediately defining cohomology of quasicoherent sheaves, we first discuss why we care, and what properties it should have.

As $\Gamma(X, \cdot)$ is a left-exact functor, if $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is a short exact sequence of sheaves on X, then

$$0 \to \mathscr{F}(X) \to \mathscr{G}(X) \to \mathscr{H}(X)$$

is exact. We dream that this sequence continues to the right, giving a long exact sequence. More explicitly, there should be some covariant functors H^i ($i \geq 0$) from quasicoherent sheaves on X to groups such that H^0 is the global section functor Γ, and so that there is a "long exact sequence in cohomology".

(18.1.0.1)
$$\begin{array}{l} 0 \longrightarrow H^0(X, \mathscr{F}) \longrightarrow H^0(X, \mathscr{G}) \longrightarrow H^0(X, \mathscr{H}) \\ \longrightarrow H^1(X, \mathscr{F}) \longrightarrow H^1(X, \mathscr{G}) \longrightarrow H^1(X, \mathscr{H}) \longrightarrow \cdots \end{array}$$

(In general, whenever we see a left-exact or right-exact functor, we should hope for this, and in good cases our dreams will come true. The machinery behind this usually involves *derived functors*, which we will discuss in Chapter 23.)

Before defining cohomology groups of quasicoherent sheaves explicitly, we first describe their important properties, which are in some ways more important than the formal definition. The boxed properties will be the important ones.

Suppose X is a separated and quasicompact A-scheme. For each quasicoherent sheaf $\mathscr{F}$ on X, we will define A-modules $H^i(X, \mathscr{F})$. In particular, if $A = k$, they are k-vector spaces. In this case, we define $h^i(X, \mathscr{F}) = \dim_k H^i(X, \mathscr{F})$ (where k is left implicit on the left side).

(i) Each $\boxed{H^i \text{ is a covariant functor } \mathit{QCoh}_X \to \mathit{Mod}_A}$.

(ii) The functor H^0 is identified with functor Γ: $\boxed{H^0(X, \mathscr{F}) = \Gamma(X, \mathscr{F}),}$ and the covariance of **(i)** for $i = 0$ is just the usual covariance for Γ ($\mathscr{F} \to \mathscr{G}$ induces $\Gamma(X, \mathscr{F}) \to \Gamma(X, \mathscr{G})$).

(iii) If $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is a short exact sequence of quasicoherent sheaves on X, then we have a $\boxed{\text{long exact sequence}}$ (18.1.0.1). The maps $H^i(X, \mathscr{F}) \to H^i(X, \mathscr{G})$ come from covariance, and similarly for $H^i(X, \mathscr{G}) \to H^i(X, \mathscr{H})$. The *connecting homomorphisms* $H^i(X, \mathscr{H}) \to H^{i+1}(X, \mathscr{F})$ will have to be defined.

(iv) If $\pi : X \to Y$ is any morphism of quasicompact separated A-schemes, and $\mathscr{F}$ is a quasicoherent sheaf on X, then there is a natural morphism $\boxed{H^i(Y, \pi_*\mathscr{F}) \to H^i(X, \mathscr{F})}$ extending $\Gamma(Y, \pi_*\mathscr{F}) \to \Gamma(X, \mathscr{F})$. (Note that π is quasicompact and separated by the Cancellation Theorem 10.1.19 for quasicompact and separated morphisms, taking $Z = \operatorname{Spec} A$ in the statement of the Cancellation Theorem, so $\pi_*\mathscr{F}$ is indeed a quasicoherent sheaf by Exercise 13.3.F.) We will later see this as part of a larger story, the *Leray spectral sequence* (Theorem 23.4.4). If $\mathscr{G}$ is a quasicoherent sheaf on Y, then setting $\mathscr{F} := \pi^*\mathscr{G}$ and using the adjunction map $\mathscr{G} \to \pi_*\pi^*\mathscr{G}$ and covariance of (ii) gives a natural **pullback map** $H^i(Y, \mathscr{G}) \to H^i(X, \pi^*\mathscr{G})$ (via $H^i(Y, \mathscr{G}) \to H^i(Y, \pi_*\pi^*\mathscr{G}) \to H^i(X, \pi^*\mathscr{G})$) extending $\Gamma(Y, \mathscr{G}) \to \Gamma(X, \pi^*\mathscr{G})$. In this way, H^i is a "contravariant functor in the space".

(v) If $\pi : X \to Y$ is an affine morphism, and $\mathscr{F}$ is a quasicoherent sheaf on X, the natural map of (iv) is an isomorphism: $\boxed{H^i(Y, \pi_*\mathscr{F}) \xrightarrow{\sim} H^i(X, \mathscr{F})\ .}$ When π is a closed embedding and $Y = \mathbb{P}^N_A$, this isomorphism translates calculations on arbitrary projective A-schemes to calculations on $\mathbb{P}^N_A$.

(vi) If X can be covered by n affine open sets, then $\boxed{H^i(X, \mathscr{F}) = 0}$ for $i \geq n$ for all $\mathscr{F}$. In particular, on affine schemes, all higher ($i > 0$) quasicoherent cohomology groups vanish. The vanishing of H^1 in this case, along with the long exact sequence **(iii)** implies that Γ is an exact functor for quasicoherent sheaves on affine schemes, something we already knew (Exercise 13.4.A). It is also true that if $\dim X = n$, then $H^i(X, \mathscr{F}) = 0$ for all $i > n$ and for all $\mathscr{F}$ (**dimensional vanishing**). We will prove this for projective A-schemes (Theorem 18.2.6) and even quasiprojective A-schemes (Exercise 22.4.T). See §18.2.7 for discussion of the general case.

18.1.1. *Side Remark: the cohomological criterion for affineness.* The converse to **(vi)** in the case when $n = 1$ is Serre's *cohomological criterion for affineness*, [**GW**, Thm. 12.35]: any quasicompact quasiseparated scheme X, such that $H^i(X, \mathscr{F}) = 0$ for all $i > 0$ and all quasicoherent $\mathscr{F}$, must be affine. (In fact, it suffices that H^1 of every quasicoherent sheaf of ideals vanishes.) We will not use this, and thus will not prove it.

18.1.2. Let's get back to our list.

(vii) The functor H^i behaves well under direct sums, and more generally under filtered colimits: $\boxed{H^i(X, \varinjlim \mathscr{F}_j) = \varinjlim H^i(X, \mathscr{F}_j).}$

(viii) We will also identify the cohomology of all $\mathscr{O}(m)$ on $\mathbb{P}^n_A$:

18.1.3. Theorem. —

- $H^0(\mathbb{P}^n_A, \mathscr{O}_{\mathbb{P}^n_A}(m))$ *is a free* A*-module of rank* $\binom{n+m}{m}$ *if* $m \geq 0$.
- $H^n(\mathbb{P}^n_A, \mathscr{O}_{\mathbb{P}^n_A}(m))$ *is a free* A*-module of rank* $\binom{-m-1}{-n-m-1}$ *if* $m \leq -n-1$.
- $H^i(\mathbb{P}^n_A, \mathscr{O}_{\mathbb{P}^n_A}(m)) = 0$ *otherwise.*

We have already shown the first statement in Essential Exercise 14.1.C.

Theorem 18.1.3 has a number of features that will be the first appearances of facts that we will prove later.

- The cohomology of these bundles vanish above n (**(vi)** above)
- These cohomology groups are always *finitely generated* A-modules. This will be true for all coherent sheaves on projective A-schemes (Theorem 18.1.4(i)), and indeed (with more work) on proper A-schemes (Theorem 18.9.1).
- The top cohomology group vanishes for $m > -n-1$. (We will later see this as an example of *Kodaira vanishing*, see §21.5.7.)
- The top cohomology group is one-dimensional for $m = -n-1$ if $A = k$. This is the first appearance of the *dualizing sheaf*.
- There is a natural duality

$$H^i(X, \mathscr{O}(m)) \times H^{n-i}(X, \mathscr{O}(-n-1-m)) \to H^n(X, \mathscr{O}(-n-1))$$

  This is the first appearance of *Serre duality*.
- The alternating sum $\sum (-1)^i h^i(X, \mathscr{O}(m))$ is a polynomial. This is a first example of a *Hilbert polynomial*.

Before proving these facts, let's first use them to prove interesting things, as motivation.

By Theorem 15.3.1, for any coherent sheaf $\mathscr{F}$ on $\mathbb{P}^n_A$ we can find a surjection $\mathscr{O}(m)^{\oplus j} \to \mathscr{F}$, which yields the exact sequence

$$0 \to \mathscr{G} \to \mathscr{O}(m)^{\oplus j} \to \mathscr{F} \to 0 \tag{18.1.3.1}$$

for some coherent sheaf $\mathscr{G}$. We can use this to prove the following.

18.1.4. Theorem. — *(i) For any coherent sheaf* $\mathscr{F}$ *on a projective* A*-scheme* X *where* A *is Noetherian,* $H^i(X, \mathscr{F})$ *is a coherent (finitely generated)* A*-module.*
(ii) (Serre vanishing) Furthermore, for $m \gg 0$, $H^i(X, \mathscr{F}(m)) = 0$ *for all* $i > 0$ *(even without Noetherian hypotheses).*

(A slightly fancier version of Serre vanishing will be given in Theorem 18.8.G.)

Proof. Because cohomology of a closed scheme can be computed on the ambient space (**(v)** above), we may immediately reduce to the case $X = \mathbb{P}^n_A$.

(i) Consider the long exact sequence:

$$\begin{tikzcd} 0 \arrow[r] & H^0(\mathbb{P}^n_A, \mathscr{G}) \arrow[r] & H^0(\mathbb{P}^n_A, \mathscr{O}(m)^{\oplus j}) \arrow[r] & H^0(\mathbb{P}^n_A, \mathscr{F}) \arrow[r] & {} \\ & H^1(\mathbb{P}^n_A, \mathscr{G}) \arrow[r] & H^1(\mathbb{P}^n_A, \mathscr{O}(m)^{\oplus j}) \arrow[r] & H^1(\mathbb{P}^n_A, \mathscr{F}) \arrow[r] & \cdots \\ \cdots \arrow[r] & H^{n-1}(\mathbb{P}^n_A, \mathscr{G}) \arrow[r] & H^{n-1}(\mathbb{P}^n_A, \mathscr{O}(m)^{\oplus j}) \arrow[r] & H^{n-1}(\mathbb{P}^n_A, \mathscr{F}) \arrow[r] & {} \\ & H^n(\mathbb{P}^n_A, \mathscr{G}) \arrow[r] & H^n(\mathbb{P}^n_A, \mathscr{O}(m)^{\oplus j}) \arrow[r] & H^n(\mathbb{P}^n_A, \mathscr{F}) \arrow[r] & 0 \end{tikzcd}$$

The exact sequence ends here because $\mathbb{P}^n_A$ is covered by $n+1$ affine open sets (**(vi)** above). Then $H^n(\mathbb{P}^n_A, \mathscr{O}(m)^{\oplus j})$ is finitely generated by Theorem 18.1.3, hence $H^n(\mathbb{P}^n_A, \mathscr{F})$ is finitely generated for all coherent sheaves $\mathscr{F}$. Hence in particular, $H^n(\mathbb{P}^n_A, \mathscr{G})$ is finitely generated. As $H^{n-1}(\mathbb{P}^n_A, \mathscr{O}(m)^{\oplus j})$ is finitely generated, and $H^n(\mathbb{P}^n_A, \mathscr{G})$ is too, we have that $H^{n-1}(\mathbb{P}^n_A, \mathscr{F})$ is finitely generated for all coherent sheaves $\mathscr{F}$. We continue inductively downwards.

(ii) Twist (18.1.3.1) by $\mathscr{O}(N)$ for $N \gg 0$. Then

$$H^n(\mathbb{P}^n_A, \mathscr{O}(m+N)^{\oplus j}) = \oplus_j H^n(\mathbb{P}^n_A, \mathscr{O}(m+N)) = 0$$

(by **(vii)** above), so $H^n(\mathbb{P}^n_A, \mathscr{F}(N)) = 0$. Translation: for any coherent sheaf, its top cohomology vanishes once you twist by $\mathscr{O}(N)$ for N sufficiently large. Hence this is true for $\mathscr{G}$ as well. Hence from the long exact sequence, $H^{n-1}(\mathbb{P}^n_A, \mathscr{F}(N)) = 0$ for $N \gg 0$. As in (i), we induct downwards, until we get that $H^1(\mathbb{P}^n_A, \mathscr{F}(N)) = 0$. (The induction stops here, as it is *not* true that $H^0(\mathbb{P}^n_A, \mathscr{O}(m+N)^{\oplus j}) = 0$ for large N — quite the opposite.) $\square$

18.1.A. $\star\star$ EXERCISE FOR THOSE WHO LIKE NON-NOETHERIAN RINGS. Prove part (i) in the above result without the Noetherian hypotheses, assuming only that A is a coherent A-module (A is "coherent over itself"). (Hint: induct downwards as before. Show the following in order: $H^n(\mathbb{P}^n_A, \mathscr{F})$ finitely generated, $H^n(\mathbb{P}^n_A, \mathscr{G})$ finitely generated, $H^n(\mathbb{P}^n_A, \mathscr{F})$ coherent, $H^n(\mathbb{P}^n_A, \mathscr{G})$ coherent, $H^{n-1}(\mathbb{P}^n_A, \mathscr{F})$ finitely generated, $H^{n-1}(\mathbb{P}^n_A, \mathscr{G})$ finitely generated, etc.)

In particular, we have proved the following, that we would have cared about even before we knew about cohomology.

18.1.5. Corollary. — *Any projective k-scheme has a finite-dimensional space of global sections. More generally, if A is Noetherian and $\mathscr{F}$ is a coherent sheaf on a projective A-scheme, then $H^0(X, \mathscr{F})$ is a coherent A-module.*

(We will generalize this in Theorem 18.8.1.) I want to emphasize how remarkable this proof is. It is a question about global sections, i.e., H^0, which we think of as the most down to earth cohomology group, yet the proof is by downward induction for H^n, starting with n large.

Corollary 18.1.5 is true more generally for proper k-schemes, not just projective k-schemes (see Theorem 18.9.1).

Here are some important consequences. They can also be shown directly, without the use of cohomology, but with much more elbow grease.

18.1.6. As a partial converse, if $h^0(X, \mathscr{O}_X) = 1$, then X is connected (why?), but need not be reduced: witness the subscheme in $\mathbb{P}^2$ cut out by $x^2 = 0$. (For experts: the geometrically connected hypothesis is necessary, as $X = \operatorname{Spec} \mathbb{C}$ is a projective integral $\mathbb{R}$-scheme, with $h^0(X, \mathscr{O}_X) = 2$. Similarly, a nontrivial purely inseparable field extension can be used to show that the geometrically reduced hypothesis is also necessary.)

18.1.B. CRUCIAL EXERCISE (PUSHFORWARDS OF COHERENT SHEAVES ARE COHERENT). Suppose $\pi : X \to Y$ is a projective morphism of locally Noetherian schemes. Show that the pushforward of a coherent sheaf on X is a coherent sheaf on Y. (See Grothendieck's Coherence Theorems 18.8.1 and 18.9.1 for generalizations.)

18.1.7. *Unimportant remark, promised in Exercise 16.2.C.* As a consequence, if $\pi : X \to Y$ is a finite morphism, and $\mathscr{O}_Y$ is coherent over itself, then π_* sends coherent sheaves on X to coherent sheaves on Y.

Finite morphisms are affine (from the definition) and projective (17.3.D). We can now show that this is a characterization of finiteness.

18.1.8. Corollary. — *Suppose Y is locally Noetherian. Then a morphism $\pi : X \to Y$ is projective and affine if and only if π is finite.*

We will see in Exercise 18.9.A that the projective hypotheses can be relaxed to proper.

Proof. We already know that finite morphisms are affine (by definition) and projective (Exercise 17.3.D), so we show the converse. Suppose π is projective and affine. By Exercise 18.1.B, $\pi_* \mathscr{O}_X$ is coherent and hence finite type. □

The following result was promised in §17.3.6, and has a number of useful consequences.

18.1.9. Theorem (projective + finite fibers = finite). — *Suppose $\pi : X \to Y$ with Y Noetherian. Then π is projective and finite fibers if and only if it is finite. Equivalently, π is projective and quasifinite if and only it is finite.*

(Recall that quasifinite = finite fibers + finite type. But projective includes finite type.) It is true more generally that (with Noetherian hypotheses) proper + finite fibers = finite, see Theorem 29.6.2.

Proof. We show π is finite near a point $q \in Y$. Fix an affine open neighborhood $\operatorname{Spec} A$ of q in Y. Pick a hypersurface H in $\mathbb{P}^n_A$ missing the preimage of q, so $H \cap X$ is closed. Let $H' = \pi(H \cap X)$, which is closed, and doesn't contain q. Let $U = \operatorname{Spec} A - H'$, which is an open set containing q. Then above U, π is projective and affine, so we are done by Corollary 18.1.8. □

A similar trick was used in the proof of the Rigidity Lemma 10.3.12.

18.1.C. EXERCISE (UPPER SEMICONTINUITY OF FIBER DIMENSION ON THE TARGET, FOR PROJECTIVE MORPHISMS). Use a similar argument as in Theorem 18.1.9 to prove *upper semicontinuity of fiber dimension of projective morphisms*: suppose $\pi : X \to Y$ is a projective morphism where Y is locally Noetherian (or more generally $\mathscr{O}_Y$ is coherent over itself). Show that $\{y \in Y : \dim f^{-1}(y) > k\}$ is a Zariski-closed subset of Y. In other words, the dimension of the fiber "jumps over Zariski-closed subsets" of the target. (You can interpret the case $k = -1$ as the fact that projective morphisms are closed, which is basically the Fundamental Theorem of Elimination Theory 7.4.7, cf. §17.3.4.) This exercise is rather important for having a sense of how projective morphisms behave. (The case of varieties was done earlier, in Theorem 11.4.2(b). This approach is much simpler.)

The final exercise of the section is on a different theme.

18.1.D. EXERCISE. Suppose $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is an exact sequence of coherent sheaves on projective X with $\mathscr{F}$ coherent. Show that for $n \gg 0$,

$$0 \to H^0(X, \mathscr{F}(n)) \to H^0(X, \mathscr{G}(n)) \to H^0(X, \mathscr{H}(n)) \to 0$$

is also exact. (Hint: for $n \gg 0$, $H^1(X, \mathscr{F}(n)) = 0$.)

18.2 Definitions and proofs of key properties

This section could be read much later; the facts we will use are all stated in the previous section. However, the arguments are not complicated, so you want to read this right away. As you read this, you should go back and check off all the facts in the previous section, to assure yourself that you understand everything promised.

18.2.1. Čech cohomology. Čech cohomology in general settings is defined using a limit over finer and finer covers of a space. In our algebro-geometric setting, the situation is much cleaner, and we can use a single cover.

Suppose X is quasicompact and separated, which is true for example if X is quasiprojective over A. In particular, X may be covered by a finite number of affine open sets, and the intersection of any two affine open sets is also an affine open set (by separatedness, Proposition 10.1.8). We will use quasicompactness and separatedness only in order to ensure these two nice properties.

Suppose $\mathscr{F}$ is a quasicoherent sheaf, and $\mathscr{U} = \{U_i\}_{i=1}^n$ is a *finite* collection of affine open sets covering X. For $I \subset \{1, \dots, n\}$ define $U_I = \cap_{i \in I} U_i$, which is affine by the separated hypothesis. (Here is a strong analogy for those who have seen cohomology in other contexts: cover a topological space X with a finite number of open sets U_i, such that all intersections $\cap_{i \in I} U_i$ are contractible.) Consider the **Čech complex**

(18.2.1.1)
$$0 \to \prod_{\substack{|I|=1 \\ I \subset \{1,\dots,n\}}} \mathscr{F}(U_I) \to \cdots \to$$
$$\prod_{\substack{|I|=i \\ I \subset \{1,\dots,n\}}} \mathscr{F}(U_I) \to \prod_{\substack{|I|=i+1 \\ I \subset \{1,\dots,n\}}} \mathscr{F}(U_I) \to \cdots .$$

The maps are defined as follows. The map from $\mathscr{F}(U_I) \to \mathscr{F}(U_J)$ is 0 unless $I \subset J$, i.e., $J = I \cup \{j\}$. If j is the kth element of J, then the map is $(-1)^{k-1}$ times the restriction map res_{U_I, U_J}.

18.2.A. EASY EXERCISE (FOR THOSE WHO HAVEN'T SEEN ANYTHING LIKE THE ČECH COMPLEX BEFORE). Show that the Čech complex is indeed a complex, i.e., that the composition of two consecutive arrows is 0.

Define $H^i_{\mathscr{U}}(X, \mathscr{F})$ to be the ith cohomology group of the complex (18.2.1.1). Note that if X is an A-scheme, then $H^i_{\mathscr{U}}(X, \mathscr{F})$ is an A-module. We have almost succeeded in defining the Čech cohomology group H^i, except our definition seems to depend on a choice of a cover $\mathscr{U}$. Note that $H^i_{\mathscr{U}}(X, \cdot)$ is clearly a covariant functor $\mathit{QCoh}_X \to \mathit{Mod}_A$.

18.2.B. EASY EXERCISE. Identify $H^0_{\mathscr{U}}(X, \mathscr{F})$ with $\Gamma(X, \mathscr{F})$. (Hint: use the sheaf axioms for $\mathscr{F}$.)

18.2.C. EXERCISE. Suppose

$$0 \to \mathscr{F}_1 \to \mathscr{F}_2 \to \mathscr{F}_3 \to 0 \tag{18.2.1.2}$$

is a short exact sequence of sheaves of abelian groups on a topological space, and $\mathscr{U}$ is a finite open cover such that on any intersection U_I of open subsets in $\mathscr{U}$, the map $\Gamma(U_I, \mathscr{F}_2) \to \Gamma(U_I, \mathscr{F}_3)$ is surjective. Show that we get a "long exact sequence of cohomology for $H^i_{\mathscr{U}}$" (where we take the same definition of $H^i_{\mathscr{U}}$). In our situation, where X is a quasicompact separated A-scheme, and (18.2.1.2) is a short exact sequence of quasicoherent sheaves on X, show that we get a long exact sequence for the A-modules $H^i_{\mathscr{U}}$.

In the proof of Theorem 18.8.1, we will make use of the fact that your construction of the connecting homomorphism will "commute with localization of A". More precisely, we will need the following.

18.2.D. EXERCISE. Suppose we are given a short exact sequence (18.2.1.2) of quasicoherent sheaves on a quasicompact separated A-scheme $\pi : X \to \operatorname{Spec} A$, a cover $\mathscr{U}$ of X by affine open sets, and some $f \in A$. The restriction of the sets of $\mathscr{U}$ to X_f yields an affine open cover $\mathscr{U}'$ of $X_f = \pi^{-1}(D(f))$. Identify the long exact sequence associated to (18.2.1.2) using $H^i_{\mathscr{U}}$, localized at f, with the long exact sequence associated to the restriction of (18.2.1.2) to X_f, using the affine open cover $\mathscr{U}'$. (First check that the maps such as $H^i(\mathscr{F})_{\mathscr{U}} \to H^i(\mathscr{G})_{\mathscr{U}}$ given by covariance "commute with localization", and then check that the connecting homomorphisms do as well.)

18.2.2. Theorem/Definition. — *Our standing assumption is that X is quasicompact and separated. $H^i_{\mathscr{U}}(X, \mathscr{F})$ is independent of the choice of (finite) cover $\{U_i\}$. More precisely, for any two covers $\{U_i\} \subset \{V_i\}$, the maps $H^i_{\{V_i\}}(X, \mathscr{F}) \to H^i_{\{U_i\}}(X, \mathscr{F})$ induced by the natural map of Čech complexes (18.2.1.1) are isomorphisms. Define the Čech cohomology group $H^i(X, \mathscr{F})$ to be this group.*

If you are unsure of what the "natural map of Čech complexes" is, by (18.2.3.1) it should become clear.

18.2.3. For experts: maps of complexes inducing isomorphisms on cohomology groups are called *quasiisomorphisms*. We are actually getting a finer invariant than cohomology out of this construction; we are getting an element of the *derived category of* A*-modules*.

Proof. We need only prove the result when $|\{V_i\}| = |\{U_i\}| + 1$. We will show that if $\{U_i\}_{1 \leq i \leq n}$ is a cover of X, and U_0 is any other open set, then the map $H^i_{\{U_i\}_{0 \leq i \leq n}}(X, \mathscr{F}) \to H^i_{\{U_i\}_{1 \leq i \leq n}}(X, \mathscr{F})$ is an isomorphism. Consider the exact sequence of complexes

(18.2.3.1)

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
\cdots \arrow[r] & \prod_{\substack{|I| = i-1 \\ 0 \in I}} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{\substack{|I| = i \\ 0 \in I}} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{\substack{|I| = i+1 \\ 0 \in I}} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] & \prod_{|I| = i-1} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{|I| = i} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{|I| = i+1} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] & \prod_{\substack{|I| = i-1 \\ 0 \notin I}} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{\substack{|I| = i \\ 0 \notin I}} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{\substack{|I| = i+1 \\ 0 \notin I}} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \cdots \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

Throughout, $I \subset \{0, \dots, n\}$. The bottom two rows are Čech complexes with respect to two covers, and the map between them induces the desired map on cohomology. We get a long exact sequence of cohomology from this short exact sequence of complexes (Exercise 1.6.C). Thus we wish to show that the top row is exact and thus has vanishing cohomology. (Note that $U_0 \cap U_j$ is affine by our separatedness hypothesis, Proposition 10.1.8.) But the ith cohomology of the top row is precisely $H^i_{\{U_i \cap U_0\}_{i>0}}(U_i, \mathscr{F})$ except at step 0, where we get 0 (because the complex starts off $0 \to \mathscr{F}(U_0) \to \prod_{j=1}^n \mathscr{F}(U_0 \cap U_j)$). So it suffices to show that higher Čech groups of affine schemes are 0. Hence we are done by the following result. □

18.2.4. Theorem. — *The higher Čech cohomology* $H^i_{\mathscr{U}}(X, \mathscr{F})$ *of an affine* A*-scheme* X *vanishes (for any affine cover* $\mathscr{U}$*,* $i > 0$*, and quasicoherent* $\mathscr{F}$*).*

Serre describes this as a partition of unity argument.

Proof. (The following argument can be made shorter using spectral sequences, but we avoid this for the sake of clarity.) We want to show that the "extended" complex

$$0 \to \mathscr{F}(X) \to \prod_{|I|=1} \mathscr{F}(U_I) \to \prod_{|I|=2} \mathscr{F}(U_I) \to \cdots \tag{18.2.4.1}$$

(where the global sections $\mathscr{F}(X)$ have been appended to the start) has no cohomology, i.e., is exact. We do this with a trick.

Suppose first that some U_i, say U_0, is X. Then the complex is the middle row of the following short exact sequence of complexes

(18.2.4.2)

$$\begin{tikzcd} 0 \arrow[r] & 0 \arrow[r] \arrow[d] & \prod_{|I|=1, 0\in I} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{|I|=2, 0\in I} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \cdots \\ 0 \arrow[r] & \mathscr{F}(X) \arrow[r] \arrow[d] & \prod_{|I|=1} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \prod_{|I|=2} \mathscr{F}(U_I) \arrow[r] \arrow[d] & \cdots \\ 0 \arrow[r] & \mathscr{F}(X) \arrow[r] & \prod_{|I|=1, 0\notin I} \mathscr{F}(U_I) \arrow[r] & \prod_{|I|=2, 0\notin I} \mathscr{F}(U_I) \arrow[r] & \cdots \end{tikzcd}$$

The top row is the same as the bottom row, slid over by 1. The corresponding long exact sequence of cohomology shows that the central row has vanishing cohomology. (You should show that the "connecting homomorphism" on cohomology is indeed an isomorphism.) This might remind you of the *mapping cone* construction (Exercise 1.7.E).

We next prove the general case by sleight of hand. Say $X = \operatorname{Spec} R$. We wish to show that the complex of A-modules (18.2.4.1) is exact. It is also a complex of R-modules, so we wish to show that the complex of R-modules (18.2.4.1) is exact. To show that it is exact, it suffices to show that for a cover of $\operatorname{Spec} R$ by distinguished open sets $D(f_i)$ ($1 \leq i \leq r$) (i.e., $(f_1, \dots, f_r) = 1$ in R) the complex is exact. (Translation: exactness of a sequence of sheaves may be checked locally.) We choose a cover so that each $D(f_i)$ is contained in some $U_j = \operatorname{Spec} A_j$. Consider the complex localized at f_i. As

$$\Gamma(\operatorname{Spec} A, \mathscr{F})_f = \Gamma(\operatorname{Spec}(A_j)_f, \mathscr{F})$$

(by quasicoherence of $\mathscr{F}$, Exercise 13.3.D), as $U_j \cap D(f_i) = D(f_i)$, we are in the situation where one of the U_i's is X, so we are done. □

We have now proved properties **(i)–(iii)** of the previous section. Property **(vi)** is also straightforward: if X is covered by n affine open sets, use these as the cover $\mathscr{U}$, and notice that the Čech complex ends by the nth step.

18.2.E. Exercise (property **(v)**). Suppose $\pi : X \to Y$ is an affine morphism, and Y is a quasicompact and separated A-scheme (and hence X is too, as affine morphisms are both quasicompact and separated). If $\mathscr{F}$ is a quasicoherent sheaf on X, describe a natural isomorphism $H^i(Y, \pi_*\mathscr{F}) \cong H^i(X, \mathscr{F})$. (Hint: if $\mathscr{U}$ is an affine cover of Y, "$\pi^{-1}(\mathscr{U})$" is an affine cover X. Use these covers to compute the cohomology of $\mathscr{F}$.)

18.2.F. Exercise (property **(iv)**). Suppose $\pi : X \to Y$ is any quasicompact separated morphism, $\mathscr{F}$ is a quasicoherent sheaf on X, and Y is a quasicompact separated A-scheme. The hypotheses on π ensure that $\pi_*\mathscr{F}$ is a quasicoherent sheaf on Y. Describe a natural morphism $H^i(Y, \pi_*\mathscr{F}) \to H^i(X, \mathscr{F})$ extending $\Gamma(Y, \pi_*\mathscr{F}) \to \Gamma(X, \mathscr{F})$. (Aside: this morphism is an isomorphism for $i = 0$, but need not be an isomorphism for higher i: consider $i = 1$, $X = \mathbb{P}^1_k$, $\mathscr{F} = \mathscr{O}(-2)$, and let Y be a point $\operatorname{Spec} k$.)

18.2.G. EXERCISE. Prove Property **(vii)** of the previous section. (This can be done by hand. Hint: in the category of modules over a ring, taking the colimit over a filtered sets is an exact functor, §1.6.12.)

We have now proved all of the properties of the previous section, except for **(viii)**, which we will get to in §18.3.

18.2.5. Useful facts about cohomology for k-schemes.

18.2.H. EXERCISE (COHOMOLOGY AND CHANGE OF BASE FIELD). Suppose X is a quasicompact separated k-scheme, and $\mathscr{F}$ is a coherent sheaf on X. Give an isomorphism

$$H^i(X, \mathscr{F}) \otimes_k K \cong H^i(X \times_{\operatorname{Spec} k} \operatorname{Spec} K, \mathscr{F} \otimes_k K)$$

for all i, where K/k is any field extension. Here $\mathscr{F} \otimes_k K$ means the pullback of $\mathscr{F}$ to $X \times_{\operatorname{Spec} k} \operatorname{Spec} K$. Hence $h^i(X, \mathscr{F}) = h^i(X \times_{\operatorname{Spec} k} \operatorname{Spec} K, \mathscr{F} \otimes_k K)$. If $i = 0$ (taking $H^0 = \Gamma$), show the result without the quasicompact and separated hypotheses. (This is useful for relating facts about k-schemes to facts about schemes over algebraically closed fields. Your proof might use vector spaces — i.e., linear algebra — in a fundamental way. If it doesn't, you may prove something more general, if $k \to K$ is replaced by a flat ring map $B \to A$. Recall that $B \to A$ is flat if $\otimes_B A$ is an exact functor $\mathit{Mod}_B \to \mathit{Mod}_A$. A hint for this harder exercise: the FHHF theorem, Exercise 1.6.H. See Exercise 18.8.B(b) for the next generalization of this.)

18.2.I. EXERCISE (BASE-POINT-FREENESS IS INDEPENDENT OF EXTENSION OF BASE FIELD). Suppose X is a scheme over a field k, $\mathscr{L}$ is an invertible sheaf on X, and K/k is a field extension. Show that $\mathscr{L}$ is base-point-free if and only if its pullback to $X \times_{\operatorname{Spec} k} \operatorname{Spec} K$ is base-point-free. (Hint: Exercise 18.2.H with $i = 0$ implies that a basis of sections of $\mathscr{L}$ over k becomes, after tensoring with K, a basis of sections of $\mathscr{L} \otimes_k K$.)

18.2.6. Theorem (dimensional vanishing for quasicoherent sheaves on projective k-schemes). — *Suppose X is a projective k-scheme, and $\mathscr{F}$ is a quasicoherent sheaf on X. Then $H^i(X, \mathscr{F}) = 0$ for $i > \dim X$.*

In other words, cohomology vanishes above the dimension of X.

Proof. Suppose $X \hookrightarrow \mathbb{P}^N$, and let $n = \dim X$. We show that X may be covered by $n + 1$ affine open sets. Exercise 11.3.C shows that there are $n + 1$ effective Cartier divisors on $\mathbb{P}^N$ such that their complements $U_0, \dots, U_n$ cover X. Then U_i is affine, so $U_i \cap X$ is affine, and thus we have covered X with $n + 1$ affine open sets. □

(It turns out that $n + 1$ affine open sets are always necessary. One way of proving this is by showing that the complement of a dense affine open subset is always pure codimension 1, see for example [**RV**, Lem. 2.3].)

18.2.7. ⋆ **Dimensional vanishing more generally.** Using the theory of blowing up, Theorem 18.2.6 can be extended to quasiprojective k-schemes, see §22.4.15. Dimensional vanishing is even true in much greater generality. To state it, we need to define cohomology with the more general machinery of derived functors (Chapter 23). If X is a Noetherian topological space (§3.6.14) and $\mathscr{F}$ is any sheaf of abelian groups on X, we have $H^i(X, \mathscr{F}) = 0$ for all $i > \dim X$. (Grothendieck sketches his elegant proof in [**GrS**, p. 29-30]; see [**Ha1**, Theorem III.2.7] for a more detailed explanation.) In particular, if X is a k-variety of dimension n, we *always* have dimensional vanishing, even for crazy varieties that can't be covered with $n+1$ affine open subsets (see §22.4.15).

18.2.8. The Kunneth formula.

Suppose X and Y are quasicompact separated k-schemes, and $\mathscr{F}$ and $\mathscr{G}$ are quasicoherent sheaves on X and Y respectively. Let $\pi_X : X \times_k Y \to X$ and $\pi_Y : X \times_k Y \to Y$ be the two projections. Recall the definition $\mathscr{F} \boxtimes \mathscr{G} := \pi_X^* \mathscr{F} \otimes \pi_Y^* \mathscr{G}$ (§16.4.8). Then we have an isomorphism

$$H^m(X \times_k Y, \mathscr{F} \boxtimes \mathscr{G}) \cong \oplus_{p+q=m} H^p(X, \mathscr{F}) \otimes_k H^q(Y, \mathscr{G}).$$

To show this, choose affine covers of X and Y, and produce the Čech complexes for $\mathscr{F}$ and $\mathscr{G}$. Show that the tensor product of these two complexes (the total complex associated to the double complex) is the Čech complex for $\mathscr{F} \boxtimes \mathscr{G}$ (with respect to the products of the affine covers of X and Y). Finally, show that the cohomology of the tensor product of two complexes over k is the tensor products of the cohomologies, a result known as the *Eilenberg-Zilber Theorem*.

18.2.9. The cup product. The cup product in Čech cohomology can be defined in a simple way; see [**Liu**, Exer. 2.17] for a particularly elegant description. We will not need this construction.

18.3 Cohomology of line bundles on projective space

We now finally prove the last promised basic fact about cohomology, property **(viii)** of §18.1, Theorem 18.1.3, on the cohomology of line bundles on projective space. More correctly, we will do one case and you will do the rest.

We begin with a warm-up that will let you (implicitly) see some of the structure that will arise in the proof. It also gives good practice in computing cohomology groups.

18.3.A. EXERCISE. Compute the cohomology groups $H^i(\mathbb{A}_k^2 \setminus \{(0,0)\}, \mathscr{O})$. (Hint: the case $i=0$ was done in Example 4.4.1. The case $i > 1$ is clear from property **(vi)** above.) In particular, show that $H^1(\mathbb{A}_k^2 \setminus \{(0,0)\}, \mathscr{O}) \neq 0$, and thus give another proof (see §4.4.3) of the fact that $\mathbb{A}_k^2 \setminus \{(0,0)\}$ is not affine. (Cf. Serre's cohomological criterion for affineness, Remark 18.1.1.)

18.3.1. *Remark.* Essential Exercise 14.1.C and the ensuing discussion showed that $H^0(\mathbb{P}_A^n, \mathscr{O}_{\mathbb{P}_A^n}(m))$ should be interpreted as the homogeneous degree m polynomials in $x_0, \dots, x_n$ (with A-coefficients). Similarly, $H^n(\mathbb{P}_A^n, \mathscr{O}_{\mathbb{P}_A^n}(m))$ should be interpreted as the homogeneous degree m Laurent polynomials in $x_0, \dots, x_n$, where in each monomial, each x_i appears with degree at most -1.

18.3.2. *Proof of Theorem 18.1.3 for* $n = 2$. We take the standard cover $U_0 = D(x_0)$, $\dots, U_n = D(x_n)$ of $\mathbb{P}_A^n$.

18.3.B. EXERCISE (ESSENTIAL FOR THE PROOF OF THEOREM 18.1.3). If $I \subset \{0, \dots, n\}$, then give an isomorphism (of A-modules) of $\Gamma(\mathscr{O}(m), U_I)$ with the homogeneous degree m Laurent monomials (in x_0, …, x_n, with coefficients in A) where each x_i for $i \notin I$ appears with non-negative degree. Your construction should be such that the restriction map $\Gamma(U_I, \mathscr{O}(m)) \to \Gamma(U_J, \mathscr{O}(m))$ ($I \subset J$) corresponds to the natural inclusion: a Laurent polynomial in $\Gamma(U_I, \mathscr{O}(m))$ maps to the *same* Laurent polynomial in $\Gamma(U_J, \mathscr{O}(m))$.

The Čech complex for $\mathscr{O}(m)$ is the degree m part of
(18.3.2.1)

$$0 \longrightarrow A[x_0, x_1, x_2, x_0^{-1}] \times A[x_0, x_1, x_2, x_1^{-1}] \times A[x_0, x_1, x_2, x_2^{-1}] \longrightarrow$$
$$A[x_0, x_1, x_2, x_0^{-1}, x_1^{-1}] \times A[x_0, x_1, x_2, x_1^{-1}, x_2^{-1}] \times A[x_0, x_1, x_2, x_0^{-1}, x_2^{-1}]$$
$$\longrightarrow A[x_0, x_1, x_2, x_0^{-1}, x_1^{-1}, x_2^{-1}] \longrightarrow 0.$$

Rather than consider $\mathscr{O}(m)$ for each m independently, it is notationally simpler to consider them all at once, by considering $\mathscr{F} = \oplus_{m \in \mathbb{Z}} \mathscr{O}(m)$: the Čech complex for $\mathscr{F}$ is (18.3.2.1). It is useful to write which U_I corresponds to which factor (see (18.3.2.2) below). The maps (from one factor of one term to one factor of the next) are all natural inclusions, or negative of natural inclusions, and in particular preserve degree.

We extend (18.3.2.1) by replacing the $0 \to$ on the left by $0 \to A[x_0, x_1, x_2] \to$:
(18.3.2.2)

$$\begin{array}{ccccccccc} & & H^0 & & U_0\ U_1\ U_2 & & & & U_{012} \\ 0 & \longrightarrow & A[x_0, x_1, x_2] & \longrightarrow & \cdots & \longrightarrow & \cdots & \longrightarrow & A[x_0, x_1, x_2, x_0^{-1}, x_1^{-1} x_2^{-1}] \longrightarrow 0. \end{array}$$

18.3.C. EXERCISE. Show that if (18.3.2.2) is exact, except that at U_{012} the cohomology/cokernel is

$$x_0^{-1} x_1^{-1} x_2^{-1} A[x_0^{-1}, x_1^{-1}, x_2^{-1}],$$

then Theorem 18.1.3 holds for $n = 2$. (Hint: Remark 18.3.1.)

Because the maps in (18.3.2.2) preserve multidegree (degrees of each x_i independently), we can study exactness of (18.3.2.2) monomial by monomial.

The "3 negative exponents" case. Consider first the monomial $x_0^{a_0} x_1^{a_1} x_2^{a_2}$, where the exponents a_i are all negative. Then (18.3.2.2) in this multidegree is:

$$0 \longrightarrow 0_{H^0} \longrightarrow 0_0 \times 0_1 \times 0_2 \longrightarrow 0_{01} \times 0_{12} \times 0_{02} \longrightarrow A_{012} \longrightarrow 0.$$

Here the subscripts serve only to remind us which "Čech" terms the factors correspond to. (For example, A_{012} corresponds to the coefficient of $x_0^{a_0} x_1^{a_1} x_2^{a_2}$ in $A[x_0, x_1, x_2, x_0^{-1}, x_1^{-1}, x_2^{-1}]$.) Clearly this complex only has (co)homology at the U_{012} spot, as desired.

The "2 negative exponents" case. Consider next the case where *two* of the exponents, say a_0 and a_1, are negative. Then the complex in this multidegree is

$$0 \longrightarrow 0_{H^0} \longrightarrow 0_0 \times 0_1 \times 0_2 \longrightarrow A_{01} \times 0_{12} \times 0_{02} \longrightarrow A_{012} \longrightarrow 0,$$

which is clearly exact.

The "1 negative exponent" case. We next consider the case where *one* of the exponents, say a_0, is negative. Then the complex in this multidegree is

$$0 \longrightarrow 0_{H^0} \longrightarrow A_0 \times 0_1 \times 0_2 \longrightarrow A_{01} \times 0_{12} \times A_{02} \longrightarrow A_{012} \longrightarrow 0$$

With a little thought (paying attention to the signs on the arrows $A \to A$), you will see that it is exact. (The subscripts, by reminding us of the subscripts in the original Čech complex, remind us what signs to take in the maps.)

The "0 negative exponent" case. Finally, consider the case where *none* of the exponents are negative. Then the complex in this multidegree is

$$0 \longrightarrow A_{H^0} \longrightarrow A_0 \times A_1 \times A_2 \longrightarrow A_{01} \times A_{12} \times A_{02} \longrightarrow A_{012} \longrightarrow 0$$

We wish to show that this is exact. We write this complex as the middle of a short exact sequence of complexes:

(18.3.2.3)

$$\begin{tikzcd}
0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & A_2 \arrow[r] \arrow[d] & A_{02} \times A_{12} \arrow[r] \arrow[d] & A_{012} \arrow[r] \arrow[d] & 0 \arrow[d] \\
0 \arrow[r] \arrow[d] & A_{H^0} \arrow[r] \arrow[d] & A_0 \times A_1 \times A_2 \arrow[r] \arrow[d] & A_{01} \times A_{12} \times A_{02} \arrow[r] \arrow[d] & A_{012} \arrow[r] \arrow[d] & 0 \arrow[d] \\
0 \arrow[r] & A_{H^0} \arrow[r] & A_0 \times A_1 \arrow[r] & A_{01} \arrow[r] & 0 \arrow[r] & 0
\end{tikzcd}$$

Thus we get a long exact sequence in cohomology (Theorem 1.6.6). But the top and bottom rows are exact (basically from the "1-negative" case), i.e., cohomology-free, so the middle row must be exact too.

18.3.D. EXERCISE. Prove Theorem 18.1.3 for general n. (I could of course just have given you the proof for general n, but seeing the argument in action may be enlightening. In particular, your argument may be much shorter. For example, the "2-negative" case could be done in the same way as the "1-negative" case, so you will not need $n+1$ separate cases if you set things up carefully.)

18.3.3. *Remarks. (i)* In fact we don't really need the exactness of the top and bottom rows of (18.3.2.3); we just need that they are the same, just as with (18.2.4.2).

(ii) This argument is basically the proof that the reduced homology of the boundary of a simplex S (known in some circles as a "sphere") is 0, unless S is the empty set, in which case it is one-dimensional. The "empty set" case corresponds to the "3-negative" case.

18.3.E. EXERCISE. Show that $H^i(\mathbb{P}^m_k \times_k \mathbb{P}^n_k, \mathscr{O}(a,b)) = \sum_{j=0}^{i} H^j(\mathbb{P}^m_k, \mathscr{O}(a)) \otimes_k H^{i-j}(\mathbb{P}^n_k, \mathscr{O}(b))$. (Can you generalize this Kunneth-type formula further?)

18.4 Riemann-Roch, degrees of coherent sheaves, arithmetic genus, and Serre duality

We have seen some powerful uses of Čech cohomology, to prove things about spaces of global sections, and to prove Serre vanishing. We will now see some classical constructions come out very quickly and cheaply.

In this section, we will work over a field k. Suppose $\mathscr{F}$ is a coherent sheaf on a projective k-scheme X. Recall the notation (§18.1) $h^i(X, \mathscr{F}) := \dim_k H^i(X, \mathscr{F})$. By Theorem 18.1.4, $h^i(X, \mathscr{F})$ is finite. (The arguments in this section will extend without change to proper X once we have this finiteness for proper morphisms, by Grothendieck's Coherence Theorem 18.9.1.) Define the **Euler characteristic**

$$\chi(X, \mathscr{F}) := \sum_{i=0}^{\dim X} (-1)^i h^i(X, \mathscr{F}).$$

We will see repeatedly here and later that Euler characteristics behave better than individual cohomology groups. As one sign, notice that for fixed n, and $m \geq 0$,

$$h^0(\mathbb{P}^n_k, \mathscr{O}(m)) = \binom{n+m}{m} = \frac{(m+1)(m+2)\cdots(m+n)}{n!}.$$

Notice that the expression on the right is a polynomial in m of degree n. (For later reference, notice also that the leading coefficient is $m^n/n!$.) But it is not true that

$$h^0(\mathbb{P}^n_k, \mathscr{O}(m)) = \frac{(m+1)(m+2)\cdots(m+n)}{n!}$$

for *all* m — it breaks down for $m \leq -n-1$. Still, you can check (using Theorem 18.1.3) that

$$\chi(\mathbb{P}^n_k, \mathscr{O}(m)) = \frac{(m+1)(m+2)\cdots(m+n)}{n!}.$$

So one lesson is this: if one cohomology group (usual the top or bottom) behaves well in a certain range, and then messes up, likely it is because (i) it is actually the Euler characteristic which behaves well *always*, and (ii) the other cohomology groups vanish in that certain range.

In fact, we will see that it is often hard to calculate cohomology groups (even h^0), but it can be easier calculating Euler characteristics. So one important way of getting a hold of cohomology groups is by computing the Euler characteristics, and then showing that all the *other* cohomology groups vanish. Hence the ubiquity and importance of *vanishing theorems*. (A vanishing theorem usually states that a certain cohomology group vanishes under certain conditions.) We will see this in action when discussing curves. (One of the first applications will be (19.2.5.1).)

The following exercise shows another way in which Euler characteristic behaves well: it is *additive in exact sequences*.

18.4.A. EXERCISE. Show that if $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is an exact sequence of coherent sheaves on a projective k-scheme X, then $\chi(X, \mathscr{G}) = \chi(X, \mathscr{F}) + \chi(X, \mathscr{H})$. (Hint: consider the long exact sequence in cohomology.) More generally, if

$$0 \to \mathscr{F}_1 \to \cdots \to \mathscr{F}_n \to 0$$

is an exact sequence of coherent sheaves, show that

$$\sum_{i=1}^{n} (-1)^i \chi(X, \mathscr{F}_i) = 0.$$

(This remark both generalizes the "exact" case of Exercise 1.6.B — consider the case where $X = \operatorname{Spec} k$ — and uses it in the proof.)

18.4.1. The Riemann-Roch Theorem for line bundles on a regular projective curve. Suppose $D := \sum_{p \in C} a_p [p]$ is a divisor on a regular projective curve C over a field k (where $a_p \in \mathbb{Z}$, and all but finitely many a_p are 0). Define the **degree of** D by

$$\deg D = \sum a_p \deg p.$$

(The degree of a point p was defined in §5.3.8, as the degree of the field extension of the residue field over k.)

18.4.B. ESSENTIAL EXERCISE: THE RIEMANN-ROCH THEOREM FOR LINE BUNDLES ON A REGULAR PROJECTIVE CURVE. Show that

$$\chi(C, \mathscr{O}_C(D)) = \deg D + \chi(C, \mathscr{O}_C)$$

by induction on $\sum |a_p|$ (where $D = \sum a_p[p]$ as above). Hint: to show that $\chi(C, \mathscr{O}_C(D)) = \deg p + \chi(C, \mathscr{O}_C(D - p))$, tensor the closed subscheme exact sequence

$$0 \to \mathscr{O}_C(-p) \to \mathscr{O}_C \to \mathscr{O}|_p \to 0$$

(where $\mathscr{O}|_p$ is the structure sheaf of the scheme p, not the stalk $\mathscr{O}_{C,p}$) by $\mathscr{O}_C(D)$, and use additivity of Euler characteristics in exact sequences (Exercise 18.4.A).

As every invertible sheaf $\mathscr{L}$ is of the form $\mathscr{O}_C(D)$ for some D (see §14.2), this exercise is very powerful.

18.4.C. IMPORTANT EXERCISE. Suppose $\mathscr{L}$ is an invertible sheaf on a regular projective curve C over k. Define the **degree** of $\mathscr{L}$ (denoted $\deg \mathscr{L}$ or $\deg_C \mathscr{L}$) as $\chi(C, \mathscr{L}) - \chi(C, \mathscr{O}_C)$. Let s be a nonzero rational section on C. Let D be the divisor of zeros and poles of s:

$$D := \sum_{p \in C} \nu_p(s)[p].$$

Show that $\deg \mathscr{L} = \deg D$. In particular, the degree can be computed by counting zeros and poles of *any* section not vanishing on a component of C.

18.4.D. EXERCISE. Give a new solution to Exercise 17.4.E (a nonzero rational function on a projective curve has the same number of zeros and poles, counted appropriately) using the ideas above.

18.4.E. EXERCISE. If $\mathscr{L}$ and $\mathscr{M}$ are two line bundles on a regular projective curve C, show that $\deg \mathscr{L} \otimes \mathscr{M} = \deg \mathscr{L} + \deg \mathscr{M}$. (Hint: choose nonzero rational sections of $\mathscr{L}$ and $\mathscr{M}$.)

18.4.F. EXERCISE. Suppose $\pi : C \to C'$ is a degree d morphism of integral projective regular curves, and $\mathscr{L}$ is an invertible sheaf on C'. Show that $\deg_C \pi^* \mathscr{L} = d \deg_{C'} \mathscr{L}$. Hint: compute $\deg_{C'} \mathscr{L}$ using any nonzero rational section s of $\mathscr{L}$, and compute $\deg_C \pi^* \mathscr{L}$ using the rational section $\pi^* s$ of $\pi^* \mathscr{L}$. Note that zeros pull back to zeros, and poles pull back to poles. Reduce to the case where $\mathscr{L} = \mathscr{O}(p)$ for a single point p. Use Exercise 17.4.D.

18.4.G. ⋆⋆ EXERCISE (COMPLEX-ANALYTIC INTERPRETATION OF DEGREE; ONLY FOR THOSE WITH SUFFICIENT ANALYTIC BACKGROUND). Suppose X is a connected regular projective complex curve. Show that the degree map is the composition of group homomorphisms

$$\operatorname{Pic} X \longrightarrow \operatorname{Pic} X_{an} \xrightarrow{c_1} H^2(X_{an}, \mathbb{Z}) \xrightarrow{\cap [X_{an}]} H_0(X_{an}, \mathbb{Z}) \cong \mathbb{Z}.$$

Hint: show it for a generator $\mathscr{O}(p)$ of the group $\operatorname{Pic} X$, using explicit transition functions. (The first map was discussed in Exercise 13.1.L. The second map is takes a line bundle to its first Chern class, and can be interpreted as follows. The transition functions for a line bundle yield a Čech 1-cycle for $\mathscr{O}^*_{X_{an}}$; this yields a map $\operatorname{Pic} X_{an} \to H^1(X_{an}, \mathscr{O}^*_{X_{an}})$. Combining this with the map $H^1(X_{an}, \mathscr{O}^*_{X_{an}}) \to H^2(X_{an}, \mathbb{Z})$ from the long exact sequence in cohomology corresponding to the exponential exact sequence (2.4.10.1) yields the first Chern class map.)

18.4.2. Arithmetic genus.
Motivated by geometry (Miracle 18.4.3 below), we define the **arithmetic genus** of a scheme X as $1 - \chi(X, \mathscr{O}_X)$. This is sometimes denoted $p_a(X)$. For integral projective curves over an algebraically closed field, $h^0(X, \mathscr{O}_X) = 1$ (§10.3.7), so $p_a(X) = h^1(X, \mathscr{O}_X)$. (In higher dimension, this is a less natural notion.)

We can restate the Riemann-Roch formula for curves (Exercise 18.4.B) as:

$$h^0(C, \mathscr{L}) - h^1(C, \mathscr{L}) = \deg \mathscr{L} - p_a(C) + 1.$$

This is the most common formulation of the Riemann-Roch formula.

18.4.3. Miracle. If C is a regular irreducible projective complex curve, then the corresponding complex-analytic object, a compact *Riemann surface*, has a notion called the *genus* g, which is (informally speaking) the number of holes (see Figure 18.1). Miraculously, $g = p_a$ in this case (see Exercise 21.7.I), and for this reason, we will often write g for p_a when discussing regular (projective irreducible) curves, over any field. We will discuss genus further in §18.6.6, when we will be able to compute it in many interesting cases. (Warning: the arithmetic genus of $\mathbb{P}^1_{\mathbb{C}}$ as an $\mathbb{R}$-variety is -1!)

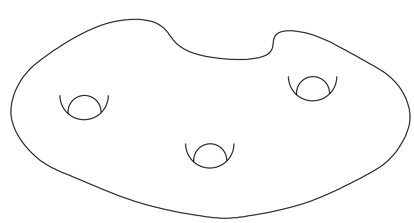


FIGURE 18.1. A genus 3 Riemann surface

18.4.4. Degree and rank of a coherent sheaf.

Suppose C is an irreducible reduced projective curve (pure dimension 1, over a field k). If $\mathscr{F}$ is a coherent sheaf on C, recall (from §13.7.4) that the **rank** of $\mathscr{F}$, denoted $\operatorname{rank} \mathscr{F}$, is its rank at the generic point of C.

18.4.H. EASY EXERCISE. Show that the rank is additive in exact sequences: if $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is an exact sequence of coherent sheaves, show that $\operatorname{rank} \mathscr{F} - \operatorname{rank} \mathscr{G} + \operatorname{rank} \mathscr{H} = 0$. Hint: localization is exact. (Caution: your argument will use the fact that the rank is at the generic point; the example

$$0 \longrightarrow \widetilde{k[t]} \xrightarrow{\times t} \widetilde{k[t]} \longrightarrow \widetilde{k[t]/(t)} \longrightarrow 0$$

on $\mathbb{A}^1_k$ shows that rank at a closed point is not additive in exact sequences.)

Define the **degree of** $\mathscr{F}$ by

$$\deg \mathscr{F} = \chi(C, \mathscr{F}) - (\operatorname{rank} \mathscr{F}) \cdot \chi(C, \mathscr{O}_C). \tag{18.4.4.1}$$

If $\mathscr{F}$ is an invertible sheaf (or if more generally the rank is the same on each irreducible component), we can drop the irreducibility hypothesis. Thus this generalizes the notion of the degree of a line bundle on a regular curve (Important Exercise 18.4.C). We now study the behavior of this notion. (In Exercise 21.7.B, you will show that if $\mathscr{F}$ is supported at a finite number of points, the degree of $\mathscr{F}$ splits up into a contribution from each point.)

18.4.I. EASY EXERCISE. Show that degree (as a function of coherent sheaves on a fixed curve C) is additive in exact sequences.

18.4.J. EXERCISE. Show that the degree of a vector bundle is the degree of its determinant bundle. Hint: Exercise 13.5.H.

The statement (18.4.4.1) is often called Riemann-Roch for coherent sheaves (or vector bundles) on a projective curve.

18.4.K. EXERCISE. If C is a projective curve, and $\mathscr{L}$ is an ample line bundle on C, show that $\deg \mathscr{L} > 0$. (Hint: show it if $\mathscr{L}$ is *very* ample.)

18.4.L. EXERCISE. Suppose $\mathscr{L}$ is a base-point-free invertible sheaf on a proper variety X, and hence induces some morphism $\phi : X \to \mathbb{P}^n$. Then $\mathscr{L}$ is ample if and only if ϕ is finite. (Hint: if ϕ is finite, use Exercise 16.6.G. If ϕ is not finite, show that there is a curve C contracted by π, using Theorem 18.1.9. Show that $\mathscr{L}$ has degree 0 on C.)

18.4.5. Extended example: the universal plane conic has no rational sections.

We use the theory of the degree to get an interesting consequence. We work over a fixed field k. We consider the following diagram.

$$\begin{tikzcd} \mathscr{C} \arrow[r, hook, "\text{cl. emb.}"] \arrow[dr] & \mathbb{P}^2 \times \mathbb{P}^5 \arrow[r] \arrow[d] & \mathbb{P}^2 \\ & \mathbb{P}^5 & \end{tikzcd}$$

If the $\mathbb{P}^2$ has projective coordinates x_0, x_1, x_2, then $\mathbb{P}^5$ has coordinates a_{00}, a_{01}, a_{11}, a_{02}, a_{12}, a_{22}, and $\mathscr{C}$ is cut out by the single equation

$$a_{00}x_0^2 + a_{01}x_0x_1 + \cdots + a_{22}x_2^2 = 0.$$

We interpret $\mathbb{P}^5$ as the parameter space of conics (in $\mathbb{P}^2$), and $\mathscr{C}$ as the universal conic over $\mathbb{P}^5$ (parametrizing a conic C along with a point $p \in C$), which comes with a canonical projection $\mathscr{C} \to \mathbb{P}^2$.

18.4.M. EXERCISE. By interpreting $\mathscr{C}$ as a $\mathbb{P}^4$-bundle over $\mathbb{P}^2$, show that $\mathscr{C}$ is a smooth sixfold, and that $\operatorname{Pic}\mathscr{C} \cong \mathbb{Z} \times \mathbb{Z}$.

18.4.N. EXERCISE. Fix a line $\ell \subset \mathbb{P}^2$ and a point $q \in \mathbb{P}^2$. Let D_ℓ be the divisor on $\mathscr{C}$ corresponding to (C, p) with p lying on ℓ. Let D_q be the divisor on $\mathscr{C}$ corresponding to (C, p) with $q \in C$. Using your description of $\mathscr{C}$ as a $\mathbb{P}^4$-bundle over $\mathbb{P}^2$, show that D_ℓ and D_q generate $\operatorname{Pic}\mathscr{C}$.

18.4.O. EXERCISE. Suppose K is a fiber of $\mathscr{C} \to \mathbb{P}^5$ over a point $r \in \mathbb{P}^5$ — i.e., a conic in $\mathbb{P}^2$ over the field $\kappa(r)$. Suppose further that neither q nor ℓ are contained in K. (This hypothesis is unnecessary, but simplifies the problem.) Show that $D_q \cdot K = 0$ and $D_\ell \cdot K = 2$. Hence show that if $\mathscr{L}$ is any invertible sheaf on $\mathscr{C}$, then $\mathscr{L} \cdot K$ is even.

18.4.P. EXERCISE. Show that the is no rational section to the projection $\pi : \mathscr{C} \to \mathbb{P}^5$. Hint: if there were a regular section over an open subset U of $\mathbb{P}^5$, it would be a divisor on $\pi^{-1}(U)$; let D be its closure in $\mathscr{C}$. Show that D meets any fiber of π over U in multiplicity 1. Use Exercise 18.4.O to obtain a contradiction.

We can restate Exercise 18.4.P in the following dramatic way: there is no way to write down three rational functions X_0, X_1, X_2 in $a_{00}, \dots, a_{22}$ such that

$$a_{00}X_0^2 + a_{01}X_0X_1 + \cdots + a_{22}X_2^2 = 0$$

without $X_0 = X_1 = X_2 = 0$.

The question of a rational point on a conic is one of arithmetic. (Think: $x^2 + y^2 = z^2$.) Our solution was topological. The unification of topology and arithmetic in this example is the beginning of a long and fruitful story in algebraic geometry.

18.4.6. ⋆ **Riemann-Roch for nonreduced curves.**

In order to state Riemann-Roch for nonreduced curves (Exercise 18.4.S), we need the notion of the *length* of a module, which is far more fundamental than might be suggested by our small discussion here. Length is a measure of a module's size, generalizing the notion of dimension of a vector space over a field. Modules of finite length share many properties with finite-dimensional vector spaces. (They play an important but implicit role in the proof of Krull's Theorem given in §11.5.)

18.4.7. *Definition.* The **length** of an A-module M, denoted $\ell(M)$, is the length of the longest strictly increasing chain of submodules of M, where the indexing (as usual) starts with 0. For example, the length of the 0-module is 0. And if $A = \mathbb{Z}$ and $n \neq 0$, then the length of $\mathbb{Z}/(n)$ is the number of prime factors of n.

18.4.8. A maximal strictly increasing chain of submodules of M is called a **composition series** for M. Clearly the subquotients of a composition series are all **simple** modules (i.e., they contain no nontrivial submodules, and are thus isomorphic to $A/\mathfrak{m}$ for some maximal ideal $\mathfrak{m}$).

18.4.Q. EXERCISE. Suppose M has a finite composition series, of length n. Show that *every* composition series of M has length n, and in particular $\ell(M) = n$. Possible hint: this parallels other composition series results you have seen in other contexts, such as for finite groups. (If absolutely necessary, see [**E**, Thm. 2.13] for an argument.)

18.4.R. EXERCISE. Suppose $(A, \mathfrak{m}, k)$ is a Noetherian local ring, and M is a finitely generated A-module. Show that M has finite length if and only if $\mathfrak{m}^n M = 0$ for some n. Thus by §18.4.8, if M has finite length, then the quotients in any composition series are all isomorphic (as A-modules) to $k = A/\mathfrak{m}$. (Informal translation: "An A-module M has finite length if and only if it can be built out of finitely many copies of k.") Hint: show that $\mathfrak{m}^n M = \mathfrak{m}^{n+1} M$ if and only if $\mathfrak{m}^n M = 0$.

You can use Exercise 18.4.R to show that M has finite length if and only if it is Artinian (i.e., satisfies the descending chain condition, §11.5). (More generally, an arbitrary module M over an arbitrary ring A has finite length if and only if it is Artinian and Noetherian, [**E**, Thm. 2.13].)

18.4.S. EXERCISE (RIEMANN-ROCH FOR NONREDUCED CURVES). Suppose C is a projective curve over a field k, and $\mathscr{F}$ is a coherent sheaf on C. Show that $\chi(\mathscr{L} \otimes \mathscr{F}) - \chi(\mathscr{F})$ is the sum over the irreducible components C_i of C of the degree $\mathscr{L}$ on C_i^{red} times the length of $\mathscr{F}$ at the generic point η_i of C_i (the length of $\mathscr{F}_{\eta_i}$ as an $\mathscr{O}_{\eta_i}$-module). Hints: (1) First reduce to the case where $\mathscr{F}$ is scheme-theoretically supported on C^{red}, by showing that both sides of the alleged equality are additive in short exact sequences, and using the filtration

$$0 = \mathscr{I}^r \mathscr{F} \subset \mathscr{I}^{r-1} \mathscr{F} \subset \cdots \subset \mathscr{I} \mathscr{F} \subset \mathscr{F}$$

of $\mathscr{F}$, where $\mathscr{I}$ is the ideal sheaf cutting out C^{red} in C. Thus we need only consider the case where C is reduced. (2) As $\mathscr{L}$ is projective, we can write $\mathscr{L} \cong \mathscr{O}(\sum n_i p_i)$ where the p_i are regular points distinct from the associated points of $\mathscr{F}_i$. Use this avatar of $\mathscr{L}$, and perhaps induction on the number of p_i.

In fact, all proper curves over k are projective (Remark 18.7.2), so "projective" can be replaced by "proper" in Exercise 18.4.S. In this guise, we will use Exercise 18.4.S when discussing intersection theory in Chapter 20.

18.4.9. $\star$ **Numerical equivalence, the Néron-Severi group, nef line bundles, and the nef and ample cones.**

The notion of a degree on a line bundle leads to important and useful notions. Suppose X is a proper k-variety, and $\mathscr{L}$ is an invertible sheaf on X. If $i : C \hookrightarrow X$ is a one-dimensional closed subscheme of X, define the degree of $\mathscr{L}$ on C by $\deg_C \mathscr{L} := \deg_C i^* \mathscr{L}$. If $\deg_C \mathscr{L} = 0$ for all C, we say that $\mathscr{L}$ is **numerically trivial**.

18.4.T. EASY EXERCISE.
(a) Show that $\mathscr{L}$ is numerically trivial if and only if $\deg_C \mathscr{L} = 0$ for all *integral*

curves C in X.
(b) Show that if $\pi : X \to Y$ is a proper morphism, and $\mathscr{L}$ is a numerically trivial invertible sheaf on Y, then $\pi^*\mathscr{L}$ is numerically trivial on X.
(c) Show that $\mathscr{L}$ is numerically trivial if and only if $\mathscr{L}$ is numerically trivial on each of the irreducible components of X.
(d) Show that if $\mathscr{L}$ and $\mathscr{L}'$ are numerically trivial, then $\mathscr{L} \otimes \mathscr{L}'$ and $\mathscr{L}^\vee$ are both numerically trivial.

18.4.10. *Numerical equivalence.* By part (d), the numerically trivial invertible sheaves form a subgroup of $\operatorname{Pic} X$, denoted $\operatorname{Pic}^\tau X$. The resulting equivalence on line bundles is called **numerical equivalence**. Two lines bundles equivalent modulo the subgroup of numerically trivial line bundles are called **numerically equivalent**. A property of invertible sheaves stable under numerical equivalence is said to be a *numerical property*. We will see that "nefness" and ampleness are numerical properties (Definition 18.4.11 and Remark 20.4.2 respectively).

We will later define the *Néron-Severi group* $NS(X)$ of X as $\operatorname{Pic} X$ modulo algebraic equivalence (Exercise 24.7.5). (We will define algebraic equivalence once we have discussed flatness.) The highly nontrivial **Néron-Severi Theorem** (or **Theorem of the Base**) states that $NS(X)$ is a finitely generated group. (The proof is quite difficult; see [**Kl1**, p. 334, Prop. 3]. For a simpler proof over $\mathbb{C}$, see [**GH1**, p. 462].) The group $\operatorname{Pic} X / \operatorname{Pic}^\tau X$ is denoted $N^1(X)$. We will see (in §24.7.5) that it is a quotient of $NS(X)$, so it is also finitely generated. As the group $N^1(X)$ is clearly abelian and torsion-free, it is finite free $\mathbb{Z}$-module (by the classification of finitely generated modules over a principal ideal domain, see §0.3). The rank of $N^1(X)$ is called the **Picard number**, and is denoted $\rho(X)$ (although we won't have need of this notion, except in our discussion of the Hodge Index Theorem in §20.2.9). For example, $\rho(\mathbb{P}^n) = 1$ and $\rho((\mathbb{P}^1)^n) = n$. We define $N^1_{\mathbb{Q}}(X) := N^1(X) \otimes_{\mathbb{Z}} \mathbb{Q}$ (so $\rho(X) = \dim_{\mathbb{Q}} N^1_{\mathbb{Q}}(X)$), and call the elements of this group **$\mathbb{Q}$-line bundles**, for lack of any common term in the literature.

18.4.U. ⋆⋆ EXERCISE (FINITENESS OF PICARD NUMBER IN THE COMPLEX CASE, ONLY FOR THOSE WITH SUFFICIENT BACKGROUND). Show (without the Néron-Severi Theorem) that if X is a complex proper variety, then $\rho(X)$ is finite, by interpreting it as a subquotient of $H^2(X, \mathbb{Z})$. Hint: show that the image of $(\mathscr{L}, C)$ under the map $H^2(X, \mathbb{Z}) \times H_2(X, \mathbb{Z}) \to H_0(X, \mathbb{Z}) \to \mathbb{Z}$ is $\deg_C \mathscr{L}$. Hint: figure out how to reduce to the case where C is a smooth projective curve, then use Exercise 18.4.G.

18.4.11. *Definition.* We say that an invertible sheaf $\mathscr{L}$ is **numerically effective**, or **nef** if for all such C, $\deg_C \mathscr{L} \geq 0$. Clearly nefness is a numerical property.

18.4.V. EASY EXERCISE (CF. EXERCISE 18.4.T).
(a) Show that $\mathscr{L}$ is nef if and only if $\deg_C \mathscr{L} \geq 0$ for all *integral* curves C in X.
(b) Show that if $\pi : X \to Y$ is a proper morphism, and $\mathscr{L}$ is a nef invertible sheaf on Y, then $\pi^*\mathscr{L}$ is nef on X. (Hint: Exercise 18.4.F will be needed.)
(c) Show that $\mathscr{L}$ is nef if and only if $\mathscr{L}$ is nef on each of the irreducible components of X.
(d) Show that if $\mathscr{L}$ and $\mathscr{L}'$ are nef, then $\mathscr{L} \otimes \mathscr{L}'$ is nef. Thus the nef elements of $\operatorname{Pic} X$ form a semigroup.

(e) Show that ample invertible sheaves are nef.
(f) Suppose $n \in \mathbb{Z}^{\geq 0}$. Show that $\mathscr{L}$ is nef if and only if $\mathscr{L}^{\otimes n}$ is nef.

18.4.W. EXERCISE. Define what it means for a $\mathbb{Q}$-line bundle to be nef. Show that the nef $\mathbb{Q}$-line bundles form a closed cone in $N^1_{\mathbb{Q}}(X)$. This is called the **nef cone**.

18.4.X. EXERCISE. Describe the nef cones of $\mathbb{P}^2_k$ and $\mathbb{P}^1_k \times_k \mathbb{P}^1_k$. (Notice in the latter case that the two boundaries of the cone correspond to linear series contracting one of the $\mathbb{P}^1$'s. This is true in general: informally speaking, linear series corresponding to the boundaries of the cone give interesting contractions. Another example will be given in Exercise 20.2.F.)

It is a surprising fact that whether an invertible sheaf $\mathscr{L}$ on X is ample depends only on its class in $N^1_{\mathbb{Q}}(X)$, i.e., on how it intersects the curves in X. Because of this (as for any $n \in \mathbb{Z}^{\geq 0}$, $\mathscr{L}$ is ample if and only if $\mathscr{L}^{\otimes n}$ is ample, see Theorem 16.6.2), it makes sense to define when a $\mathbb{Q}$-line bundle is ample. Then by Exercise 16.6.H, the ample divisors form a cone in $N^1_{\mathbb{Q}}(X)$, necessarily contained in the nef cone by Exercise 18.4.V(e). It turns out that if X is projective, the ample divisors are precisely the interior of the nef cone. The new facts in this paragraph are a consequence of Kleiman's numerical criterion for ampleness, Theorem 20.4.6.

18.5 A first glimpse of Serre duality

A common version of Riemann-Roch involves Serre duality, which unlike Riemann-Roch is *hard*.

18.5.1. Theorem (Serre duality for smooth projective varieties). — *Suppose X is a geometrically irreducible smooth projective k-variety, of dimension n. Then there is an invertible sheaf ω_X (or simply ω) on X such that*

$$h^i(X, \mathscr{F}) = h^{n-i}(X, \omega_X \otimes \mathscr{F}^\vee)$$

for all $i \in \mathbb{Z}$ and all finite rank locally free sheaves $\mathscr{F}$.

The invertible sheaf ω_X is an example of a *dualizing sheaf*, which will be formally defined in §30.1.4. We will see in Chapter 30 that Theorem 18.5.1 is a consequence of a perfect pairing

$$H^i(X, \mathscr{F}) \times H^{n-i}(X, \omega_X \otimes \mathscr{F}^\vee) \to H^n(X, \omega_X) \cong k, \tag{18.5.1.1}$$

and that smoothness can be relaxed somewhat.

18.5.2. *Further miracle: the sheaf of algebraic volume forms is Serre-dualizing.* The invertible sheaf ω_X turns out to be the *canonical sheaf* (or *canonical bundle*) $\mathscr{K}_X$, which is defined as the determinant of the cotangent bundle $\Omega_{X/k}$ of X (see §21.5.3); we will define the cotangent in Chapter 21. We connect the dualizing sheaf to the canonical bundle in §30.4; see Desideratum 30.1.1.

18.5.3. *Back to Riemann-Roch.* For the purposes of restating Riemann-Roch for a curve C, it suffices to note that $h^1(C, \mathscr{L}) = h^0(C, \omega_C \otimes \mathscr{L}^\vee)$. Then the Riemann-Roch formula can be rewritten as

$$h^0(C, \mathscr{L}) - h^0(C, \omega_C \otimes \mathscr{L}^\vee) = \deg \mathscr{L} - p_a(C) + 1.$$

18.5.A. EXERCISE (ASSUMING SERRE DUALITY). Suppose C is a geometrically integral smooth curve over k.
(a) Show that $h^0(C, \omega_C)$ is the genus g of C.
(b) Show that $\deg \omega_C = 2g - 2$. (Hint: Riemann-Roch for $\mathscr{L} = \omega_C$.)

18.5.4. *Example.* If $C = \mathbb{P}^1_k$, Exercise 18.5.A implies that $\omega_C \cong \mathscr{O}(-2)$. Indeed, $h^1(\mathbb{P}^1, \mathscr{O}(-2)) = 1$. Moreover, we also have a natural perfect pairing (cf. (18.5.1.1))

$$H^0(\mathbb{P}^1, \mathscr{O}(n)) \times H^1(\mathbb{P}^1, \mathscr{O}(-2-n)) \to k.$$

We can interpret this pairing as follows. If $n < 0$, both factors on the left are 0, so we assume $n > 0$. Then $H^0(\mathbb{P}^1, \mathscr{O}(n))$ corresponds to homogeneous degree n polynomials in x and y, and $H^1(\mathbb{P}^1, \mathscr{O}(-2-n))$ corresponds to homogeneous degree $-2-n$ Laurent polynomials in x and y so that the degrees of x and y are both at most $n-1$ (see Remark 18.3.1). You can quickly check that the dimension of both vector spaces are $n+1$. The pairing is given as follows: multiply the polynomial by the Laurent polynomial, to obtain a Laurent polynomial of degree -2. Read off the coefficient of $x^{-1}y^{-1}$. (This works more generally for $\mathbb{P}^n_k$; see the discussion after the statement of Theorem 18.1.3.)

18.5.B. EXERCISE (AMPLE DIVISORS ON A CONNECTED SMOOTH PROJECTIVE VARIETY ARE CONNECTED). Suppose X is a connected smooth projective $\overline{k}$-variety of dimension at least 2, and D is an effective ample divisor. Show that D is connected. (Hint: Suppose $D = V(s)$, where s is a section of an ample invertible sheaf. Then $V(s^n) = V(s)$ for all $n > 0$, so we may replace $\mathscr{L}$ with a high power of our choosing. Use the long exact sequence for $0 \to \mathscr{O}_X(-nD) \to \mathscr{O}_X \to \mathscr{O}_{V(s^n)} \to 0$ to show that for $n \gg 0$, $h^0(\mathscr{O}_{V(s^n)}) = 1$.)

Once we know that Serre duality holds for Cohen-Macaulay projective schemes (§30.3), this result will automatically extend to these schemes. (A related result is Exercise 18.6.Q, which doesn't use Serre duality.) On the other hand, the result is false if X is the union of two 2-planes in $\mathbb{P}^4$ meeting at a point (why?), so this will imply that this X is not Cohen-Macaulay. (We will show this in another way in Counterexample 26.2.2.)

18.5.5. $\star$ **Classification of vector bundles on $\mathbb{P}^1_k$.**

As promised in Example 17.2.4, we classify the vector bundles on $\mathbb{P}^1_k$. We discuss this in §18.5 because you might use Serre duality at one step (although you might not).

18.5.6. Theorem. — *If $\mathscr{E}$ is a rank r vector bundle on $\mathbb{P}^1_k$, then $\mathscr{E} \cong \mathscr{O}(a_1) \oplus \cdots \oplus \mathscr{O}(a_r)$, for a unique nondecreasing sequence of integers $a_1, \ldots, a_r$.*

This result was proved independently many times (see for example Dedekind and Weber's article [**DW**, §22]), and is a special case of a theorem of Grothendieck, [**Gr2**]. It is sometimes called Grothendieck's Theorem, because Grothendieck doesn't have enough theorems named after him.

18.5.7. For $\mathbb{P}^n_k$ more generally, the case $r = 1$ was shown in §14.2.9, but the statement is false for $r > 1$ and $n > 1$. A counterexample is given in Exercise 21.4.G. One true generalization is a theorem of Horrocks, which states that a finite rank locally free sheaf $\mathscr{E}$ on $\mathbb{P}^n_k$ splits precisely when "all the middle cohomology of all of its twists is zero" — when $H^i(\mathbb{P}^n_k, \mathscr{E}(m)) = 0$ for all $m \in \mathbb{Z}$. This has the surprising consequence that if $n \geq 2$, then a finite rank locally free sheaf on $\mathbb{P}^n_k$) splits if and only if its restriction to some previously chosen 2-plane ($\mathbb{P}^2_k$) splits — all of the additional complication turns up already in dimension 2. See [**OSS**, §2.3] for more.

Proof. Note that the classification makes no reference to cohomology, but the proof uses cohomology in an essential way. It is possible to prove Theorem 18.5.6 with no cohomological machinery (see for example [**HM**] or [**GW**, p. 314-5]), or with more cohomological machinery (see for example [**Ha1**, Ex. V.2.6]).

18.5.C. EXERCISE. Suppose that $\mathscr{E} \cong \mathscr{O}(a_1)\oplus\cdots\oplus\mathscr{O}(a_r)$, with $a_1 \leq \cdots \leq a_r$. Show that the a_i can be determined from the numbers $\dim_k \operatorname{Hom}(\mathscr{O}(m), \mathscr{E})$, as m ranges over the positive integers. Use this to show the uniqueness part of Theorem 18.5.6.

We now begin the proof, by induction on r. Fix a rank r locally free sheaf $\mathscr{E}$. The case $r = 1$ was established in §14.2.9, so we assume $r > 1$.

18.5.D. EXERCISE. Show that for $m \ll 0$, $\operatorname{Hom}(\mathscr{O}(m), \mathscr{E}) > 0$, and that for $m \gg 0$, $\operatorname{Hom}(\mathscr{O}(m), \mathscr{E}) = 0$. Hint: show that $\operatorname{Hom}(\mathscr{O}(m), \mathscr{E}) = \mathscr{E}(-m)$ using Exercise 13.1.F. Use Serre vanishing (Theorem 18.1.4(ii)) for the first part. Feel free to use Serre duality and Serre vanishing for the second part. But you may prefer to come up with an argument without Serre duality, to avoid invoking something we have not yet proved.

Thus there is some a_r for which $\operatorname{Hom}(\mathscr{O}(a_r), \mathscr{E}) > 0$, but for which $\operatorname{Hom}(\mathscr{O}(m), \mathscr{E}) = 0$ for all $m > a_r$. Choose a nonzero map $\phi : \mathscr{O}(a_r) \to \mathscr{E}$.

18.5.E. EXERCISE. Show that ϕ is an injection. (Hint: $\mathscr{O}(a_r)$ is torsion-free, and thus the kernel is torsion-free.)

18.5.F. EXERCISE. Let $\mathscr{F}$ be the cokernel of ϕ. Show that $\mathscr{F}$ is locally free. Hint: Exercise 13.7.G(c) gives an exact sequence $0 \to \mathscr{F}_{\text{tors}} \to \mathscr{F} \to \mathscr{F}_{\text{lf}} \to 0$, where $\mathscr{F}_{\text{lf}}$ is locally free. Let $\mathscr{L}$ be the kernel of the surjection $\mathscr{E} \to \mathscr{F}_{\text{lf}}$. Show that $\mathscr{L}$ is locally free, and thus is isomorphic to $\mathscr{O}(N)$ for some N. Show that there is a nonzero map $\phi' : \mathscr{O}(a_r) \to \mathscr{O}(N)$. Show (using the same idea as the previous exercise) that this nonzero map ϕ' must be an injection, so $a_r \leq N$. Show that $N \leq a_r$ because there is a nonzero map $\mathscr{O}(N) \to \mathscr{E}$ (recall how a_r was chosen). Show that ϕ' is an isomorphism, and thus that $\mathscr{F} = \mathscr{F}_{\text{lf}}$.

By our inductive hypothesis, we have $\mathscr{F} = \mathscr{O}(a_1) \oplus \cdots \oplus \mathscr{O}(a_{r-1})$, where $a_1 \leq \cdots \leq a_{r-1}$, so we have a short exact sequence

$$0 \to \mathscr{O}(a_r) \to \mathscr{E} \to \mathscr{O}(a_1) \oplus \cdots \oplus \mathscr{O}(a_{r-1}) \to 0. \tag{18.5.7.1}$$

We next show that $a_r \geq a_i$ for $i < r$. We are motivated by the fact that for a quasicoherent sheaf $\mathscr{G}$ on $\mathbb{P}^1$,

$$\begin{aligned} \operatorname{Hom}_{\mathbb{P}^1}(\mathscr{O}(a_r+1), \mathscr{G}) &= \operatorname{Hom}_{\mathbb{P}^1}(\mathscr{O}, \mathscr{G}(-a_r-1)) \quad \text{(Exercise 13.7.B)} \\ &= H^0(\mathbb{P}^1, \mathscr{G}(-a_r-1)). \end{aligned}$$

Tensor (18.5.7.1) with $\mathscr{O}(-a_r-1)$ (preserving exactness, by Exercise 13.1.E), and take the long exact sequence in cohomology. Part of the long exact sequence is

$$H^0(\mathbb{P}^1, \mathscr{E}(-a_r-1)) \to H^0(\mathbb{P}^1, \mathscr{F}(-a_r-1)) \to H^1(\mathbb{P}^1, \mathscr{O}(-1)).$$

Notice that $H^0(\mathbb{P}^1, \mathscr{E}(-a_r-1)) = 0$ (as $\operatorname{Hom}(\mathscr{O}(a_r+1), \mathscr{E}) = 0$, by the definition of a_r), and $H^1(\mathbb{P}^1, \mathscr{O}(-1)) = 0$ (Theorem 18.1.3). Thus

$$\begin{aligned} 0 &= H^0(\mathbb{P}^1, \mathscr{F}(-a_r-1)) \\ &= H^0(\mathbb{P}^1, \oplus_i \mathscr{O}(a_i - a_r - 1)) \\ &= \oplus_i H^0(\mathbb{P}^1, \mathscr{O}(a_i - a_r - 1)). \end{aligned}$$

Hence $a_i - a_r - 1 < 0$ (by Exercise 14.1.A), so $a_r \geq a_i$ as desired.

Finally, we wish to show that exact sequence (18.5.7.1) expresses $\mathscr{E}$ as a direct sum (of the subsheaf and quotient sheaf). For simplicity, we focus on the case $r = 2$, and return to the general case in Exercise 18.5.I.

18.5.G. EXERCISE. Show that the transition functions for the vector bundle, in appropriate coordinates, are given by

$$\begin{pmatrix} t^{-a_2} & \alpha(t) \\ 0 & t^{-a_1} \end{pmatrix},$$

where $\alpha(t)$ is a Laurent polynomial in t. Hint/reminder: Recall that all vector bundles on $\mathbb{A}^1_k$ are trivial, Exercise 13.2.C. Transition matrices for extensions of one vector bundle (with known transition matrices) by another were discussed in Exercise 13.5.A.

18.5.H. EXERCISE. Implicit in the above 2×2 matrix is a choice of a basis of a rank 2 free module M over the ring $k[t]$ corresponding to one of the standard affine open subsets $\operatorname{Spec} k[t]$, and a rank 2 free module M' over the ring $k[1/t]$ over the other standard open subset. Show that by an appropriate "upper-triangular" change of basis of M, you can arrange for $\alpha(t)$ to have no monomials of degree $\geq -a_2$. Show that by an appropriate "upper-triangular" change of basis of M', you can arrange for $\alpha(t)$ to have no monomials of degree $\leq -a_1$. Thus by choosing bases of M and M' appropriately, we can take $\alpha(t) = 0$. Show that this implies that $\mathscr{E} = \mathscr{O}(a_1) \oplus \mathscr{O}(a_2)$.

18.5.I. EXERCISE. Finish the proof of Theorem 18.5.6 for arbitrary r. Hint: there are no new ideas beyond the case $r = 2$.

□

18.6 Hilbert functions, Hilbert polynomials, and genus

If $\mathscr{F}$ is a coherent sheaf on a projective k-scheme $X \subset \mathbb{P}^n$, define the **Hilbert function of** $\mathscr{F}$ by

$$h_{\mathscr{F}}(m) := h^0(X, \mathscr{F}(m)).$$

The **Hilbert function of** X is the Hilbert function of the structure sheaf. The ancients were aware that the Hilbert function is "eventually polynomial", i.e., for large enough m, it agrees with some polynomial. This polynomial contains lots of interesting geometric information, as we will soon see. In modern language, we expect that this "eventual polynomiality" arises because the Euler characteristic should be a polynomial, and that for $m \gg 0$, the higher cohomology vanishes. This is indeed the case, as we now verify.

18.6.1. Theorem. — *If $\mathscr{F}$ is a coherent sheaf on a projective k-scheme $X \hookrightarrow \mathbb{P}^n_k$, $\chi(X, \mathscr{F}(m))$ is a polynomial of degree equal to* $\dim \operatorname{Supp} \mathscr{F}$*. Hence by Serre vanishing (Theorem 18.1.4 (ii)), for $m \gg 0$, $h^0(X, \mathscr{F}(m))$ is a polynomial $p_{\mathscr{F}}(m)$ of degree* $\dim \operatorname{Supp} \mathscr{F}$*. In particular, for $m \gg 0$, $h^0(X, \mathscr{O}_X(m))$ is polynomial with degree equal to* $\dim X$.

18.6.2. Definition. The polynomial $p_{\mathscr{F}}(m)$ defined in Theorem 18.6.1 is called the **Hilbert polynomial**. If $X \subset \mathbb{P}^n$ is a projective k-scheme, define $p_X(m) := p_{\mathscr{O}_X}(m)$.

In Theorem 18.6.1, $\mathscr{O}_X(m)$ is the restriction or pullback of $\mathscr{O}_{\mathbb{P}^n_k}(1)$. Both the degree of the 0 polynomial and the dimension of the empty set is defined to be -1. In particular, the only coherent sheaf with Hilbert polynomial 0 is the zero-sheaf.

This argument uses the notion of associated points of a coherent sheaf on a locally Noetherian scheme, §13.6.5. (The resolution given by the Hilbert Syzygy Theorem, §15.3.2, can give a shorter proof; but we haven't proved the Hilbert Syzygy Theorem.)

Proof. Define $p_{\mathscr{F}}(m) = \chi(X, \mathscr{F}(m))$. We will show that $p_{\mathscr{F}}(m)$ is a polynomial of the desired degree.

We first use Exercise 18.2.H to reduce to the case where k is algebraically closed, and in particular infinite. (This is one of those cases where even if you are concerned with potentially arithmetic questions over some non-algebraically closed field like $\mathbb{F}_p$, you are forced to consider the "geometric" situation where the base field is algebraically closed.)

The coherent sheaf $\mathscr{F}$ has a finite number of associated points. We show a useful fact that we will use again.

18.6.A. EXERCISE. Suppose X is a projective k-scheme with k infinite, and $\mathscr{F}$ is a coherent sheaf on X. Show that if $\mathscr{L}$ is a very ample invertible sheaf on X, then there is an effective divisor D on X with $\mathscr{L} \cong \mathscr{O}(D)$, and where D does not meet the associated points of $\mathscr{F}$. (Hint: show that given any finite set of points of $\mathbb{P}^n_k$, there is a hyperplane not containing any of them. This is a variant of the key step in Exercise 11.3.C(c).)

Thus there is a hyperplane $x = 0$ ($x \in \Gamma(X, \mathscr{O}(1))$) missing this finite number of points. (This is where we use the infinitude of k.)

Then the map $\mathscr{F}(-1) \xrightarrow{\times x} \mathscr{F}$ is injective (on any affine open subset, $\mathscr{F}$ corresponds to a module, and x is not a zerodivisor on that module, as it doesn't vanish at any associated point of that module, see Theorem 5.5.10(c)). Thus we have a short exact sequence

$$0 \longrightarrow \mathscr{F}(-1) \longrightarrow \mathscr{F} \longrightarrow \mathscr{G} \longrightarrow 0 \tag{18.6.2.1}$$

where $\mathscr{G}$ is a coherent sheaf.

18.6.B. EXERCISE. Show that $\operatorname{Supp}\mathscr{G} = (\operatorname{Supp}\mathscr{F}) \cap V(x)$. (Hint: show that $\mathscr{F}(-1) \to \mathscr{F}$ is an isomorphism away from $V(x)$, and hence $\mathscr{G} = 0$ on this locus. If $p \in V(x)$, show that the $\mathscr{F}(-1)|_p \to \mathscr{F}|_p$ is the 0 map, and hence $\mathscr{F}|_p \to \mathscr{G}|_p$ is an isomorphism.)

Hence $V(x)$ meets all positive-dimensional components of $\operatorname{Supp}\mathscr{G}$ (Exercise 11.3.C(a)), so $\dim \operatorname{Supp}\mathscr{G} = \dim \operatorname{Supp}\mathscr{F} - 1$ by Krull's Principal Ideal Theorem 11.3.3 unless $\mathscr{F} = 0$ (in which case we already know the result, so assume this is not the case).

Twisting (18.6.2.1) by $\mathscr{O}(m)$ yields

$$0 \longrightarrow \mathscr{F}(m-1) \longrightarrow \mathscr{F}(m) \longrightarrow \mathscr{G}(m) \longrightarrow 0$$

Euler characteristics are additive in exact sequences, from which $p_{\mathscr{F}}(m) - p_{\mathscr{F}}(m-1) = p_{\mathscr{G}}(m)$. Now $p_{\mathscr{G}}(m)$ is a polynomial of degree $\dim \operatorname{Supp}\mathscr{F} - 1$.

The result is then a consequence from the following elementary fact about polynomials in one variable.

18.6.C. EXERCISE. Suppose f and g are functions on the integers, $f(m+1) - f(m) = g(m)$ for all m, and $g(m)$ is a polynomial of degree $d \geq 0$. Show that f is a polynomial of degree $d+1$.

$\square$

Example 1. The Hilbert polynomial of projective space is $p_{\mathbb{P}^n}(m) = \binom{n+m}{m}$, where we interpret this as the polynomial $(m+1)\cdots(m+n)/n!$.

Example 2. Suppose H is a degree d hypersurface in $\mathbb{P}^n$. Then from the closed subscheme exact sequence

$$0 \longrightarrow \mathscr{O}_{\mathbb{P}^n}(-d) \longrightarrow \mathscr{O}_{\mathbb{P}^n} \longrightarrow \mathscr{O}_H \longrightarrow 0,$$

we have

$$p_H(m) = p_{\mathbb{P}^n}(m) - p_{\mathbb{P}^n}(m-d) = \binom{n+m}{n} - \binom{m+n-d}{n}. \tag{18.6.2.2}$$

(Note: implicit in this argument is the fact that if $i : H \hookrightarrow \mathbb{P}^n$ is the closed embedding, then $(i_*\mathscr{O}_H) \otimes \mathscr{O}_{\mathbb{P}^n}(m) \cong i_*(\mathscr{O}_H \otimes i^*\mathscr{O}_{\mathbb{P}^n}(m))$. This follows from the projection formula, Exercise 16.3.H(b).)

18.6.D. EXERCISE. Show that the twisted cubic (in $\mathbb{P}^3$) has Hilbert polynomial $3m+1$. (The twisted cubic was defined in Exercise 8.2.A.)

18.6.E. EXERCISE. More generally, find the Hilbert polynomial for the dth Veronese embedding of $\mathbb{P}^n$ (i.e., the closed embedding of $\mathbb{P}^n$ in a bigger projective space by way of the line bundle $\mathscr{O}(d)$, §8.2.6).

18.6.F. EXERCISE (TO BE USED SEVERAL TIMES IN CHAPTER 19). Suppose $X \subset Y \subset \mathbb{P}^n_k$ are a sequence of closed embeddings.
(a) Show that $p_X(m) \leq p_Y(m)$ for $m \gg 0$. Hint: let $\mathscr{I}_{X/Y}$ be the ideal sheaf of X in Y. Consider the exact sequence

$$0 \longrightarrow \mathscr{I}_{X/Y}(m) \longrightarrow \mathscr{O}_Y(m) \longrightarrow \mathscr{O}_X(m) \longrightarrow 0.$$

(b) If $p_X(m) = p_Y(m)$ for $m \gg 0$, show that $X = Y$. Hint: Show that if the Hilbert polynomial of a coherent sheaf $\mathscr{F}$ is 0, then $\mathscr{F} = 0$. (Handy trick: For $m \gg 0$, $\mathscr{F}(m)$ is generated by global sections.) Apply this to $\mathscr{F} = \mathscr{I}_{X/Y}$.

From the Hilbert polynomial, we can extract many invariants, of which two are particularly important. The first is the *degree*, and the second is the arithmetic genus (§18.6.6). The **degree of a projective k-scheme of dimension n** is defined to be leading coefficient of the Hilbert polynomial (the coefficient of m^n) times $n!$.

Using the examples above, we see that the degree of $\mathbb{P}^n$ in itself is 1. The degree of the twisted cubic is 3.

18.6.G. EXERCISE. Show that the degree is always an integer. Hint: by induction, show that any polynomial in m of degree k taking on only integer values must have coefficient of m^k an integral multiple of $1/k!$. Hint for this: if $f(x)$ takes on only integral values and is of degree k, then $f(x+1) - f(x)$ takes on only integral values and is of degree $k - 1$.

18.6.H. EXERCISE. Show that the degree of a degree d hypersurface (Definition 8.2.2) is d (preventing a notational crisis).

18.6.I. EXERCISE. Suppose a curve C is embedded in projective space via an invertible sheaf of degree d (as defined in §18.4.4). In other words, this line bundle determines a closed embedding. Show that the degree of C under this embedding is d, preventing another notational crisis. Hint: Riemann-Roch, Exercise 18.4.B. (An earlier notation crisis was also averted in Exercise 17.4.F.)

18.6.J. EXERCISE. Show that the degree of the dth Veronese embedding of $\mathbb{P}^n$ is d^n.

18.6.K. EXERCISE (BÉZOUT'S THEOREM, GENERALIZING EXERCISES 8.2.E AND 16.4.H). Suppose X is a projective scheme of dimension at least 1, and H is a hypersurface not containing any associated points of X. (For example, if X is reduced and thus has no embedded points, we are just requiring H not to contain any irreducible components of X.) Show that $\deg(H \cap X) = (\deg H)(\deg X)$. (As an example, we have *Bézout's theorem for plane curves*: if C and D are plane curves of degrees m and n respectively, with no common components, then C and D meet at mn points, counted with appropriate multiplicity.)

18.6.3. This is a very handy result! For example: if two projective plane curves of degree m and degree n share no irreducible components, then they intersect in mn points, counted with appropriate multiplicity. (To apply Exercise 18.6.K, you need to know that plane curves have no embedded points. You can either do this using Exercise 5.5.I, or save time by assuming that one of the curves is reduced.)

The notion of multiplicity of intersection is just the degree of the intersection as a k-scheme.

18.6.L. EXERCISE. Suppose C is a degree 1 curve in $\mathbb{P}^3_k$ (or more precisely, a degree 1 pure-one-dimensional closed subscheme of $\mathbb{P}^3_k$). Show that C is a line. Hint: reduce to the case $k = \overline{k}$. Suppose p and q are distinct closed points on C. Use Bézout's Theorem (Exercise 18.6.K) to show that any hyperplane containing p and q must contain C, and thus that $C \subset \overline{pq}$.

18.6.M. FUN EXERCISE. Let k be a field, which we assume to be algebraically closed for convenience (although you are free to remove this hypothesis if you wish). Suppose C is a degree d integral cuve in $\mathbb{P}^N_k$ with $N \geq d$. Show that C is contained in a linear $\mathbb{P}^d_k \subset \mathbb{P}^N_k$. (Exercise 18.6.L is not quite a special case of this problem, but the hint may still be helpful.)

18.6.N. EXERCISE (A FORM OF BÉZOUT'S THEOREM). Classically, the degree of a complex projective variety of dimension n was defined as follows. We slice the variety with n generally chosen hyperplanes. Then the intersection will be a finite number of reduced points, by Exercise 12.4.C (a consequence of Bertini's Theorem 12.4.2). The degree is this number of points. Use Bézout's theorem to make sense of this in a way that agrees with our definition of degree. You will need to assume that k is infinite.

Thus the classical definition of the degree, which involved making a choice and then showing that the result is independent of choice, has been replaced by making a cohomological definition involving Euler characteristics. This should remind you of how we got around to "correctly" understanding the degree of a line bundle. It was traditionally defined as the degree of a divisor of any nonzero rational section (Important Exercise 18.4.C), and we found a better definition in terms of Euler characteristics (§18.4.4).

18.6.4. $\star\star$ *Aside: Connection to the topological definition of degree.* Another definition of degree of a dimension d complex projective variety $X \subset \mathbb{P}^n_{\mathbb{C}}$ is as the number d such that $[X]$ is d times the "positive" generator of $H_{2d}(\mathbb{P}^n_{\mathbb{C}}, \mathbb{Z})$. over $\mathbb{C}$. You can show this by induction on d as follows. Suppose X is a complex projective variety of dimension d. For a generally chosen hyperplane, $H \cap X$ is a complex projective variety of (complex) dimension $d-1$. (Do you see why?) Show that $[H \cap X] = c_1(\mathscr{O}(1)) \cup [X]$ in $H_{2(d-1)}(X, \mathbb{Z})$, by suitably generalizing the solution to Exercise 18.4.G. (A further generalization is given in §20.1.7.) For this reason, $c_1(\mathscr{O}(1))$ is often called the "hyperplane class".

18.6.5. *Revisiting an earlier example.* We revisit the enlightening example of Example 9.3.3 and §17.4.8: let $k = \mathbb{Q}$, and consider the parabola $x = y^2$. We intersect it with the four lines, $x = 1$, $x = 0$, $x = -1$, and $x = 2$, and see that we get 2 each time (counted with the same convention as with the last time we saw this example).

If we intersect it with $y = 2$, we only get one point — but that's because this isn't a projective curve, and we really should be doing this intersection on $\mathbb{P}^2_k$, and in this case, the conic meets the line in two points, one of which is "at ∞".

18.6.O. EXERCISE. Show that the degree of the d-fold Veronese embedding of $\mathbb{P}^n$ is d^n in a different way from Exercise 18.6.J as follows. Let $\nu_d : \mathbb{P}^n \to \mathbb{P}^N$ be the Veronese embedding. To find the degree of the image, we intersect it with n hyperplanes in $\mathbb{P}^N$ (scheme-theoretically), and find the number of intersection points (counted with multiplicity). But the pullback of a hyperplane in $\mathbb{P}^N$ to $\mathbb{P}^n$ is a degree d hypersurface. Perform this intersection in $\mathbb{P}^n$, and use Bézout's theorem (Exercise 18.6.K).

18.6.P. EXERCISE (DEGREE IS ADDITIVE FOR UNIONS). Suppose X and Y are two d-dimensional closed subschemes of $\mathbb{P}^n_k$, with no d-dimensional irreducible components in common. Show that $\deg X \cup Y = \deg X + \deg Y$.

18.6.6. Arithmetic genus, again.

There is another central piece of information residing in the Hilbert polynomial. Notice that $1 - p_X(0) = 1 - \chi(X, \mathscr{O}_X)$ is the arithmetic genus (§18.4.2), an *intrinsic* invariant of the scheme X, independent of the projective embedding.

Imagine how amazing this must have seemed to the ancients: they defined the Hilbert function by counting how many "functions of various degrees" there are; then they noticed that when the degree gets large, it agrees with a polynomial; and then when they plugged 0 into the polynomial — extrapolating backwards, to where the Hilbert function and Hilbert polynomials didn't agree — they found a magic invariant! Furthermore, in the case when X is a complex curve, this invariant was basically the topological genus!

We can now see a large family of curves over an algebraically closed field that is provably not $\mathbb{P}^1$! Note that the Hilbert polynomial of $\mathbb{P}^1$ is $(m+1)/1 = m+1$, so $\chi(\mathscr{O}_{\mathbb{P}^1}) = 1$. Suppose C is a degree d curve in $\mathbb{P}^2$. Then the Hilbert polynomial of C is

$$p_{\mathbb{P}^2}(m) - p_{\mathbb{P}^2}(m-d) = (m+1)(m+2)/2 - (m-d+1)(m-d+2)/2.$$

Plugging in $m = 0$ gives us $-(d^2 - 3d)/2$. Thus when $d > 2$, we have a curve that cannot be isomorphic to $\mathbb{P}^1$! (And it is not hard to show that there exists a *regular* degree d curve, Exercise 12.3.E.)

Now from $0 \to \mathscr{O}_{\mathbb{P}^2}(-d) \to \mathscr{O}_{\mathbb{P}^2} \to \mathscr{O}_C \to 0$, using $h^1(\mathscr{O}_{\mathbb{P}^2}(d)) = 0$, we have that $h^0(C, \mathscr{O}_C) = 1$. As $h^0 - h^1 = \chi$, we have

$$h^1(C, \mathscr{O}_C) = (d-1)(d-2)/2. \tag{18.6.6.1}$$

We now revisit an interesting question we first saw in §6.5.10. If k is an algebraically closed field, is every finitely generated transcendence degree 1 extension of k isomorphic to $k(x)$? In that section, we found ad hoc (but admittedly beautiful) examples showing that the answer is "no". But we now have a better answer. The question initially looks like an algebraic question, but we now recognize it as a fundamentally geometric one. There is an integer-valued cohomological invariant of such field extensions that is has good geometric meaning: the genus.

Equation (18.6.6.1) yields examples of curves of genus $0, 1, 3, 6, 10, \dots$ (corresponding to degree 1 or $2, 3, 4, 5, \dots$). This begs some questions, such as: are there curves of other genera? (We will see soon, in §19.5.5, that the answer is yes.) Are there other genus 0 curves? (Not if k is algebraically closed, but sometimes yes otherwise — consider $x^2+y^2+z^2 = 0$ in $\mathbb{P}^2_{\mathbb{R}}$, which has no $\mathbb{R}$-points and hence is not isomorphic to $\mathbb{P}^1_{\mathbb{R}}$ — we will discuss this more in §19.3.) Do we have all the curves of genus 3?

(Almost all, but not quite. We will see more in §19.7.) Do we have all the curves of genus 6? (We are missing "most of them", as is suggested by §19.8.2.)

Caution: The Euler characteristic of the structure sheaf is in incomplete invariant. In particular, it doesn't distinguish between isomorphism classes of irreducible smooth projective varieties. For example, $\mathbb{P}^2$ and $\mathbb{P}^1 \times \mathbb{P}^1$ both have Euler characteristic 1 (see Theorem 18.1.3 and Exercise 18.3.E), but are not isomorphic — $\operatorname{Pic} \mathbb{P}^2 \cong \mathbb{Z}$ (§14.2.7) while $\operatorname{Pic}(\mathbb{P}^1 \times \mathbb{P}^1) \cong \mathbb{Z} \oplus \mathbb{Z}$ (Exercise 14.2.O).

18.6.7. Complete intersections.

18.6.Q. EXERCISE (POSITIVE-DIMENSIONAL COMPLETE INTERSECTIONS ARE CONNECTED). Show that complete intersections of *positive* dimension are connected. (Hint: show that $h^0(X, \mathscr{O}_X) = 1$.) For experts: this argument will even show that they are geometrically connected (§9.5), as h^0 is preserved by field extension (Exercise 18.2.H).

18.6.R. EXERCISE. Find the genus of the complete intersection of 2 quadrics in $\mathbb{P}^3_k$.

18.6.S. EXERCISE. More generally, find the genus of the complete intersection of a degree m surface with a degree n surface in $\mathbb{P}^3_k$. (If $m = 2$ and $n = 3$, you should get genus 4. We will see in §19.8 that in some sense most genus 4 curves arise in this way. Note that Bertini's Theorem 12.4.2 ensures that there *are* regular curves of this form.)

18.6.T. EXERCISE. Show that the rational normal curve of degree d in $\mathbb{P}^d$ is *not* a complete intersection if $d > 2$. (Hint: If it *were* the complete intersection of $d-1$ hypersurfaces, what would the degree of the hypersurfaces be? Why could none of the degrees be 1?)

18.6.U. EXERCISE. Show that the union of two planes in $\mathbb{P}^4$ meeting at a point is not a complete intersection. Hint: it is connected, but you can slice with another plane and get something not connected (see Exercise 18.6.Q).

18.7 ⋆ Serre's cohomological characterization of ampleness

Theorem 16.6.2 gave a number of characterizations of ampleness, in terms of projective geometry, global generation, and the Zariski topology. Here is another characterization, this time cohomological, under Noetherian hypotheses. Because (somewhat surprisingly) we won't use this result, this section is starred.

18.7.1. Theorem (Serre's cohomological criterion for ampleness). — *Suppose A is a Noetherian ring, X is a proper A-scheme, and $\mathscr{L}$ is an invertible sheaf on X. Then the following are equivalent.*

(a-c) *The invertible sheaf $\mathscr{L}$ is ample on X (over A).*

(e) *For all coherent sheaves $\mathscr{F}$ on X, there is an n_0 such that for $n \geq n_0$, $H^i(X, \mathscr{F} \otimes \mathscr{L}^{\otimes n}) = 0$ for all $i > 0$.*

The label (a-c) is intended to reflect the statement of Theorem 16.6.2. We avoid the label (d) because it appeared in Theorem 16.6.6. (Aside: the "properness" assumption cannot be removed, as can be shown by the example of Exercise 18.8.F.) Before getting to the proof, we motivate this result by giving some applications. (As a warm-up, you can give a second solution to Exercise 16.6.G in the Noetherian case, using the affineness of π to show that $H^i(Y, \mathscr{F} \otimes \mathscr{L}^{\otimes m}) = H^i(X, \pi_* \mathscr{F} \otimes \mathscr{L}^{\otimes m})$.)

18.7.A. EXERCISE. Suppose X is a proper A-scheme, and $\mathscr{L}$ is an invertible sheaf on X. Show that $\mathscr{L}$ is ample on X if and only if $\mathscr{L}|_{X^{\text{red}}}$ is ample on X^{red}. Hint: for the "only if" direction, use Exercise 16.6.G. For the "if" direction, let $\mathscr{I}$ be the ideal sheaf cutting out the closed subscheme X^{red} in X. Filter $\mathscr{F}$ by powers of $\mathscr{I}$:

$$0 = \mathscr{I}^r \mathscr{F} \subset \mathscr{I}^{r-1} \mathscr{F} \subset \cdots \subset \mathscr{I} \mathscr{F} \subset \mathscr{F}.$$

(Essentially the same filtration appeared in Exercise 18.4.S, for similar reasons.) Show that each quotient $\mathscr{I}^n \mathscr{F} / \mathscr{I}^{n-1} \mathscr{F}$, twisted by a high enough power of $\mathscr{L}$, has no higher cohomology. Use descending induction on n to show each part $\mathscr{I}^n \mathscr{F}$ of the filtration (and hence in particular $\mathscr{F}$) has this property as well.

18.7.B. EXERCISE. Suppose X is a proper A-scheme, and $\mathscr{L}$ is an invertible sheaf on X. Show that $\mathscr{L}$ is ample on X if and only if $\mathscr{L}$ is ample on each component. Hint: follow the outline of the solution to the previous exercise, taking instead $\mathscr{I}$ as the ideal sheaf of one component. Perhaps first reduce to the case where $X = X^{red}$.

18.7.C. EXERCISE. (In Exercise 19.2.E, we will show that on a smooth projective integral curve, an invertible sheaf is ample if and only if it has positive degree. Use that fact in this exercise. There will be no logical circularity.) Show that a line bundle on a projective curve is ample if and only if it has positive degree on each component.

18.7.2. *Remark: Proper curves are projective.* Serre's criterion for ampleness is the key ingredient for showing that every proper curve over a field is projective. The steps are as follows. (i) Recall that every regular integral proper curve is projective, Exercise 17.4.A. (ii) The hardest step is showing that every reduced integral proper curve C is projective. This is done by choosing a regular point on each irreducible component of C, and letting $\mathscr{L}$ be the corresponding invertible sheaf. Because of Exercise 18.7.C, we hope that $\mathscr{L}$ will be ample. Show that a line bundle on C is ample if its pullback to the normalization of C is ample (a partial converse to Exercise 16.6.G, see for example [**Ha1**, Ex. III.5.7(d)]). Thus our $\mathscr{L}$ is ample. (iii) Show that every reduced proper curve is projective using Exercise 18.7.B. (iv) Show that every proper curve C is projective, using Exercise 18.7.A, after first finding an invertible sheaf on C that will be shown to be ample.

18.7.3. *Very ample versus ample.* The previous exercises don't work with "ample" replaced by "very ample", which shows again how the notion of ampleness is better-behaved than very ampleness.

18.7.4. *Proof of Theorem 18.7.1.* For the fact that (a-c) implies (e), use the fact that $\mathscr{L}^{\otimes N}$ is very ample for some N (Theorem 16.6.2(a)), and apply Serre vanishing (Theorem 18.1.4(ii)) to $\mathscr{F}$, $\mathscr{F} \otimes \mathscr{L}$, ..., and $\mathscr{F} \otimes \mathscr{L}^{\otimes(N-1)}$.

So we now assume (e), and show that $\mathscr{L}$ is ample by criterion (b) of Theorem 16.6.2: we will show that for any coherent sheaf $\mathscr{F}$ on X, $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated for $n \gg 0$.

We begin with a special case: we will show that $\mathscr{L}^{\otimes n}$ is globally generated (i.e., base-point-free) for $n \gg 0$. To do this, it suffices to show that every closed point p has a neighborhood U so that there exists some N_p so that $n \geq N_p$, $\mathscr{L}^{\otimes n}$ is globally generated for all points of U_p. (Reason: by quasicompactness, every closed subset of X contains a closed point, by Exercise 5.1.E. So as p varies over the closed points of X, these U_p cover X. By quasicompactness again, we can cover X by a finite number of these U_p. Let N be the maximum of the corresponding N_p. Then for $n \geq N$, $\mathscr{L}^{\otimes n}$ is globally generated in each of these U_p, and hence on all of X.)

Let p be a closed point of X. For all n, $\mathfrak{m}_p \otimes \mathscr{L}^{\otimes n}$ is coherent (by our Noetherian hypotheses). By (e), there exists some n_0 so that for $n \geq n_0$, $H^1(X, \mathfrak{m}_p \otimes \mathscr{L}^{\otimes n}) = 0$. By the long exact sequence arising from the closed subscheme exact sequence

$$0 \to \mathfrak{m}_p \otimes \mathscr{L}^{\otimes n} \to \mathscr{L}^{\otimes n} \to \mathscr{L}^{\otimes n}|_p \to 0,$$

we have that $\mathscr{L}^{\otimes n}$ is globally generated at p for $n \geq n_0$. By Exercise 15.3.C(b), there is an open neighborhood V_0 of p such that $\mathscr{L}^{\otimes n_0}$ is globally generated at all points of V_0. Thus $\mathscr{L}^{\otimes k n_0}$ is globally generated at all points of V_0 for all positive integers k (using Easy Exercise 15.3.B). For each $i \in \{1, \dots, n_0 - 1\}$, there is an open neighborhood V_i of p such that $\mathscr{L}^{\otimes(n_0+i)}$ is globally generated at all points of V_i (again by Exercise 15.3.C(b)). We may take each V_i to be contained in V_0. By Easy Exercise 15.3.B, $\mathscr{L}^{\otimes(kn_0+n_0+i)}$ is globally generated at every point of V_i (as this is the case for $\mathscr{L}^{\otimes k n_0}$ and $\mathscr{L}^{\otimes(n_0+i)}$). Thus in the open neighborhood $U_p := \cap_{i=0}^{n-1} V_i$, $\mathscr{L}^{\otimes n}$ is globally generated for $n \geq N_p := 2n_0$.

We have now shown that there exists some N such that for $n \geq N$, $\mathscr{L}^{\otimes n}$ is globally generated. Now suppose $\mathscr{F}$ is a coherent sheaf. To conclude the proof, we will show that $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated for $n \gg 0$. This argument has a similar flavor to what we have done so far, so we give it as an exercise.

18.7.D. EXERCISE. Suppose p is a closed point of X.
(a) Show that for $n \gg 0$, $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated at p.
(b) Show that there exists an open neighborhood U_p of p such that for $n \gg 0$, $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated at every point of U_p. Caution: while it is true that by Exercise 15.3.C(b), for each $n \gg 0$, there is some neighborhood V_n of p such that $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated there, it need not be true that

$$\cap_{n \gg 0} V_n \qquad (18.7.4.1)$$

is an open set. You may need to use the fact that $\mathscr{L}^{\otimes n}$ is globally generated for $n \geq N$ to replace (18.7.4.1) by a finite intersection.

18.7.E. EXERCISE. Conclude the proof of Theorem 18.7.1 by showing that $\mathscr{F} \otimes \mathscr{L}^{\otimes n}$ is globally generated for $n \gg 0$. $\square$

18.7.5. *Aside: Serre's cohomological characterization of affineness.* Serre gave a characterization of affineness similar in flavor to Theorem 18.7.1.

18.7.6. Theorem (Serre's cohomological characterization of affineness). — *Suppose X is a Noetherian separated scheme. Then the following are equivalent.*

(a) *The scheme X is affine.*
(b) *For any quasicoherent sheaf $\mathscr{F}$ on X, $H^i(X, \mathscr{F}) = 0$ for all $i > 0$.*
(c) *For any coherent sheaf of ideals $\mathscr{I}$ on X, $H^1(X, \mathscr{I}) = 0$.*

Because we won't use it, we omit the proof. (One is given in [**Ha1**, Thm. III.3.7].) Clearly (a) implies (b) implies (c) (the former from Property **(vi)** of §18.1) without any Noetherian assumptions, so the real substance is in the implication from (c) to (a).

Serre proved an analogous result in complex analytic geometry: Stein spaces are also characterized by the vanishing of cohomology of coherent sheaves.

18.8 Higher direct image sheaves

Cohomology groups were defined for $X \to \operatorname{Spec} A$ where the structure morphism is quasicompact and separated; for any quasicoherent $\mathscr{F}$ on X, we defined $H^i(X, \mathscr{F})$. We will now define a "relative" version of this notion, for quasicompact and separated morphisms $\pi : X \to Y$: for any quasicoherent $\mathscr{F}$ on X, we will define $R^i\pi_*\mathscr{F}$, a quasicoherent sheaf on Y. (Now would be a good time to do Exercise 1.6.H, the FHHF Theorem, if you haven't done it before.)

We have many motivations for doing this. In no particular order:

(1) It "globalizes" what we did before with cohomology.
(2) If $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ is a short exact sequence of quasicoherent sheaves on X, then we know that $0 \to \pi_*\mathscr{F} \to \pi_*\mathscr{G} \to \pi_*\mathscr{H}$ is exact, and higher pushforwards will extend this to a long exact sequence.
(3) We will later see that this will show how cohomology groups vary in families, especially in "nice" situations. Intuitively, if we have a nice family of varieties, and a family of sheaves on them, we could hope that the cohomology varies nicely in families, and in fact in "nice" situations, this is true. (As always, "nice" usually means "flat", whatever that means. We will see that Euler characteristics are locally constant in proper flat families in §24.7, and the Cohomology and Base Change Theorem 28.1.6 will show that in particularly good situations, dimensions of cohomology groups are constant.)

All of the important properties of cohomology described in §18.1 will carry over to this more general situation. Best of all, there will be no extra work required.

In the notation $R^i\pi_*\mathscr{F}$ for higher pushforward sheaves, the "R" stands for "right derived functor", and corresponds to the fact that we get a long exact sequence in cohomology extending to the right (from the 0th terms). In Chapter 23, we will see that in good circumstances, if we have a left-exact functor, there is a long exact sequence going off to the right, in terms of right derived functors. Similarly, if we have a right-exact functor (e.g. if M is an A-module, then $\otimes_A M$ is a right-exact functor from the category of A-modules to itself), there may be a long exact sequence going off to the left, in terms of left derived functors.

Suppose $\pi : X \to Y$, and $\mathscr{F}$ is a quasicoherent sheaf on X. For each $\operatorname{Spec} A \subset Y$, we have A-modules $H^i(\pi^{-1}(\operatorname{Spec} A), \mathscr{F})$. We now show that these patch together to form a quasicoherent sheaf, in the sense of §13.3.3. We need check only one fact: that this behaves well with respect to taking distinguished open sets. In other words, we must check that for each $f \in A$, the natural map $H^i(\pi^{-1}(\operatorname{Spec} A), \mathscr{F}) \to H^i(\pi^{-1}(\operatorname{Spec} A), \mathscr{F})_f$ (induced by the map of spaces in the opposite direction — H^i is contravariant in the space) is precisely the localization $\otimes_A A_f$. But this can be verified easily: let $\{U_i\}$ be an affine cover of $\pi^{-1}(\operatorname{Spec} A)$. We can compute $H^i(\pi^{-1}(\operatorname{Spec} A), \mathscr{F})$ using the Čech complex (18.2.1.1). But this induces a cover $\operatorname{Spec} A_f$ in a natural way: If $U_i = \operatorname{Spec} A_i$ is an affine open for $\operatorname{Spec} A$, we define $U_i' = \operatorname{Spec}(A_i)_f$. The resulting Čech complex for $\operatorname{Spec} A_f$ is the localization of the Čech complex for $\operatorname{Spec} A$. As taking cohomology of a complex commutes with localization (as discussed in the FHHF Theorem, Exercise 1.6.H), we have defined a quasicoherent sheaf on Y by the characterization of quasicoherent sheaves in §13.3.3.

Define the **ith higher direct image sheaf** or the **ith (higher) pushforward sheaf** to be this quasicoherent sheaf.

18.8.1. Theorem. — *Suppose $\pi : X \to Y$ is a quasicompact separated morphism of schemes. Then:*

(a) $R^i\pi_*$ *is a covariant functor* $QCoh_X \to QCoh_Y$.
(b) *We can identify* $R^0\pi_*$ *with* $\pi_*\mathscr{F}$.
(c) *(the* **long exact sequence of higher pushforward sheaves***) A short exact sequence* $0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$ *of sheaves on* X *induces a long exact sequence*

$$\begin{tikzcd} 0 \arrow[r] & R^0\pi_*\mathscr{F} \arrow[r] & R^0\pi_*\mathscr{G} \arrow[r] & R^0\pi_*\mathscr{H} \arrow[r] & {} \\ & R^1\pi_*\mathscr{F} \arrow[r] & R^1\pi_*\mathscr{G} \arrow[r] & R^1\pi_*\mathscr{H} \arrow[r] & \cdots \end{tikzcd} \tag{18.8.1.1}$$

of sheaves on Y.
(d) *(projective pushforwards of coherent are coherent: Grothendieck's coherence theorem for projective morphisms) If π is a projective morphism and $\mathscr{O}_Y$ is coherent on Y (this hypothesis is automatic for Y locally Noetherian), and $\mathscr{F}$ is a coherent sheaf on X, then for all i, $R^i\pi_*\mathscr{F}$ is a coherent sheaf on Y.*

18.8.2. *Unimportant Remark.* If X and Y are Noetherian, the hypothesis "separated" can be relaxed to "quasiseparated"; see Unimportant Remark 23.5.8.

18.8.3. *Proof of Theorem 18.8.1.* We first show covariance: if $\mathscr{F} \to \mathscr{G}$ is a morphism of quasicoherent sheaves on X, we define a map $R^i\pi_*\mathscr{F} \to R^i\pi_*\mathscr{G}$. (It will be clear we will have shown that $R^i\pi_*$ is a functor.) It suffices to define this map on the "distinguished affine base" of Y (Definition 13.3.1). Thus it suffices to show the following: if X' is a quasicompact separated A-scheme, and $\mathscr{F} \to \mathscr{G}$ is a morphism of quasicoherent sheaves on X, then the map $H^i(X', \mathscr{F}) \to H^i(X', \mathscr{G})$ constructed in §18.2 (property **(i)** of §18.1) "commutes with localization at $f \in A$". But this was shown in Exercise 18.2.D.

In a similar way, we construct the connecting homomorphism $R^i\pi_* \mathscr{H} \to R^{i+1}\pi_*\mathscr{F}$ in the long exact sequence (18.8.1.1), by showing that the construction in the case where $Y = \operatorname{Spec} A$ "commutes with localization at $f \in A$". Again, this was shown in Exercise 18.2.D.

It suffices to check all other parts of this statement on affine open subsets of Y, so they all follow from the analogous statements in Čech cohomology (§18.1). □

The following result is handy, and essentially immediate from our definition.

18.8.A. EASY EXERCISE. Show that if π is affine, then for $i > 0$, $R^i\pi_*\mathscr{F} = 0$.

18.8.4. *How higher pushforwards behave with respect to base change.*

18.8.B. EXERCISE (HIGHER PUSHFORWARDS AND BASE CHANGE).
(a) (easy) Suppose $\psi : Z \to Y$ is any morphism, and $\pi : X \to Y$ is quasicompact and separated. Suppose $\mathscr{F}$ is a quasicoherent sheaf on X. Let

(18.8.4.1)
$$\begin{tikzcd} W \arrow[r, "\psi'"] \arrow[d, "\pi'"'] & X \arrow[d, "\pi"] \\ Z \arrow[r, "\psi"] & Y \end{tikzcd}$$

be a fiber diagram. Describe a natural morphism $\psi^*(R^i\pi_*\mathscr{F}) \to R^i\pi'_*(\psi')^*\mathscr{F}$ of sheaves on Z. (Hint: the FHHF Theorem, Exercise 1.6.H. You may want to compare the $i = 0$ case to the push-pull formula of Exercise 16.3.G.)
(b) (cohomology commutes with affine flat base change) If $\psi : Z \to Y$ is an affine morphism, and for a cover $\operatorname{Spec} A_i$ of Y, where $\psi^{-1}(\operatorname{Spec} A_i) = \operatorname{Spec} B_i$, B_i is a *flat* A-algebra (§1.6.11: $\otimes_A B_i$ is exact), and the diagram in (a) is a fiber diagram, show that the natural morphism of (a) is an isomorphism. (Exercise 18.2.H was a special case of this exercise. You can likely generalize this to non-affine morphisms — and thus the Cohomology and Flat Base Change Theorem 24.2.8 — but we wait until Chapter 24 to discuss flatness at length.)

18.8.C. EXERCISE (CF. EXERCISE 16.3.G). Prove Exercise 18.8.B(a) *without* the hypothesis that (18.8.4.1) is a fiber diagram, but adding the requirement that π' is quasicompact and separated (just so our definition of $R^i\pi'_*$ applies). In the course of the proof, you will see a map arising in the Leray spectral sequence (Theorem 23.4.4). (Hint: *use* Exercise 18.8.B(a).)

A useful special case of Exercise 18.8.B(a) is the following.

18.8.D. EXERCISE. If $q \in Y$, describe a natural morphism $(R^i\pi_*\mathscr{F}) \otimes \kappa(q) \to H^i(\pi^{-1}(q), \mathscr{F}|_{\pi^{-1}(q)})$. (Hint: the FHHF Theorem, Exercise 1.6.H.)

Thus the fiber of the pushforward may not be the cohomology of the fiber, but at least it always maps to it. We will later see that in good situations this map is an isomorphism, and thus the higher direct image sheaf indeed "patches together" the cohomology on fibers (the Cohomology and Base Change Theorem 28.1.6).

18.8.E. EXERCISE (PROJECTION FORMULA, GENERALIZING EXERCISE 16.3.H). Suppose $\pi : X \to Y$ is quasicompact and separated, and $\mathscr{F}$ and $\mathscr{G}$ are quasicoherent sheaves on X and Y respectively.

$$\begin{tikzcd}[column sep=small]
\mathscr{F} \arrow[dr, dash] & & \pi^*\mathscr{G} \arrow[dl, dash] \\
\pi_*\mathscr{F} \arrow[dr, dash] & X \arrow[d] & \mathscr{G} \arrow[dl, dash] \\
& Y &
\end{tikzcd}$$

(a) Describe a natural morphism

$$(R^i\pi_*\mathscr{F}) \otimes \mathscr{G} \to R^i\pi_*(\mathscr{F} \otimes \pi^*\mathscr{G}).$$

(Hint: the FHHF Theorem, Exercise 1.6.H.)
(b) If $\mathscr{G}$ is locally free, show that this natural morphism is an isomorphism.

The following fact uses the same trick as Theorem 18.1.9 and Exercise 18.1.C.

18.8.5. Theorem (relative dimensional vanishing). — *If $\pi : X \to Y$ is a projective morphism and Y is locally Noetherian (or more generally $\mathscr{O}_Y$ is coherent over itself), then the higher pushforwards vanish in degree higher than the maximum dimension of the fibers.*

This is false without the projective hypothesis (see Exercise 18.8.F below). In particular, you might hope that just as dimensional vanishing generalized from projective varieties to quasiprojective varieties or more general settings (§18.2.7) that relative dimensional vanishing would generalize from projective morphisms to quasiprojective morphisms, but this is not the case.

18.8.F. EXERCISE. Consider the open embedding $\pi : \mathbb{A}^n - \{0\} \to \mathbb{A}^n$. By direct calculation, show that $R^{n-1}\pi_*\mathscr{O}_{\mathbb{A}^n - \{0\}} \neq 0$. (This calculation will remind you of the proof of the H^n part of Theorem 18.1.3; see also Remark 18.3.1.)

Proof of Theorem 18.8.5. Let m be the maximum dimension of all the fibers.

The question is local on Y, so we will show that the result holds near a point p of Y. We may assume that Y is affine, and hence that $X \hookrightarrow \mathbb{P}^n_Y$.

Let k be the residue field at p. Then $\pi^{-1}(p)$ is a projective k-scheme of dimension at most m. By Exercise 11.3.C we can find affine open sets $D(f_1)$, …, $D(f_{m+1})$ that cover $\pi^{-1}(p)$. In other words, the intersection of $V(f_i)$ does not intersect $\pi^{-1}(p)$.

If $Y = \operatorname{Spec} A$ and $p = [\mathfrak{p}]$ (so $k = A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p}$), then arbitrarily lift each f_i from an element of $k[x_0, \dots, x_n]$ to an element f_i' of $A_\mathfrak{p}[x_0, \dots, x_n]$. Let F be the product of the denominators of the f_i'; note that $F \notin \mathfrak{p}$, i.e., $p = [\mathfrak{p}] \in D(F)$. Then $f_i' \in A_F[x_0, \dots, x_n]$. The intersection of their zero loci $\cap V(f_i') \subset \mathbb{P}^n_{A_F}$ is a closed subscheme of $\mathbb{P}^n_{A_F}$. Intersect it with X to get another closed subscheme of $\mathbb{P}^n_{A_F}$. Take its image under π; as projective morphisms are closed, we get a closed subset of $D(F) = \operatorname{Spec} A_F$. But this closed subset does not include p; hence we can find an affine neighborhood $\operatorname{Spec} B$ of p in Y missing the image. But if f_i'' are the restrictions of f_i' to $B[x_0, \dots, x_n]$, then $D(f_i'')$ cover $\pi^{-1}(\operatorname{Spec} B)$; in other words, over

$\pi^{-1}(\operatorname{Spec} B)$ is covered by $m+1$ affine open sets, so by the affine-cover vanishing theorem, its cohomology vanishes in degree at least $m+1$. But the higher-direct image sheaf is computed using these cohomology groups, hence the higher direct image sheaf $R^i\pi_*\mathscr{F}$ vanishes on $\operatorname{Spec} B$ too. $\square$

18.8.G. EXERCISE (RELATIVE SERRE VANISHING, CF. THEOREM 18.1.4(II)). Suppose $\pi : X \to Y$ is a proper morphism of Noetherian schemes, and $\mathscr{L}$ is a π-ample invertible sheaf on X. Show that for any coherent sheaf $\mathscr{F}$ on X, for $m \gg 0$, $R^i\pi_*\mathscr{F} \otimes \mathscr{L}^{\otimes m} = 0$ for all $i > 0$.

18.9 ⋆ Chow's Lemma and Grothendieck's Coherence Theorem

The proofs in this section are starred because the results aren't absolutely necessary in the rest of our discussions, and may not be worth reading right now. But just knowing the statement Grothendieck's Coherence Theorem 18.9.1, (generalizing Theorem 18.8.1(d)) will allow you to immediately translate many of our arguments about projective schemes and morphisms to proper schemes and morphisms, and Chow's Lemma is a multi-purpose tool to extend results from the projective situation to the proper situation in general.

18.9.1. Grothendieck's Coherence Theorem. — *Suppose $\pi : X \to Y$ is a proper morphism of locally Noetherian schemes. Then for any coherent sheaf $\mathscr{F}$ on X, $R^i\pi_*\mathscr{F}$ is coherent on Y.*

The case $i = 0$ has already been mentioned a number of times.

18.9.A. EXERCISE. Recall that finite morphisms are affine (by definition) and proper. Use Theorem 18.9.1 to show that if $\pi : X \to Y$ is proper and affine and Y is Noetherian, then π is finite. (Hint: mimic the proof of the weaker result where proper is replaced by projective, Corollary 18.1.8.)

The proof of Theorem 18.9.1 requires two sophisticated facts. The first is the Leray spectral sequence (Theorem 23.4.4, which applies in this situation because of Exercise 23.5.F). Suppose $\pi : X \to Y$ and $\rho : Y \to Z$ are quasicompact separated morphisms. Then for any quasicoherent sheaf $\mathscr{F}$ on X, there is a spectral sequence with E_2 term given by $R^p\rho_*(R^q\pi_*\mathscr{F})$ converging to $R^{p+q}(\rho \circ \pi)_*\mathscr{F}$. Because this would be a reasonable (but hard) exercise in the case we need it (where Z is affine), we will feel comfortable using it. But because we will later prove it, we won't prove it now.

We will also need Chow's Lemma.

18.9.2. Chow's Lemma. — *Suppose $\pi : X \to \operatorname{Spec} A$ is a proper morphism, and A is Noetherian. Then there exists $\mu : X' \to X$ which is surjective and projective, such that*

$\pi \circ \mu$ *is also projective, and such that* μ *is an isomorphism on a dense open subset of* X*:*

$$\begin{tikzcd} X' \arrow[rr, "\mu"] \arrow[dr, "\pi\circ\mu"'] & & X \arrow[dl, "\pi"] \\ & \operatorname{Spec} A. & \end{tikzcd}$$

In particular, if X is a projective k-variety, it admits a projective birational morphism from a projective k-variety.

Many generalizations of results from projective to proper situations go through Chow's Lemma. We will prove this version, and state other versions of Chow's Lemma, in §18.9.3. Assuming these two facts, we now prove Theorem 18.9.1 in a series of exercises.

Proof. The question is local on Y, so we may assume Y is affine, say $Y = \operatorname{Spec} A$. We work by induction on $\dim \operatorname{Supp} \mathscr{F}$, with the base case when $\dim \operatorname{Supp} \mathscr{F} = -1$ (i.e., $\operatorname{Supp} \mathscr{F} = \varnothing$, i.e., $\mathscr{F} = 0$), which is obvious. So fix $\mathscr{F}$, and assume the result is known for all coherent sheaves with support of smaller dimension.

18.9.B. EXERCISE. Show that we may assume that $\operatorname{Supp} \mathscr{F} = X$. (Hint: the idea is to replace X by the **scheme-theoretic support** of $\mathscr{F}$, the smallest closed subscheme of X on which $\operatorname{Supp} \mathscr{F}$ "lives". More precisely, it is the smallest closed subscheme $i : W \hookrightarrow X$ such that there is a coherent sheaf $\mathscr{F}'$ on W, with $\mathscr{F} \cong i_* \mathscr{F}'$. Show that this notion makes sense, using the ideas of §8.3, by defining it on each affine open subset.)

We now invoke Chow's Lemma to construct a projective morphism $\mu : X' \to X$ that is an isomorphism on a dense open subset U of X (so $\dim X \setminus U < \dim X$), and such that $\pi \circ \mu : X' \to \operatorname{Spec} A$ is projective.

Then $\mathscr{G} = \mu^* \mathscr{F}$ is a coherent sheaf on X', $\mu_* \mathscr{F}$ is a coherent sheaf on X (by the projective case, Theorem 18.8.1(d)) and the adjunction map $\mathscr{F} \to \mu_* \mathscr{G} = \mu_* \mu^* \mathscr{F}$ is an isomorphism on U. The kernel $\mathscr{E}$ and cokernel $\mathscr{H}$ are coherent sheaves on X that are supported in smaller dimension:

$$0 \to \mathscr{E} \to \mathscr{F} \to \mu_* \mathscr{G} \to \mathscr{H} \to 0.$$

18.9.C. EXERCISE. By the inductive hypothesis, the higher pushforwards of $\mathscr{E}$ and $\mathscr{H}$ are coherent. Show that if all the higher pushforwards of $\mu_* \mathscr{G}$ are coherent, then the higher pushforwards of $\mathscr{F}$ are coherent.

So we are reduced to showing that the higher pushforwards of $\mu_* \mathscr{G}$ are coherent for any coherent $\mathscr{G}$ on X'.

The Leray spectral sequence for $X' \xrightarrow{\mu} X \xrightarrow{\pi} \operatorname{Spec} A$ has E_2 page given by $R^p \pi_* (R^q \mu_* \mathscr{G})$ converging to $R^{p+q} (\pi \circ \mu)_* \mathscr{G}$. Now $R^q \mu_* \mathscr{G}$ is coherent by Theorem 18.8.1(d). Furthermore, as μ is an isomorphism on a dense open subset U of X, $R^q \mu_* \mathscr{G}$ is zero on U, and is thus supported on the complement of U, whose dimension is *less than* that of X. Hence by our inductive hypothesis, $R^p \pi_* (R^q \mu_* \mathscr{G}')$ is coherent for all p, and all $q \geq 1$. The only possibly noncoherent sheaves on the E_2 page are in the row $q = 0$ — precisely the sheaves we are interested in. Also, by Theorem 18.8.1(d) applied to $\pi \circ \mu$, $R^{p+q} (\pi \circ \mu)_* \mathscr{F}$ is coherent.

18.9.D. EXERCISE. Show that $E_n^{p,q}$ is always coherent for any $n \geq 2$, $q > 0$. Show that $E_n^{p,0}$ is coherent for a given $n \geq 2$ if and only if $E_2^{p,0}$ is coherent. Show that $E_\infty^{p,q}$ is coherent, and hence that $E_2^{p,0}$ is coherent, thereby completing the proof of Theorem 18.9.1.

$\square$

18.9.3. $\star\star$ **Proof of Chow's Lemma.**

We use the properness hypothesis on π through each of its three constituent parts: finite type, separated, universally closed. The parts using separatedness are particularly tricky.

As X is Noetherian, it has finitely many irreducible components. Cover X with affine open sets $U_1, \dots, U_n$. We may assume that each U_i meets each irreducible component. (If some U_i does not meet an irreducible component Z, then take any affine open subset Z' of $Z - \overline{X - Z}$, and replace U_i by $U_i \cup Z'$.) Then $U := \cap_i U_i$ is a dense open subset of X. As each U_i is finite type over A, we can choose a closed embedding $U_i \subset \mathbb{A}_A^{n_i}$. Let $\overline{U}_i$ be the (scheme-theoretic) closure of U_i in $\mathbb{P}_A^{n_i}$.

Now we have the diagonal morphism $U \to X \times_A \prod \overline{U}_i$ (where the product is over $\operatorname{Spec} A$), which is a locally closed embedding (the composition of the closed embedding $U \hookrightarrow U^{n+1}$ with the open embedding $U^{n+1} \hookrightarrow X \times_A \prod \overline{U}_i$). Let X' be the scheme-theoretic closure of U in $X \times_A \prod \overline{U}_i$. Let μ be the composed morphism $X' \to X \times_A \prod \overline{U}_i \to X$, so we have a diagram

$$\begin{tikzcd}
X' \arrow[d, hook, "\text{cl. emb.}"'] \arrow[dr, "\mu"] & \\
X \times_A \prod \overline{U}_i \arrow[r, "\text{proj.}"'] \arrow[d, "\text{proper}"'] & X \arrow[d, "\text{proper}"] \\
\prod \overline{U}_i \arrow[r, "\text{proj.}"'] & \operatorname{Spec} A
\end{tikzcd}$$

(where the square is Cartesian). The morphism μ is projective (as it is the composition of two projective morphisms and X is quasicompact, Exercise 17.3.B). We will conclude the argument by showing that $\mu^{-1}(U) = U$ (or more precisely, μ is an isomorphism above U), and that $X' \to \prod \overline{U}_i$ is a closed embedding (from which the composition

$$X \to \prod \overline{U}_i \to \operatorname{Spec} A$$

is projective).

18.9.E. EXERCISE. Suppose $T_0, \dots, T_n$ are *separated* schemes over A with isomorphic open sets, which we sloppily call V in each case. Then V is a locally closed subscheme of $T_0 \times \cdots \times T_n$. Let $\overline{V}$ be the closure of this locally closed subscheme. Show that

$$\begin{aligned}
V &\cong \overline{V} \cap (V \times_A T_1 \times_A \cdots \times_A T_n) \\
&= \overline{V} \cap (T_0 \times_A V \times_A T_2 \times_A \cdots \times_A T_n) \\
&= \cdots \\
&= \overline{V} \cap (T_0 \times_A \cdots \times_A T_{n-1} \times_A V).
\end{aligned}$$

(Hint for the first isomorphism: the graph of the morphism $V \to T_1 \times_A \cdots \times_A T_n$ is a closed embedding, as $T_1 \times_A \cdots \times_A T_n$ is separated over A, by Proposition 10.1.18. Thus the scheme-theoretic closure of V in $V \times_A T_1 \times_A \cdots \times_A T_n$ is V itself. Finally, the scheme-theoretic closure can be computed locally, essentially by Theorem 8.3.4.)

18.9.F. EXERCISE. Using (the idea behind) the previous exercise, show that $\mu^{-1}(U) = U$.

It remains to show that $X' \to \prod \overline{U}_i$ is a closed embedding. Now $X' \to \prod \overline{U}_i$ is closed (it is the composition of two closed maps), so it suffices to show that $X' \to \prod \overline{U}_i$ is a locally closed embedding.

18.9.G. EXERCISE. Let A_i be the closure of U in

$$B_i := X \times_A \overline{U}_1 \times_A \cdots \times_A U_i \times_A \cdots \times_A \overline{U}_n$$

(only the ith term is missing the bar), and let C_i be the closure of U in

$$D_i := \overline{U}_1 \times_A \cdots \times_A U_i \times_A \cdots \times_A \overline{U}_n.$$

Show that there is an isomorphism $A_i \to C_i$ induced by the projection $B_i \to D_i$. Hint: note that the section $D_i \to B_i$ of the projection $B_i \to D_i$, given informally by $(t_1, \ldots, t_n) \mapsto (t_i, t_1, \ldots, t_n)$, is a closed embedding, as it can be interpreted as the graph of a map to a separated scheme (over A). So U can be interpreted as a locally closed subscheme of D_i, which in turn can be interpreted as a closed subscheme of B_i. Thus the closure of U in D_i may be identified with its closure in B_i.

As the U_i cover X, the $\mu^{-1}(U_i)$ cover $\overline{X}$. But $\mu^{-1}(U_i) = A_i$ (closure can be be computed locally — the closure of U in B_i is the intersection of B_i with the closure $\overline{X}$ of U in $X \times_A \overline{U}_1 \times_A \cdots \overline{U}_n$).

Hence over each U_i, we get a closed embedding of $A_i \hookrightarrow D_i$, and thus $X' \to \prod \overline{U}_i$ is a locally closed embedding as desired. ☐

18.9.4. Other versions of Chow's Lemma. We won't use these versions, but their proofs are similar to what we have already shown.

18.9.5. *Remark.* Notice first that if X is reduced (resp. irreducible, integral), then X' can be taken to be reduced (resp. irreducible, integral) as well.

18.9.H. EXERCISE. By suitably crossing out lines in the proof above, weaken the hypothesis "π is proper" to "π finite type and separated", at the expense of weakening the conclusion "$\pi \circ \mu$ is projective" to "$\pi \circ \mu$ is quasiprojective".

18.9.6. *Remark.* The target Spec A can be generalized to a scheme S that is (i) Noetherian, or (ii) separated and quasicompact with finitely many irreducible components. This can be combined with Remark 18.9.5 and Exercise 18.9.H. See [**Gr-EGA**, II.5.6.1] for a proof. See also [**GW**, Thm. 13.100] for a version that is slightly more general.

CHAPTER 19

Application: Curves

We now use what we have developed to study something explicit — curves. Throughout this chapter, we will assume that all curves are projective, geometrically integral, regular curves over a field k. We will sometimes add the hypothesis that k is algebraically closed. Most people are happy with working over algebraically closed fields, and those people should ignore the adverb "geometrically".

We certainly don't need the massive machinery we have developed in order to understand curves, but with the perspective we have gained, the development is quite clean. The key ingredients we will need are as follows. We use a criterion for a morphism to be a closed embedding, that we prove in §19.1. We use the "black box" of Serre duality (to be proved in Chapter 30). In §19.2, we use this background to observe a very few useful facts, which we will use repeatedly. Finally, in the course of applying them to understand curves of various genera, we develop the theory of hyperelliptic curves in a hands-on way (§19.5), in particular proving a special case of the Riemann-Hurwitz formula.

If you are jumping into this chapter without reading much beforehand, you should skip §19.1, taking Theorem 19.1.1 as a black box. Depending on your background, you may want to skip §19.2 as well, taking the crucial observations as a black box.

19.1 A criterion for a morphism to be a closed embedding

We will repeatedly use a criterion for when a morphism is a closed embedding, which is not special to curves. Before stating it, we recall some facts about closed embeddings. Suppose $\pi : X \to Y$ is a closed embedding. Then π is projective, and it is injective on points. This is not enough to ensure that it is a closed embedding, as the example of the normalization of the cusp shows (Figure 9.4). Another example is the following.

19.1.A. EXERCISE (FROBENIUS, CF. §7.3.16). Suppose $\operatorname{char} k = p$, and π is the map $\pi : \mathbb{A}^1_k \to \mathbb{A}^1_k$ given by $x \mapsto x^p$. Show that π is a bijection on points, and even induces an isomorphism of residue fields on closed points, yet is not a closed embedding.

The additional information you need is that the tangent map is an isomorphism at all closed points.

19.1.B. EXERCISE. Show (directly, not invoking Theorem 19.1.1) that in the two examples described above (the normalization of a cusp and the Frobenius morphism), the tangent map is *not* an isomorphism at all closed points.

19.1.1. Theorem. — *Suppose $k = \overline{k}$, and $\pi : X \to Y$ is a projective morphism of finite type k-schemes that is injective on closed points and injective on tangent vectors at closed points. Then π is a closed embedding.*

Remark: "injective on closed points and tangent vectors at closed points" means that π is unramified (under these hypotheses). (We will defined *unramified* in §21.6; in general unramified morphisms need not be injective.)

The example $\operatorname{Spec} \mathbb{C} \to \operatorname{Spec} \mathbb{R}$ shows that we need the hypothesis that k is algebraically closed in Theorem 19.1.1. Those allergic to algebraically closed fields should still pay attention, as we will use this to prove things about curves over k where k is *not* necessarily algebraically closed (see also Exercises 9.2.J and 19.1.E).

We need the hypothesis that the morphism be projective, as shown by the example of Figure 19.1. It is the normalization of the node, except we erase one of the preimages of the node. We map $\mathbb{A}^1$ to the plane, so that its image is a curve with one node. We then consider the morphism we get by discarding one of the preimages of the node. Then this morphism is an injection on points, and is also injective on tangent vectors, but it is not a closed embedding. (In the world of differential geometry, this fails to be an embedding because the map doesn't give a homeomorphism onto its image.)

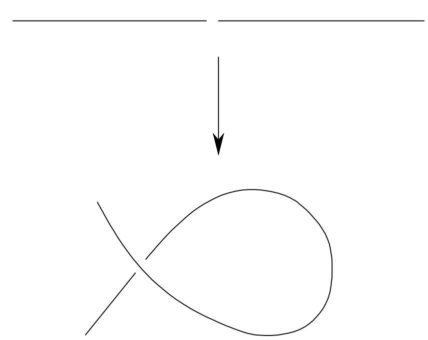


FIGURE 19.1. The projectivity hypothesis in Theorem 19.1.1 cannot be dropped

Theorem 19.1.1 appears to be fundamentally a statement about varieties, but it isn't. We will reduce it to the following result.

19.1.2. Theorem. — *Suppose $\pi : X \to Y$ is a finite morphism of Noetherian schemes whose degree at every point of Y (§13.7.5) is 0 or 1. Then π is a closed embedding.*

Once we know the meaning of "unramified", this will translate to: "unramified + finite = closed embedding" for Noetherian schemes.

19.1.C. EXERCISE. Suppose $\pi : X \to Y$ is a finite morphism whose degree at every point of Y is 0 or 1. Show that π is injective on points (easy). If $p \in X$ is any point, show that π induces an isomorphism of residue fields $\kappa(\pi(p)) \to \kappa(p)$. Show that π induces an injection of tangent spaces. Thus key hypotheses of Theorem 19.1.1 are implicitly in the hypotheses of Theorem 19.1.2.

19.1.3. *Reduction of Theorem 19.1.1 to Theorem 19.1.2.* The property of being a closed embedding is local on the base, so we may assume that Y is affine, say $\operatorname{Spec} B$.

I next claim that π has finite fibers, not just finite fibers above closed points: the fiber dimension for projective morphisms is upper semicontinuous (Exercise 18.1.C, or Theorem 11.4.2(b)), so the locus where the fiber dimension is at least 1 is a closed subset, so if it is nonempty, it must contain a closed point of Y. Thus the fiber over any point is a dimension 0 finite type scheme over that point, hence a finite set.

Hence π is a projective morphism with finite fibers, thus finite by Corollary 18.1.9.

But the degree of a finite morphism is upper semicontinuous, (§13.7.5), and is at most 1 at closed points of Y, hence is at most 1 at all points.

19.1.4. *Proof of Theorem 19.1.2.* The problem is local on Y, so we may assume Y is affine, say $Y = \operatorname{Spec} B$. Thus X is affine too, say $\operatorname{Spec} A$, and π corresponds to a ring morphism $B \to A$. We wish to show that this is a surjection of rings, or (equivalently) of B-modules. Let K be the cokernel of this morphism of B-modules:

$$B \to A \to K \to 0. \tag{19.1.4.1}$$

We wish to show that $K = 0$. It suffices to show that for any maximal ideal $\mathfrak{n}$ of B, $K_{\mathfrak{n}} = 0$. (Do you remember why?) Localizing (19.1.4.1) at $\mathfrak{n}$, we obtain the exact sequence

$$B_{\mathfrak{n}} \to A_{\mathfrak{n}} \to K_{\mathfrak{n}} \to 0.$$

Applying the right-exact functor $\otimes_{B_{\mathfrak{n}}} (B_{\mathfrak{n}}/\mathfrak{n}B_{\mathfrak{n}})$, we obtain

$$B_{\mathfrak{n}}/\mathfrak{n}B_{\mathfrak{n}} \xrightarrow{\alpha} A_{\mathfrak{n}}/\mathfrak{n}A_{\mathfrak{n}} \longrightarrow K_{\mathfrak{n}}/\mathfrak{n}K_{\mathfrak{n}} \longrightarrow 0. \tag{19.1.4.2}$$

Now $\operatorname{Spec} A_{\mathfrak{n}}/\mathfrak{n}A_{\mathfrak{n}}$ is the scheme theoretic preimage of $[\mathfrak{n}] \in \operatorname{Spec} B$, so by hypothesis, it is either empty, or the map of residue fields α is an isomorphism. In the first case, $A_{\mathfrak{n}}/\mathfrak{n}A_{\mathfrak{n}} = 0$, from which $K_{\mathfrak{n}}/\mathfrak{n}K_{\mathfrak{n}} = 0$. In the second case, $K_{\mathfrak{n}}/\mathfrak{n}K_{\mathfrak{n}} = 0$ as well. Applying Nakayama's Lemma 7.2.9 (noting that A is a finitely generated B-module, hence K is too, hence $K_{\mathfrak{n}}$ is a finitely generated $B_{\mathfrak{n}}$-module), $K_{\mathfrak{n}} = 0$ as desired. $\square$

19.1.D. EXERCISE. Use Theorem 19.1.1 to show that the dth Veronese embedding from $\mathbb{P}^n_k$, corresponding to the complete linear series $|\mathscr{O}_{\mathbb{P}^n_k}(d)|$, is a closed embedding. Do the same for the Segre embedding from $\mathbb{P}^m_k \times_{\operatorname{Spec} k} \mathbb{P}^n_k$. (This is just for practice for using this criterion. It is a weaker result than we had before; we have earlier checked both of these statements over an arbitrary base ring in Remark 8.2.8 and §9.6 respectively, and we are now checking it only over algebraically closed fields. However, see Exercise 19.1.E below.)

Exercise 9.2.J can be used to extend Theorem 19.1.1 to general fields k, not necessarily algebraically closed.

19.1.E. LESS IMPORTANT EXERCISE. Using the ideas from this section, prove that the dth Veronese embedding from $\mathbb{P}^n_{\mathbb{Z}}$ (over the integers!), is a closed embedding. (Again, we have done this before. This exercise is simply to show that these methods can easily extend to work more generally.)

19.2 A series of crucial tools

We are now ready to start understanding curves in a hands-on way. We will repeatedly make use of the following series of crucial remarks, and it will be important to have them at the tip of your tongue.

19.2.1. In what follows, C *will be a projective, geometrically regular, geometrically integral curve over a field* k*, and* $\mathscr{L}$ *is an invertible sheaf on* C*.*

19.2.2. Reminder: Serre duality. Serre duality (Theorem 18.5.1) on a geometrically irreducible regular genus g curve C over k involves an invertible sheaf ω_C (of degree $2g-2$, with g sections, Exercise 18.5.A), such that for any coherent sheaf $\mathscr{F}$ on C, $h^i(C, \mathscr{F}) = h^{1-i}(C, \omega_C \otimes \mathscr{F}^\vee)$ for $i = 0, 1$. (Better: there is a duality between the two cohomology groups.)

19.2.3. Negative degree line bundles have no nonzero section. $h^0(C, \mathscr{L}) = 0$ if $\deg \mathscr{L} < 0$. Reason: $\deg \mathscr{L}$ is the number of zeros minus the number of poles (suitably counted) of any rational section (Important Exercise 18.4.C). If there is a regular section (i.e., with no poles), then this is necessarily non-negative. Refining this argument yields the following.

19.2.4. Degree 0 line bundles, and recognizing when they are trivial. $h^0(C, \mathscr{L}) = 0$ or 1 if $\deg \mathscr{L} = 0$, and if $h^0(C, \mathscr{L}) = 1$ then $\mathscr{L} \cong \mathscr{O}_C$. Reason: if there is a section s, it has no poles, and hence no zeros, because $\deg \mathscr{L} = 0$. Then $\operatorname{div} s = 0$, so $\mathscr{L} \cong \mathscr{O}_C(\operatorname{div} s) = \mathscr{O}_C$. (Recall how this works, cf. Important Exercise 14.2.E: s gives a trivialization for the invertible sheaf. We have a natural bijection for any open set $\Gamma(U, \mathscr{L}) \leftrightarrow \Gamma(U, \mathscr{O}_U)$, where the map from left to right is $s' \mapsto s'/s$, and the map from right to left is $f \mapsto sf$.) Conversely, for a geometrically integral projective variety, $h^0(\mathscr{O}) = 1$. (§10.3.7 shows this for k algebraically closed — this is where geometric integrality is used — and Exercise 18.2.H shows that cohomology commutes with base field extension.)

Serre duality turns these statements about line bundles of degree at most 0 into statements about line bundles of degree at least $2g-2$, as follows.

19.2.5. We know $h^0(C, \mathscr{L})$ if the degree is sufficiently high. If $\deg \mathscr{L} > 2g-2$, then

$$\boxed{h^0(C, \mathscr{L}) = \deg \mathscr{L} - g + 1.} \tag{19.2.5.1}$$

So we know $h^0(C, \mathscr{L})$ if $\deg \mathscr{L} \gg 0$. *(This is important — remember this!)* Reason: $h^1(C, \mathscr{L}) = h^0(C, \omega_C \otimes \mathscr{L}^\vee)$; but $\omega_C \otimes \mathscr{L}^\vee$ has negative degree (as ω_C has degree $2g-2$), and thus this invertible sheaf has no sections. The result then follows from the Riemann-Roch Theorem 18.4.B.

19.2.A. Useful exercise (recognizing ω_C among degree $2g-2$ line bundles). Suppose $\mathscr{L}$ is a degree $2g-2$ invertible sheaf. Show that it has $g-1$ or g sections, and it has g sections if and only if $\mathscr{L} \cong \omega_C$.

19.2.6. Twisting $\mathscr{L}$ by a (degree 1) point changes h^0 by at most 1. Suppose p is any closed point of degree 1 (i.e., the residue field of p is k). Then $h^0(C, \mathscr{L}) - h^0(C, \mathscr{L}(-p)) = 0$ or 1. (The twist of $\mathscr{L}$ by a divisor, such as $\mathscr{L}(-p)$, was defined in §14.2.11.) Reason: consider $0 \to \mathscr{O}_C(-p) \to \mathscr{O}_C \to \mathscr{O}|_p \to 0$, tensor with $\mathscr{L}$ (this is exact as $\mathscr{L}$ is locally free) to get

$$0 \to \mathscr{L}(-p) \to \mathscr{L} \to \mathscr{L}|_p \to 0.$$

Then $h^0(C, \mathscr{L}|_p) = 1$, so as the long exact sequence of cohomology starts off

$$0 \to H^0(C, \mathscr{L}(-p)) \to H^0(C, \mathscr{L}) \to H^0(C, \mathscr{L}|_p),$$

we are done.

19.2.7. A numerical criterion for $\mathscr{L}$ to be base-point-free. Suppose for this remark that k is algebraically closed, so *all* closed points have degree 1 over k. Then if $h^0(C, \mathscr{L}) - h^0(C, \mathscr{L}(-p)) = 1$ for *all* closed points p, then $\mathscr{L}$ is base-point-free, and hence induces a morphism from C to projective space (Theorem 16.4.1). Reason: given any p, our equality shows that there exists a section of $\mathscr{L}$ that does not vanish at p — so by definition, p is not a base-point of $\mathscr{L}$.

19.2.8. Next, suppose p and q are distinct (closed) points of degree 1. Then $h^0(C, \mathscr{L}) - h^0(C, \mathscr{L}(-p-q)) = 0$, 1, or 2 (by repeating the argument of Remark 19.2.6 twice). If $h^0(C, \mathscr{L}) - h^0(C, \mathscr{L}(-p-q)) = 2$, then necessarily
(19.2.8.1)

$$h^0(C, \mathscr{L}) = h^0(C, \mathscr{L}(-p)) + 1 = h^0(C, \mathscr{L}(-q)) + 1 = h^0(C, \mathscr{L}(-p-q)) + 2.$$

Then the linear series $\mathscr{L}$ separates points p and q, i.e., the corresponding map f to projective space satisfies $f(p) \neq f(q)$. Reason: there is a hyperplane of projective space passing through p but not passing through q, or equivalently, there is a section of $\mathscr{L}$ vanishing at p but not vanishing at q. This is because of the last equality in (19.2.8.1).

19.2.9. By the same argument as above, if p is a (closed) point of degree 1, then $h^0(C, \mathscr{L}) - h^0(C, \mathscr{L}(-2p)) = 0$, 1, or 2. I claim that if this is 2, then map corresponds to $\mathscr{L}$ (which is already seen to be base-point-free from the above) separates the tangent vectors at p. To show this, we need to show that the cotangent map is *surjective*. To show surjectivity onto a one-dimensional vector space, we just need to show that the map is nonzero. So we need a function on the target vanishing at the image of p that pulls back to a function that vanishes at p to order 1 but not 2. In other words, we want a section of $\mathscr{L}$ vanishing at p to order 1 but not 2. But that is the content of the statement $h^0(C, \mathscr{L}(-p)) - h^0(C, \mathscr{L}(-2p)) = 1$.

19.2.10. Criterion for $\mathscr{L}$ to be very ample. Combining some of our previous comments: suppose C is a curve over an *algebraically closed* field k, and $\mathscr{L}$ is an invertible sheaf such that for *all* closed points p and q, *not necessarily distinct*, $h^0(C, \mathscr{L}) - h^0(C, \mathscr{L}(-p-q)) = 2$, then $\mathscr{L}$ gives a *closed embedding into projective space*, as it separates points and tangent vectors, by Theorem 19.1.1.

19.2.B. EXERCISE. Suppose that k is algebraically closed, so the previous remark applies. Show that $C \setminus \{p\}$ is affine. Hint: Show that if $k \gg 0$, then $\mathscr{O}(kp)$ is base-point-free and has at least two linearly independent sections, one of which has divisor kp. Use these two sections to map to $\mathbb{P}^1$ so that the set-theoretic preimage of ∞ is p. Argue that the map is finite, and that $C \setminus \{p\}$ is the preimage of $\mathbb{A}^1$. (A trivial variation of this argument shows that $C \setminus \{p_1, \dots, p_n\}$ is affine if $n > 0$.)

19.2.11. Conclusion. We can combine much of the above discussion to give the following useful fact. If k is algebraically closed, then $\deg \mathscr{L} \geq 2g$ implies that $\mathscr{L}$ is base-point-free (and hence determines a morphism to projective space). Also, $\deg \mathscr{L} \geq 2g+1$ implies that this is in fact a closed embedding (so $\mathscr{L}$ is very ample). Remember this!

19.2.C. EXERCISE (ON A SMOOTH PROJECTIVE INTEGRAL CURVE OVER $\overline{k}$, AMPLE = POSITIVE DEGREE). Show that an invertible sheaf $\mathscr{L}$ on a projective, regular integral curve over $\overline{k}$ is ample if and only if $\deg \mathscr{L} > 0$.

(This can be extended to curves over general fields using Exercise 19.2.D below.) Thus there is a blunt purely numerical criterion for ampleness of line bundles on curves. This generalizes to projective varieties of higher dimension; this is called *the Nakai-Moishezon criterion for ampleness*, Theorem 20.4.1.

19.2.D. EXERCISE (EXTENSION TO NON-ALGEBRAICALLY CLOSED FIELDS). Show that the statements in §19.2.11 hold even without the hypothesis that k is algebraically closed. (Hint: to show one of the facts about some curve C and line bundle $\mathscr{L}$, consider instead $C \otimes_{\operatorname{Spec} k} \operatorname{Spec} \overline{k}$. Then show that if the pullback of $\mathscr{L}$ here has sections giving you one of the two desired properties, then there are sections downstairs with the same properties. You may want to use facts that we have used, such as the fact that base-point-freeness is independent of extension of base field, Exercise 18.2.I, or that the property of an affine morphism over k being a closed embedding holds if and only if it does after an extension of k, Exercise 9.2.J.)

19.2.E. EXERCISE (EXTENDING EXERCISE 19.2.C). Suppose $\mathscr{L}$ is an invertible sheaf on a projective, smooth integral curve C (over a field k). Show that $\mathscr{L}$ is ample if and only if $\deg \mathscr{L} > 0$. Hint: reduce to the case where k is algebraically closed, with the help of Exercise 9.2.J. (This was promised in Exercise 18.7.C.)

We are now ready to take these facts and go to the races.

19.3 Curves of genus 0

We are now ready to (in some form) answer the question: what are the curves of genus 0?

In §6.5.9, we saw a genus 0 curve (over a field k) that was *not* isomorphic to $\mathbb{P}^1$: $x^2+y^2+z^2=0$ in $\mathbb{P}^2_{\mathbb{R}}$. (It has genus 0 by (18.6.6.1).) We have already observed that this curve is *not* isomorphic to $\mathbb{P}^1_{\mathbb{R}}$, because it doesn't have an $\mathbb{R}$-valued point. On the other hand, we haven't seen a genus 0 curve over an algebraically closed field with this property. This is no coincidence: the lack of an existence of a k-valued point is the only obstruction to a genus 0 curve being $\mathbb{P}^1$.

19.3.1. Proposition. — *Suppose C is genus 0, and C has a k-valued (degree 1) point. Then $C \cong \mathbb{P}^1_k$.*

Thus we see that all genus 0 (integral, regular) curves over an algebraically closed field are isomorphic to $\mathbb{P}^1$.

Proof. Let p be the point, and consider $\mathscr{L} = \mathscr{O}(p)$. Then $\deg \mathscr{L} = 1$, so we can apply what we know above: first, $h^0(C, \mathscr{L}) = 2$ (Remark 19.2.5), and second, these two sections give a closed embedding into $\mathbb{P}^1_k$ (Remark 19.2.11). But the only closed embedding of a curve into the integral curve $\mathbb{P}^1_k$ is an isomorphism! $\square$

As a bonus, Proposition 19.3.1 implies that $x^2+y^2+z^2 = 0$ in $\mathbb{P}^2_{\mathbb{R}}$ has no *line bundles* of degree 1 over $\mathbb{R}$; otherwise, we could just apply the above argument to the corresponding line bundle. This example shows us that over a non-algebraically closed field, there can be genus 0 curves that are not isomorphic to $\mathbb{P}^1_k$. The next result lets us get our hands on them as well.

19.3.2. Claim. — *All genus* 0 *curves can be described as conics in* $\mathbb{P}^2_k$.

Proof. Any genus 0 curve has a degree -2 line bundle — the dualizing sheaf ω_C. Thus any genus 0 curve has a degree 2 line bundle: $\mathscr{L} = \omega_C^\vee$. We apply Remark 19.2.11: $\deg \mathscr{L} = 2 \geq 2g+1$, so this line bundle gives a closed embedding into $\mathbb{P}^2$. $\square$

19.3.A. EXERCISE. Suppose C is a genus 0 curve (projective, geometrically integral and regular). Show that C has a point of degree at most 2. (The degree of a point was defined in §5.3.8.)

The geometric means of finding Pythagorean triples presented in §6.5.8 looked quite different, but was really the same. There was a genus 0 curve C (a plane conic) with a k-valued point p, and we proved that it was isomorphic to $\mathbb{P}^1_k$. The line bundle used to show the isomorphism wasn't the degree 1 line bundle $\mathscr{O}_C(p)$; it was the degree 1 line bundle $\mathscr{O}_{\mathbb{P}^2}(1)|_C \otimes \mathscr{O}_C(-p)$.

19.4 Classical geometry arising from curves of positive genus

We will use the following Proposition and Corollary later, and we take this as an excuse to revisit some very classical geometry from a modern standpoint.

19.4.1. Proposition. — *Recall our standing assumptions for this chapter (§19.2.1), that* C *is a projective, geometrically regular, geometrically integral curve over a field* k*. Suppose* C *is not isomorphic to* $\mathbb{P}^1_k$ *(with no assumptions on the genus of* C*), and* $\mathscr{L}$ *is an invertible sheaf of degree* 1*. Then* $h^0(C, \mathscr{L}) < 2$.

Proof. Otherwise, let s_1 and s_2 be two (independent) sections. As the divisor of zeros of s_i is the degree of $\mathscr{L}$, each vanishes at a single point p_i (to order 1). But $p_1 \neq p_2$ (or else s_1/s_2 has no poles or zeros, i.e., is a constant function, i.e., s_1 and s_2 are linearly dependent). Thus we get a map $C \to \mathbb{P}^1$ which is base-point-free. This is a finite degree 1 map of regular curves (Exercise 17.4.F), which hence induces a degree 1 extension of function fields, i.e., an isomorphism of function fields, which means that the curves are isomorphic (cf. Theorem 17.4.3). But we assumed that C is not isomorphic to $\mathbb{P}^1_k$, so we have a contradiction. $\square$

19.4.2. Corollary. — *If C is a projective regular geometrically integral curve over k, not isomorphic to $\mathbb{P}^1_k$ and p and q are degree 1 points, then $\mathscr{O}_C(p) \cong \mathscr{O}_C(q)$ if and only if $p = q$.*

19.4.A. EXERCISE. Show that if k is algebraically closed, then C has genus 0 if and only if all degree 0 line bundles are trivial.

19.4.B. EXERCISE. Suppose C is a regular plane curve of degree $e > 2$, and D_1 and D_2 are two plane curves of the same degree d not containing C. By Bézout's theorem for plane curves (§18.6.3), D_i meets C at de points, counted "correctly". Suppose D_1 and D_2 meet C at $de - 1$ of the "same points", plus one more. Show that the remaining points are the same as well. More precisely, suppose there is a divisor E on C of degree $de - 1$, and degree 1 (k-valued) points p_1 and p_2 such that $D_i \cap C = E + p_i$ (as divisors on C). Show that $p_1 = p_2$. (The case $d = e = 3$ is Chasles' Theorem, and is the first case of the Cayley-Bacharach Theorem, see [**EGH**].)

As an entertaining application of Exercise 19.4.B, we can prove two classical results. For convenience, in this discussion we will assume k is algebraically closed, although this assumption can be easily removed.

19.4.3. Pappus's Theorem (Pappus of Alexandria), see Figure 19.2. — *Suppose ℓ and m are distinct lines in the plane, and α, β, γ are distinct points on ℓ, and α', β', and γ' are distinct points on m, and all six points are distinct from $\ell \cap m$. Then points $x = \overline{\alpha\beta'} \cap \overline{\alpha'\beta}$, $y = \overline{\alpha\gamma} \cap \overline{\alpha\gamma'}$, and $z = \overline{\beta\gamma'} \cap \overline{\beta'\gamma}$ are collinear.*

[figure will be made later]

FIGURE 19.2. Pappus' Theorem

Pascal's "Mystical Hexagon" Theorem was discovered by Pascal at age 16. (What were you doing at age 16?)

19.4.4. Pascal's "Mystical Hexagon" Theorem, see Figure 19.3. — *If a hexagon $\alpha\gamma'\beta\alpha'\gamma\beta'$ is inscribed in a smooth conic X, and opposite pairs of sides are extended until they meet, the three intersection points $x = \overline{\alpha\beta'} \cap \overline{\alpha'\beta}$, $y = \overline{\alpha'\gamma} \cap \overline{\alpha\gamma'}$, and $z = \overline{\beta\gamma'} \cap \overline{\beta'\gamma}$ are collinear.*

[figure will be made later]

FIGURE 19.3. Pascal's "Mystical Hexagon" Theorem

Pappus's Theorem can be seen as a degeneration of Pascal's Theorem: the conic degenerates into the union of two lines, and the six points degenerate so that α, β, and γ are on one, and α', β', and γ' are on the other. We thus prove Pascal's Theorem, and you should check that the proof readily applies to prove Pappus's Theorem.

19.4.C. EXERCISE. Suppose α, β, γ, α', β' are five points in $\mathbb{P}^2_{\mathbb{R}}$, no three on a line. Show that there is a unique conic C passing through the five points. Show that C is regular. Explain how to construct the tangent to C at α using only a straightedge.

(Hint for the last part: apply Pascal's Theorem, taking $\gamma' = \alpha$. You may need to first figure out why you can apply Pascal's Theorem in this degenerate case.)

In particular, given an ellipse $C \subset \mathbb{R}^2$ and a point $\alpha \in C$, you will now be able to construct the tangent to C at α using only a straightedge.

19.4.5. *Proof of Pascal's Theorem 19.4.4.*

We wish to show that the line $\overline{xy}$ meets $\overline{\beta\gamma'}$ and $\overline{\beta'\gamma}$ at the same point. Let C be the curve that is the union of X and the line $\overline{xy}$. (Warning, cf. §19.2.1: for one time in this chapter, we are not assuming C to be regular!)

Let D_1 be the union of the three lines $\overline{\alpha\beta'}$, $\overline{\beta\gamma'}$, and $\overline{\gamma\alpha'}$, and let D_2 be the union of the three lines $\overline{\alpha'\beta}$, $\overline{\beta'\gamma}$, and $\overline{\gamma'\alpha}$. Note that D_1 meets C at the nine points α, β, γ, α', β', γ', x, y, and $\overline{xy} \cap \overline{\beta\gamma'}$, and D_2 meets C at the same nine points, except $\overline{xy} \cap \overline{\beta\gamma'}$ is replaced by $\overline{xy} \cap \overline{\beta'\gamma}$. Thus if we knew Corollary 19.4.2 so that it applied to our C (which is singular), then we would be done (cf. Exercise 19.4.B).

So we extend Corollary 19.4.2 to our situation. So to do this, we extend Proposition 19.4.1 to our situation. We have a plane cubic C which is the union of a line L and a conic K, and two points $p = \overline{xy} \cap \overline{\beta\gamma'}$ and $q = \overline{xy} \cap \overline{\beta'\gamma}$, with $\mathscr{O}(p) \cong \mathscr{O}(q)$. (Reason: both are the restriction of $\mathscr{O}_{\mathbb{P}^2}(3)$ to C, twisted by $-(\alpha + \beta + \gamma + \alpha' + \beta' + \gamma' + x + y)$.) Call this invertible sheaf $\mathscr{L}$. Suppose $p \neq q$. Then the two sections of $\mathscr{L}$ with zeros at p and q give a base-point-free linear series, and thus a morphism $\pi : C \to \mathbb{P}^1$, with $\pi^{-1}(0) = p$ and $\pi^{-1}(\infty) = q$.

By the argument in the proof of Proposition 19.4.1, π gives an isomorphism of L with $\mathbb{P}^1$. As π is proper, $\pi(K)$ is closed, and as K is irreducible, $\pi(K)$ is irreducible. As $\pi(K)$ does not contain 0 (or ∞), it can't be all of $\mathbb{P}^1$. Hence $\pi(K)$ is a point.

The conic K meets the line L in two points with multiplicity (by Bézout's Theorem, §18.6.3, or by simple algebra). If $K \cap C$ consists of two points a and b, we have a contradiction: K can't be contracted to both $\pi(a)$ and $\pi(b)$. But what happens if K meets L at one point, i.e., if the conic is tangent to the line?

19.4.D. EXERCISE. Finish the proof of Pascal's Theorem by dealing with this case. Hint: nilpotents will come to the rescue. The intuition is as follows: $K \cap L$ is a subscheme of L of length 2, but the scheme theoretic image $\pi(K)$ can be shown to be a reduced closed point.

□

19.4.6. *Remark.* The key motivating fact that makes our argument work, Proposition 19.4.1, is centrally about curves not of genus 0, yet all the curves involved in Pappus's Theorem and Pascal's Theorem have genus 0. The insight to keep in mind is that union of curves of genus 0 need not have genus 0. In our case, it mattered that cubic curves have genus 1, even if they are union of $\mathbb{P}^1$'s.

19.5 Hyperelliptic curves

We next discuss an important class of curves, the hyperelliptic curves. In this section, *we assume k is algebraically closed of characteristic not 2.* (These hypotheses can be relaxed, at some cost.)

A (projective regular irreducible) genus g curve C is **hyperelliptic** if it admits a double cover of (i.e., degree 2, necessarily finite, morphism to) $\mathbb{P}^1_k$. For convenience, when we say C is hyperelliptic, we will implicitly have in mind a *choice* of double cover $\pi : C \to \mathbb{P}^1$. (We will later see that if $g \geq 2$, then there is at most one such double cover, Proposition 19.5.7, so this is not a huge assumption.) The map π is called the **hyperelliptic map**.

By Exercise 17.4.D, the preimage of any closed point p of $\mathbb{P}^1$ consists of either one or two points. If $\pi^{-1}(p)$ is a single point, we say p is a **branch point**, and $\pi^{-1}(p)$ is a **ramification point** of π. (The notion of ramification will be defined more generally in §21.6.)

19.5.1. Theorem (hyperelliptic Riemann-Hurwitz formula). — *Suppose* $k = \overline{k}$ *and* $\operatorname{char} k \neq 2$, $\pi : C \to \mathbb{P}^1_k$ *is a double cover by a projective regular irreducible genus* g *curve over* k. *Then* π *has* $2g + 2$ *branch points.*

This is a special case of the Riemann-Hurwitz formula, which we will state and prove in §21.7. You may have already heard about genus 1 complex curves double covering $\mathbb{P}^1$, branched over 4 points.

To prove Theorem 19.5.1, we first prove the following.

19.5.2. Proposition. — *Assume* $\operatorname{char} k \neq 2$ *and* $k = \overline{k}$. *Given* r *distinct points* $p_1, \dots, p_r \in \mathbb{P}^1$, *there is precisely one double cover branched at precisely these points if* r *is even, and none if* r *is odd.*

Proof. Pick points 0 and ∞ of $\mathbb{P}^1$ distinct from the r branch points. All r branch points are in $\mathbb{P}^1 - \infty = \mathbb{A}^1 = \operatorname{Spec} k[x]$. Suppose we have a double cover of $\mathbb{A}^1$, $C' \to \mathbb{A}^1$, where x is the coordinate on $\mathbb{A}^1$. This induces a quadratic field extension K over $k(x)$. As $\operatorname{char} k \neq 2$, this extension is Galois. Let $\sigma : K \to K$ be the Galois involution. Let y be a nonzero element of K such that $\sigma(y) = -y$, so 1 and y form a basis for K over the field $k(x)$, and are eigenvectors of σ. Now $\sigma(y^2) = y^2$, so $y^2 \in k(x)$. We can replace y by an appropriate $k(x)$-multiple so that y^2 is a polynomial, with no repeated factors, and monic. (This is where we use the hypothesis that k is algebraically closed, to get leading coefficient 1.)

Thus $y^2 = x^N + a_{N-1}x^{N-1} + \cdots + a_0$, where the polynomial on the right (call it $f(x)$) has no repeated roots. The Jacobian criterion (in the guise of Exercise 12.2.D) implies that this curve C'_0 in $\mathbb{A}^2 = \operatorname{Spec} k[x, y]$ is regular. Then C'_0 is normal and has the same function field as C. Thus C'_0 and C' are both normalizations of $\mathbb{A}^1$ in the finite field extension generated by y, and hence are isomorphic. Thus we have identified C' in terms of an explicit equation.

The branch points correspond to those values of x for which there is exactly one value of y, i.e., the roots of $f(x)$. In particular, $N = r$, and $f(x) = (x-p_1)\cdots(x-p_r)$, where the p_i are interpreted as elements of $\overline{k}$.

Having mastered the situation over $\mathbb{A}^1$, we return to the situation over $\mathbb{P}^1$. We will examine the branched cover over the affine open set $\mathbb{P}^1 \setminus \{0\} = \operatorname{Spec} k[u]$, where $u = 1/x$. The previous argument applied to $\operatorname{Spec} k[u]$ rather than $\operatorname{Spec} k[x]$ shows that any such double cover must be of the form

$$\begin{aligned} C'' &= \operatorname{Spec} k[z, u]/(z^2 - (u - 1/p_1)\cdots(u - 1/p_r)) = \operatorname{Spec} k[z, u]/(z^2 - u^r f(1/u)) \\ &\to \operatorname{Spec} k[u] = \mathbb{A}^1. \end{aligned}$$

So if there is a double cover over all of $\mathbb{P}^1$, it must be obtained by gluing C'' to C', "over" the gluing of $\operatorname{Spec} k[x]$ to $\operatorname{Spec} k[u]$ to obtain $\mathbb{P}^1$.

Thus in $K(C)$, we must have

$$z^2 = u^r f(1/u) = f(x)/x^r = y^2/x^r$$

from which $z^2 = y^2/x^r$.

If r is even, considering $K(C)$ as generated by y and x, there are two possible values of z: $z = \pm y/x^{r/2}$. After renaming z by $-z$ if necessary, there is a single way of gluing these two patches together (we choose the positive square root).

If r is odd, the result follows from Exercise 19.5.A below.

19.5.A. EXERCISE. Suppose char $k \neq 2$. Show that x does not have a square root in the field $k(x)[y]/(y^2 - f(x))$, where f is a polynomial with nonzero roots $p_1, \dots, p_r$. (Possible hint: why is $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$?) Explain how this proves Proposition 19.5.2 in the case where r is odd.

□

For future reference, we collect here our explicit (two-affine) description of the hyperelliptic cover $C \to \mathbb{P}^1$.

$$\begin{tikzcd}[column sep=large]
\operatorname{Spec} k[x,y]/(y^2 - f(x)) \arrow[r, dash, "z=y/x^{r/2}", "y=z/u^{r/2}"'] \arrow[d] & \operatorname{Spec} k[u,z]/(z^2 - u^r f(1/u)) \arrow[d] \\
\operatorname{Spec} k[x] \arrow[r, dash, "u=1/x", "x=1/u"'] & \operatorname{Spec} k[u]
\end{tikzcd} \tag{19.5.2.1}$$

19.5.3. *If k is not algebraically closed.* If k is not algebraically closed (but of characteristic not 2), the above argument shows that if we have a double cover of $\mathbb{A}^1$, then it is of the form $y^2 = af(x)$, where f is monic (and $a \neq 0$). Furthermore, if a and a' differ (multiplicatively) by an element of $(k^\times)^2$, then $y^2 = af(x)$ is isomorphic to $y^2 = a'f(x)$. You may be able to use this to show that (assuming that $k^\times \neq (k^\times)^2$) a double cover is *not* determined by its branch points. Moreover, this failure is classified by $k^\times/(k^\times)^2$. Thus we have lots of curves that are not isomorphic over k, but become isomorphic over $\overline{k}$. These are often called *twists* of each other.

(In particular, once we define elliptic curves, you will be able to show that there exist two elliptic curves over $\mathbb{Q}$ with the same j-invariant, that are not isomorphic, see Exercise 19.9.D.)

19.5.4. *Back to proving the hyperelliptic Riemann-Hurwitz formula, Theorem 19.5.1.* Our explicit description of the unique double cover of $\mathbb{P}^1$ branched over r different points will allow us to compute the genus, thereby completing the proof of Theorem 19.5.1.

We continue the notation (19.5.2.1) of the proof of Proposition 19.5.2. Suppose $\mathbb{P}^1$ has affine cover by $\operatorname{Spec} k[x]$ and $\operatorname{Spec} k[u]$, with $u = 1/x$, as usual. Suppose $C \to \mathbb{P}^1$ is a double cover, given by $y^2 = f(x)$ over $\operatorname{Spec} k[x]$, where f has degree r, and $z^2 = u^r f(1/u)$. Then C has an affine open cover by $\operatorname{Spec} k[x,y]/(y^2 - f(x))$

and Spec $k[u,z]/(z^2 - u^r f(1/u))$. The corresponding Čech complex for $\mathscr{O}_C$ is

$$0 \longrightarrow k[x,y]/(y^2 - f(x)) \times k[u,z]/(z^2 - u^r f(1/u)) \xrightarrow{d}$$

$$\left(k[x,y]/(y^2 - f(x))\right)_x \longrightarrow 0.$$

The degree 1 part of the complex has basis consisting of monomials $x^n y^\epsilon$, where $n \in \mathbb{Z}$ and $\epsilon = 0$ or 1. To compute the genus $g = h^1(C, \mathscr{O}_C)$, we must compute coker d. We can use the first factor $k[x,y]/(y^2 - f(x))$ to hit the monomials $x^n y^\epsilon$ where $n \in \mathbb{Z}^{\geq 0}$, and $\epsilon = 0$ or 1. The image of the second factor is generated by elements of the form $u^m z^\epsilon$, where $m \geq 0$ and $\epsilon = 0$ or 1. But $u^m z^\epsilon = x^{-m}(y/x^{r/2})^\epsilon$. By inspection, the cokernel has basis generated by monomials $x^{-1}y$, $x^{-2}y$, ..., $x^{-r/2+1}y$, and thus has dimension $r/2 - 1$. Hence $g = r/2 - 1$, from which Theorem 19.5.1 follows. □

19.5.5. *Curves of every genus.* As a consequence of the hyperelliptic Riemann-Hurwitz formula (Theorem 19.5.1), we see that there are curves of every genus $g \geq 0$ over an algebraically closed field of characteristic not 2: to get a curve of genus g, consider the branched cover branched over $2g+2$ distinct points. The unique genus 0 curve is of this form, and we will see in §19.6.2 that every genus 2 curve is of this form. We will soon see that every genus 1 curve (reminder: over an algebraically closed field!) is too (§19.9.5). But it is too much to hope that all curves are of this form, and we will soon see (§19.7.2) that there are genus 3 curves that are *not* hyperelliptic, and we will get heuristic evidence that "most" genus 3 curves are not hyperelliptic. We will later give vague evidence (that can be made precise) that "most" genus g curves are not hyperelliptic if $g > 2$ (§19.8.2).

19.5.B. EXERCISE. Verify that a curve C of genus at least 1 admits a degree 2 cover of $\mathbb{P}^1$ if and only if it admits a degree 2 invertible sheaf $\mathscr{L}$ with $h^0(C, \mathscr{L}) = 2$. Possibly in the course of doing this, verify that if C is a curve, and $\mathscr{L}$ has a degree 2 invertible sheaf with at least 2 (linearly independent) sections, then $\mathscr{L}$ has precisely two sections, and that this $\mathscr{L}$ is base-point-free and gives a hyperelliptic map.

19.5.6. Proposition. — *If $\mathscr{L}$ corresponds to a hyperelliptic cover $C \to \mathbb{P}^1$, then $\mathscr{L}^{\otimes(g-1)} \cong \omega_C$.*

Proof. Compose the hyperelliptic map with the $(g-1)$th Veronese map:

$$C \xrightarrow{|\mathscr{L}|} \mathbb{P}^1 \xrightarrow{|\mathscr{O}_{\mathbb{P}^1}(g-1)|} \mathbb{P}^{g-1}. \tag{19.5.6.1}$$

The composition corresponds to $\mathscr{L}^{\otimes(g-1)}$. This invertible sheaf has degree $2g-2$ (by Exercise 18.4.F). The pullback $H^0(\mathbb{P}^{g-1}, \mathscr{O}(1)) \to H^0(C, \mathscr{L}^{\otimes(g-1)})$ is injective because the image of C in $\mathbb{P}^{g-1}$ (a rational normal curve) is nondegenerate: if there were a hyperplane $s \in H^0(\mathbb{P}^{g-1}, \mathscr{O}(1))$ that pulled back to 0 on C, then the image of C would lie in that hyperplane, yet a rational normal curve cannot. Thus $\mathscr{L}^{\otimes(g-1)}$ has at least g sections. But by Exercise 19.2.A, the only invertible sheaf of degree $2g-2$ with (at least) g sections is the dualizing sheaf. □

As an added bonus, we see that the composition of (19.5.6.1) is the *complete* linear series $|\mathscr{L}^{\otimes(g-1)}|$ — all sections of $\mathscr{L}^{\otimes(g-1)}$ come up in this way.

19.5.7. Proposition (a genus ≥ 2 curve can be hyperelliptic in only one way). — *Any curve C of genus at least 2 admits at most one double cover of $\mathbb{P}^1$. More precisely, if $\mathscr{L}$ and $\mathscr{M}$ are two degree two line bundles yielding maps $C \to \mathbb{P}^1$, then $\mathscr{L} \cong \mathscr{M}$.*

Proof. If C is hyperelliptic, then we can recover the hyperelliptic map by considering the canonical linear series given by ω_C (the *canonical map*, which we will use again repeatedly in the next few sections): it is a double cover of a degree $g-1$ rational normal curve (by the previous proposition), which is isomorphic to $\mathbb{P}^1$. This double cover is the hyperelliptic cover (also by the proof of the previous proposition). Thus we have uniquely recovered the map $C \to \mathbb{P}^1$, and this map must be induced by both $\mathscr{L}$ and $\mathscr{M}$, from which we have $\mathscr{L} \cong \mathscr{M}$ (using Theorem 16.4.1, relating maps to projective space and line bundles). □

19.5.8. *The "space of hyperelliptic curves".* Thanks to Proposition 19.5.7, we can now classify hyperelliptic curves of genus at least 2. Hyperelliptic curves of genus $g \geq 2$ correspond to precisely $2g+2$ distinct points on $\mathbb{P}^1$ modulo S_{2g+2}, and modulo automorphisms of $\mathbb{P}^1$. Thus "the space of hyperelliptic curves" has dimension

$$2g+2-\dim \operatorname{Aut} \mathbb{P}^1 = 2g-1.$$

This is not a well-defined statement, because we haven't rigorously defined "the space of hyperelliptic curves" — it is an example of a *moduli space*. For now, take this as a plausibility statement. It is also plausible that this space is irreducible and reduced — it is the image of something irreducible and reduced.

19.6 Curves of genus 2

19.6.1. *The reason for leaving genus 1 for later.* It might make most sense to jump to genus 1 at this point, but the theory of elliptic curves is especially rich and subtle, so we will leave it for §19.9.

In general, curves have quite different behaviors (topologically, arithmetically, geometrically) depending on whether $g=0$, $g=1$, or $g \geq 2$. This trichotomy extends to varieties of higher dimension. We already have some inkling of it in the case of curves. Arithmetically, genus 0 curves can have lots and lots of rational points, genus 1 curves can have lots of rational points, and by Faltings' Theorem (Mordell's Conjecture) any curve of genus at least 2 has at most finitely many rational points. (Thus even before Wiles' proof of the Taniyama-Shimura conjecture, we knew that $x^n+y^n=z^n$ in $\mathbb{P}^2$ has at most finitely many rational solutions for $n \geq 4$, as such curves have genus $\binom{n-1}{2} > 1$, see (18.6.6.1).) In the language of differential geometry, Riemann surfaces of genus 0 are positively curved, Riemann surfaces of genus 1 are flat, and Riemann surfaces of genus 1 are negatively curved. It is a fact that curves of genus at least 2 have finite automorphism groups (see §21.7.8), while curves of genus 1 have one-dimensional automorphism groups, see Question 19.10.6), and the unique curve of genus 0 over an algebraically closed field has a three-dimensional automorphism group (see Exercises 16.4.B and 16.4.C). (See Exercise 21.5.G for more on this issue.)

19.6.2. *Back to curves of genus* 2.

Over an algebraically closed field, we saw in §19.3 that there is only one genus 0 curve. In §19.5 we established that there are hyperelliptic curves of genus 2. How can we get a hold of curves of genus 2? For example, are they all hyperelliptic? "How many" are there? We now tackle these questions.

Fix a curve C of genus $g = 2$. Then ω_C is degree $2g - 2 = 2$, and has 2 sections (Exercise 19.2.A). I claim that ω_C is base-point-free. We may assume k is algebraically closed, as base-point-freeness is independent of field extension of k (Exercise 18.2.I). If p is a base point of ω_C, then $\omega_C(-p)$ is a degree 1 invertible sheaf with 2 sections, which Proposition 19.4.1 shows is impossible. Thus we canonically constructed a double cover $C \to \mathbb{P}^1$ (unique up to automorphisms of $\mathbb{P}^1$, which we studied in Exercises 16.4.B and 16.4.C). Conversely, any double cover $C \to \mathbb{P}^1$ arises from a degree 2 invertible sheaf with at least 2 sections, so if $g(C) = 2$, this invertible sheaf must be the dualizing sheaf (by the easiest case of Proposition 19.5.6).

Hence we have a natural bijection between genus 2 curves and genus 2 double covers of $\mathbb{P}^1$ (up to automorphisms of $\mathbb{P}^1$). If the characteristic is not 2, the hyperelliptic Riemann-Hurwitz formula (Theorem 19.5.1) shows that the double cover is branched over $2g + 2 = 6$ geometric points. In particular, we have a "three-dimensional space of genus 2 curves". This isn't rigorous, but we can certainly show that there are an infinite number of genus 2 curves. Precisely:

19.6.A. EXERCISE. Fix an algebraically closed field k of characteristic not 2. Show that there are an infinite number of (pairwise) non-isomorphic genus 2 curves k.

19.6.B. EXERCISE. Show that every genus 2 curve (over any field of characteristic not 2) has finite automorphism group.

19.7 Curves of genus 3

Suppose C is a curve of genus 3. Then ω_C has degree $2g-2 = 4$, and has $g = 3$ sections.

19.7.1. Claim. — *The invertible sheaf ω_C is base-point-free, and hence gives a map to $\mathbb{P}^2$.*

Proof. We check base-point-freeness by working over the algebraic closure $\overline{k}$ (which we can, by Exercise 18.2.I). For any point p, by Riemann-Roch,

$$h^0(C, \omega_C(-p)) - h^0(C, \mathscr{O}(p)) = \deg(\omega_C(-p)) - g + 1 = 3 - 3 + 1 = 1.$$

But $h^0(C, \mathscr{O}(p)) = 1$ by Proposition 19.4.1, so

$$h^0(C, \omega_C(-p)) = 2 = h^0(C, \omega_C) - 1.$$

Thus p is not a base-point of ω_C for any p, so by Criterion 19.2.7 ω_C is base-point-free. $\square$

The next natural question is: Is this a closed embedding? Again, we can check over algebraic closure. We use our "closed embedding test" (again, see our useful facts). If it *isn't* a closed embedding, then we can find two points p and q (possibly identical) such that

$$h^0(C, \omega_C) - h^0(C, \omega_C(-p-q)) = 1 \text{ or } 0,$$

i.e., $h^0(C, \omega_C(-p-q)) = 2$. But by Serre duality, this means that $h^0(C, \mathscr{O}(p+q)) = 2$. We have found a degree 2 divisor with 2 sections, so C is hyperelliptic. Conversely, if C is hyperelliptic, then we already know that ω_C gives a double cover of a regular conic in $\mathbb{P}^2$, and hence ω_C does not give a closed embedding.

Thus we conclude that if (and only if) C is not hyperelliptic, then the canonical map describes C as a degree 4 curve in $\mathbb{P}^2$.

Conversely, any quartic plane curve is canonically embedded. Reason: the curve has genus 3 (see (18.6.6.1)), and is mapped by an invertible sheaf of degree 4 with 3 sections. But by Exercise 19.2.A, the only invertible sheaf of degree $2g-2$ with g sections is ω_C.

In particular, each non-hyperelliptic genus 3 curve can be described as a quartic plane curve in only one way (up to automorphisms of $\mathbb{P}^2$).

In conclusion, there is a bijection between non-hyperelliptic genus 3 curves, and plane quartics up to projective linear transformations.

19.7.2. *Remark.* In particular, as there exist regular plane quartics (Exercise 12.3.E), there exist non-hyperelliptic genus 3 curves.

19.7.A. EXERCISE. Give a heuristic (non-rigorous) argument that the nonhyperelliptic curves of genus 3 form a family of dimension 6. (Hint: Count the dimension of the family of regular quartics, and quotient by $\operatorname{Aut}\mathbb{P}^2 = PGL(3)$.)

The genus 3 curves thus seem to come in two families: the hyperelliptic curves (a family of dimension 5), and the nonhyperelliptic curves (a family of dimension 6). This is misleading — they actually come in a single family of dimension 6.

In fact, hyperelliptic curves are naturally limits of nonhyperelliptic curves. We can write down an explicit family. (This explanation necessarily requires some hand-waving, as it involves topics we haven't seen yet.) Suppose we have a hyperelliptic curve branched over $2g+2 = 8$ points of $\mathbb{P}^1$. Choose an isomorphism of $\mathbb{P}^1$ with a conic in $\mathbb{P}^2$. There is a regular quartic meeting the conic at precisely those 8 points. (This requires a short argument using Bertini's Theorem 12.4.2, which we omit.) Then if f is the equation of the conic, and g is the equation of the quartic, then $f^2 + t^2 g$ is a family of quartics that are smooth for most t (smoothness is an open condition in t, as we will see in Theorem 25.3.3 on generic smoothness). The $t = 0$ case is a double conic. Then it is a fact that if you normalize the family, the central fiber (above $t = 0$) turns into our hyperelliptic curve. Thus we have expressed our hyperelliptic curve as a limit of nonhyperelliptic curves.

19.7.B. UNIMPORTANT EXERCISE. A (projective) curve (over a field k) admitting a degree 3 cover of $\mathbb{P}^1$ is called **trigonal**. Show that every non-hyperelliptic genus 3 complex curve is trigonal, by taking the quartic model in $\mathbb{P}^2$, and projecting to $\mathbb{P}^1$ from any point on the curve. Do this by choosing coordinates on $\mathbb{P}^2$ so that p is at $[0,0,1]$. (The same idea, applied to cubics rather than quartics, will be used in §19.9.9.)

19.7.3. $\star$ **A genus** 3 **curve with no nontrivial automorphisms.**

We have seen that a (smooth projective integrable) curve of genus at most 2 always has nontrivial automorphisms. It turns out that there are genus 3 curves with no nontrivial automorphisms.

19.7.C. EXERCISE. Suppose $C' \subset \mathbb{P}^2$ is a smooth plane quartic curve (over any field k). Show that there is bijection between automorphisms of C' and automorphisms of $\mathbb{P}^2$ preserving C' (as a set).

Thus to find a genus 3 curve with no nontrivial automorphisms, we need only find a smooth quartic plane curve C' such that the only automorphism of $\mathbb{P}^2$ fixing C' as a set must be the identity. Your intuition may (correctly) tell you that most quartics are of this form. But exhibiting a specific C' (with proof) requires rolling up our sleeves and getting to work. Poonen gives automorphism free curves over any field in [**P1**]; an example in characteristic 0 is

$$y^3z - 3yz^3 = 3x^4 - 4x^3z + z^4.$$

19.7.D. EXERCISE. Suppose C is a smooth projective curve with no nontrivial automorphisms. Show that no two open subsets of C are isomorphic.

19.7.4. *Genus* 3 *curves with nontrivial automorphisms.*

Certainly genus 3 curves can have automorphisms: witness hyperelliptic curves. More impressive is the Klein quartic

$$x^3y + y^3z + z^3x = 0$$

in $\mathbb{P}^2_{\mathbb{C}}$, which has 168 automorphisms. (Can you find them all?) In fact, the automorphism group of the Klein quartic is the unique finite simple group of order 168 (the second-smallest nonabelian finite simple group). In §21.7.8, we will see that no genus 3 curve can have more (assuming the hard fact that curves of genus greater than 1 have finite automorphism groups).

19.8 Curves of genus 4 and 5

We begin with two exercises in general genus, then specialize to genus 4.

19.8.A. EXERCISE. Assume $k = \overline{k}$ (purely to avoid distraction — feel free to remove this hypothesis). Suppose C is a genus g curve. Show that if C is not hyperelliptic, then the dualizing invertible sheaf gives a closed embedding $C \hookrightarrow \mathbb{P}^{g-1}$. (In the hyperelliptic case, we have already seen that the dualizing sheaf gives us a double cover of a rational normal curve.) Hint: follow the genus 3 case. Such a curve is called a **canonical curve**, and this closed embedding is called the **canonical embedding** of C.

19.8.B. EXERCISE. Suppose C is a curve of genus $g > 1$, over a field k that is not algebraically closed. Show that C has a closed point of degree at most $2g - 2$ over the base field. (For comparison: if $g = 1$, for any n, there is a genus 1 curve over $\mathbb{Q}$ with no point of degree less than n!)

We next consider nonhyperelliptic curves C of genus 4. Note that $\deg \omega_C = 6$ and $h^0(C, \omega_C) = 4$, so the canonical map expresses C as a sextic curve in $\mathbb{P}^3$. We shall see that all such C are complete intersections of quadric surfaces and cubic surfaces, and conversely all regular complete intersections of quadrics and cubics are genus 4 non-hyperelliptic curves, canonically embedded.

By (19.2.5.1) (Riemann-Roch and Serre duality),

$$h^0(C, \omega_C^{\otimes 2}) = \deg \omega_C^{\otimes 2} - g + 1 = 12 - 4 + 1 = 9.$$

We have the restriction map $H^0(\mathbb{P}^3, \mathscr{O}(2)) \to H^0(C, \omega_C^{\otimes 2})$, and $\dim \operatorname{Sym}^2 \Gamma(C, \omega_C) = \binom{4+1}{2} = 10$. Thus there is at least one quadric in $\mathbb{P}^3$ that vanishes on our curve C. Translation: C lies on at least on quadric Q. Now quadrics are either double planes, or the union of two planes, or cones, or regular quadrics. (They corresponds to quadric forms of rank 1, 2, 3, and 4 respectively.) But C can't lie in a plane, so Q must be a cone or regular. In particular, Q is irreducible.

Now C can't lie on *two* (distinct) such quadrics, say Q and Q'. Otherwise, as Q and Q' have no common components (they are irreducible and not the same!), $Q \cap Q'$ is a curve (not necessarily reduced or irreducible). By Bézout's theorem (Exercise 18.6.K), $Q \cap Q'$ is a curve of degree 4. Thus our curve C, being of degree 6, cannot be contained in $Q \cap Q'$. (If you don't see why directly, Exercise 18.6.F might help.)

We next consider cubic surfaces. By (19.2.5.1) again, $h^0(C, \omega_C^{\otimes 3}) = \deg \omega_C^{\otimes 3} - g + 1 = 18 - 4 + 1 = 15$. Now $\operatorname{Sym}^3 \Gamma(C, \omega_C)$ has dimension $\binom{4+2}{3} = 20$. Thus C lies on at least a 5-dimensional vector space of cubics. Now a 4-dimensional subspace come from multiplying the quadric Q by a linear form $(?w + ?x + ?y + ?z)$. But hence there is still one cubic K whose underlying form is not divisible by the quadric form Q (i.e., K doesn't contain Q.) Then K and Q share no component, so $K \cap Q$ is a complete intersection containing C as a closed subscheme. Now $K \cap Q$ and C are both degree 6 (the former by Bézout's theorem, Exercise 18.6.K, and the latter because C is embedded by a degree 6 line bundle, Exercise 18.6.I). Also, $K \cap Q$ and C both have arithmetic genus 4 (the former by Exercise 18.6.S). These two invariants determine the (linear) Hilbert polynomial, so $K \cap Q$ and C have the same Hilbert polynomial. Hence $C = K \cap Q$ by Exercise 18.6.F.

We now show the converse, and that any regular complete intersection C of a quadric surface with a cubic surface is a canonically embedded genus 4 curve. By Exercise 18.6.S, such a complete intersection has genus 4.

19.8.C. EXERCISE. Show that $\mathscr{O}_C(1)$ has at least 4 sections. (Translation: C doesn't lie in a hyperplane.)

The only degree $2g-2$ invertible sheaf with (at least) g sections is the dualizing sheaf (Exercise 19.2.A), so $\mathscr{O}_C(1) \cong \omega_C$, and C is indeed canonically embedded.

19.8.D. EXERCISE. Give a heuristic argument suggesting that the nonhyperelliptic curves of genus 4 "form a family of dimension 9".

On to genus 5!

19.8.E. EXERCISE. Suppose C is a nonhyperelliptic genus 5 curve. Show that the canonical curve is degree 8 in $\mathbb{P}^4$. Show that it lies on a three-dimensional vector space of quadrics (i.e., it lies on 3 linearly independent independent quadrics).

Show that a regular complete intersection of 3 quadrics is a canonical(ly embedded) genus 5 curve.

Unfortunately, not all canonical genus 5 curves are the complete intersection of 3 quadrics in $\mathbb{P}^4$. But in the same sense that most genus 3 curves can be described as plane quartics, most canonical genus 5 curves are complete intersections of 3 quadrics, and most genus 5 curves are non-hyperelliptic. The correct way to say this is that there is a dense Zariski-open locus in the moduli space of genus 5 curves consisting of nonhyperelliptic curves whose canonical embedding is cut out by 3 quadrics.

(Those nonhyperelliptic genus 5 canonical curves not cut out by a three-dimensional vector space of quadrics are precisely the trigonal curves, see Exercise 19.7.B. The triplets of points mapping to the same point of $\mathbb{P}^1$ under the trigonal map turn out to lie on a line in the canonical map. Any quadric vanishing along those 3 points must vanish along the line — basically, any quadratic polynomial with three zeros must be the zero polynomial.)

19.8.F. EXERCISE. Assuming the discussion above, count complete intersections of three quadrics to give a heuristic argument suggesting that the curves of genus 5 "form a family of dimension 12".

19.8.1. We have now understood curves of genus 3 through 5 by thinking of canonical curves as complete intersections. Sadly our luck has run out.

19.8.G. EXERCISE. Show that if $C \subset \mathbb{P}^{g-1}$ is a canonical curve of genus $g \geq 6$, then C is *not* a complete intersection. (Hint: Bézout's theorem, Exercise 18.6.K.)

19.8.2. *Some discussion on curves of general genus.* However, we still have some data. If $\mathscr{M}_g$ is this ill-defined "moduli space of genus g curves", we have heuristics to find its dimension for low g. In genus 0, over an algebraically closed field, there is only genus 0 curve (Proposition 19.3.1), so it appears that $\dim \mathscr{M}_0 = 0$. In genus 1, over an algebraically closed field, we will soon see that the elliptic curves are classified by the j-invariant (Exercise 19.9.C), so it appears that $\dim \mathscr{M}_1 = 1$. We have also informally computed $\dim \mathscr{M}_2 = 3$, $\dim \mathscr{M}_3 = 6$, $\dim \mathscr{M}_4 = 9$, $\dim \mathscr{M}_5 = 12$. What is the pattern? In fact in some strong sense it was known by Riemann that $\dim \mathscr{M}_g = 3g - 3$ for $g > 1$. What goes wrong in genus 0 and genus 1? As a clue, recall our insight when discussing Hilbert functions (§18.6) that whenever some function is "eventually polynomial", we should assume that it "wants to be polynomial", and there is some better function (usually an Euler characteristic) that *is* polynomial, and that cohomology-vanishing ensures that the original function and the better function "eventually agree". Making sense of this in the case of $\mathscr{M}_g$ is far beyond the scope of our current discussion, so we will content ourselves by observing the following facts. *Every* regular curve of genus greater than 1 has a finite number of automorphisms — a zero-dimensional automorphism group. *Every* regular curve of genus 1 has a one-dimensional automorphism group (see Question 19.10.6). And the only regular curve of genus 0 has a three-dimensional automorphism group (Exercise 16.4.C). (See Aside 21.5.12 for more discussion.) So notice that for all $g \geq 0$,

$$\dim \mathscr{M}_g - \dim \operatorname{Aut} C_g = 3g - 3$$

where Aut C_g means the automorphism group of any curve of genus g.

In fact, in the language of stacks (or orbifolds), it makes sense to say that the dimension of the moduli space of (smooth projective geometrically irreducible) genus 0 curves is -3, and the dimension of the moduli space of genus 1 curves is 0.

19.9 Curves of genus 1

Finally, we come to the very rich case of curves of genus 1. We will present the theory by thinking about line bundles of steadily increasing degree.

19.9.1. Line bundles of degree 0.
Suppose C is a genus 1 curve. Then $\deg \omega_C = 2g - 2 = 0$ and $h^0(C, \omega_C) = g = 1$ (by Exercise 19.2.A). But the only degree 0 invertible sheaf with a section is the structure sheaf (§19.2.4), so we conclude that $\omega_C \cong \mathscr{O}_C$. (If you know that complex genus 1 curves are of the form $\mathbb{C}$ modulo a lattice, and Miracle 18.5.2 that the sheaf of differentials is dualizing, then you might not be surprised.)

We move on to line bundles of higher degree. Next, note that if $\deg \mathscr{L} > 0$, then Riemann-Roch and Serre duality (19.2.5.1) give

$$h^0(C, \mathscr{L}) = \deg \mathscr{L} - g + 1 = \deg \mathscr{L}.$$

19.9.2. Line bundles of degree 1.
Each degree 1 (k-valued) point q determines a line bundle $\mathscr{O}(q)$, and two distinct points determine two distinct line bundles (as a degree 1 line bundle has only one section, up to scalar multiples). Conversely, any degree 1 line bundle $\mathscr{L}$ is of the form $\mathscr{O}(q)$ (as $\mathscr{L}$ has a section — then just take its divisor of zeros).

Thus we have a canonical bijection between degree 1 line bundles and degree 1 (closed) points. (If k is algebraically closed, as all closed points have residue field k, this means that we have a canonical bijection between degree 1 line bundles and closed points.)

Define an **elliptic curve** to be a genus 1 curve E with a choice of k-valued point p. The choice of this point should always be considered part of the definition of an elliptic curve — "elliptic curve" is not a synonym for "genus 1 curve". (Note: a genus 1 curve need not have any k-valued points at all! For example, the genus 1 curve "compactifying" $y^2 = -x^4 - 1$ in $\mathbb{A}^2_{\mathbb{Q}}$ has no $\mathbb{R}$-points, and hence no $\mathbb{Q}$-points. Of course, if $k = \overline{k}$, then any closed point is k-valued, by the Nullstellensatz 3.2.4.) We will often denote elliptic curves by E rather than C.

If (E, p) is an elliptic curve, then there is a canonical bijection between the set of degree 0 invertible sheaves (up to ismorphism) and the set of degree 1 points of E: simply the twist the degree 1 line bundles by $\mathscr{O}(-p)$. Explicitly, the bijection is given by

$$\mathscr{L} \longmapsto \operatorname{div}(\mathscr{L}(p))$$

$$\mathscr{O}(q - p) \longleftarrow\!\shortmid\ q$$

But the degree 0 invertible sheaves form a group (under tensor product), so we have proved:

19.9.3. Proposition (the group law on the degree 1 points of an elliptic curve). — *The above bijection defines an abelian group structure on the degree 1 points of an elliptic curve, where p is the identity.*

From now on, we will identify closed points of E with degree 0 invertible sheaves on E without comment.

For those familiar with the complex analytic picture, this isn't surprising: E is isomorphic to the complex numbers modulo a lattice: $E \cong \mathbb{C}/\Lambda$.

Proposition 19.9.3 is currently just a bijection of sets. Given that E has a much richer structure (it has a generic point, and the structure of a variety), this is a sign that there should be a way of defining some *scheme* $\operatorname{Pic}^0(E)$, and that this should be an isomorphism of schemes. We will soon show (Theorem 19.10.4) that this group structure on the degree 1 points of E comes from a group variety structure on E.

19.9.4. *Aside: The Mordell-Weil Theorem, group, and rank.* This is a good excuse to mention the *Mordell-Weil Theorem*: for any elliptic curve E over $\mathbb{Q}$, the $\mathbb{Q}$-points of E form a *finitely generated* abelian group, often called the *Mordell-Weil group*. By the classification of finitely generated abelian groups (a special case of the classification of finitely generated modules over a principal ideal domain, Remark 12.5.15), the $\mathbb{Q}$-points are a direct sum of a torsion part, and of a free $\mathbb{Z}$-module. The rank of the $\mathbb{Z}$-module is called the *Mordell-Weil rank*.

19.9.5. Line bundles of degree 2.

Note that $\mathscr{O}_E(2p)$ has 2 sections, so E admits a double cover of $\mathbb{P}^1$ (Exercise 19.5.B). One of the branch points is $2p$: one of the sections of $\mathscr{O}_E(2p)$ vanishes to p of order 2, so there is a point of $\mathbb{P}^1$ consists of p (with multiplicity 2). Assume now that $k = \overline{k}$ and $\operatorname{char} k \neq 2$, so we can use the hyperelliptic Riemann-Hurwitz formula (Theorem 19.5.1), which implies that E has 4 branch points (p and three others). Conversely, given 4 points in $\mathbb{P}^1$, there exists a unique double cover branched at those 4 points (Proposition 19.5.2). Thus elliptic curves correspond to 4 distinct points in $\mathbb{P}^1$, where one is marked p, up to automorphisms of $\mathbb{P}^1$. Equivalently, by placing p at ∞, elliptic curves correspond to 3 points in $\mathbb{A}^1$, up to affine maps $x \mapsto ax + b$.

19.9.A. EXERCISE. Show that the other three branch points are precisely the (non-identity) 2-torsion points in the group law. (Hint: if one of the points is q, show that $\mathscr{O}(2q) \cong \mathscr{O}(2p)$, but $\mathscr{O}(q)$ is not congruent to $\mathscr{O}(p)$.)

Thus (if $\operatorname{char} k \neq 2$ and $k = \overline{k}$) every elliptic curve has precisely four 2-torsion points. If you are familiar with the complex picture $E \cong \mathbb{C}/\Lambda$, this isn't surprising.

19.9.6. *Follow-up remark.* An elliptic curve with *full level n-structure* is an elliptic curve with an isomorphism of its n-torsion points with $(\mathbb{Z}/n)^2$. (This notion has problems if n is divisible by $\operatorname{char} k$.) Thus an elliptic curve with *full level 2 structure* is the same thing as an elliptic curve with an ordering of the three other branch points in its degree 2 cover description. Thus (if $k = \overline{k}$) these objects are parametrized by the λ-line, which we discuss below.

Follow-up to the follow-up. There is a notion of moduli spaces of elliptic curves with full level n structure. Such moduli spaces are smooth curves (where this is interpreted appropriately — they are stacks), and have smooth compactifications. A *weight* k *level* n *modular form* is a section of $\omega_M^{\otimes k}$ where ω_M is the dualizing (or canonical) sheaf of this moduli space ("modular curve").

19.9.7. *The cross-ratio and the* j*-invariant.* If the three other points are temporarily labeled q_1, q_2, q_3, there is a unique automorphism of $\mathbb{P}^1$ taking p, q_1, q_2 to $(\infty, 0, 1)$ respectively (as $\operatorname{Aut}\mathbb{P}^1$ is three-transitive, Exercise 16.4.C). Suppose that q_3 is taken to some number λ under this map, where necessarily $\lambda \neq 0, 1, \infty$.

The value λ is called the **cross-ratio** of the four-points (p, q_1, q_2, q_3) of $\mathbb{P}^1$ (first defined by Clifford, but implicitly known since the time of classical Greece).

19.9.B. EXERCISE. Show that isomorphism class of four ordered distinct points on $\mathbb{P}^1$, up to projective equivalence (automorphisms of $\mathbb{P}^1$), are classified by the cross-ratio.

We have not defined the notion of *moduli space*, but the previous exercise illustrates the fact that $\mathbb{P}^1 - \{0, 1, \infty\}$ (the image of the cross-ratio map) is the moduli space for four ordered distinct points of $\mathbb{P}^1$ up to projective equivalence.

Notice:

- If we had instead sent p, q_2, q_1 to $(\infty, 0, 1)$, then q_3 would have been sent to $1 - \lambda$.
- If we had instead sent p, q_1, q_3 to $(\infty, 0, 1)$, then q_2 would have been sent to $1/\lambda$.
- If we had instead sent p, q_3, q_1 to $(\infty, 0, 1)$, then q_2 would have been sent to $1 - 1/\lambda = (\lambda - 1)/\lambda$.
- If we had instead sent p, q_2, q_3 to $(\infty, 0, 1)$, then q_2 would have been sent to $1/(1 - \lambda)$.
- If we had instead sent p, q_3, q_2 to $(\infty, 0, 1)$, then q_2 would have been sent to $1 - 1/(1 - \lambda) = \lambda/(\lambda - 1)$.

Thus these six values (which correspond to S_3) yield the same elliptic curve, and this elliptic curve will (upon choosing an ordering of the other 3 branch points) yield one of these six values.

This is fairly satisfactory already. To check if two elliptic curves (E, p), (E', p') over $k = \overline{k}$ are isomorphic, we write both as double covers of $\mathbb{P}^1$ ramified at p and p' respectively, then order the remaining branch points, then compute their respective λ's (say λ and λ' respectively), and see if they are related by one of the six expressions above:

$$\lambda' = \lambda,\ 1 - \lambda,\ (\lambda - 1)/\lambda,\ 1/(1 - \lambda),\ \text{or}\ \lambda/(\lambda - 1). \tag{19.9.7.1}$$

It would be far more convenient if, instead of a "six-valued invariant" λ, there were a single invariant (let's call it j), such that $j(\lambda) = j(\lambda')$ if and only if one of the equalities of (19.9.7.1) holds. This j-function should presumably be algebraic, so it would give a map j from the λ-line $\mathbb{A}^1 - \{0, 1\}$ to the $\mathbb{A}^1$. By the Curve-to-Projective Extension Theorem 16.5.1, this would extend to a morphism $j : \mathbb{P}^1 \to \mathbb{P}^1$. By Exercise 17.4.D, because this is (for most λ) a 6-to-1 map, the degree of this cover is 6 (or more correctly, at least 6).

We can make this dream more precise as follows. The elliptic curves over k corresponds to k-valued points of $\mathbb{P}^1 - \{0, 1, \lambda\}$, modulo the action of S_3 on λ given above. Consider the subfield K of $k(\lambda)$ fixed by S_3. Then $k(\lambda)/K$ is necessarily Galois (see for example [**DF**, §14.2, Thm. 9]), and a degree 6 extension. We are hoping that this subfield is of the form $k(j)$, and if so, we would obtain the j-map $\mathbb{P}^1 \to \mathbb{P}^1$ as described above. One could show that K is finitely generated over k, and then invoke Lüroth's theorem, which we will soon prove in Example 21.7.6; but we won't need this.

Instead, we will just hunt for such a j. Note that λ should satisfy a sextic polynomial over $k(\lambda)$ (or more precisely given what we know right now, a polynomial of degree at least six), as for each j-invariant, there are six values of λ in general.

As you are undoubtedly aware, there is such a j-invariant. Here is the formula for the j-invariant that everyone uses:

$$j = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}. \tag{19.9.7.2}$$

You can readily check that $j(\lambda) = j(1/\lambda) = j(1 - \lambda) = \cdots$, and that as j has a degree 6 numerator and degree < 6 denominator, j indeeds determines a degree 6 map from $\mathbb{P}^1$ (with coordinate λ) to $\mathbb{P}^1$ (with coordinate j). But this complicated-looking formula begs the question: where did this formula come from? How did someone think of it? We will largely answer this, but we will ignore the 2^8. This, as you might imagine, arises from characteristic 2 issues, and in order to invoke the results of §19.5 we have been assuming $\operatorname{char} k \neq 2$. (To see why the 2^8 is forced upon us by characteristic 2, see [**De**, p. 64]. From a different, complex-analytic, point of view, the 2^8 comes from the fact that the q-expansion of j begins $j = q^{-1} + 744 + \cdots$: the q^{-1} term has coefficient 1; see [**Se4**, p. 90, Rem. 2]. These seemingly different explanations are related by the theory of the Tate curve, see [**Si2**, Thm. V.3.1].)

Rather than using the formula handed to us, let's try to guess what j is. We won't expect to get the same formula as (19.9.7.2), but our answer should differ by an automorphism of the j-line ($\mathbb{P}^1$) — we will get $j' = (aj + b)/(cj + d)$ for some a, b, c, d (Exercise 16.4.B).

We are looking for some $j'(\lambda)$ such that $j'(\lambda) = j'(1/\lambda) = \cdots$. Hence we want some expression in λ that is invariant under this S_3-action. A first possibility would be to take the product of the six numbers

$$\lambda \cdot (1 - \lambda) \cdot \frac{1}{\lambda} \cdot \frac{\lambda - 1}{\lambda} \cdot \frac{1}{1 - \lambda} \cdot \frac{\lambda}{\lambda - 1}$$

This is silly, as the product is obviously 1.

A better idea is to add them all together:

$$\lambda + (1 - \lambda) + \frac{1}{\lambda} + \frac{\lambda - 1}{\lambda} + \frac{1}{1 - \lambda} + \frac{\lambda}{\lambda - 1}$$

This also doesn't work, as they add to 3 — the six terms come in pairs adding to 1.

(Another reason you might realize this can't work: if you look at the sum, you will realize that you will get something of the form "degree at most 3" divided by "degree at most 2". Then if $j' = p(\lambda)/q(\lambda)$, then λ is a root of a cubic over j. But we said that λ should satisfy a sextic over j'. The only way we avoid a contradiction is if $j' \in k$.)

But you will undoubtedly have another idea immediately. One good idea is to take the second symmetric function in the six roots. An equivalent one that is easier to do by hand is to add up the squares of the six terms. Even before doing the calculation, we can see that this will work: it will clearly produce a fraction whose numerator and denominator have degree at most 6, and it is not constant, as when λ is some fixed small number (say 1/2), the sum of squares is some small real number, while when λ is a large real number, the sum of squares will have to be some large real number (different from the value when $\lambda = 1/2$).

When you add up the squares by hand (which is not hard), you will get

$$j' = \frac{2\lambda^6 - 6\lambda^5 + 9\lambda^4 - 8\lambda^3 + 9\lambda^2 - 6\lambda + 2}{\lambda^2(\lambda-1)^2}.$$

Indeed $k(j) \cong k(j')$: you can check (again by hand) that

$$2j/2^8 = \frac{2\lambda^6 - 6\lambda^5 + 12\lambda^4 - 14\lambda^3 + 12\lambda^2 - 6\lambda + 2}{\lambda^2(\lambda-1)^2}.$$

Thus $2j/2^8 - j' = 3$.

19.9.C. EXERCISE. Explain why genus 1 curves over an algebraically closed field are classified by j-invariant.

19.9.D. EXERCISE. Give (with proof) two genus 1 curves over $\mathbb{Q}$ with the same j-invariant that are not isomorphic. (Hint: §19.5.3.)

19.9.8. Line bundles of degree 3.

In the discussion of degree 2 line bundles 19.9.5, we assumed $\operatorname{char} k \neq 2$ and $k = \overline{k}$, in order to invoke the Riemann-Hurwitz formula. In this section, we will start with no assumptions, and add them as we need them. In this way, you will see what partial results hold with weaker assumptions.

Consider the degree 3 invertible sheaf $\mathscr{O}_E(3p)$. By Riemann-Roch (19.2.5.1), $h^0(E, \mathscr{O}_E(3p)) = \deg(3p) - g + 1 = 3$. As $\deg E > 2g$, this gives a closed embedding (Remark 19.2.11 and Exercise 19.2.D). Thus we have a closed embedding $E \hookrightarrow \mathbb{P}^2_k$ as a cubic curve. Moreover, there is a line in $\mathbb{P}^2_k$ meeting E at point p with multiplicity 3, corresponding to the section of $\mathscr{O}(3p)$ vanishing precisely at p with multiplicity 3. (A line in the plane meeting a smooth curve with multiplicity at least 2 is a *tangent line*, see Definition 12.3.2. A line in the plane meeting a smooth curve with multiplicity at least 3 is said to be a **flex line**, and that point is a **flex point** of the curve.)

Choose projective coordinates on $\mathbb{P}^2_k$ so that p maps to $[0,1,0]$, and the flex line is the line at infinity $z=0$. Then the cubic is of the following form:

$$\begin{array}{ccccccccccc} ?\,x^3 & + & 0\,x^2y & + & 0\,xy^2 & + & 0\,y^3 & & \\ + & ?\,x^2z & + & ?\,xyz & + & ?\,y^2z & & = 0 \\ & + & ?\,xz^2 & + & ?\,yz^2 & & & & \\ & & + & ?\,z^3 & & & & & \end{array}$$

The coefficient of x is not 0 (or else this cubic is divisible by z). Dividing the entire equation by this coefficient, we can assume that the coefficient of x^3 is 1. The coefficient of y^2z is not 0 either (or else this cubic is singular at $x=z=0$). We can scale z (i.e., replace z by a suitable multiple) so that the coefficient of y^2z is -1. If the characteristic of k is not 2, then we can then replace y by $y+?x+?z$ so that the coefficients of xyz and yz^2 are 0, and if the characteristic of k is not 3, we can replace x by $x+?z$ so that the coefficient of x^2z is also 0. In conclusion, if $\operatorname{char} k \neq 2,3$, the elliptic curve may be written as

$$y^2z = x^3 + ax^2z + bz^3. \tag{19.9.8.1}$$

This is called the **Weierstrass normal form** of the curve.

19.9.9. *From the Weierstrass cubic to the double cover.* We see the hyperelliptic (double cover) description of the curve by setting $z=1$, or more precisely, by working in the distinguished open set $z \neq 0$ and using inhomogeneous coordinates. Here is the geometric explanation of why the double cover description is visible in the cubic description. Project the cubic from $p=[0,1,0]$. This is a map $E-p \to \mathbb{P}^1$ (given by $[x,y,z] \dashrightarrow [x,z]$), and is basically the vertical projection of the cubic to the x-axis. (Figure 19.4 may help you visualize this.) By the Curve-to-Projective Extension Theorem 16.5.1, the morphism $E-p \to \mathbb{P}^1$ extends over p. If $\mathscr{L} := \mathscr{O}(3p)$ is the line bundle giving the morphism $E \hookrightarrow \mathbb{P}^2$ (via Theorem 16.4.1 describing maps to projective space in terms of line bundles), then the two sections x and z giving the map to $\mathbb{P}^1$ vanish at p to order 1 and 3 respectively. To "resolve" the rational map into an honest morphism, we interpret x and z as sections of $\mathscr{L}(-p) = \mathscr{O}(2p)$, and now they generate a base-point-free linear series, and thus a morphism $\mathbb{P}^1$. (You may be able to interpret this as implementing the proof of Theorem 16.4.1 in a different language.) A similar idea, applied to quartics rather than cubics, was used in Exercise 19.7.B.

As a consequence, with a little sweat, we can compute the j-invariant:

$$j(a,b) = \frac{2^8 3^3 a^3}{4a^3+27b^2}.$$

19.9.E. Exercise. Show that the flexes of the cubic are the 3-torsion points in the group E. ("Flex" was defined in §19.9.8: it is a point where the tangent line meets the curve with multiplicity at least 3 at that point. In fact, if k is algebraically closed and char $k \neq 3$, there are nine of them. This won't be surprising if you are familiar with the complex story, $E \cong \mathbb{C}/\Lambda$.)

19.9.10. The group law, geometrically.

The group law has a beautiful classical description in terms of the Weierstrass form. Consider Figure 19.4. In the Weierstrass coordinates, the origin p is the only point of E meeting the line at infinity ($z = 0$); in fact the line at infinity corresponds to the tautological section of $\mathscr{O}(3p)$. If a line meets E at three points p_1, p_2, p_3, then

$$\mathscr{O}(p_1 + p_2 + p_3) \cong \mathscr{O}(3p)$$

from which (in the group law) $p_1 + p_2 + p_3 = 0$.

Hence to find the inverse of a point s, we consider the intersection of E with the line sp; $-s$ is the third point of intersection. To find the sum of two points q and r, we consider the intersection of E with the line qr, and call the third points s. We then compute $-s$ by connecting s to p, obtaining $q + r$.

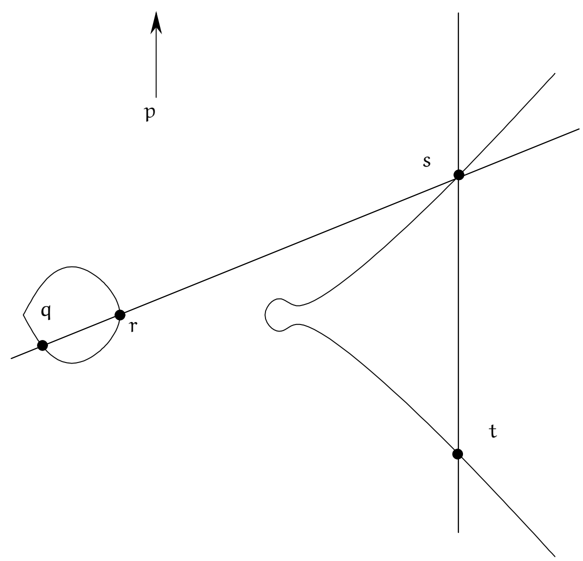


FIGURE 19.4. The group law on the elliptic curve, geometrically

We could give this description of a group law on a cubic curve in Weierstrass normal form to anyone familiar with the notion of projective space, and the notion of a group, but we would then have to prove that the construction we are giving indeed defines a group. In particular, we would have to prove associativity, which is not a priori clear. But in this case, we have already established that the degree 1 points form a group, by giving a bijection to $\operatorname{Pic}^0 E$, and we are merely interpreting the group law on $\operatorname{Pic}^0 E$.

Note that this description works even in characteristic 2 and 3; we don't need the cubic to be in Weierstrass normal form, and we need only that $\mathscr{O}(3p)$ gives a closed embedding into $\mathbb{P}^2$.

19.9.11. Degree 4 line bundles. You have probably forgotten that we began by studying line bundles degree by degree. The story doesn't stop in degree 3. In the same way that we showed that a canonically embedded nonhyperelliptic curve of genus 4 is the complete intersection in $\mathbb{P}^3_k$ of a quadric and a cubic (§19.8), we can show the following.

19.9.F. EXERCISE. Show that the complete linear series for $\mathscr{O}(4p)$ embeds E in $\mathbb{P}^3$ as the complete intersection of two quadrics. (Hint: Show the image of E is contained in at least 2 linearly independent quadrics. Show that neither can be reducible, so they share no components. Use Bézout's theorem, Exercise 18.6.K.)

The beautiful structure doesn't stop with degree 4, but it gets more complicated. For example, the degree 5 embedding is not a complete intersection (of hypersurfaces), but is the complete intersection of $G(2,5)$ under its Plücker embedding with a five hyperplanes (or perhaps better, a codimension 5 linear space). In seemingly different terminology, its equations are 4×4 Pfaffians of a general 5×5 skew-symmetric matrix of linear forms, although I won't say what this means.

19.10 Elliptic curves are group varieties

We initially described the group law on the degree 1 points of an algebraic curve in a rather abstract way. From that definition, it was not clear that over $\mathbb{C}$ the group operations (addition, inverse) are continuous. But the explicit description in terms of the Weierstrass cubic makes this clear. In fact we can observe even more: addition and inverse are algebraic in general. Better yet, elliptic curves are group varieties. (Thus they are abelian varieties, see Definition 10.3.11.)

19.10.1. (This is a clue that $\operatorname{Pic}^0(E)$ really wants to be a scheme, and not just a group. Once the notion of "moduli space of line bundles on a variety" is made precise, this can be shown.)

We begin with the inverse case, as a warm-up.

19.10.2. Proposition. — *If* $\operatorname{char} k \neq 2, 3$, *there is a morphism of* k*-varieties* $E \to E$ *sending a (degree* 1*) point to its inverse, and this construction behaves well under field extension of* k.

In other words, the "inverse map" in the group law actually arises from a morphism of schemes — it isn't just a set map. (You are welcome to think through the two remaining characteristics, and to see that essentially the same proof applies. But the proof of Theorem 19.10.4 will give you a better sense of how to proceed.)

Proof. Consider the map (the hyperelliptic involution) $y \mapsto -y$ of the Weierstrass normal form. □

The algebraic description of addition would be a big mess if we were to write it down. We will be able to show algebraicity by a trick — not by writing it down explicitly, but by thinking through how we *could* write it down explicitly. The main part of the trick is the following proposition. We give it in some generality just because it can be useful, but you may prefer to assume that $k = \overline{k}$ and C is a regular cubic.

19.10.3. Proposition. — *Suppose $C \subset \mathbb{P}^2_k$ is a geometrically integral cubic curve (so in particular C contains no lines). Let C^{ns} be the regular points of C. There is a unique morphism $t : C^{ns} \times C^{ns} \to C^{ns}$ such that*

(a) *if p and q are distinct regular k-valued points of C, then $t(p, q)$ is obtained by intersecting the line $\overline{pq}$ with C, and taking the third "residual" point of intersection with C. More precisely, $\overline{pq}$ will meet C at three points with multiplicity (Exercise 8.2.E), including p and q; $t(p, q)$ is the third point.*
(b) *this property remains true after extension to $\overline{k}$.*

Furthermore, if p is a k-valued point of C^{ns}, then $t(2p)$ is where the tangent line ℓ to C at p meets C again. More precisely, ℓ will meet C at three points with multiplicity, which includes p with multiplicity 2; $t(p, p)$ is the third point.

We will need property (b) because C may have few enough k-valued points (perhaps none!) that the morphism t can not be determined by its behavior on them. In the course of the proof, we will see that (b) can be extended to "this property remains true after any field extension of k".

Proof. We first show (in this paragraph) that if p and q are distinct regular points, then the third point r of intersection of $\overline{pq}$ with C is also regular. If $r = p$ or $r = q$, we are done. Otherwise, the cubic obtained by restricting C to $\overline{pq}$ has three distinct (hence reduced, i.e., multiplicity 1) roots, p, q, and r. Thus $C \cap \overline{pq}$ is regular at r, so r is a regular point of C by the slicing criterion for regularity, Exercise 12.2.B.

We now assume that $k = \overline{k}$, and leave the general case to the end. Fix p, q, and r, where $p \neq q$, and r is the "third" point of intersection of $\overline{pq}$ with C. We will describe a morphism $t_{p,q}$ in a neighborhood of $(p, q) \in C^{ns} \times C^{ns}$. By Exercise 10.2.B, showing that morphisms of varieties over $\overline{k}$ are determined by their behavior on closed ($\overline{k}$-valued) points, that these morphisms glue together (uniquely) to give a morphism t, completing the proof in the case $k = \overline{k}$.

Choose projective coordinates on $\mathbb{P}^2$ in such a way that $U_0 \cong \operatorname{Spec} k[x_1, x_2]$ contains p, q, and r, and the line $\overline{pq}$ is not "vertical". More precisely, in $\operatorname{Spec} k[x_1, x_2]$, say $p = (p_1, p_2)$ (in terms of "classical coordinates" — more pedantically, $p = [(x_1 - p_1, x_2 - p_2)]$), $q = (q_1, q_2)$, $r = (r_1, r_2)$, and $p_1 \neq q_1$. In these coordinates, the curve C is cut out by some cubic, which we also sloppily denote C: $C(x_1, x_2) = 0$.

Now suppose $P = (P_1, P_2)$ and $Q = (Q_1, Q_2)$ are two points of in $C \cap U_0$ (not necessarily our p and q). We attempt to compute the third point of intersection of $\overline{PQ}$ with C, in a way that works on an open subset of $C \times C$ that includes (p, q). To do this explicitly requires ugly high school algebra, but because we know how it looks, we will be able to avoid dealing with any details!

The line $\overline{PQ}$ is given by $x_2 = mx_1 + b$, where $m = \frac{P_2 - Q_2}{P_1 - Q_1}$ and $b = P_2 - mP_1$ are both rational functions of P and Q. Then m and b are defined for all P and Q such that $P_1 \neq Q_1$ (and hence for a neighborhood of (p, q), as $p_1 \neq q_1$, and as $P_1 \neq Q_1$ is an open condition).

Now we solve for $C \cap \overline{PQ}$, by substituting $x_2 = mx_1 + b$ into C, to get $C(x_1, mx_1 + b)$. This is a cubic in x_1, say

$$\gamma(x_1) = Ax_1^3 + Bx_1^2 + Cx_1 + D = 0.$$

The coefficients of γ are rational functions of P_1, P_2, Q_1, and Q_2. The cubic γ has 3 roots (with multiplicity) so long as $A \neq 0$, which is a Zariski-open condition on m and b, and hence a Zariski-open condition on P_1, P_2, Q_1, Q_2. As $P, Q \in C \cap \overline{PQ} \cap U_0$, P_1 and Q_1 are two of the roots of $\gamma(x_1) = 0$. The sum of the roots of $\gamma(x_1) = 0$ is $-B/A$ (by Viète's formula), so the third root of γ is $R_1 := -B/A - P_1 - Q_1$. Thus if we take $R_2 = mR_1 + b$, we have found the third points of intersection of $\overline{PQ}$ with C (which happily lies in U_0) We have thus described a morphism from the open subset of $(C^{ns} \cap U_0) \times (C^{ns} \cap U_0)$, containing (p, q), that does what we want. (Precisely, the open subset is defined by $A \neq 0$, which can be explicitly unwound.) We have thus completed the proof of Proposition 19.10.3 (except for the last paragraph) for $k = \overline{k}$. (Those who believe they are interested only in algebraically closed fields can skip ahead.)

We extend this to Proposition 19.10.3 for every field k except $\mathbb{F}_2$. Suppose $U_0^{[x_1,x_2]} = \operatorname{Spec} k[x_1, x_2]$ is any affine open subset of $\mathbb{P}^2_k$, along with choice of coordinates. (The awkward superscript "$[x_1, x_2]$" is there to emphasize that the particular coordinates are used in the construction.) Then the construction above gives a morphism *defined over* k from an open subset of $(C^{ns} \cap U_0^{[x_1,x_2]}) \times (C^{ns} \cap U_0^{[x_1,x_2]})$ (note that all of the hypothetical algebra was done over k), that sends P and Q to the third points of intersection of $\overline{PQ}$ with C. Note that this construction commutes with any field extension, as the construction is insensitive to the field we are working over. Thus after base change to the algebraic closure, the map also has the property that it takes as input two points, and spits out the third point of intersection of the line with the cubic. Furthermore, all of these maps (as $U_0^{[x_1,x_2]}$ varies over all complements U_0 of lines "with k-coefficients", and choices of coordinates on U_0) can be glued together: they agree on their pairwise overlaps (as after base change to $\overline{k}$ they are the same, by our previous discussion, and two maps that are the same after base change to $\overline{k}$ were the same to begin with by Exercise 9.2.I), and this is what is required to glue them together (Exercise 6.2.A).

We can geometrically interpret the open subset $(C^{ns} \cap U_0^{[x_1,x_2]}) \times (C^{ns} \cap U_0^{[x_1,x_2]})$ by examining the construction: it is defined in the locus $\{P = (P_1, P_2), Q = (Q_1, Q_2)\}$ where (i) $P_1 \neq Q_1$, and (ii) the third point of intersection R of $\overline{PQ}$ with C also lies in U_0.

So which points (P, Q) of $C^{ns} \times C^{ns}$ are missed? Condition (i) isn't important; if (P, Q) satisfies (ii) but not (i), we can swap the roles of x_1 and x_2, and (P, Q) will then satisfy (i). The only way (P, Q) can not be covered by one of these open sets is if there is *no* U_0 (a complement of a line defined over k) that includes P, Q, and R.

19.10.A. EXERCISE. Use $|k| > 2$ to show that there is a linear form on $\mathbb{P}^2$ with coefficients in k that misses P, Q, and R. (This is sadly *not* true if $k = \mathbb{F}_2$ — do you see why?)

19.10.B. EXERCISE. Prove the last statement of Proposition 19.10.3.

19.10.C. ⋆⋆ UNIMPORTANT EXERCISE. Complete the proof by dealing with the case $k = \mathbb{F}_2$. Hint: first produce the morphism t over $\mathbb{F}_4$. The goal is then to show that this t is really "defined over" $\mathbb{F}_2$ ("descends to" $\mathbb{F}_2$). The morphism t is initially described locally by considering the complement of a line defined over $\mathbb{F}_4$ (and then letting the line vary). Instead, look at the map by looking at the complement of a line and its "conjugate". The complement of the line and its conjugate is an affine $\mathbb{F}_2$-variety. The partially-defined map t on this affine variety is a priori defined over $\mathbb{F}_4$, and is preserved by conjugation. Show that this partially defined map is "really" defined over $\mathbb{F}_2$. (If you figure out what all of this means, you will have an important initial insight into the theory of "descent".)

□

We can now use this to define the group variety structure on E.

19.10.4. Theorem. — *Suppose* (E, p) *is an elliptic curve (a regular genus* 1 *curve over* k*, with a* k*-valued point* p*). Take the Weierstrass embedding of* E *in* $\mathbb{P}^2_k$*, via the complete linear series* $|\mathscr{O}_E(3p)|$*. Define the* k*-morphism* $e : \operatorname{Spec} k \to E$ *by sending* $\operatorname{Spec} k$ *to* p*. Define the* k*-morphism* $i : E \to E$ *via* $q \mapsto t(p, q)$*, or more precisely, as the composition*

$$E \xrightarrow{(id,e)} E \times E \xrightarrow{t} E.$$

Define the k*-morphism* $m : E \times E \to$ *via* $(q, r) \mapsto t(p, t(q, r))$*. Then* (E, e, i, m) *is a group variety over* k*.*

By the construction of t, all of these morphisms "commute with arbitrary base extension".

Proof. We need to check that various pairs of morphisms described in axioms (i)–(iii) of §6.6.4 are equal. For example, in axiom (iii), we need to show that $m \circ (i, id) = m \circ (id, i)$; all of the axioms are clearly of this sort.

Assume first that $k = \overline{k}$. Then each of these pairs of morphisms agree as maps of $\overline{k}$-points: $\operatorname{Pic} E$ is a group, and under the bijection between $\operatorname{Pic} E$ and E of Proposition 19.9.3, the group operations translate into the maps described in the statement of Theorem 19.10.4 by the discussion of §19.9.10.

But morphisms of $\overline{k}$-varieties are determined by their maps on the level of $\overline{k}$-points (Exercise 10.2.B), so each of these pairs of morphisms are the same.

For general k, we note that from the $\overline{k}$ case, these morphisms agree after base change to the algebraic closure. Then by Exercise 9.2.I, they must agree to begin with.

19.10.5. *Features of this construction.* The most common derivation of the properties of an elliptic curve are to describe it as a cubic, and describe addition using the explicit construction with lines. Then one has to work to prove that the multiplication described is associative.

Instead, we started with something that was patently a group (the degree 0 line bundles). We interpreted the maps used in the definition of the group (addition and inverse) geometrically using our cubic interpretation of elliptic curves. This allowed us to see that these maps were algebraic.

As a bonus, we see that in some vague sense, the Picard group of an elliptic curve wants to be an algebraic variety.

19.10.D. EXERCISE. Suppose p and q are k-points of a genus 1 curve E. Show that there is an automorphism of E sending p to q.

19.10.E. EXERCISE. Suppose (E, p) is an elliptic curve over an algebraically closed field k of characteristic not 2. Show that the automorphism group of (E, p) is isomorphic to $\mathbb{Z}/2$, $\mathbb{Z}/4$, or $\mathbb{Z}/6$. (An automorphism of an elliptic curve (E, p) over $k = \overline{k}$ is an automorphism of E fixing p scheme-theoretically, or equivalently, fixing the k-valued points by Exercise 10.2.B.) Hint: reduce to the question of automorphisms of $\mathbb{P}^1$ fixing a point ∞ and a set of distinct three points $\{p_1, p_2, p_3\} \in \mathbb{P}^1 \setminus \{\infty\}$. (The algebraic closure of k is not essential, so feel free to remove this hypothesis, using Exercise 9.2.I.)

19.10.6. *Vague question.* What are the possible automorphism groups of a genus 1 curve over an algebraically closed k of characteristic not 2? You should be able to convince yourself that the group has "dimension 1".

19.10.F. IMPORTANT EXERCISE: A DEGENERATE ELLIPTIC CURVE. Consider the genus 1 curve $C \subset \mathbb{P}^2_k$ given by $y^2z = x^3 + x^2z$, with the point $p = [0, 1, 0]$. Emulate the above argument to show that $C \setminus \{[0, 0, 1]\}$ is a group variety. Show that it is isomorphic to $\mathbb{G}_m$ (the multiplicative group scheme $\operatorname{Spec} k[t, t^{-1}]$, see Exercise 6.6.E) with coordinate $t = y/x$, by showing an isomorphism of schemes, and showing that multiplication and inverse in both group varieties agree under this isomorphism.

19.10.G. EXERCISE: AN EVEN MORE DEGENERATE ELLIPTIC CURVE. Consider the genus 1 curve $C \subset \mathbb{P}^2_k$ given by $y^2z = x^3$, with the point $p = [0, 1, 0]$. Emulate the above argument to show that $C \setminus \{[0, 0, 1]\}$ is a group variety. Show that it is isomorphic to $\mathbb{A}^1$ (with additive group structure) with coordinate $t = y/x$, by showing an isomorphism of schemes, and showing that multiplication/addition and inverse in both group varieties agree under this isomorphism.

19.10.7. $\star$ **Towards Poncelet's Porism.**

These ideas lead to a beautiful classical fact, Poncelet's Porism. (A *porism* is an archaic name for a type of mathematical result.)

[figure to be made later]

FIGURE 19.5. Poncelet's Porism 19.10.8

19.10.8. Poncelet's Porism. — *Suppose C and D are two ellipses in $\mathbb{R}^2$, with C containing D. Choose any point p_0 on C. Choose one of the two tangents ℓ_1 from p to D. Then ℓ_1 meets C at two points in total: p_0 and another point p_1. From p_1, there are two tangents to D, ℓ_1 and another line ℓ_2. The line ℓ_2 meets C at some other point p_2. Continue this to get a sequence of points p_0, p_1, p_2, ... Suppose this sequence starting with p_0 is periodic, i.e., $p_0 = p_n$ for some n. Then it is periodic with* any *starting point $p \in C$ (see Figure 19.5).*

It is possible to prove Poncelet's Porism in an elementary manner, but a proof involving elliptic curves is quite beautiful, and gives connections to more sophisticated ideas. Rather than proving Poncelet's Porism, we discuss some related facts.

19.10.H. EXERCISE. Suppose E is a smooth degree 3 projective plane curve over an algebraically closed field, $q, r \in E$, and n is a positive integer. Let $F : E \to E$

be the morphism $x \mapsto t(r, t(q, x))$. Suppose that $F^n(p) = p$ for some closed point $p \in E$. Show that $F^n(x) = x$ for all points $x \in E$.

You should feel the urge to improve this result. (Can you show that if $F^n(p) = p$ for some $p \in E$, then F^n is the identity morphism on E? Can you extend this to the case where the base field is not algebraically closed?)

19.10.I. EXERCISE. Suppose E is an ellipse in $\mathbb{R}^2$. For each point $(x, y) \in E$, there is another point $G(x, y) := (x, y')$ (possibly the same), with the same first coordinate, corresponding to the "other" intersection of the vertical line through (x, y) with E. Similarly, there is another point $H(x, y) := (x', y)$, with the same second coordinate. Let $F = G \circ H$. Show that if $F^{17}(p) = p$ for some point $p \in E$, then $F^{17}(x) = x$ for all points $x \in E$. (See Figure 19.6.) Hint: this might be best done not using any fancy methods from algebraic geometry.

Again, you should wish to improve this. To what extent is this dependent on the real numbers? What if the ellipse was instead a different conic section? (And what does this have to do with elliptic curves? Why is the conic called E?!)

[figure to be made later; caution, use smaller number than 17]

FIGURE 19.6. Exercise 19.10.I

19.10.J. ⋆⋆ EXERCISE. Let k be a field, algebraically closed purely for convenience. Suppose Q is a smooth quadric surface in $\mathbb{P}^3_k$, and K is a cone (a rank 3 quadric) in $\mathbb{P}^3$, such that $E = Q \cap K$ is a smooth curve.
(a) Show that E has genus 1.
(b) If ℓ is a line on Q, show that E meets ℓ at two points (with multiplicity).
(c) Fix one of the rulings (family of lines) of Q. Show that there exists a morphism $G : E \to E$ that takes a closed point p of E, and sends it to the other point of E on the line ℓ in the ruling containing p.
(d) Define $H : E \to E$ similarly, using the other ruling of Q. Fix a positive integer n. Let $F = G \circ H$. Show that if $F^n(p) = p$ for one closed point $p \in E$, then $F^n(x) = x$ for all $x \in E$.

We now connect Exercise 19.10.J to Poncelet's Porism 19.10.8. Let v be the cone point of K. Then $v \notin Q$, as otherwise v would be a singular point of $Q \cap K$ (do you see why?). It is a fact that projection from v gives a morphism $Q \to \mathbb{P}^2_k$ that is branched over a smooth conic D. (You may be able to make this precise, and prove it, after reading Chapter 21.) The lines on Q project to tangent lines to D in $\mathbb{P}^2_k$. The projection from v contracts $K \setminus \{v\}$ to a conic C. You may be able to interpret the statement of Poncelet's Porism 19.10.8 (for this C and D) in terms of Exercise 19.10.J. (Caution: the field is not algebraically closed in Poncelet's Porism 19.10.8, and *is* algebraically closed in Exercise 19.10.J.) You may even be able to take an arbitrary C and D as in Poncelet's Porism 19.10.8, and reverse-engineer K and Q as in Exercise 19.10.J, thereby proving Poncelet's Porism (and discovering why there is an elliptic curve hidden in its statement).

19.11 Counterexamples and pathologies using elliptic curves

We now give some fun counterexamples using our understanding of elliptic curves. The main extra juice elliptic curves give us comes from the fact that elliptic curves are the simplest varieties with "continuous Picard groups".

19.11.1. An example of a scheme that is factorial, but such that no affine open neighborhood of any point has ring that is a unique factorization domain.

Suppose E is an elliptic curve over $\mathbb{C}$ (or some other uncountable algebraically closed field). Consider $p \in E$. The local ring $\mathscr{O}_{E,p}$ is a discrete valuation ring and hence a unique factorization domain (Theorem 12.5.8). Then an open neighborhood of E is of the form $E - q_1 - \cdots - q_n$. I claim that its Picard group is nontrivial. Recall the exact sequence:

$$\mathbb{Z}^{\oplus n} \xrightarrow{(a_1,\dots,a_n)\mapsto a_1 q_1+\cdots+a_n q_n} \operatorname{Pic} E \longrightarrow \operatorname{Pic}(E - q_1 - \cdots - q_n) \longrightarrow 0 .$$

But the group on the left is countable, and the group in the middle is uncountable, so the group on the right, $\operatorname{Pic}(E - q_1 - \cdots - q_n)$, is nonzero. We have shown that every nonempty open subset of E has nonzero line bundles, as promised in Remark 13.1.8.

If $n > 0$, then $E - q_1 - \cdots - q_n$ is affine (Exercise 19.2.B). Thus by Exercise 14.2.T, the corresponding ring is not a unique factorization domain. To summarize: complex elliptic curves are factorial, but no affine open subset has a ring that has unique factorization, as promised in §5.4.5.

19.11.A. EXERCISE. The above argument shows that over an uncountable field, $\operatorname{Pic} E$ is not a finitely generated group. Show that even over the countable field $\overline{\mathbb{Q}}$, $\operatorname{Pic} E$ is not a finitely generated group, as follows. If the elliptic curve E is generated by $q_1, \dots, q_n$, then there is a finite field extension K of $\mathbb{Q}$ over which all q_i are defined (the compositum of the residue fields of the q_i). Show that any point in the subgroup of E generated by the q_i must also be defined over K. Show that E has a point not defined over K. Use this to show that $\operatorname{Pic} E$ is not finitely generated. (The same argument works with $\overline{\mathbb{Q}}$ replaced by $\overline{\mathbb{F}}_p$.)

19.11.2. *Remark.* In contrast to the above discussion, over $\mathbb{Q}$, the Mordell-Weil Theorem states that $\operatorname{Pic} E$ *is* finitely generated (Aside 19.9.4).

19.11.3. $\star\star$ **A complex surface with infinitely many algebraic structures (for those with complex geometric background).**

As remarked in Exercise 6.3.K, a complex manifold may have many algebraic structures. The following example of M. Kim gives an example with *infinitely* many algebraic structures. Suppose E is an elliptic curve over $\mathbb{C}$ with origin p, and let $\mathscr{L}$ be a nontrivial line bundle on $E - p$. Then $\mathscr{L}$ is *analytically* trivial because $E - p$ is a Stein space, so the analytification of the total space is independent of the choice of $\mathscr{L}$. However, you know enough to show (with work) that there are infinitely many pairwise (algebraically) nonisomorphic $\mathscr{L}$, and also that their total spaces are likewise pairwise (algebraically) nonisomorphic. This gives a complex surface with infinitely many algebraic structures (indeed a continuum of them). See [**MO68421**] for more details.

19.11.4. Counterexamples using a non-torsion point.

We next give a number of counterexamples using the existence of a non-torsion point of a complex elliptic curve. We first show the existence of such a point.

19.11.5. We have a "multiplication by n" map $[n] : E \to E$, which sends p to np. If $n = 0$, the map $[n]$ has degree 0 (even though $[n]$ isn't dominant, degree is still defined, see Definition 11.2.2). If $n = 1$, the map $[n]$ has degree 1. Given the complex picture of a torus, you might not be surprised that the degree of $[n]$ is n^2. (Unimportant quibble: we have defined degree only for finite morphisms, so for $n = 0$ the degree hasn't been defined.) If $n = 2$, we have almost shown that it has degree 4, as we have checked that there are precisely 4 points q such that $2p = 2q$ (Exercise 19.9.A). All that really shows (using Exercise 17.4.D(b)) is that the degree is at least 4. (We could check by hand that the degree is 4 is we really wanted to.)

19.11.6. Proposition. — *Suppose E is an elliptic curve over a field k of characteristic not 2. For each $n > 0$, the "multiplication by n" morphism $[n]$ has positive degree, so there are only a finite number of n-torsion points.*

Proof. We may assume $k = \overline{k}$, as the degree of a map of curves is independent of field extension.

We prove the result by induction; it is true for $n = 1$ and $n = 2$.

If n is odd ($2k + 1$, say), then assume otherwise that $nq = 0$ for all closed points q. Let r be a non-trivial 2-torsion point, so $2r = 0$. But $nr = 0$ as well, so $r = (n - 2k)r = 0$, contradicting $r \neq 0$.

If n is even, then $[n] = [2] \circ [n/2]$ (degree is multiplicative under composition of rational maps, §11.2.2), and by our inductive hypothesis both $[2]$ and $[n/2]$ have positive degree. $\square$

In particular, the total number of torsion points on E is countable. If k is an uncountable field, then E has an uncountable number of closed points (consider an open subset of the curve as $y^2 = x^3 + ax + b$; there are uncountably many choices for x, and each of them has 1 or 2 choices for y). Thus we have the following.

19.11.7. Corollary. — *If E is a curve over an uncountable algebraically closed field of characteristic not 2 (e.g. $\mathbb{C}$), then E has a non-torsion point.*

Proof. For each n, there are only finitely many n-torsion points. Thus there are (at most) countably many torsion points. The curve E has uncountably many closed points. (One argument for this: take a double cover $\pi : E \to \mathbb{P}^1$. Then $\mathbb{P}^1$ has uncountably many closed points, and π is surjective on closed points.) $\square$

19.11.8. *Remark.* In a sense we can make precise using cardinalities, almost all points on E are non-torsion. You will notice that this argument breaks down over countable fields. In fact, over $\overline{\mathbb{F}}_p$, all points of an elliptic curve E are torsion. (Any point x is defined over some finite field $\mathbb{F}_{p^r}$. The points defined over $\mathbb{F}_{p^r}$ form a subgroup of E, using the explicit geometric construction of the group law, and there are finite number of points over $\mathbb{F}_{p^r}$ — certainly no more than the number of $\mathbb{F}_{p^r}$-points of $\mathbb{P}^2$.) But over $\overline{\mathbb{Q}}$, there are elliptic curves with non-torsion points. Even better, there are examples over $\mathbb{Q}$: $[2, 1, 8]$ is a $\mathbb{Q}$-point of the elliptic curve $y^2z = x^3 + 4xz^2 - z^3$ that is not torsion. The proof would carry us too far afield, but one method is to use the Nagell-Lutz Theorem (see for example [**Si1**, Cor. 7.2]).

We now use the existence of a non-torsion point to create some interesting pathologies.

19.11.9. A map of projective varieties not arising from a map of graded rings, even after regrading.

19.11.B. EXERCISE. Suppose $E \hookrightarrow \mathbb{P}^2_{\mathbb{C}}$ is a smooth complex plane cubic (hence genus 1), yielding a graded ring $S_\bullet$. Let $t : E \to E$ be translation by a non-torsion point. Show that t does not correspond to a map of graded rings $S_\bullet \to S_\bullet$, even after regrading (cf. Remark 16.4.7). (If you wish, you can show the following. If $u : E \to E$ is a translation, show that u corresponds to some map of regraded rings $S_{n\bullet} \to S_{n\bullet}$ if and only if u is translation by a torsion point.)

19.11.10. An affine open subset of an affine scheme that is not a distinguished open set.

We can use this to construct an example of an affine scheme X and an affine open subset Y that is not distinguished in X. Let $X = E - p$, which is affine (see Exercise 19.2.B, or better, note that the linear series $\mathscr{O}(3p)$ sends E to $\mathbb{P}^2$ in such a way that the "line at infinity" meets E only at p; then $E - p$ has a closed embedding into the affine scheme $\mathbb{A}^2$).

Let q be another point on E so that $q - p$ is non-torsion. Then $E - p - q$ is affine (Exercise 19.2.B). Assume that it is distinguished. Then there is a function f on $E - p$ that vanishes on q (to some positive order d). Thus f is a rational function on E that vanishes at q to order d, and (as the total number of zeros minus poles of f is 0) has a pole at p of order d. But then $d(p - q) = 0$ in $\operatorname{Pic}^0 E$, contradicting our assumption that $p - q$ is non-torsion.

In particular, $E - p$ is an affine scheme, and q is locally cut out by one equation, but it is not globally cut out *even set-theoretically* by one equation. This was promised in Exercise 7.3.F.

19.11.11. A proper (nonprojective) surface with no nontrivial line bundles.

We next use a non-torsion point p on an elliptic curve E to construct a proper nonprojective surface with no nontrivial line bundles. Let X_1 be $\mathbb{P}^2$, and Z_1 a smooth cubic in X_1, identified with E. Let $Z_2 \subset X_2$ be exactly the same. Glue X_1 to X_2 along the isomorphism $Z_1 \cong Z_2$ given, not by the identity, but by translation by p. (This gluing construction was described in §16.4.11.) Call the result X. By Exercise 16.4.O (and the sentence thereafter), X is proper.

19.11.C. EXERCISE. Show that every line bundle on X is trivial. Hint: Suppose $\mathscr{L}$ is an invertible sheaf on X. Then $\mathscr{L}|_{X_i}$ is $\mathscr{O}(d_i)$ for some d_i (for $i = 1, 2$). The restriction of $\mathscr{L}_{X_1}$ to Z_1 must agree with the restriction of $\mathscr{L}_{X_2}$ to Z_2. Use this to show that $d_1 = d_2 = 0$. Explain why gluing two trivial bundles (on X_1 and X_2 respectively) together in this way yields a trivial bundle on X.

19.11.D. EXERCISE. Show that X is not projective.

See Exercise 20.2.G for a somewhat simpler example of a proper nonprojective surface.

19.11.12. A Picard group that has no chance of being a scheme.

We informally observed that the Picard group of an elliptic curve "wants to be" a scheme (see §19.10.5). This is true of projective (and even proper) varieties in general (see [**FGIKNV**, Ch. 9]). On the other hand, if we work over $\mathbb{C}$, the affine scheme $E - p - q$ (in the language of §19.11.10 above) has a Picard group that can be interpreted as $\mathbb{C}$ modulo a lattice modulo a non-torsion point (e.g. $\mathbb{C}/\langle 1, i, \pi \rangle$). This has no reasonable interpretation as a manifold, let alone a variety. So the fact that the Picard group of proper varieties turns out to be a scheme should be seen as quite remarkable.

19.11.13. Example of a variety with non-finitely-generated ring of global sections.

We next show an example of a complex variety whose ring of global sections is not finitely generated. (An example over $\mathbb{Q}$ can be constructed in the same way using the curve of Remark 19.11.8.) This is related to Hilbert's fourteenth problem (see [**Mu5**, §3]).

19.11.E. PRELIMINARY EXERCISE. Suppose X is a scheme, and L is the total space of a line bundle corresponding to invertible sheaf $\mathscr{L}$, so $L = \operatorname{Spec} \oplus_{n \geq 0} (\mathscr{L}^\vee)^{\otimes n}$. (This construction first appeared in Definition 17.1.4.) Show that $H^0(L, \mathscr{O}_L) = \oplus H^0(X, (\mathscr{L}^\vee)^{\otimes n})$. (Possible hint: choose a trivializing cover for $\mathscr{L}$. Rhetorical question: can you figure out the more general statement if $\mathscr{L}$ is a rank r locally free sheaf?)

Let E be an elliptic curve over some ground field k, $\mathscr{N}$ a degree 0 non-torsion invertible sheaf on E, and $\mathscr{P}$ a positive-degree invertible sheaf on E. Then $H^0(E, \mathscr{N}^m \otimes \mathscr{P}^n)$ is nonzero if and only if either (i) $n > 0$, or (ii) $m = n = 0$ (in which case the sections are elements of k).

19.11.F. EASY EXERCISE. Show that the ring $R = \oplus_{m,n \geq 0} H^0(E, \mathscr{N}^m \otimes \mathscr{P}^n)$ is not finitely generated.

19.11.G. EXERCISE. Let X be the total space of the vector bundle associated to $(\mathscr{N} \oplus \mathscr{P})^\vee$ over E. Show that the ring of global sections of X is R, and hence is not finitely generated. (Hint: interpret X as a line bundle over a line bundle over E.)

19.11.H. EXERCISE. Show that X (as in the above exercise) is a Noetherian variety whose ring of global sections is not Noetherian.

CHAPTER 20

$\star$ Application: A glimpse of intersection theory

The only reason this chapter appears after Chapter 19 is because we will use Exercise 19.2.E in the proof of the Nakai-Moishezon criterion for ampleness (see §20.4.5).

20.1 Intersecting n line bundles with an n-dimensional variety

Throughout this chapter, X will be a k-variety; in most applications, X will be projective. The central tool in this chapter is the following.

20.1.1. *Definition: intersection product, or intersection number.* Suppose $\mathscr{F}$ is a coherent sheaf on X with proper support (automatic if X is proper) *of dimension at most* n, and $\mathscr{L}_1, \dots, \mathscr{L}_n$ are invertible sheaves on X. Let $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$ be the signed sum over the 2^n subsets of $\{1, \dots, n\}$

$$\sum_{\{i_1,\dots,i_m\}\subset\{1,\dots,n\}} (-1)^m \chi(X, \mathscr{L}_{i_1}^\vee \otimes \cdots \otimes \mathscr{L}_{i_m}^\vee \otimes \mathscr{F}). \tag{20.1.1.1}$$

We call this the *intersection of* $\mathscr{L}_1, \dots, \mathscr{L}_n$ *with* $\mathscr{F}$. (Never forget that whenever we write $(\mathscr{L}_1 \cdots \mathscr{L}_n \cdot \mathscr{F})$, we are implicitly assuming that $\dim \operatorname{Supp} \mathscr{F} \leq n$.) The case we will find most useful is if $\mathscr{F}$ is the structure sheaf of a closed subscheme Y (of dimension at most n). In this case, we may write it $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot Y)$. If the $\mathscr{L}_i$ are all the same, say $\mathscr{L}$, one often writes $(\mathscr{L}^n \cdot \mathscr{F})$ or $(\mathscr{L}^n \cdot Y)$. (Be careful with this confusing notation: $\mathscr{L}^n$ does not mean $\mathscr{L}^{\otimes n}$.) In some circumstances the convention is to omit the parentheses.

We will prove many things about the intersection product in this chapter. One fact will be left until we study flatness (Exercise 24.7.D): that it is "deformation-invariant" — that it is constant in "nice" families.

20.1.A. EXERCISE (REALITY CHECK). Show that if $\mathscr{L}_1 \cong \mathscr{O}_X$ then $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) = 0$.

The following exercise suggests that the intersection product might be interesting, as it "interpolates" between two useful notions: the degree of a line bundle on a curve, and Bézout's theorem.

20.1.B. EXERCISE.
(a) If X is a curve, and $\mathscr{L}$ is an invertible sheaf on X, show that $(\mathscr{L} \cdot X) = \deg_X \mathscr{L}$.
(b) Suppose k is an infinite field, $X = \mathbb{P}^N$, and Y is a dimension n subvariety of X. If $H_1, \dots, H_n$ are generally chosen hypersurfaces of degrees $d_1, \dots, d_n$ respectively (so $\dim(H_1 \cap \cdots \cap H_n \cap Y) = 0$ by Exercise 11.3.C(d)), then by Bézout's theorem

(Exercise 18.6.K),

$$\deg(H_1 \cap \cdots \cap H_n \cap Y) = d_1 \cdots d_n \deg(Y).$$

Show that

$$(\mathscr{O}_X(H_1) \cdots \mathscr{O}_X(H_n) \cdot Y) = d_1 \cdots d_n \deg(Y).$$

We now describe some of the properties of the intersection product. In the course of proving Exercise 20.1.B(b) you will in effect solve the following exercise.

20.1.C. EXERCISE. Suppose D is an effective Cartier divisor on X that restricts to an effective Cartier divisor $D|_Y$ on Y (i.e., remains locally not a zerodivisor on Y). Show that

$$(\mathscr{L}_1 \cdots \mathscr{L}_{n-1} \cdot \mathscr{O}(D) \cdot Y) = (\mathscr{L}_1 \cdots \mathscr{L}_{n-1} \cdot D|_Y).$$

More generally, if D is an effective Cartier divisor on X that does not contain any associated point of $\mathscr{F}$, show that

$$(\mathscr{L}_1 \cdots \mathscr{L}_{n-1} \cdot \mathscr{O}(D) \cdot \mathscr{F}) = (\mathscr{L}_1 \cdots \mathscr{L}_{n-1} \cdot \mathscr{F}|_D).$$

(A similar idea came up in the proof that the Hilbert polynomial is actually polynomial; see the discussion around (18.6.2.1).)

20.1.2. *Definition.* For this reason, if D is an effective Cartier divisor, in the symbol for the intersection product, we often write D instead of $\mathscr{O}(D)$. We interchangeably think of intersecting divisors rather than line bundles. For example, we will discuss the special case of intersection theory on a surface in §20.2, and when we intersect two curves C and D, we will write the intersection as $(C \cdot D)$ or even $C \cdot D$.

20.1.D. EXERCISE. Show that the intersection product (20.1.1.1) is preserved by field extension of k.

20.1.3. Proposition. — *Assume X is projective. For fixed $\mathscr{F}$, the intersection product $(\mathscr{L}_1 \cdots \mathscr{L}_n \cdot \mathscr{F})$ is a symmetric multilinear function of the $\mathscr{L}_1, \dots, \mathscr{L}_n$.*

Proposition 20.1.3 is actually true with "projective" replaced by "proper", see [**Kl1**, Prop. 2] or [**Ko1**, Prop. VI.2.7]. Unlike most extensions to the proper case, this is not just an application of Chow's lemma. It involves a different approach, involving a beautiful trick called *dévissage*.

Proof. Symmetry is clear. By Exercise 20.1.D, we may assume that k is infinite (e.g. algebraically closed). We now prove the result by induction on n.

20.1.E. EXERCISE (BASE CASE). Prove the result when $n = 1$. Hint: Exercise 18.4.S. (In fact, you can take the base case to be $n = 0$, if this doesn't confuse you.)

We now assume the result for when the support of the coherent sheaf has dimension less than n.

We now use a trick. We wish to show that (for arbitrary $\mathscr{L}_1, \mathscr{L}_1', \mathscr{L}_2, \dots, \mathscr{L}_n$),

(20.1.3.1) $$(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) + (\mathscr{L}_1' \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) - ((\mathscr{L}_1 \otimes \mathscr{L}_1') \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$$

is 0.

20.1.F. EXERCISE. Rewrite (20.1.3.1) as

$$(20.1.3.2) \qquad (\mathscr{L}_1 \cdot \mathscr{L}_1' \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}).$$

(There are now $n+1$ line bundles appearing in the product, but this does not contradict the definition of the intersection product, as $\dim \operatorname{Supp} \mathscr{F} \leq n < n+1$.)

20.1.G. EXERCISE. Use the inductive hypothesis to show that (20.1.3.1) is 0 if $\mathscr{L}_n \cong \mathscr{O}(D)$ for D an effective Cartier divisor missing the associated points of $\mathscr{F}$.

In particular, if $\mathscr{L}_n$ is very ample, then (20.1.3.1) is 0, as Exercise 18.6.A shows that there exists a section of $\mathscr{L}_n$ missing the associated points of $\mathscr{F}$.

By the symmetry of its incarnation as (20.1.3.2), expression (20.1.3.1) vanishes if $\mathscr{L}_1$ is very ample. Let $\mathscr{A}$ and $\mathscr{B}$ be any two very ample line bundles on X. Then by substituting $\mathscr{L}_1 = \mathscr{B}$ and $\mathscr{L}_1' = \mathscr{A} \otimes \mathscr{B}^\vee$, using the vanishing of (20.1.3.1), we have

$$(20.1.3.3) \quad (\mathscr{A} \otimes \mathscr{B}^\vee \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) = (\mathscr{A} \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) - (\mathscr{B} \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$$

Both summands on the right side of (20.1.3.3) are linear in $\mathscr{L}_n$, so the same is true of the left side. But by Exercise 16.6.C, *any* invertible sheaf on X may be written in the form $\mathscr{A} \otimes \mathscr{B}^\vee$ ("as the difference of two very amples"), so $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$ is linear in $\mathscr{L}_n$, and thus (by symmetry) in each of the $\mathscr{L}_i$. (An interesting feature of this argument is that we intended to show linearity in $\mathscr{L}_1$, and ended up showing linearity in $\mathscr{L}_n$.) □

We have an added bonus arising from the proof.

20.1.H. EXERCISE. Suppose X is projective. Show that if $\dim \operatorname{Supp} \mathscr{F} < n+1$, and $\mathscr{L}_1, \mathscr{L}_1', \mathscr{L}_2, \dots, \mathscr{L}_n$ are invertible sheaves on X, then (20.1.3.2) vanishes. In other words, the intersection product of $n+1$ invertible sheaves with a coherent sheaf $\mathscr{F}$ vanishes if the $\dim \operatorname{Supp} \mathscr{F} < n+1$. (In fact, the result holds with "projective" replaced by "proper", as the results it relies on hold in this greater generality.)

20.1.4. Proposition. — *Suppose X is projective. The intersection product depends only on the numerical equivalence classes of the $\mathscr{L}_i$.*

(Numerical equivalence was defined in §18.4.10.) Again, the result remains true with "projective" replaced by "proper". But in the proof here, we use the fact that every line bundles is the difference two very ample line bundles in both the proof of Proposition 20.1.3 and in the proof of Proposition 20.1.4 itself. For the proof of the Proposition in the proper case, see [**FGIKNV**, Prop. B.20].

Proof. Suppose $\mathscr{L}_1$ is numerically equivalent to $\mathscr{L}_1'$, and $\mathscr{L}_2, \dots, \mathscr{L}_n$, and $\mathscr{F}$ are arbitrary. We wish to show that $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) = (\mathscr{L}_1' \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$. By Exercise 20.1.D, we may assume that k is infinite (e.g. algebraically closed). We proceed by induction on n. The case $n = 1$ follows from Exercise 18.4.S. We assume that $n > 1$, and assume the result for "smaller n". By multilinearity of the intersection product, and the fact that each $\mathscr{L}_n$ maybe written as the "difference" of two very ample invertible sheaves (Exercise 16.6.C), it suffices to prove the result in the case when $\mathscr{L}_n$ is very ample. We may then write $\mathscr{L}_n = \mathscr{O}(D)$, where D is an effective Cartier divisor missing the associated points of $\mathscr{F}$ (Exercise 18.6.A). Then

$$\begin{aligned}(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) &= (\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_{n-1} \cdot \mathscr{F}|_D) \quad \text{(Ex. 20.1.C)} \\ &= (\mathscr{L}_1' \cdot \mathscr{L}_2 \cdots \mathscr{L}_{n-1} \cdot \mathscr{F}|_D) \quad \text{(inductive hyp.)} \\ &= (\mathscr{L}_1' \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) \quad \text{(Ex. 20.1.C).}\end{aligned}$$

□

20.1.5. *Asymptotic Riemann-Roch.*

If Y is a proper curve, $\chi(Y, \mathscr{L}^{\otimes m}) = m \deg_Y \mathscr{L} + \chi(Y, \mathscr{O}_Y)$ (see (18.4.4.1)) is a linear polynomial in m, whose leading term is an intersection product. This generalizes.

20.1.I. EXERCISE (ASYMPTOTIC RIEMANN-ROCH). Suppose X is projective. (As usual, the result will remain true with "projective" replaced by "proper".) Suppose $\mathscr{F}$ is a coherent sheaf on X with $\dim \operatorname{Supp} \mathscr{F} \leq n$, and $\mathscr{L}$ is a line bundle on X. Show that $\chi(X, \mathscr{L}^{\otimes m} \otimes \mathscr{F})$ is a polynomial in m of degree at most n. Show that the coefficient of m^n in this polynomial (the "leading term") is $(\mathscr{L}^n \cdot \mathscr{F})/n!$. Hint: Exercise 20.1.H implies that $(\mathscr{L}^{n+1} \cdot (\mathscr{L}^{\otimes i} \otimes \mathscr{F})) = 0$. (Careful with this notation: $\mathscr{L}^{n+1}$ doesn't mean $\mathscr{L}^{\otimes(n+1)}$, it means $\mathscr{L} \cdot \mathscr{L} \cdots \mathscr{L}$ with $n+1$ factors.) Expand this out using (20.1.1.1) to get a recursion for $\chi(X, \mathscr{L}^{\otimes m} \otimes \mathscr{F})$. Your argument may resemble the proof of polynomiality of the Hilbert polynomial, Theorem 18.6.1, so you may find further hints there. Exercise 18.6.C in particular might help.

We know all the coefficients of this polynomial if X is a curve, by Riemann-Roch (see (18.4.4.1)), or basically by definition. We will know/interpret all the coefficients if X is a regular projective surface and $\mathscr{F}$ is an invertible sheaf when we prove Riemann-Roch for surfaces (Exercise 20.2.B(b)). To understand the general case, we need the theory of Chern classes. The result is the Hirzebruch-Riemann-Roch Theorem, which can be further generalized to the celebrated Grothendieck-Riemann-Roch Theorem (see [**F**, §15.2]).

20.1.J. EXERCISE (THE PROJECTION FORMULA). Suppose $\pi : X_1 \to X_2$ is a (projective) morphism of integral projective schemes (over a field k) of the same dimension n, and $\mathscr{L}_1, \dots, \mathscr{L}_n$ are invertible sheaves on X_2. Show that $(\pi^* \mathscr{L}_1 \cdots \pi^* \mathscr{L}_n) = \deg(X_1/X_2)(\mathscr{L}_1 \cdots \mathscr{L}_n)$. (The first intersection is on X_1, and the second is on X_2.) Hint: Let $d = \deg \pi$, and assume $d > 0$. (Deal with the case where π is not dominant separately, so $d = 0$ by convention, using Chevalley's Theorem 7.4.2.) Argue that by the multilinearity of the intersection product, it suffices to deal with the case where the $\mathscr{L}_i$ are very ample. Then choose sections of each $\mathscr{L}_i$, all of whose intersection lies in an open subset U where π has "genuine degree deg d". To find U: first use Exercise 9.3.G to find a dense open subset $U' \subset X_2$ over which π is finite. Then use Useful Exercise 13.7.F to show that there exists a dense open subset $U \subset U'$ on which $\pi_* \mathscr{O}$ is a locally free sheaf of rank d. In the language of Chapter 24, you are showing that there is a dense open subset U of X_2 over which π is finite and flat (and hence has "constant degree", see Exercise 24.4.G). (As usual, the result holds with "projective" replaced with "proper"; see [**Ko1**, Prop. VI.2.11].)

20.1.6. *Remark: A more general projection formula.* Suppose $\pi : X_1 \to X_2$ is a proper morphism of proper varieties, and $\mathscr{F}$ is a coherent sheaf on X_1 with $\dim \operatorname{Supp} \mathscr{F} \leq$

n (so $\dim \operatorname{Supp} \pi_* \mathscr{F} \leq n$, using Exercise 11.2.C). Suppose also that $\mathscr{L}_1, \dots, \mathscr{L}_n$ are invertible sheaves on X_2. Then

$$(\pi^* \mathscr{L}_1 \cdots \pi^* \mathscr{L}_n \cdot \mathscr{F}) = (\mathscr{L}_1 \cdots \mathscr{L}_n \cdot \pi_* \mathscr{F}).$$

This is called the **projection formula** (and generalizes, in a nonobvious way, Exercise 20.1.J). Proofs are given in [**FGIKNV**, B.15] and [**Ko1**, Prop. VI.2.11]. Because we won't use this version of the projection formula, we do not give a proof here.

20.1.K. EXERCISE (INTERSECTING WITH AMPLE LINE BUNDLES). Suppose X is a projective k-variety, and $\mathscr{L}$ is an ample line bundle on X. Show that for any subvariety Y of X of dimension n, $(\mathscr{L}^n \cdot Y) > 0$. (Hint: use Proposition 20.1.3 and Theorem 16.6.2 to reduce to the case where $\mathscr{L}$ is very ample. Then show that $(\mathscr{L}^n \cdot Y) = \deg Y$ in the embedding into projective space induced by the linear series $|\mathscr{L}|$.)

The Nakai-Moishezon criterion (Theorem 20.4.1) states that this characterizes ampleness.

20.1.7. $\star\star$ *Cohomological interpretation in the complex projective case, generalizing Exercise 18.4.G and §18.6.4.* If $k = \mathbb{C}$, we can interpret $(\mathscr{L}_1 \cdots \mathscr{L}_n \cdot Y)$ as the degree of

$$c_1((\mathscr{L}_1)_{an}) \cup \cdots \cup c_1((\mathscr{L}_n)_{an}) \cap [Y_{an}] \tag{20.1.7.1}$$

in $H_0(Y_{an}, \mathbb{Z})$. (Recall $c_1((\mathscr{L}_i)_{an}) \in H^2(X_{an}, \mathbb{Z})$, as discussed in Exercise 18.4.G.) One way of proving this is to use multilinearity of both the intersection product and (20.1.7.1) to reduce to the case where the $\mathscr{L}_n$ is very ample, so $\mathscr{L}_n \cong \mathscr{O}(D)$, where D restricts to an effective Cartier divisor E on Y. Then show that if $\mathscr{L}$ is an analytic line bundle on Y_{an} with nonzero section E_{an}, then $c_1(\mathscr{L}) \cap [Y_{an}] = [E_{an}]$. Finally, use induction on n and Exercise 20.1.C.

20.1.8. Important remark: additive notation for line bundles. There is a standard, useful, but confusing convention suggested by the multilinearity of the intersection product: we write tensor product of invertible sheaves *additively*. This is convenient for working with multilinearity. (Some people try to avoid confusion by using divisors rather than line bundles, as we add divisors when we "multiply" the corresponding line bundles. This is psychologically helpful, but may add more confusion, as one then has to worry about the whether and why and how and when line bundles correspond to divisors.) We will use this, for example, in Exercises 20.2.B–20.2.G, §20.2.9, and §20.4.

20.2 Intersection theory on a surface

We now apply the general machinery of §20.1 to the case of a regular projective surface X. (What matters is that is X is Noetherian and factorial, so $\operatorname{Pic} X \to \operatorname{Cl} X$ is an isomorphism, Proposition 14.2.10. Recall that regular schemes are factorial by the Auslander-Buchsbaum Theorem 12.8.5.)

20.2.A. EXERCISE/DEFINITION. Suppose C and D are effective divisors (i.e., curves) on X.

(a) Show that

$$\deg_C \mathscr{O}_X(D)|_C \tag{20.2.0.1}$$
$$= (\mathscr{O}(C)\cdot\mathscr{O}(D)\cdot X) \tag{20.2.0.2}$$
$$= \deg_D \mathscr{O}_X(C)|_D. \tag{20.2.0.3}$$

We call this the **intersection number** of C and D, and denote it $C \cdot D$.
(b) If C does not contain any associated point of D (so in particular, if C and D have no components in common), show that

$$C\cdot D = h^0(C\cap D, \mathscr{O}_{C\cap D}) \tag{20.2.0.4}$$

where $C \cap D$ is the scheme-theoretic intersection of C and D on X.

20.2.1. *Important aside.* The hypothesis in Exercise 20.2.A(b), that C not contain any associated point of D, is a red herring. In fact, D can never have any embedded points, as we will see in §26.2.5 when discussing Cohen-Macaulay rings. (More generally, we will see that local complete intersections have no embedded points — this should help motivate you to learn about Cohen-Macaulayness.) The case of $\mathbb{A}^2$ (and hence $\mathbb{P}^2$) can be done by hand (Exercise 5.5.I).

Thus the hypothesis in Exercise 20.2.A(b) can be replaced by the more simple "C and D have no common components".

20.2.2. *Advantages and disadvantages.* We thus have four descriptions of the intersection number (20.2.0.1)–(20.2.0.4), each with advantages and disadvantages. The Euler characteristic description (20.2.0.2) is remarkably useful (for example, in the exercises below), but the geometry is obscured. The definition $\deg_C \mathscr{O}_X(D)|_C$, (20.2.0.1), is not obviously symmetric in C and D. The definition $h^0(C\cap D, \mathscr{O}_{C\cap D})$, (20.2.0.4), is clearly local — to each point of $C \cap D$, we have a vector space. For example, we know that in $\mathbb{A}^2_k$, $y - x^2 = 0$ meets the x-axis with multiplicity 2, because h^0 of the scheme-theoretic intersection $(k[x,y]/(y-x^2, y))$ has dimension 2. (This h^0 is also the *length* of the dimension 0 scheme, whose definition you may be able to figure out given Definition 18.4.7 of the length of a module. But we won't use this terminology.)

By Proposition 20.1.3, the intersection number induces a bilinear "intersection form"

$$\operatorname{Pic} X \times \operatorname{Pic} X \to \mathbb{Z}. \tag{20.2.2.1}$$

By Asymptotic Riemann-Roch (Exercise 20.1.I), $\chi(X, \mathscr{O}(nD))$ is a quadratic polynomial in n.

20.2.3. You can verify that Exercise 20.2.A recovers Bézout's theorem for plane curves (see Exercise 18.6.K), using $\chi(\mathbb{P}^2, \mathscr{O}(n)) = (n+2)(n+1)/2$ (from Theorem 18.1.3), and the fact that effective Cartier divisors on $\mathbb{P}^2_k$ have no embedded points (Exercise 5.5.I).

Before getting to a number of interesting explicit examples, we derive a couple of fundamental theoretical facts.

20.2.B. EXERCISE. Assuming Serre duality for X (Theorem 18.5.1), prove the following for a smooth projective surface X. (We are mixing divisor and invertible sheaf notation, as described in Remark 20.1.8, so be careful. Here K_X is a divisor corresponding to ω_X.)

(a) *(sometimes called the adjunction formula)* $C \cdot (K_X + C) = 2p_a(C) - 2$ for any curve $C \subset X$. Hint: compute $(C \cdot (-K_X - C)$ instead. (See Exercise 21.5.B and §30.4 for other versions of the adjunction formula.)
(b) *(Riemann-Roch for surfaces)*

$$\chi(X, \mathscr{O}_X(D)) = D \cdot (D - K_X)/2 + \chi(X, \mathscr{O}_X)$$

for any Weil divisor D (cf. Riemann-Roch for curves, Exercise 18.4.B).

20.2.4. Two explicit examples: $\mathbb{P}^1 \times \mathbb{P}^1$ and $\mathrm{Bl}_p \mathbb{P}^2$.

20.2.C. EXERCISE: $X = \mathbb{P}^1 \times \mathbb{P}^1$. Recall from Exercise 14.2.O that $\operatorname{Pic}(\mathbb{P}^1 \times \mathbb{P}^1) = \mathbb{Z}\ell \times \mathbb{Z}m$, where ℓ is the curve $\mathbb{P}^1 \times \{0\}$ and m is the curve $\{0\} \times \mathbb{P}^1$. Show that the intersection form (20.2.2.1) is given by $\ell \cdot \ell = m \cdot m = 0$, $\ell \cdot m = 1$. (Hint: You can compute the cohomology groups of line bundles on $\mathbb{P}^1 \times \mathbb{P}^1$ using Exercise 18.3.E, but it is much faster to use Exercise 20.2.A(b).) What is the class of the diagonal in $\mathbb{P}^1 \times \mathbb{P}^1$ in terms of these generators?

20.2.D. EXERCISE: THE BLOWN UP PROJECTIVE PLANE. (You needn't have read Chapter 22 to do this exercise!) Let $X = \mathrm{Bl}_p \mathbb{P}^2$ be the blow-up of $\mathbb{P}^2_k$ at a k-valued point (the origin, say) p — see Exercise 9.3.F, which describes the blow-up of $\mathbb{A}^2_k$, and "compactify". Interpret $\operatorname{Pic} X$ as generated (as an abelian group) by ℓ and e, where ℓ is a line not passing through the origin, and e is the exceptional divisor. Show that the intersection form (20.2.2.1) is given by $\ell \cdot \ell = 1$, $e \cdot e = -1$, and $\ell \cdot e = 0$. Hence show that $\operatorname{Pic} X \cong \mathbb{Z}\ell \times \mathbb{Z}e$ (as promised in the aside in Exercise 14.2.P). In particular, the exceptional divisor has negative self-intersection. (This exercise will be generalized in §22.4.13.)

20.2.5. *Hint.* Here is a possible hint to get the intersection form in Exercise 20.2.D. The scheme-theoretic preimage in $\mathrm{Bl}_p \mathbb{P}^2$ of a line through the origin is the scheme-theoretic union of the exceptional divisor e and the "proper transform" m of the line through the origin. Show that $\ell = e + m$ in $\operatorname{Pic}(\mathrm{Bl}_p \mathbb{P}^2)$ (writing the Picard group law additively, cf. Remark 20.1.8). Show that $\ell \cdot m = e \cdot m = 1$ and $m \cdot m = 0$.

20.2.6. *Definition: (-1)-curve.* Notice that the exceptional divisor e has self-intersection -1. We will see more generally in Exercise 22.4.O that this is the case for all exceptional divisors of blow-ups of smooth surfaces (over k) blown up at a k-valued point. We give such curves a name. If X is a surface over k, and $C \subset X$ is a curve in X consisting of smooth points, with $C \cong \mathbb{P}^1_k$, and $C \cdot C = -1$, we say that C is a (-1)**-curve**. (Once we know more, we will be able to restate "$C \cdot C = -1$" as "C has normal bundle $\mathscr{O}(-1)$": combine Exercise 21.2.H with Exercise 20.2.A(a).)

20.2.E. EXERCISE. Show that the blown up projective plane $\mathrm{Bl}_p \mathbb{P}^2$ in Exercise 20.2.D is not isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$, perhaps considering their (isomorphic) Picard groups, and identifying which classes are effective (represented by effective divisors). (This is an example of a pair of smooth projective birational surfaces that have isomorphic Picard groups, but which are not isomorphic. This exercise shows that $\mathbb{F}_0$ is not isomorphic to $\mathbb{F}_1$. (The method and result are generalized in Exercise 20.2.Q.)

20.2.F. EXERCISE (CF. EXERCISE 18.4.X). Show that the nef cone (Exercise 18.4.W) of $\mathrm{Bl}_p \mathbb{P}^2$ is generated by ℓ and m. Hint: show that ℓ and m are nef. By intersecting line bundles with the *curves* e and ℓ, show that nothing outside the cone spanned by ℓ and m is nef. (Side Remark: note that as in Exercise 18.4.X, linear series corresponding to the boundaries of the cone give "interesting contractions".)

20.2.G. EXERCISE: A PROPER NONPROJECTIVE SURFACE. Show the existence of a proper nonprojective surface over a field as follows, parallelling the construction of a proper nonprojective threefold in §16.4.10. Take two copies of the blown up projective plane $\mathrm{Bl}_p \mathbb{P}^2$, gluing ℓ on the first to e on the second, and e on the second to ℓ on the first. Hint: show that if $\mathscr{L}$ is a line bundle having positive degree on each effective curve, then $\mathscr{L} \cdot \ell > \mathscr{L} \cdot e$, using $\ell = e + m$ from Hint 20.2.5. (See §19.11.11 for another example of a proper nonprojective surface.)

20.2.7. Fibrations.

Suppose $\pi : X \to B$ is a morphism from a regular projective surface to a regular curve and $b \in B$ is a closed point. Let $F = \pi^* b$. Then $\mathscr{O}_X(F) = \pi^* \mathscr{O}_B(b)$, which is isomorphic to $\mathscr{O}$ on F. Thus $F \cdot F = \deg_F \mathscr{O}_X(F) = 0$: "the self-intersection of a fiber is 0". The same argument works without X being regular, as long as you phrase it properly: $(\pi^* \mathscr{O}_X(b))^2 = 0$.

20.2.H. EXERCISE. Suppose E is an elliptic curve, with origin p. On $E \times E$, let Δ be the diagonal. By considering the "difference" map $E \times E \to E$, for which $\pi^*(p) = \Delta$, show that $\Delta^2 = 0$. Show that $N^1_{\mathbb{Q}}(X)$ has rank at least 3. Show that in general for schemes X and Y, $\operatorname{Pic} X \times \operatorname{Pic} Y \to \operatorname{Pic}(X \times Y)$ (defined by pulling back and tensoring) need not be isomorphism; the case of $X = Y = \mathbb{P}^1$ is misleading.

Remark: $\dim_{\mathbb{Q}} N^1_{\mathbb{Q}}(E \times E)$ is always 3 or 4. It is 4 if there is a nontrivial endomorphism from E to itself (i.e., not just multiplication $[n]$ by some n, §19.11.5, followed by a translation); the additional class comes from the graph of this endomorphism. (See [**Mu3**, §21, App. III] for an introduction to the tools needed to show this.)

Our next goal is to describe the self-intersection of a curve on a ruled surface (Exercise 20.2.J). To set this up, we have a useful preliminary result.

20.2.I. EXERCISE (THE NORMAL BUNDLE TO A SECTION OF $\mathcal{P}roj$ OF A RANK 2 VECTOR BUNDLE). Suppose X is a scheme, and and $\mathscr{V}$ is a rank 2 locally free sheaf on C. Explain how the short exact sequences

$$0 \to \mathscr{S} \to \mathscr{V} \to \mathscr{Q} \to 0 \tag{20.2.7.1}$$

on X, where $\mathscr{S}$ and $\mathscr{Q}$ have rank 1, correspond to the sections $\sigma : X \to \mathbb{P}\mathscr{V}$ to the projection $\mathbb{P}\mathscr{V} \to X$. Show that the normal bundle to $\sigma(X)$ in $\mathbb{P}\mathscr{V}$ is $\mathscr{Q} \otimes \mathscr{S}^\vee$. (A generalization is stated in §21.4.10.) Hint: (i) For simplicity, it is convenient to assume $\mathscr{S} = \mathscr{O}_X$, by replacing $\mathscr{V}$ by $\mathscr{V} \otimes \mathscr{S}^\vee$, as the statement of the problem respects tensoring by an invertible sheaf (see Exercise 17.2.G). (ii) Assume now (*with* loss of generality) that $\mathscr{Q} \cong \mathscr{O}_X$. Then describe the section as $\sigma : X \to \mathbb{P}^1 \times X$, with X mapping to the 0 section. Describe an isomorphism of $\mathscr{O}_X$ with the normal bundle to $\sigma(X) \to \mathbb{P}^1 \times X$. (Do *not* just say that the normal bundle "is trivial".) (iii) Now consider the case where $\mathscr{Q}$ is general. Choose trivializing neighborhoods U_i of $\mathscr{Q}$, and let g_{ij} be the the transition function for $\mathscr{Q}$. On the overlap between two trivializing neighborhoods $U_i \cap U_j$, determine how your two isomorphisms of $\mathscr{O}_X$

with $N_{\sigma(X)/\mathbb{P}^1_X}$ with $\mathscr{O}_X$ from (ii) (one for U_i, one for U_j) are related. In particular, show that they differ by g_{ij}.

20.2.J. EXERCISE (SELF-INTERSECTIONS OF SECTIONS OF RULED SURFACES). Suppose C is a regular curve, and $\mathscr{V}$ is a rank 2 locally free sheaf on C. Then $\mathbb{P}\mathscr{V}$ is a ruled surface (Example 17.2.4). Fix a section σ of $\mathbb{P}\mathscr{V}$ corresponding to a filtration (20.2.7.1). Show that $\sigma(C) \cdot \sigma(C) = \deg_C \mathscr{Q} \otimes \mathscr{S}^\vee$.

20.2.8. The Hirzebruch surfaces $\mathbb{F}_n = \mathcal{P}roj_{\mathbb{P}^1}(\mathscr{O}_{\mathbb{P}^1} \oplus \mathscr{O}_{\mathbb{P}^1}(n))$.

Recall the definition of the Hirzebruch surface $\mathbb{F}_n = \mathcal{P}roj_{\mathbb{P}^1}(\mathscr{O}_{\mathbb{P}^1} \oplus \mathscr{O}_{\mathbb{P}^1}(n))$ in Example 17.2.4. It is a $\mathbb{P}^1$-bundle over $\mathbb{P}^1$; let $\pi : \mathbb{F}_n \to \mathbb{P}^1$ be the structure morphism. Using Exercise 20.2.J, corresponding to

$$0 \to \mathscr{O}(n) \to \mathscr{O} \oplus \mathscr{O}(n) \to \mathscr{O} \to 0,$$

we have a section of π of self-intersection $-n$; call it $E \subset \mathbb{F}_n$. Similarly, corresponding to

$$0 \to \mathscr{O} \to \mathscr{O} \oplus \mathscr{O}(n) \to \mathscr{O}(n) \to 0,$$

we have a section $C \subset \mathbb{F}_n$ of self-intersection n. Let p be any k-valued point of $\mathbb{P}^1$, and let $F = \pi^*(p)$.

20.2.K. EXERCISE. Show that $\mathscr{O}(F)$ is independent of the choice of p.

20.2.L. EXERCISE. Show that $\operatorname{Pic} \mathbb{F}_n$ is generated by E and F. In the course of doing this, you will develop "local charts" for $\mathbb{F}_n$, which will help you solve later exercises.

20.2.M. EXERCISE. Compute the intersection matrix on $\operatorname{Pic} \mathbb{F}_n$. Show that E and F are independent, and thus $\operatorname{Pic} \mathbb{F}_n \cong \mathbb{Z}E \oplus \mathbb{Z}F$. Calculate C in terms of E and F.

20.2.N. EXERCISE. Show how to identify $\mathbb{F}_n \setminus E$, along with the structure map π, with the total space of the line bundle $\mathscr{O}(n)$ on $\mathbb{P}^1$, with C as the 0-section. Similarly show how to identify $\mathbb{F}_n \setminus C$ with the total space of the line bundle $\mathscr{O}(-n)$ on $\mathbb{P}^1$; with E as the 0-section.

20.2.O. EXERCISE. Show that $h^0(\mathbb{F}_n, \mathscr{O}_{\mathbb{F}_n}(C)) > 1$. Hint: As $\mathscr{O}_{\mathbb{F}_n}(C)$ has a section — namely C — we have that $h^0(\mathbb{F}_n, \mathscr{O}_{\mathbb{F}_n}(C)) \geq 1$. One way to proceed is to write down another section using local charts for $\mathbb{F}_n$.

20.2.P. EXERCISE. Show that every effective curve on $\mathbb{F}_n$ is a non-negative linear combination of E and F. (Conversely, it is clear that for every nonnegative a and b, $\mathscr{O}(aE + bF)$ has a section, corresponding to the effective curve "$aE + bF$". The extension of this to $N^1_{\mathbb{Q}}$ is called the *effective cone*, and this notion, extended to proper varieties more general, can be very useful. This exercise shows that E and F generate the effective cone of $\mathbb{F}_n$.) Hint: show that because "F moves", any effective curve must intersect F nonnegatively, and similarly because "C moves" (Exercise 20.2.O), any effective curve must intersect C nonnegatively. If $\mathscr{O}(aE+bF)$ has a section corresponding to an effective curve D, what does this say about a and b?

20.2.Q. EXERCISE. By comparing effective cones, and the intersection pairing, show that the $\mathbb{F}_n$ are pairwise nonisomorphic. (This result was promised in Example 17.2.4. Exercise 20.2.E is a special case.)

Exercise 20.2.Q is difficult to do otherwise, and foreshadows the fact that nef and effective cones are useful tools in classifying and understanding varieties general. In particular, they are central to the minimal model program.

20.2.R. EXERCISE. If $n = 0$, show that there are no curves on $\mathbb{F}_n$ of negative self-intersection. If $n < 0$, show that E is the unique curve on $\mathbb{F}_n$ with self-intersection $-n$, and there are no curves on $\mathbb{F}_n$ of smaller self-intersection. This again gives another (related) means of showing that the $\mathbb{F}_n$ are pairwise nonisomorphic.

20.2.S. EXERCISE. Show that the nef cone of $\mathbb{F}_n$ is generated by C and F. (We will soon see that by Kleiman's criterion for ampleness, Theorem 20.4.7, that the ample cone is the interior of this cone, so we have now identified the ample line bundles on $\mathbb{F}_n$.)

20.2.T. EXERCISE. We have seen earlier (Exercises 20.2.F and 18.4.X) that the boundary of the nef cone give "interesting contractions". What are the maps given by the two linear series corresponding to $\mathscr{O}(F)$ and $\mathscr{O}(C)$?

After this series of exercises, you may wish to revisit Exercises 20.2.C-20.2.F, and interpret them as special cases: $\mathbb{F}_0 \cong \mathbb{P}^1 \times \mathbb{P}^1$ and $\mathbb{F}_1 \cong \mathrm{Bl}_p \mathbb{P}^2$.

20.2.9. The Hodge Index Theorem.

We use what we have learned to prove the following celebrated result.

20.2.10. The Hodge Index Theorem. — *Suppose X is an irreducible smooth surface (over a field k) with $\mathscr{L}, \mathscr{H} \in \operatorname{Pic} X$, with $\mathscr{H} \cdot \mathscr{H} > 0$ and $\mathscr{L} \cdot \mathscr{H} = 0$. Then (a) $\mathscr{L} \cdot \mathscr{L} \leq 0$, and (b) equality holds if and only if $\mathscr{L}$ is numerically trivial.*

Before reading on, you should review the definitions of $N^1(X)$ and $\rho(X)$ in §18.4.10.

20.2.U. EXERCISE (REASON FOR THE NAME). (We will not use this.) By arguments similar to the classification of quadratic forms over algebraically closed fields (Exercise 5.4.J), one can show that given a symmetric bilinear form on a *real* vector space of (finite) dimension n, there are integers a_+, a_0, a_- such that then given any orthogonal basis $v_1, \dots, v_n$, the number of $v_i \cdot v_i$ that are positive (resp. zero, negative) is a_+ (resp. a_0, a_-). The ordered pair (a_+, a_-) is often called the **index**, and is sometimes called the **signature**. (See, for example, [**Lan**, §XV.4] for a proof of this theorem of Sylvester.) Show that the Hodge index theorem implies that $N^1_{\mathbb{R}}(X) := N^1(X) \otimes_{\mathbb{Z}} \mathbb{R}$, along with the bilinear form coming from the intersection product, has index $(1, \rho(X) - 1)$. You should think of the positive "term" as coming from the existence of an ample class; the "rest" is negative.

Proof.

20.2.11. We begin with the following fact. If $\mathscr{H}$ is an ample invertible sheaf, $\mathscr{M}$ is any invertible sheaf, and $\mathscr{M} \cdot \mathscr{H} > \omega_X \cdot \mathscr{H}$, then $h^2(X, \mathscr{M}) = 0$. (Translation: for all invertible sheaves "more positive than ω_X", there is no top cohomology;

compare this to §19.2.5, the corresponding statement in dimension 1.) Reason: by Serre duality (Theorem 18.5.1), we wish to show that $h^0(X, \omega_X \otimes \mathscr{M}^\vee) = 0$. But if otherwise $\omega_X \otimes \mathscr{M}^\vee$ had a nonzero section, then its (effective nonzero) divisor would intersect $\mathscr{H}$ nonnegatively (Exercise 20.1.K), so (using additive notation, see Remark 20.1.8) $(\omega_X - \mathscr{M}) \cdot \mathscr{H} \geq 0$, contradicting $\mathscr{M} \cdot \mathscr{H} > \omega_X \cdot \mathscr{H}$.

20.2.12. We next show that if $\mathscr{H}$ is ample, and $\mathscr{L}$ is any invertible sheaf with $\mathscr{L} \cdot \mathscr{L} > 0$ and $\mathscr{L} \cdot \mathscr{H} > 0$, then for $n \gg 0$, $\mathscr{L}^{\otimes n}$ has a section. Here is why. For $n \gg 0$,

$$\mathscr{L}^{\otimes n} \cdot \mathscr{H} = n\mathscr{L} \cdot \mathscr{H} > \omega_X \cdot \mathscr{H},$$

so $h^2(X, \mathscr{L}^{\otimes n}) = 0$ by §20.2.11 (taking $\mathscr{M} = \mathscr{L}^{\otimes n}$). Then for $n \gg 0$, by Riemann-Roch for surfaces (Exercise 20.2.B(b)),

$$\begin{aligned} h^0(X, \mathscr{L}^{\otimes n}) - h^1(X, \mathscr{L}^{\otimes n}) &= \chi(X, \mathscr{L}^{\otimes n}) \\ &= (n^2/2)\mathscr{L} \cdot \mathscr{L} - n\mathscr{L} \cdot \omega_X + \chi(X, \mathscr{O}_X). \end{aligned}$$

As $\mathscr{L} \cdot \mathscr{L} > 0$, the right side is positive for $n \gg 0$, so

$$h^0(X, \mathscr{L}^{\otimes n}) \geq h^0(X, \mathscr{L}^{\otimes n}) - h^1(X, \mathscr{L}^{\otimes n}) > 0$$

as desired.

20.2.13. We are now ready to prove the Hodge Index Theorem. Assume first that $\mathscr{L} \cdot \mathscr{L} > 0$ (and $\mathscr{L} \cdot \mathscr{H} = 0$). Then for $n \gg 0$, $\mathscr{H}' := \mathscr{L} \otimes \mathscr{H}^{\otimes n}$ is ample (indeed very ample) by Exercise 16.6.E. Then $\mathscr{L} \cdot \mathscr{H}' = \mathscr{L} \cdot \mathscr{L} > 0$, so by §20.2.12, $\mathscr{L}^{\otimes n}$ is effective, which implies that $\mathscr{L} \cdot \mathscr{H} = (\mathscr{L}^{\otimes n} \cdot \mathscr{H})/n > 0$ (using Exercise 20.1.K again), contradicting our hypothesis.

20.2.14. Assume finally that $\mathscr{L} \cdot \mathscr{H} = 0$, and (i) $\mathscr{L} \cdot \mathscr{L} = 0$, but that (ii) $\mathscr{L}$ is not numerically trivial. By (ii), we can find an invertible sheaf $\mathscr{Q}$ such that $\mathscr{Q} \cdot \mathscr{L} \neq 0$. Then we can find an invertible sheaf $\mathscr{R}$ such that $\mathscr{R} \cdot \mathscr{L} \neq 0$ and $\mathscr{R} \cdot \mathscr{H} = 0$: take

$$\mathscr{R} = (\mathscr{H} \cdot \mathscr{H})\mathscr{Q} - (\mathscr{Q} \cdot \mathscr{H})\mathscr{H}.$$

Then take $\mathscr{L}' = \mathscr{L}^{\otimes n} \otimes \mathscr{R}$, so $\mathscr{L}' \cdot \mathscr{H} = 0$, but because $\mathscr{R} \cdot \mathscr{L} \neq 0$, we can find some $n \in \mathbb{Z}$ so that

$$\mathscr{L}' \cdot \mathscr{L}' = n\mathscr{L} \cdot \mathscr{R} + \mathscr{R} \cdot \mathscr{R} > 0.$$

Then the argument of §20.2.13 applies, with $\mathscr{L}'$ in the place of $\mathscr{L}$. □

20.2.15. *Generalizations and variations.* The hypotheses can be weakened considerably. We used smoothness only because we need Serre duality, with an invertible sheaf ω_X. We will see in §30.4 that we need less than smoothness for this. But we can do much better, and a proof where X is merely required to be geometrically irreducible and proper requires only a minor modification of our argument, see [**FGIKNV**, Thm. B.27].

20.3 The Grothendieck group of coherent sheaves, and an algebraic version of homology

The construction of the intersection product (20.1.1.1) may leave you hungry for something more, especially in light of the cohomological interpretation of §20.1.7. You may want some sort of homology-like theory which is a repository for cycles of different directions, on which (Chern classes of) line bundles can act. We can actually do this easily, given what we know.

20.3.1. *Definition.* If X is a k-variety, we define the **Grothendieck group of coherent sheaves**, which we denote $K(X)$ (and which is often denoted $K_0(X)$), as the abelian group generated symbols of the form $[\mathscr{F}]$ where $\mathscr{F}$ is a coherent sheaf on X, subject to the relations that $[\mathscr{F}] = [\mathscr{F}']$ if $\mathscr{F} \cong \mathscr{F}'$, and $[\mathscr{F}'] + [\mathscr{F}''] = [\mathscr{F}]$ if there is a short exact sequence

$$0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0.$$

By construction, the Grothendieck group is the universal construction of an operator on the category of coherent sheaves that "behaves well in exact sequences". For example, if X is proper, then:

(i) We have a map $\chi : \mathit{Coh}_X \to \mathbb{Z}$, as the Euler characteristic of a coherent sheaf is finite by Theorem 18.1.4(i) in the projective case, and more generally by Grothendieck's Coherence Theorem 18.9.1 in the proper case. This map descends to $\chi : K(X) \to \mathbb{Z}$ by Exercise 18.4.A (which extends without change to the proper case).

(ii) If X is integral, then the rank function

$$\operatorname{rank} : \mathit{Coh}_X \to \mathbb{Z}$$

descends to $\operatorname{rank} : K(X) \to \mathbb{Z}$. (The argument of Exercise 18.4.H applies.)

20.3.2. *Definition and nonstandard notation.* The Grothendieck group is filtered by dimension: let $K(X)^{\leq d}$ be the subgroup of $K(X)$ generated by coherent sheaves supported in dimension at most d. Let $A'_d(X)$ be the dth graded piece of $K(X)$, i.e., $A'_d(X) := K(X)^{\leq d}/K(X)^{\leq d-1}$.

If $\mathscr{L}$ is an invertible sheaf on X, define $\mathscr{L}\cdot : K(X) \to K(X)$ by $\mathscr{L} \cdot [\mathscr{F}] = [\mathscr{F}] - [\mathscr{L}^\vee \otimes \mathscr{F}]$. (Do you see why this operator is well-defined?)

20.3.A. Exercise. If X is projective and k is infinite, show that $\mathscr{L}\cdot$ sends $K(X)^{\leq d}$ to $K(X)^{\leq d-1}$. (Hint: for each fixed $\mathscr{F}$ supported on a subset of dimension at most d, write $\mathscr{L}$ is a difference of two very ample invertible sheaves, and choose sections of those two, missing the associated points of $\mathscr{F}$.)

20.3.3. *Remark.* The previous exercise holds true without k being infinite or X being proper; see [**Ko1**, Prop. VI.2.5].

20.3.4. For the rest of this section, we assume X is projective and k is infinite. (But in light of Remark 20.3.3, these hypotheses can be removed.) By Exercise 20.3.A, $\mathscr{L}\cdot$ descends to a map $A'_d(X) \to A'_{d-1}(X)$; we denote this operator $c_1(\mathscr{L})\cap$. (It is the action of the first Chern class.)

20.3.B. Exercise ("c_1 is additive"). Show that $c_1(\mathscr{L} \otimes \mathscr{L}')\cap = (c_1(\mathscr{L})\cap) + (c_1(\mathscr{L}')\cap)$. Hint: show that $((\mathscr{L} \otimes \mathscr{L}')\cdot) - (\mathscr{L}\cdot) - (\mathscr{L}'\cdot) = (\mathscr{L}\cdot) \circ (\mathscr{L}'\cdot)$, and thus sends $K(X)^{\leq d}$ to $K(X)^{\leq d-2}$. (This will remind you of the trick in the proof of Proposition 20.1.3, and indeed is the motivation for that trick. Caution: the action of Pic is not additive on $K(X)$; it only becomes additive once we pass to the associated graded ring $A'_\bullet(X)$.)

If $\mathscr{F}$ has support of dimension at most n, and $\mathscr{L}_1, \dots, \mathscr{L}_n$ are n invertible sheaves, we can now reinterpret the intersection product $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$ as

$$c_1(\mathscr{L}_1) \cap \cdots \cap c_1(\mathscr{L}_n) \cap [\mathscr{F}]$$

where $[\mathscr{F}]$ is interpreted as lying in either $K(X)$ or $A'_n(X)$. You can now go back and read §20.1, and reprove all the results with this starting point, sometimes obtaining interesting generalizations.

These $A'_d(X)$ behave like homology groups in a number of ways. If $Y \subset X$ is a closed subvariety of pure dimension d, we have a class $[Y] := [\mathscr{O}_Y] \in A'_d(X)$. The groups $A'_d(X)$ have appropriate functorial properties. For example, if $\pi : X_1 \to X_2$ is a proper morphism (even if the X_i are not themselves proper), we have a map $\pi_* : K(X_1) \to K(X_2)$ given by

$$\pi_*[\mathscr{F}] = [\pi_*\mathscr{F}] - [R^1\pi_*\mathscr{F}] + \cdots \tag{20.3.4.1}$$

(using the long exact sequence for higher pushforwards, Theorem 18.8.1(c)) which descends to a map $A'_d(X_1) \to A'_d(X_2)$, where the "later terms" on the right side of (20.3.4.1) disappear, so $\pi_*[\mathscr{F}] = [\pi_*\mathscr{F}]$. This pushforward interacts well with the fist Chern class of line bundles, yielding the projection formula of Remark 20.1.6.

If π is instead a *flat morphism* (Remark 16.3.8, soon to be discussed at length in Chapter 24), then π^* is exact, so we have a map $\pi^* : K(X_2) \to K(X_1)$. If the "relative dimension" of this map is r (to be properly defined in Definition 24.5.6), this yields a map $\pi^* A'_d(X_2) \to A'_{d+r}(X_1)$. This interacts well with first Chern classes and proper pushforwards.

If $k = \mathbb{C}$ and X is proper, there is a map $A'_d(X) \to H_{2d}(X, \mathbb{Q})$, which behaves as you might hope (for example, in its interaction with Chern classes of line bundles). If X is *not* proper (but $k = \mathbb{C}$), then the map is to Borel-Moore homology rather than usual homology.

Our groups $A'_d(X)$ are a good approximation of the theory of Chow groups $A_d(X)$, as developed in [**F**]. In fact, there is a surjective map

$$A_d(X) \to A'_d(X) \tag{20.3.4.2}$$

[**F**, Examp. 15.1.5], and this map is an isomorphism once tensored with $\mathbb{Q}$ [**F**, Thms. 18.2 and 18.3]. This is the beginning of a long and rich story in algebraic geometry.

20.3.5. *Side Remark.* The surjection map (20.3.4.2) need not be an isomorphism, see [**SGA6**, Exp. XIV, §4.5-4.7] (although its kernel must be torsion, as described above). As a tantalizing example, if X is the group E_8, the kernel has 2-torsion, 3-torsion, and 5-torsion, see [**KN**, **DZ**].

20.4 ⋆⋆ The Nakai-Moishezon and Kleiman criteria for ampleness

Exercise 20.1.K stated that if X is projective k-variety, and $\mathscr{L}$ is an ample line bundle on X, then for any subvariety Y of X of dimension n, $(\mathscr{L}^n \cdot Y) > 0$. The Nakai-Moishezon criterion states that this is a characterization:

20.4.1. Theorem (Nakai-Moishezon criterion for ampleness). — *If $\mathscr{L}$ is an invertible sheaf on a projective k-scheme X, and for every subvariety Y of X of dimension n, $(\mathscr{L}^n \cdot Y) > 0$, then $\mathscr{L}$ is ample.*

20.4.2. *Remarks.* We note that X need only be proper for this result to hold ([**Kl1**, Thm. III.1.1]).

Before proving the Nakai-Moishezon criterion, we point out some consequences related to our discussion of numerical equivalence in §18.4.9. By Proposition 20.1.4, $(\mathscr{L}^n \cdot Y)$ depends only on the numerical equivalence class of $\mathscr{L}$, so ampleness is a numerical property. As a result, the notion of ampleness makes sense on $N^1_{\mathbb{Q}}(X)$. As the tensor product of two ample invertible sheaves is ample (Exercise 16.6.H), the ample $\mathbb{Q}$-line bundles in $N^1_{\mathbb{Q}}(X)$ form a cone, called the **ample cone** of X.

20.4.3. Proposition. — *If X is a projective k-scheme, the ample cone is open.*

20.4.4. In the rest of this section, we often use additive notation for the tensor product of invertible sheaves, as described in Remark 20.1.8. This is because we want to deal with intersections on the $\mathbb{Q}$-vector space $N^1_{\mathbb{Q}}(X)$. For example by $((a\mathscr{L}_1 + b\mathscr{L}_1') \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$ $(a, b \in \mathbb{Q})$, we mean $a(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F}) + b(\mathscr{L}_1' \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})$.

Proof. Suppose $\mathscr{A}$ is an ample invertible sheaf on X. We will describe a small open neighborhood of $[\mathscr{A}]$ in $N^1_{\mathbb{Q}}(X)$ consisting of ample $\mathbb{Q}$-line bundles. Choose invertible sheaves $\mathscr{L}_1, \dots, \mathscr{L}_n$ on X whose classes form a basis of $N^1_{\mathbb{Q}}(X)$. By Exercise 16.6.E, there is some m such that $\mathscr{A}^{\otimes m} \otimes \mathscr{L}_i$ and $\mathscr{A}^{\otimes m} \otimes \mathscr{L}_i^\vee$ are both very ample for all n. Thus (in the additive notation of § 20.4.4), $\mathscr{A} + \frac{1}{m}\mathscr{L}_i$ and $\mathscr{A} - \frac{1}{m}\mathscr{L}_i$ are both ample. As the ample $\mathbb{Q}$-line bundles form a cone, it follows that $\mathscr{A} + \epsilon_1\mathscr{L}_1 + \cdots + \epsilon_n\mathscr{L}_n$ is ample for $\sum_i |\epsilon_i| \le 1/m$. $\square$

20.4.5. *Proof of the Nakai-Moishezon criterion, Theorem 20.4.1.* We prove the Nakai-Moishezon criterion in several steps.

20.4.A. UNIMPORTANT EXERCISE. Prove the case where $\dim X = 0$.

Step 1: initial reductions. Suppose $\mathscr{L}$ satisfies the hypotheses of the Theorem; we wish to show that $\mathscr{L}$ is ample. By Exercises 18.7.A and 18.7.B, we may assume that X is integral. Moreover, we can work by induction on dimension, so we can assume that $\mathscr{L}$ is ample on any closed subvariety. The base case is dimension 1, which was done in Exercise 19.2.E.

Step 2: sufficiently high powers of $\mathscr{L}$ have sections. We show that $H^0(X, \mathscr{L}^{\otimes m}) \neq 0$ for $m \gg 0$.

Our plan is as follows. By Asymptotic Riemann-Roch (Exercise 20.1.I), $\chi(X, \mathscr{L}^{\otimes m}) = m^n(\mathscr{L}^n)/n! + \cdots$ grows (as a function of m) without bound. A plausible means of attack is to show that $h^i(X, \mathscr{L}^{\otimes m}) = 0$ for $i > 0$ and $m \gg 0$. We won't do that, but will do something similar.

By Exercise 16.6.C, $\mathscr{L}$ is the difference of two very ample line bundles, say $\mathscr{L} \cong \mathscr{A} \otimes \mathscr{B}^{-1}$ with $\mathscr{A} = \mathscr{O}(A)$ and $\mathscr{B} = \mathscr{O}(B)$. From $0 \to \mathscr{O}(-A) \to \mathscr{O} \to \mathscr{O}_A \to 0$ we have

$$0 \to \mathscr{L}^{\otimes m}(-B) \to \mathscr{L}^{\otimes(m+1)} \to \mathscr{L}^{\otimes(m+1)}|_A \to 0. \tag{20.4.5.1}$$

From $0 \to \mathscr{O}(-B) \to \mathscr{O} \to \mathscr{O}_B \to 0$, we have

$$0 \to \mathscr{L}^{\otimes m}(-B) \to \mathscr{L}^{\otimes m} \to \mathscr{L}^{\otimes m}|_B \to 0. \tag{20.4.5.2}$$

Choose m large enough so that both $\mathscr{L}^{\otimes(m+1)}|_A$ and $\mathscr{L}^{\otimes m}|_B$ have vanishing higher cohomology (i.e., $h^{>0} = 0$ for both; use the inductive hypothesis that $\mathscr{L}$ is ample on all proper closed subvarieties, and Serre vanishing, Theorem 18.1.4(ii)). This implies that for $i \geq 2$,

$$\begin{aligned} H^i(X, \mathscr{L}^{\otimes m}) &\cong H^i(X, \mathscr{L}^{\otimes m}(-B)) \quad \text{(long exact sequence for (20.4.5.2))} \\ &\cong H^i(X, \mathscr{L}^{\otimes m+1}) \quad \text{(long exact sequence for (20.4.5.1))} \end{aligned}$$

so the higher cohomology stabilizes (is constant) for large m. From

$$\chi(X, \mathscr{L}^{\otimes m}) = h^0(X, \mathscr{L}^{\otimes m}) - h^1(X, \mathscr{L}^{\otimes m}) + \text{constant},$$

$H^0(\mathscr{L}^{\otimes m}) \neq 0$ for $m \gg 0$, completing Step 2.

So by replacing $\mathscr{L}$ by a suitably large multiple (ampleness is independent of taking tensor powers, Theorem 16.6.2), we may assume $\mathscr{L}$ has a section D. We now use D as a crutch.

Step 3: $\mathscr{L}^{\otimes m}$ is globally generated for $m \gg 0$.

As D is effective, $\mathscr{L}^{\otimes m}$ is globally generated on the complement of D: we have a section nonvanishing on that big open set. Thus any base locus must be contained in D. Consider the short exact sequence

$$0 \to \mathscr{L}^{\otimes(m-1)} \to \mathscr{L}^{\otimes m} \to \mathscr{L}^{\otimes m}|_D \to 0 \tag{20.4.5.3}$$

Now $\mathscr{L}|_D$ is ample by our inductive hypothesis. Choose m so large that $H^1(X, \mathscr{L}^{\otimes m}|_D) = 0$ (Serre vanishing, Theorem 18.1.4(b)). From the exact sequence associated to (20.4.5.3),

$$\phi_m : H^1(X, \mathscr{L}^{\otimes(m-1)}) \to H^1(X, \mathscr{L}^{\otimes m})$$

is surjective for $m \gg 0$. Using the fact that the $H^1(X, \mathscr{L}^{\otimes m})$ are finite-dimensional vector spaces, as m grows, $H^1(X, \mathscr{L}^{\otimes m})$ must eventually stabilize, so the ϕ_m are isomorphisms for $m \gg 0$.

Thus for large m, from the long exact sequence in cohomology for (20.4.5.3), $H^0(X, \mathscr{L}^{\otimes m}) \to H^0(X, \mathscr{L}^{\otimes m}|_D)$ is surjective for $m \gg 0$. But $H^0(X, \mathscr{L}^{\otimes m}|_D)$ has no base points by our inductive hypothesis (applied to D), i.e., for any point p of D there is a section of $\mathscr{L}^{\otimes m}|_D$ not vanishing at p, so $H^0(X, \mathscr{L}^{\otimes m})$ has no base points on D either, completing Step 3.

Step 4. Thus $\mathscr{L}$ is a base-point-free line bundle with positive degree on each curve (by hypothesis of Theorem 20.4.1), so by Exercise 18.4.L we are done. □

The following result is the key to proving Kleiman's numerical criterion of ampleness, Theorem 20.4.7.

20.4.6. Kleiman's Theorem. — *Suppose X is a projective k-scheme. If $\mathscr{L}$ is a nef invertible sheaf on X, then $(\mathscr{L}^k \cdot V) \geq 0$ for every irreducible subvariety $V \subset X$ of dimension k.*

As usual, this extends to the proper case, see [**Kl1**, Thm. IV.2.1]. And as usual, we postpone the proof until after we appreciate the consequences.

20.4.B. EXERCISE.
(a) *(limit of amples is nef)* If $\mathscr{L}$ and $\mathscr{H}$ are any two invertible sheaves such that

$\mathscr{L} + \epsilon\mathscr{H}$ is ample for all sufficiently small $\epsilon > 0$, show that $\mathscr{L}$ is nef. (Hint: $\lim_{\epsilon \to 0}$. This doesn't require Kleiman's Theorem.)
(b) *(nef + ample = ample)* Suppose X is a projective k-scheme, $\mathscr{H}$ is ample, and $\mathscr{L}$ is nef. Show that $\mathscr{L} + \epsilon\mathscr{H}$ is ample for all $\epsilon \in \mathbb{Q}^{\geq 0}$. (Hint: use Nakai-Moishezon: $((\mathscr{L} + \epsilon\mathscr{H})^k \cdot V) > 0$. This may help you appreciate the additive notation.)

20.4.7. Theorem (Kleiman's numerical criterion for ampleness). — *Suppose X is a projective k-scheme.*

(a) *The nef cone is the closure of the ample cone.*
(b) *The ample cone is the interior of the nef cone.*

Side remark: of course (a) is false if "projective" is relaxed to "proper" (as the ample cone is empty, but 0 is nef so the nef cone is nonempty). However, part (b) is true if X is proper and factorial (see [**Kl1**, Thm. IV.2.2] for a proof and a more general statement). Hence if X is smooth, proper, and nonprojective, then the interior of the nef cone is empty.

20.4.C. EXERCISE. See [**Kl1**, p. 326, Ex. 2] for Mumford's example of a non-ample line bundle on a smooth projective surface that meets every curve positively. Doesn't this contradict Theorem 20.4.7?

Proof. (a) Ample invertible sheaves are nef (Exercise 18.4.V(e)), and the nef cone is closed (Exercise 18.4.W), so the closure of the ample cone is contained in the cone. Conversely, each nef element of $N^1_{\mathbb{Q}}(X)$ is the limit of ample classes by Exercise 20.4.B(a), so the nef cone is contained in the closure of the ample cone.

(b) As the ample cone is open (Proposition 20.4.3), the ample cone is contained in the interior of the nef cone. Conversely, suppose $\mathscr{L}$ is in the interior of the nef cone, and $\mathscr{H}$ is any ample class. Then $\mathscr{L} - \epsilon\mathscr{H}$ is nef for all small enough positive ϵ. Then by Exercise 20.4.B(b), $\mathscr{L} = (\mathscr{L} - \epsilon\mathscr{H}) + \epsilon\mathscr{H}$ is ample. □

Suitably motivated, we prove Kleiman's Theorem 20.4.6.

Proof. We may immediately reduce to the case where X is irreducible and reduced. We work by induction on $n := \dim X$. The base case $n = 1$ is obvious. So we assume that $(\mathscr{L}^{\dim V} \cdot V) \geq 0$ for all irreducible V not equal to X. We need only show that $(\mathscr{L}^n \cdot X) \geq 0$.

Fix some very ample invertible sheaf $\mathscr{H}$ on X.

20.4.D. EXERCISE. Show that $(\mathscr{L}^k \cdot \mathscr{H}^{n-k} \cdot X) \geq 0$ for all $k < n$. (Hint: use the inductive hypothesis).

Consider $P(t) := ((\mathscr{L} + t\mathscr{H})^n \cdot X) \in \mathbb{Q}[t]$. We wish to show that $P(0) \geq 0$. Assume otherwise that $P(0) < 0$. Now for $t \gg 0$, $\mathscr{L} + t\mathscr{H}$ is ample, so $P(t)$ is positive for large t. Thus $P(t)$ has positive real roots. Let t_0 be the largest positive real root of t. (In fact there is only one positive root, as Exercise 20.4.D shows that all the nonconstant coefficients of $P(t)$ are nonnegative.

20.4.E. EXERCISE. Show that for (rational) $t > t_0$, $\mathscr{L} + t\mathscr{H}$ is ample. Hint: use the Nakai-Moishezon criterion (Theorem 20.4.1); if $V \neq X$ is an irreducible subvariety, show that $((\mathscr{L} + t\mathscr{H})^{\dim V} \cdot V) > 0$ by expanding $(\mathscr{L} + t\mathscr{H})^{\dim V}$.

Let $Q(t) := (\mathscr{L} \cdot (\mathscr{L} + t\mathscr{H})^{n-1} \cdot X)$ and $R(t) := (t\mathscr{H} \cdot (\mathscr{L} + t\mathscr{H})^{n-1} \cdot X)$, so $P(t) = Q(t) + R(t)$.

20.4.F. EXERCISE. Show that $Q(t) \geq 0$ for all rational $t > t_0$. Hint (which you will have to make sense of): $(\mathscr{L} + t\mathscr{H})$ is ample by Exercise 20.4.E, so for N sufficiently large, $N(\mathscr{L} + t\mathscr{H})$ is very ample. Use the idea of the proof of Proposition 20.1.4 to intersect X with $n-1$ divisors in the class of $N(\mathscr{L} + t\mathscr{H})$ so that "$((N(\mathscr{L} + t\mathscr{H}))^{n-1} \cdot X)$ is an effective curve C". Then $(\mathscr{L} \cdot C) \geq 0$ as $\mathscr{L}$ is nef. Thus $Q(t_0) \geq 0$.

20.4.G. EXERCISE. Show that $R(t_0) > 0$. (Hint: Exercise 20.4.D.)

Thus $P(t_0) > 0$ as desired. $\square$

CHAPTER 21

Differentials

21.1 Motivation and game plan

Differentials are an intuitive geometric notion, and we are going to figure out the right description of them algebraically. The algebraic manifestation is somewhat non-intuitive, so it is helpful to understand differentials first in terms of geometry. Also, although the algebraic statements are odd, none of the proofs are hard or long. You will notice that this topic could have been done as soon as we knew about morphisms and quasicoherent sheaves. We have usually introduced new ideas through a number of examples, but in this case we will spend a fair amount of time discussing theory, and only then get to examples.

Suppose X is a "smooth" k-variety. We would like to define a tangent bundle. We will see that the right way to do this will easily apply in much more general circumstances.

• We will see that cotangent is more "natural" for schemes than tangent bundle. This is similar to the fact that the Zariski *cotangent* space is more natural than the *tangent space* (i.e., if A is a ring and $\mathfrak{m}$ is a maximal ideal, then $\mathfrak{m}/\mathfrak{m}^2$ is "more natural" than $(\mathfrak{m}/\mathfrak{m}^2)^\vee$), as we have repeatedly discussed since §12.1. In both cases this is because we are understanding "spaces" via their (sheaf of) functions on them, which is somehow dual to the geometric pictures you have of spaces in your mind.

So we will define the cotangent sheaf first. An element of the (co)tangent space will be called a **(co)tangent vector**.

• Our construction will automatically apply for general X, even if X is not "smooth" (or even at all nice, e.g. finite type). The cotangent sheaf will not be locally free, but it will still be a quasicoherent sheaf.

• Better yet, this construction will naturally work "relatively". For *any* $\pi : X \to Y$, we will define $\Omega_\pi = \Omega_{X/Y}$, a quasicoherent sheaf on X, the sheaf of *relative differentials*. The fiber of this sheaf at a point will be the cotangent vectors of the fiber of the map. This will specialize to the earlier case by taking $Y = \operatorname{Spec} k$. The idea is that this glues together the cotangent sheaves of the fibers of the family. Figure 21.1 is a sketch of the relative tangent space of a map $X \to Y$ at a point $p \in X$ — it is the tangent to the fiber. (The tangent space is easier to draw than the cotangent space!) An element of the relative (co)tangent space is called a **vertical** or **relative (co)tangent vector**.

Thus the central concept of this chapter is the cotangent sheaf $\Omega_\pi = \Omega_{X/Y}$ for a morphism $\pi : X \to Y$ of schemes. A good picture to have in your mind is the following. If $\pi : X \to Y$ is a submersion of manifolds (a map inducing a

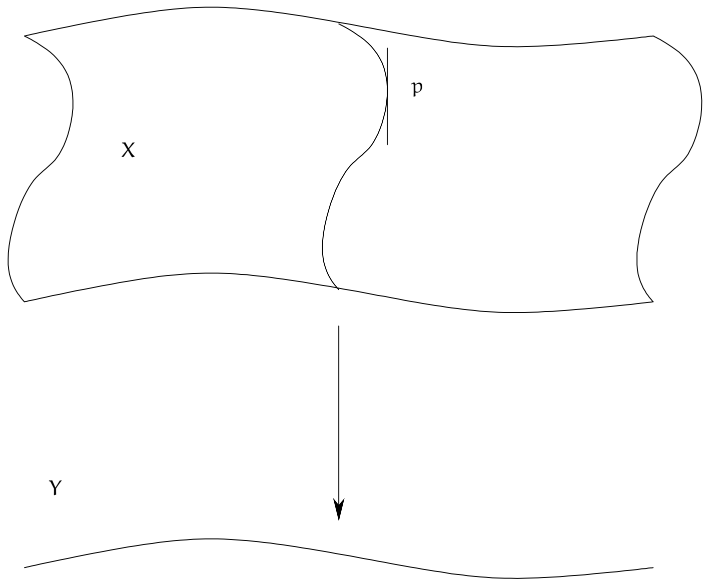


FIGURE 21.1. The relative tangent space of a morphism $X \to Y$ at a point p

surjection on tangent spaces), you might hope that the tangent spaces to the fibers at each point $p \in X$ might fit together to form a vector bundle. This is the relative tangent bundle (of π), and its dual is $\Omega_{X/Y}$ (see Figure 21.1). Even if you are not geometrically minded, you will find this useful. (For an arithmetic example, see Exercise 21.2.F.)

21.2 Definitions and first properties

21.2.1. The affine case: three definitions.

We first study the affine case. Suppose A is a B-algebra, so we have a morphism of rings $\phi : B \to A$ and a morphism of schemes $\operatorname{Spec} A \to \operatorname{Spec} B$. I will define an A-module $\Omega_{A/B}$ in three ways. This is called the **module of relative differentials** or the **module of Kähler differentials**. The module of differentials will be defined to be this module, as well as a map $d : A \to \Omega_{A/B}$ satisfying three properties. (Caution: although d sends an A-module to an A-module, it is not in general A-linear. A priori we take it as a homomorphism of abelian groups, but we will momentarily make it a homomorphism of B-modules, Exercise 21.2.A.)

(i) *additivity:* $da + da' = d(a + a')$.
(ii) *Leibniz:* $d(aa') = a\, da' + a'\, da$.
(iii) *triviality on pullbacks:* $db = 0$ for $b \in \phi(B)$.

These properties will not be surprising if you have seen differentials in any other context.

21.2.A. TRIVIAL EXERCISE. Show that d is B-linear.

21.2.B. EXERCISE. Prove the quotient rule: if $b = as$, then $da = (s\ db - b\ ds)/s^2$.

21.2.C. EXERCISE. State and prove the chain rule for $d(f(g))$ where f is a polynomial with B-coefficients, and $g \in A$. (As motivation, think of the case $B = k$. So for example, $da^n = na^{n-1}da$, and more generally, if f is a polynomial in one variable, $df(a) = f'(a)\ da$, where f' is defined formally: if $f = \sum c_i x^i$ then $f' = \sum c_i i x^{i-1}$.)

We will now see three definitions of the module of Kähler differentials, which will soon "sheafify" to the sheaf of relative differentials. The first definition is a concrete hands-on definition. The second is by universal property. And the third will globalize well, and will allow us to define $\Omega_{X/Y}$ conveniently in general.

21.2.2. First definition of differentials: explicit description. We define $\Omega_{A/B}$ to be finite A-linear combinations of symbols "da" for $a \in A$, subject to the three rules (i)–(iii) above. For example, take $A = k[x, y]$, $B = k$. Then a sample differential is $3x^2\ dy + 4\ dx \in \Omega_{A/B}$. We have identities such as $d(3xy^2) = 3y^2\ dx + 6xy\ dy$.

21.2.3. Key fact. Note that if A is generated over B (as an algebra) by $x_i \in A$ (where i lies in some index set, possibly infinite), subject to some relations r_j (where j lies in some index set, and each is a polynomial in the x_i), then the A-module $\Omega_{A/B}$ is generated by the dx_i, subject to the relations (i)–(iii) and $dr_j = 0$. In short, we needn't take every single element of A; we can take a generating set. And we needn't take every single relation among these generating elements; we can take generators of the relations.

21.2.D. EXERCISE. Verify Key fact 21.2.3. (If you wish, use the affine conormal exact sequence, Theorem 21.2.12, to verify it; different people prefer to work through the theory in different orders. Just take care not to make circular arguments.)

In particular:

21.2.4. Proposition. — *If A is a finitely generated B-algebra, then $\Omega_{A/B}$ is a finite type (i.e., finitely generated) A-module. If A is a finitely presented B-algebra, then $\Omega_{A/B}$ is a finitely presented A-module.*

Recall (§7.3.17) that an algebra A is *finitely presented* over another algebra B if it can be expressed with finite number of generators and finite number of relations:

$$A = B[x_1, \dots, x_n]/(r_1(x_1, \dots, x_n), \dots, r_j(x_1, \dots, x_n)).$$

If A is Noetherian, then finitely presented is the same as finite type, as the "finite number of relations" comes for free, so most of you will not care.

Let's now see some examples. Among these examples are three particularly important building blocks for ring maps: adding free variables; localizing; and taking quotients. If we know how to deal with these, we know (at least in theory) how to deal with any ring map. (They were similarly useful in understanding the fibered product in practice, in §9.2.)

21.2.5. Example: taking a quotient. If $A = B/I$, then $\Omega_{A/B} = 0$: $da = 0$ for all $a \in A$, as each such a is the image of an element of B. This should be believable; in this case, there are no "vertical tangent vectors".

21.2.6. Example: adding variables. If $A = B[x_1, \dots, x_n]$, then $\Omega_{A/B} = A dx_1 \oplus \cdots \oplus A dx_n$. (Note that this argument applies even if we add an arbitrarily infinite number of indeterminates.) The intuitive geometry behind this makes the answer very reasonable. The cotangent bundle of affine n-space should indeed be free of rank n.

21.2.7. Explicit example: an affine plane curve. Consider the plane curve $y^2 = x^3 - x$ in $\mathbb{A}^2_k$, where the characteristic of k is not 2. Let $A = k[x,y]/(y^2 - x^3 + x)$ and $B = k$. By Key fact 21.2.3, the module of differentials $\Omega_{A/B}$ is generated by dx and dy, subject to the relation

$$2y\ dy = (3x^2 - 1)\ dx.$$

Thus in the locus where $y \neq 0$, dx is a generator (as dy can be expressed in terms of dx). We conclude that where $y \neq 0$, $\widetilde{\Omega_{A/B}}$ is isomorphic to the trivial line bundle (invertible sheaf). Similarly, in the locus where $3x^2 - 1 \neq 0$, dy is a generator. These two loci cover the entire curve, as solving $y = 0$ gives $x^3 - x = 0$, i.e., $x = 0$ or ± 1, and in each of these cases $3x^2 - 1 \neq 0$. We have shown that $\widetilde{\Omega_{A/B}}$ is an invertible sheaf.

We can interpret dx and dy geometrically. Where does the differential dx vanish? The previous paragraph shows that it doesn't vanish on the patch where $2y \neq 0$. On the patch where $3x^2 - 1 \neq 0$, where dy is a generator, $dx = (2y/(3x^2 - 1))\ dy$ from which we see that dx vanishes precisely where $y = 0$. You should find this believable from the picture. We have shown that $dx = 0$ precisely where the curve has a vertical tangent vector (see Figure 19.4 for a picture). Once we can pull back differentials (Exercise 21.2.K(a) or Theorem 21.2.27), we can interpret dx as the pullback of a differential on the x-axis to $\operatorname{Spec} A$ (pulling back along the projection to the x-axis). When we do that, using the fact that dx doesn't vanish on the x-axis, we can interpret the locus where $dx = 0$ as the locus where the projection map branches. (Can you compute where $dy = 0$, and interpret it geometrically?)

This discussion applies to plane curves more generally. Suppose $A = k[x,y]/f(x,y)$, where for convenience $k = \overline{k}$. Then the same argument as the one given above shows that $\widetilde{\Omega_{A/k}}$ is free of rank 1 on the open set $D(\partial f/\partial x)$, and also on $D(\partial f/\partial y)$. If $\operatorname{Spec} A$ is a regular curve, then these two sets cover all of $\operatorname{Spec} A$. (Exercise 12.2.D — basically the Jacobian criterion — gives regularity at the closed point. Furthermore, the curve must be reduced, or else as the nonreduced locus is closed, it would be nonreduced at a closed point, contradicting regularity. Finally, reducedness at a generic point is equivalent to regularity (a scheme whose underlying set is a point is reduced if and only if it is regular). Alternatively, we could invoke a big result, Fact 12.8.2, to get regularity at the generic point from regularity at the closed points.)

Conversely, if the plane curve is singular, then Ω is *not* locally free of rank one. For example, consider the plane curve $\operatorname{Spec} A$ where $A = \mathbb{C}[x,y]/(y^2 - x^3)$, so

$$\Omega_{A/\mathbb{C}} = (A\ dx \oplus A\ dy)/(2y\ dy - 3x^2\ dx).$$

Then the fiber of $\Omega_{A/\mathbb{C}}$ over the origin (computed by setting $x = y = 0$) is rank 2, as it is generated by dx and dy, with no relation.

Implicit in the above discussion is the following exercise, showing that Ω can be computed using the Jacobian matrix.

21.2.E. IMPORTANT BUT EASY EXERCISE (JACOBIAN DESCRIPTION OF $\Omega_{A/B}$). Suppose $A = B[x_1, \dots, x_n]/(f_1, \dots, f_r)$. Then $\Omega_{A/B} = \{\oplus_i A\, dx_i\}/\{df_j = 0\}$ may be interpreted as the cokernel of the Jacobian matrix (12.1.6.1)

$$J : A^{\oplus r} \to A^{\oplus n}.$$

21.2.8. Example: localization. If T is a multiplicative subset of B, and $A = T^{-1}B$, then $\Omega_{A/B} = 0$. Reason: by the quotient rule (Exercise 21.2.B), if $a = b/t$, then $da = (t\ db - b\ dt)/t^2 = 0$. If $A = B_f$, this is intuitively believable; then $\operatorname{Spec} A$ is an open subset of $\operatorname{Spec} B$, so there should be no vertical (co)tangent vectors.

21.2.F. IMPORTANT EXERCISE (FIELD EXTENSIONS). This notion of relative differentials is interesting even for finite extensions of fields. In other words, even when you map a reduced point to a reduced point, something interesting can happen with differentials.
(a) Suppose K/k is a separable algebraic extension. Show that $\Omega_{K/k} = 0$. Do not assume that K/k is a finite extension! (Hint: for any $\alpha \in K$, there is a polynomial $f(x)$ such that $f(\alpha) = 0$ and $f'(\alpha) \neq 0$.)
(b) Suppose k is a field of characteristic p, $K = k(t^p)$, $L = k(t)$. Compute $\Omega_{L/K}$ (where $K \hookrightarrow L$ is the "obvious" inclusion).

We now delve a little deeper, and discuss two useful and geometrically motivated exact sequences.

21.2.9. Theorem (relative cotangent sequence, affine version). — *Suppose $C \to B \to A$ are ring morphisms. Then there is a natural exact sequence of A-modules*

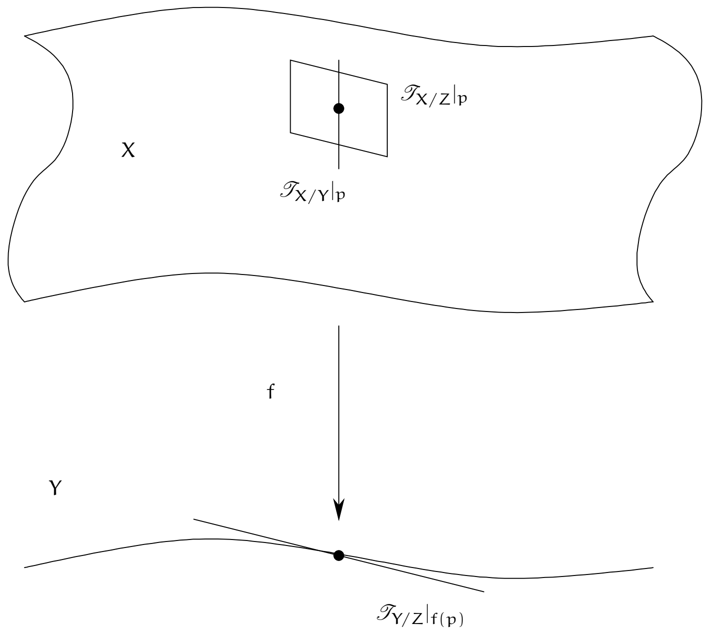
$$A \otimes_B \Omega_{B/C} \xrightarrow{a \otimes db \mapsto a\ db} \Omega_{A/C} \xrightarrow{da \mapsto da} \Omega_{A/B} \longrightarrow 0.$$

The proof will be quite straightforward algebraically, but the statement comes fundamentally from geometry, and that is how best to remember it. Figure 21.2 is a sketch of a map $\pi : X \to Y$. Here X should be interpreted as $\operatorname{Spec} A$, Y as $\operatorname{Spec} B$, and $\operatorname{Spec} C$ is a point. (If you would like a picture with a higher-dimensional $\operatorname{Spec} C$, just "take the product of Figure 21.2 with a curve".) In the Figure, Y is "smooth", and X is "smooth over Y" — which means roughly that all the fibers are smooth. Suppose p is a point of X. Then the tangent space of the fiber of π at p is certainly a subspace of the tangent space of the total space of X at p. The cokernel is naturally the pullback of the tangent space of Y at $\pi(p)$. This short exact sequence for each p should be part of a short exact sequence of "relative tangent sheaves"

$$0 \to \mathscr{T}_{X/Y} \to \mathscr{T}_{X/Z} \to \pi^* \mathscr{T}_{Y/Z} \to 0$$

on X. (We will formally define "relative tangent sheaf" in §21.2.20.) Dualizing this yields

$$0 \to \pi^* \Omega_{Y/Z} \to \Omega_{X/Z} \to \Omega_{X/Y} \to 0.$$



[Change f to π. Add p and $\pi(p)$.]

FIGURE 21.2. A sketch of the geometry behind the relative cotangent sequence

This is precisely the statement of Theorem 21.2.9, except we also have left-exactness. This discrepancy is because the statement of the theorem is more general; we will see in Exercise 21.2.S that in the "smooth" case, we indeed have left-exactness.

21.2.10. *Intriguing Remark.* As always, whenever you see something right-exact, you should suspect that there should be some sort of (co)homology theory so that this is the end of a long exact sequence. This is indeed the case, and this exact sequence involves *André-Quillen homology* (see [**E**, p. 386] for more). You should expect that the next term to the left should be the first homology corresponding to A/B, and in particular shouldn't involve C. So if you already suspect that you have exactness on the left in the case where A/B and B/C are "smooth" (whatever that means), and the intuition of Figure 21.2 applies, then you should expect further that all that is necessary is that A/B be "smooth", and that this would imply that the first André-Quillen homology should be zero. Even though you wouldn't precisely know what all the words meant, you would be completely correct! You would also be developing a vague inkling about the *cotangent complex*. We will see examples when left-exactness holds in a sufficiently "smooth" situation in Proposition 21.7.2 and Exercise 21.2.S. For a more general statement, see [**Stacks**, tag 06B6]. See also [**Liu**, Cor. 3.6.22] and [**Liu**, Prop. 3.6.11] for clean discussions of the next two terms in the sequence: "relative cotangent sequence", "regular conormal sequence", ...

21.2.11. *Proof of Theorem 21.2.9 (the relative cotangent sequence, affine version).* First, note that surjectivity of $\Omega_{A/C} \to \Omega_{A/B}$ is clear, as this map is given by $da \mapsto da$ (where $a \in A$).

Next, the composition over the middle term is clearly 0, as this composition is given by $a \otimes db \mapsto adb \mapsto 0$.

Finally, we wish to identify $\Omega_{A/B}$ as the cokernel of $A \otimes_B \Omega_{B/C} \to \Omega_{A/C}$. Now $\Omega_{A/B}$ is exactly the same as $\Omega_{A/C}$, except we have extra relations: $db = 0$ for $b \in B$. These are precisely the images of $1 \otimes db$ on the left. □

21.2.12. Theorem (conormal exact sequence, affine version). — *Suppose B is a C-algebra, I is an ideal of B, and $A = B/I$. Then there is a natural exact sequence of A-modules*

$$I/I^2 \xrightarrow{\delta} A \otimes_B \Omega_{B/C} \xrightarrow{a \otimes db \mapsto a\, db} \Omega_{A/C} \longrightarrow 0.$$

Here δ is, informally, $i \mapsto 1 \otimes di$, or more formally, $1 \otimes d : B/I \otimes_B I \to B/I \otimes_B \Omega_{B/C}$.

(You will recognize the map $A \otimes_B \Omega_{B/C} \to \Omega_{A/C}$ from the relative cotangent sequence, Theorem 21.2.9.) The proof is algebraic, so the geometric discussion thereafter may help clarify how you should really think of it.

Proof. We will identify the cokernel of $\delta : I/I^2 \to A \otimes_B \Omega_{B/C}$ with $\Omega_{A/C}$. Consider $A \otimes_B \Omega_{B/C}$. As an A-module, it is generated by db (where $b \in B$), subject to three relations: $dc = 0$ for $c \in \phi(C)$ (where $\phi : C \to B$ describes B as a C-algebra), additivity, and the Leibniz rule. Given any relation *in* B, d of that relation is 0.

Now $\Omega_{A/C}$ is defined similarly, except there are more relations *in* A; these are precisely the elements of $I \subset B$. Thus we obtain $\Omega_{A/C}$ by starting out with $A \otimes_B \Omega_{B/C}$, and adding the additional relations di where $i \in I$. But this is precisely the image of δ! (Be sure that you see how the identification of the cokernel of δ with $\Omega_{A/C}$ is precisely via the map $a \otimes db \mapsto a\, db$.) □

We now give a geometric interpretation of the conormal exact sequence, and in particular define conormal modules/sheaves/bundles.

As with the relative cotangent sequence (Theorem 21.2.9), the conormal exact sequence is fundamentally about geometry. To motivate it, consider the sketch of Figure 21.3. In the sketch, everything is "smooth", X is one-dimensional, Y is two-dimensional, j is the inclusion $j : X \hookrightarrow Y$, and Z is a point. Then at a point $p \in X$, the tangent space $\mathscr{T}_X|_p$ clearly injects into the tangent space of $j(p)$ in Y, and the cokernel is the normal vector space to X in Y at p. This should give an exact sequence of bundles on X:

$$0 \to \mathscr{T}_X \to j^*\mathscr{T}_Y \to \mathscr{N}_{X/Y} \to 0.$$

Dualizing this should give

$$0 \to \mathscr{N}^\vee_{X/Y} \to j^*\Omega_{Y/Z} \to \Omega_{X/Z} \to 0.$$

This is precisely what appears in the statement of the Theorem, except (i) the exact sequence in algebraic geometry is not necessarily exact on the left, and (ii) we see I/I^2 instead of $\mathscr{N}^\vee_{\operatorname{Spec} A/\operatorname{Spec} B}$.

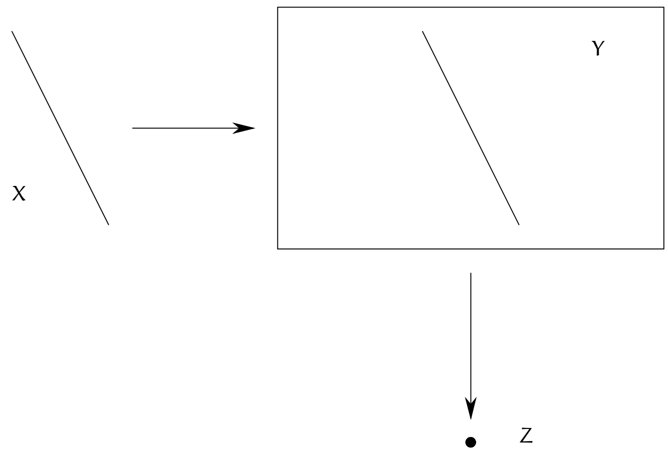


FIGURE 21.3. A sketch of the geometry behind the conormal exact sequence

21.2.13. We resolve the first issue (i) by expecting that the sequence of Theorem 21.2.12 is exact on the left in appropriately "smooth" situations, and this is indeed the case, see Theorem 21.3.8 and Remark 21.3.11. (If you enjoyed Remark 21.2.10, you might correctly guess several things. The next term on the left should be the André-Quillen homology of A/C, so we should only need that A/C is smooth, and B should be irrelevant. Also, if $A = B/I$, then we should expect that I/I^2 is the first André-Quillen homology of A/B.)

21.2.14. *Conormal modules and conormal sheaves.* We resolve the second issue (ii) by *declaring* I/I^2 to be the **conormal module**, and in Definition 21.2.15 we will define the obvious analog as the *conormal sheaf*.

Here is some geometric intuition as to why we might want to call (the sheaf associated to) I/I^2 the conormal sheaf, which will likely confuse you, but may offer some enlightenment. First, if $\operatorname{Spec} A$ is a closed point of $\operatorname{Spec} B$, we expect the conormal space to be precisely the cotangent space. And indeed if $A = B/\mathfrak{m}$, the Zariski cotangent space is $\mathfrak{m}/\mathfrak{m}^2$. (We made this subtle connection in §12.1.) In particular, at some point you will develop a sense of why the conormal (=cotangent) space to the origin in $\mathbb{A}^2_k = \operatorname{Spec} k[x, y]$ is naturally the space of linear forms $\alpha x + \beta y$. But then consider the z-axis in $\operatorname{Spec} k[x, y, z] = \mathbb{A}^3_k$, cut out by $I = (x, y)$. Elements of I/I^2 may be written as $\alpha(z)x + \beta(z)y$, where $\alpha(z)$ and $\beta(z)$ are polynomial. This reasonably should be the conormal space to the z-axis: as z varies, the coefficients of x and y vary. More generally, the same idea suggests that the conormal module/sheaf to any coordinate k-plane inside n-space corresponds to I/I^2. Now consider a k-dimensional (smooth or differential real) manifold X inside an n-dimensional manifold Y, with the classical topology. We can apply the same construction: if $\mathscr{I}$ is the ideal sheaf of X in Y, then $\mathscr{I}/\mathscr{I}^2$ can be identified with the conormal sheaf (essentially the conormal vector bundle), because analytically locally $X \hookrightarrow Y$ can be identified with $\mathbb{R}^k \hookrightarrow \mathbb{R}^n$. For this reason, you might hope that in algebraic geometry, if $\operatorname{Spec} A \hookrightarrow \operatorname{Spec} B$ is an inclusion of something "smooth" in something "smooth", I/I^2 should be the conormal module (or, after applying the functor $\sim$, the conormal sheaf). Motivated by this, we define the conormal module as I/I^2 *always*, and then notice that it has good properties (such as Theorem 21.2.12), but take care to learn what unexpected behavior it might have when we are not in the "smooth" situation, by working out examples such as that of §21.2.7.

21.2.15. *Definition.* Suppose $i : X \hookrightarrow Y$ is a closed embedding of schemes cut out by ideal sheaf $\mathscr{I}$. Define the **conormal sheaf for a closed embedding** by $\mathscr{I}/\mathscr{I}^2$, denoted by $\mathscr{N}^\vee_{X/Y}$. (The product of quasicoherent ideal sheaves was defined in Exercise 14.3.D.) Important: We interpret $\mathscr{N}^\vee_{X/Y}$ as a quasicoherent sheaf on X, **not** (just) Y — be sure you understand why we may do so. (Exercise 17.1.E is one reason, but it may confuse the issue more than help.)

Define the **normal sheaf** as its dual $\mathscr{N}_{X/Y} := \mathscr{H}om(\mathscr{N}^\vee_{X/Y}, \mathscr{O}_X)$. This is imperfect notation, because it suggests that the dual of $\mathscr{N}$ is always $\mathscr{N}^\vee$. This is not always true, as for A-modules, the natural morphism from a module to its double-dual is not always an isomorphism. (Modules for which this is true are called **reflexive**, but we won't use this notion.)

21.2.G. EASY EXERCISE. Define the **conormal sheaf** $\mathscr{N}^\vee_{X/Y}$ (and hence the normal sheaf $\mathscr{N}_{X/Y}$) *for a locally closed embedding* $i : X \hookrightarrow Y$ of schemes, a quasicoherent sheaf on X. (Make sure your definition is well-defined!)

In the good situation of a regular embedding, the conormal sheaf (and hence the normal sheaf) is locally free (or more informally, a vector bundle). In particular, the dual of $\mathscr{N}$ is indeed $\mathscr{N}^\vee$. As a warm-up we deal with the important codimension 1 case.

21.2.H. EXERCISE: NORMAL BUNDLES TO EFFECTIVE CARTIER DIVISORS. Suppose $D \subset X$ is an effective Cartier divisor (§8.4.1). Show that the conormal sheaf $\mathscr{N}^\vee_{D/X}$ is $\mathscr{O}(-D)|_D$ (and in particular is an invertible sheaf), and hence that the normal sheaf is $\mathscr{O}(D)|_D$. It may be surprising that the normal sheaf should be locally free if $X \cong \mathbb{A}^2$ and D is the union of the two axes (and more generally if X is regular but D is singular), because you may be used to thinking that a "tubular neighborhood" is isomorphic to the normal bundle.

We now treat the general case.

21.2.16. Proposition. —
(a) Suppose A is Noetherian. If I is generated by a regular sequence $x_1, \dots, x_r$, then the map $\gamma : (A/I)^{\oplus r} \to I/I^2$ given by $(a_1, \dots, a_r) \mapsto a_1x_1 + \cdots + a_rx_r$ describes I/I^2 as a free module of rank r over A/I with basis $x_1, \dots, x_r$.
(b) The (co)normal sheaf of a codimension r regular embedding is locally free of rank r.

Proof. (a) Clearly γ is surjective. We now show that it is injective. It suffices to show that it is injective upon localization to all points of $\operatorname{Spec} A/I$, because injectivity is a stalk-local condition. Then in such a localization, the regular sequence in A remains regular (Exercise 8.4.D). We are thus reduced to the local question. (Caution: Make sure you see why $(I/I^2)_\mathfrak{m} = I_\mathfrak{m}/(I_\mathfrak{m})^2$.)

Consider the a_i's instead as elements of A. Suppose now that $(a_1, \dots, a_r) \in \ker \gamma$; we will show that each a_i is in I. Now because x_r is not a zerodivisor on $A/(x_1, \dots, x_{r-1})$, since $a_r x_r = 0$ in $A/(x_1, \dots, x_{r-1})$, we have $a_r \in (x_1, \dots, x_{r-1}) \subset I$. Because the role of the a_i's is symmetric by Theorem 8.4.6, we are done.

(b) follows immediately from (a). $\square$

We will soon meet a related but harder fact, that if $\mathscr{I}$ is the ideal sheaf of a regular embedding of codimension r, then $\mathscr{I}^n/\mathscr{I}^{n+1}$ is locally free, because it is $\operatorname{Sym}^n(\mathscr{I}/\mathscr{I}^2)$ (see Theorem 22.3.8).

21.2.17. Second definition: universal property. Here is a second definition that is important philosophically, by universal property. Of course, it is a characterization rather than a definition: by universal property nonsense, it shows that if the module exists (with the d map), then it is unique up to unique isomorphism, and then one still has to construct it to make sure that it exists.

Suppose A is a B-algebra, and M is a A-module. A **B-linear derivation of A into M** is a map $d : A \to M$ of B-modules (*not necessarily a map of A-modules*) satisfying the Leibniz rule: $d(fg) = f\, dg + g\, df$. As an example, suppose $B = k$, and $A = k[x]$, and $M = A$. Then d/dx is a k-linear derivation. As a second example, if $B = k$, $A = k[x]$, and $M = k$, then $(d/dx)|_0$ (the operator "evaluate the derivative at 0") is a k-linear derivation.

A third example is $d : A \to \Omega_{A/B}$, and indeed $d : A \to \Omega_{A/B}$ is the *universal B-linear derivation of A*. Precisely, the map $d : A \to \Omega_{A/B}$ is defined by the following universal property: any other B-linear derivation $d' : A \to M$ factors uniquely through d:

$$\begin{tikzcd} A \arrow[rr, "d'"] \arrow[dr, "d"'] & & M \\ & \Omega_{A/B} \arrow[ur, "f"'] & \end{tikzcd}$$

Here f is a map of A-modules. (Note again that d and d' are not necessarily maps of A-modules — they are only B-linear.) By universal property nonsense, if it exists, it is unique up to unique isomorphism. The map $d : A \to \Omega_{A/B}$ clearly satisfies this universal property, essentially by definition.

The next result connects the cotangent module $\Omega_{A/B}$ to the cotangent space at a (rational) point.

21.2.18. Proposition (the fiber of Ω at a rational point is the cotangent space). — *Suppose B is a k-algebra, and $\mathfrak{m} \subset B$ is a maximal ideal with residue field k. Then there is an isomorphism of k-vector spaces $\delta : \mathfrak{m}/\mathfrak{m}^2 \to \Omega_{B/k} \otimes_B k$ (where the k on the right is a B-module via the isomorphism $k \cong B/\mathfrak{m}$).*

Corollary 21.3.9 will give a quite different proof, and generalize it to the case where $B/\mathfrak{m}$ is a separable extension of k.

Proof. We instead show an isomorphism of dual vector spaces

$$\operatorname{Hom}_k(\Omega_{B/k} \otimes_B k, k) \to \operatorname{Hom}_k(\mathfrak{m}/\mathfrak{m}^2, k).$$

We have (canonical) isomorphisms

$$\begin{aligned}\mathrm{Hom}_k(\Omega_{B/k}\otimes_B k, k) &= \mathrm{Hom}_B(\Omega_{B/k}\otimes_B k, k)\\ &= \mathrm{Hom}_B(\Omega_{B/k}, \mathrm{Hom}_B(k,k))\\ &= \mathrm{Hom}_B(\Omega_{B/k}, \mathrm{Hom}_k(k,k))\\ &= \mathrm{Hom}_B(\Omega_{B/k}, k),\end{aligned}$$

where in the right argument of $\mathrm{Hom}_B(\Omega_{B/k}, k)$, k is a B-module via its manifestation as $B/\mathfrak{m}$. By the universal property of $\Omega_{B/k}$ (§21.2.17), $\mathrm{Hom}_B(\Omega_{B/k}, k)$ corresponds to the k-derivations of B into $B/\mathfrak{m} \cong k$.

21.2.I. EXERCISE. Show that these are precisely the elements of $\mathrm{Hom}_k(\mathfrak{m}/\mathfrak{m}^2, k)$. (The algebra involved is essentially the same as that of Exercise 12.1.A.)

□

You can verify that this δ is the one appearing in the conormal exact sequence, Theorem 21.2.12, with $I = \mathfrak{m}$ and $A = C = k$. In fact, from the conormal exact sequence, we can immediately see that δ is a surjection, as $\Omega_{k/k} = 0$.

21.2.19. *Remark.* Proposition 21.2.18, in combination with the Jacobian exercise 21.2.E above, gives a second proof of Exercise 12.1.G, the Jacobian method for computing the Zariski tangent space at a k-valued point of a finite type k-scheme. Corollary 21.3.9 will extend this to the case of a separable closed point.

21.2.J. EXERCISE. Suppose X is a finite type scheme over an algebraically closed field. Show that the function from the closed points of X to $\mathbb{Z}^{\geq 0}$ given by $p \mapsto \dim T_p X$ is upper semicontinuous in the Zariski topology. (Be clear on what the Zariski topology is on the set of closed points.) Corollary 21.3.9 will allow you to extend this to all fields of characteristic 0 and all finite fields.

Depending on how your brain works, you may prefer using the first (constructive) or second (universal property) definition to do the next two exercises.

21.2.K. EXERCISE.
(a) *(pullback of differentials)* If

$$\begin{tikzcd} A' & A \arrow[l] \\ B' \arrow[u] & B \arrow[l] \arrow[u] \end{tikzcd}$$

is a commutative diagram, describe a natural homomorphism of A'-modules $A' \otimes_A \Omega_{A/B} \to \Omega_{A'/B'}$. An important special case is $B = B'$.
(b) *(differentials behave well with respect to base extension, affine case)* If furthermore the above diagram is a "tensor diagram" (i.e., $A' \cong B' \otimes_B A$, so the diagram is "co-Cartesian") then show that $A' \otimes_A \Omega_{A/B} \to \Omega_{A'/B'}$ is an isomorphism. (Depending on how your proceed, this may be trickier than you expect.)

21.2.L. EXERCISE: LOCALIZATION (STRONGER FORM; CF. EXAMPLE 21.2.8). Suppose $\phi : B \to A$ is a map of rings, S is a multiplicative subset of A, and T is a multiplicative subset of B with $\phi(T) \subset S$, so we have the following commutative diagram.

$$\begin{array}{ccc} S^{-1}A & \longleftarrow & A \\ \uparrow & & \uparrow \\ T^{-1}B & \longleftarrow & B \end{array}$$

Show that the pullback of differentials $S^{-1}\Omega_{A/B} \to \Omega_{S^{-1}A/B}$ of Exercise 21.2.K(a) is an isomorphism. (This should be believable from the intuitive picture of "vertical cotangent vectors".) An important case is when $T = \{1\}$. The case $\phi(T) = S$ is Example 21.2.8.

21.2.M. EXERCISE (FIELD EXTENSIONS CONTINUED, CF. EXERCISE 21.2.F).
(a) Compute $\Omega_{k(t)/k}$. (Hint: §21.2.6 followed by Exercise 21.2.L.)
(b) If K/k is **separably generated by** $t_1, \ldots, t_n \in K$ (i.e., $t_1, \ldots, t_n$ form a transcendence basis, and $K/k(t_1, \ldots, t_n)$ is algebraic and separable), show that $\Omega_{K/k}$ is a free K-module (i.e., vector space) with basis $dt_1, \ldots, dt_n$. Hint: use the relative cotangent sequence (Theorem 21.2.9) for $k \hookrightarrow k(t_1, \ldots, t_n) \hookrightarrow K$ to show that the dt_i span $\Omega_{K/k}$ as a K-vector space. The tricky part is showing that the dt_i are linearly independent. Do this by showing that there exists a unique map $\Omega_{K/k} \to K$ sending dt_1 to 1, and dt_i to 0 for $i > 1$. Do this first for the case where $K/k(t_1, \ldots, t_n)$ is generated by one element; then where it is finitely generated; then the general case by defining it on all finitely generated subextensions, and using uniqueness to show that they "all agree".

21.2.20. Third definition: global. We now want to globalize this definition for an arbitrary morphism of schemes $\pi : X \to Y$. We could do this "affine by affine"; we just need to make sure that the above notion behaves well with respect to "change of affine sets". Thus a relative differential on X would be the data of, for every affine $U \subset X$, a differential of the form $\sum a_i \, db_i$, and on the intersection of two affine open sets $U \cap U'$, with representatives $\sum a_i \, db_i$ on U and $\sum a'_i \, db'_i$ on the second, an equality on the overlap. Instead, we take a different approach. I will give the (seemingly unintuitive) definition, then tell you how to think about it, and then get back to the definition.

Let $\pi : X \to Y$ be any morphism of schemes. Recall that $\delta : X \to X \times_Y X$ is a locally closed embedding (Proposition 10.1.3). *Define* the **relative cotangent sheaf** $\Omega_{X/Y}$ (or Ω_π) as the conormal sheaf $\mathscr{N}^\vee_{X/X\times_Y X}$ of the diagonal (see §21.2.14 — and if $X \to Y$ is separated you needn't even worry about Exercise 21.2.G). (Now is also as good a time as any to define the **relative tangent sheaf** $\mathscr{T}_{X/Y}$ as the dual $\mathscr{H}om(\Omega_{X/Y}, \mathscr{O}_X)$ to the relative cotangent sheaf. If we are working in the category of k-schemes, then $\Omega_{X/k}$ and $\mathscr{T}_{X/k}$ are often called the **cotangent sheaf** and **tangent sheaf** of X respectively.)

We now define $d : \mathscr{O}_X \to \Omega_{X/Y}$. Let $pr_1 : X \times_Y X \to X$ and $pr_2 : X \times_Y X \to X$ be the two projections. Then define $d : \mathscr{O}_X \to \Omega_{X/Y}$ on the open set U as follows: $df = pr_2^* f - pr_1^* f$. (*Warning:* this is not a morphism of quasicoherent sheaves on X, although it *is* $\mathscr{O}_Y$-linear in the only possible meaning of that phrase.) We will soon see that this is indeed a derivation of the sheaf $\mathscr{O}_X$ (in the only possible meaning of the phrase), and at the same time see that our new notion of differentials agrees with our old definition on affine open sets, and hence globalizes the definition.

Note that for any open subset $U \subset Y$, d induces a map

$$\Gamma(U, \mathscr{O}_X) \to \Gamma(U, \Omega_{X/Y}), \tag{21.2.20.1}$$

which we also call d, and interpret as "taking the derivative".

21.2.21. *Motivation.* Before connecting this to our other definitions, let me try to convince you that this is a reasonable definition to make. (This discussion is informal and nonrigorous.) Say for example that Y is a point, and X a manifold. Then the tangent bundle $T_{X\times X}$ on $X \times X$ is $\mathrm{pr}_1^* T_X \oplus \mathrm{pr}_2^* T_X$, where pr_1 and pr_2 are the projections from $X \times X$ onto its two factors. Restrict this to the diagonal Δ, and look at the normal bundle exact sequence:

$$0 \to T_\Delta \to T_{X\times X}|_\Delta \to N_{\Delta/X} \to 0.$$

Now the left morphism sends v to (v, v), so the cokernel can be interpreted as $(v, -v)$. Thus $N_{\Delta/X}$ is isomorphic to T_X. Thus we can turn this on its head: we know how to find the normal bundle (or more precisely the conormal sheaf), and we can use this to define the tangent bundle (or more precisely the cotangent sheaf).

21.2.22. *Testing this out in the affine case.* Let's now see how this works for the special case $\operatorname{Spec} A \to \operatorname{Spec} B$. Then the diagonal $\operatorname{Spec} A \hookrightarrow \operatorname{Spec} A \otimes_B A$ corresponds to the ideal I of $A \otimes_B A$ that is the kernel of the ring map

$$\alpha : \sum x_i \otimes y_i \to \sum x_i y_i.$$

21.2.23. The ideal I of $A \otimes_B A$ is generated by the elements of the form $1 \otimes a - a \otimes 1$. Reason: if $\alpha(\sum x_i \otimes y_i) = 0$, i.e., $\sum x_i y_i = 0$, then

$$\sum x_i \otimes y_i = \sum (x_i \otimes y_i - x_i y_i \otimes 1) = \sum x_i (1 \otimes y_i - y_i \otimes 1).$$

The derivation is $d : A \to A \otimes_B A$, $a \mapsto 1 \otimes a - a \otimes 1$ (taken modulo I^2). (We shouldn't really call this "d" until we have verified that it agrees with our earlier definition, but we irresponsibly will anyway.)

Let's check that d is indeed a derivation (§21.2.17). Clearly d is B-linear, so we check the Leibniz rule:

$$\begin{aligned} d(aa') - a\, da' - a'\, da &= 1 \otimes aa' - aa' \otimes 1 - a \otimes a' + aa' \otimes 1 - a' \otimes a + a'a \otimes 1 \\ &= -a \otimes a' - a' \otimes a + a'a \otimes 1 + 1 \otimes aa' \\ &= (1 \otimes a - a \otimes 1)(1 \otimes a' - a' \otimes 1) \\ &\in I^2. \end{aligned}$$

Thus by the universal property of $\Omega_{A/B}$, we have a natural morphism $\Omega_{A/B} \to I/I^2$ of A-modules.

21.2.24. Theorem. — *The natural morphism $f : \Omega_{A/B} \to I/I^2$ induced by the universal property of $\Omega_{A/B}$ is an isomorphism.*

Proof. We will show this as follows. (i) We will show that f is surjective, and (ii) we will describe $g : I/I^2 \to \Omega_{A/B}$ such that $g \circ f : \Omega_{A/B} \to \Omega_{A/B}$ is the identity (showing that f is injective).

(i) The map f sends da to $1 \otimes a - a \otimes 1$, and such elements generate I (§21.2.23), so f is surjective.

(ii) Consider the map $A \otimes_B A \to \Omega_{A/B}$ defined by $x \otimes y \mapsto x\, dy$. (This is a well-defined map, by the universal property of $\otimes$, see §1.3.5.) Define $g : I/I^2 \to \Omega_{A/B}$ as the restriction of this map to I. We need to check that this is well-defined, i.e., that elements of I^2 are sent to 0, i.e., we need that

$$\left(\sum x_i \otimes y_i\right)\left(\sum x_j' \otimes y_j'\right) = \sum_{i,j} x_i x_j' \otimes y_i y_j' \mapsto 0$$

when $\sum_i x_i y_i = \sum x_j' y_j' = 0$. But by the Leibniz rule,

$$\begin{aligned}\sum_{i,j} x_i x_j'\, d(y_i y_j') &= \sum_{i,j} x_i x_j' y_i\, dy_j' + \sum_{i,j} x_i x_j' y_j'\, dy_i \\ &= \left(\sum_i x_i y_i\right)\left(\sum_j x_j'\, dy_j'\right) + \left(\sum_i x_i\, dy_i\right)\left(\sum_j x_j' y_j'\right) \\ &= 0.\end{aligned}$$

Then $g \circ f$ is indeed the identity, as

$$da \xmapsto{f} 1 \otimes a - a \otimes 1 \xmapsto{g} 1\, da - a\, d1 = da$$

as desired. □

21.2.N. Easy exercise. Suppose $\pi : X \to Y$ is a morphism of schemes, with open subschemes $\operatorname{Spec} A \subset X$ and $\operatorname{Spec} B \subset Y$, with $\operatorname{Spec} A \subset \pi^{-1}(\operatorname{Spec} B)$. Identify $\Omega_{X/Y}|_{\operatorname{Spec} A}$ with $\widetilde{\Omega_{A/B}}$, and identify $d : \Gamma(\operatorname{Spec} A, \mathscr{O}_X) \to \Gamma(\operatorname{Spec} A, \Omega_{X/Y})$ with $d : A \to \Omega_{A/B}$. Thus the global construction indeed naturally "glues together" the affine construction.

We can now use our understanding of how Ω works on affine open sets to generalize previous statements to non-affine settings.

21.2.O. Exercise. If $U \subset X$ is an open subset, show that the map (21.2.20.1) is a derivation.

21.2.P. Exercise. Suppose $\pi : X \to Y$ is locally of finite type, and Y (and hence X) is locally Noetherian. Show that $\Omega_{X/Y}$ is a coherent sheaf on X. (Feel free to weaken the Noetherian hypotheses for weaker conclusions.)

21.2.Q. Exercise. Suppose $\pi : X \to Y$ is smooth of relative dimension n. Prove that $\Omega_{X/Y}$ is locally free of rank n. Hint: the Jacobian description of Ω, Exercise 21.2.E.

The relative cotangent exact sequence and the conormal exact sequence for schemes now directly follow.

21.2.25. Theorem. —
(a) (relative cotangent exact sequence) Suppose $X \xrightarrow{\pi} Y \xrightarrow{\rho} Z$ *are morphisms of schemes. Then there is an exact sequence of quasicoherent sheaves on* X

$$\pi^*\Omega_{Y/Z} \longrightarrow \Omega_{X/Z} \longrightarrow \Omega_{X/Y} \longrightarrow 0,$$

globalizing Theorem 21.2.9.
(b) (conormal exact sequence) Suppose $\rho : Y \to Z$ is a morphism of schemes, and $i : X \hookrightarrow Y$ is a closed embedding, with conormal sheaf $\mathscr{N}^{\vee}_{X/Y}$. Then there is an exact sequence of sheaves on X:

$$\mathscr{N}^{\vee}_{X/Y} \xrightarrow{\delta} i^*\Omega_{Y/Z} \longrightarrow \Omega_{X/Z} \longrightarrow 0,$$

globalizing Theorem 21.2.12.

21.2.R. EXERCISE. Prove Theorem 21.2.25. (What needs to be checked?)

You should expect these exact sequences to be left-exact as well in the presence of appropriate smoothness (see Remark 21.2.10 and §21.2.13). The following is an important special case.

21.2.S. EXERCISE (LEFT-EXACTNESS OF THE RELATIVE COTANGENT SEQUENCE WHEN f AND g ARE SMOOTH). Show that the relative cotangent exact sequence is exact on the left if f and g are smooth. Hint: the $(m+n) \times (m+n)$ matrix you used in Exercise 12.6.D is "block upper triangular".

21.2.26. *Pulling back relative differentials.* Not surprisingly, the sheaf of relative differentials pull back, and behave well under base change.

21.2.27. Theorem (pullback of differentials). —
(a) If

$$\begin{tikzcd} X' \arrow[r, "\mu"] \arrow[d] & X \arrow[d] \\ Y' \arrow[r] & Y \end{tikzcd}$$

is a commutative diagram of schemes, there is a natural homomorphism $\mu^*\Omega_{X/Y} \to \Omega_{X'/Y'}$ of quasicoherent sheaves on X'. An important special case is $Y = Y'$.
(b) (Ω behaves well under base change) If furthermore the above diagram is a Cartesian square (so $X' \cong X \times_Y Y'$) then $\mu^*\Omega_{X/Y} \to \Omega_{X'/Y'}$ is an isomorphism.

21.2.T. EXERCISE. Derive Theorem 21.2.27 from Exercise 21.2.K. (Why does the construction of Exercise 21.2.K(a) "glue well"?)

As a particular case of Theorem 21.2.27(b), the fiber of the sheaf of relative differentials is indeed the sheaf of differentials of the fiber. Thus the sheaf of differentials notion indeed "glues together" the differentials on each fiber.

21.2.U. EXERCISE. Suppose $X \to Z$ and $Y \to Z$ are two morphisms. Describe an isomorphism $\Omega_{X \times_Z Y/Z} \cong \Omega_{X/Z} \boxtimes \Omega_{Y/Z}$.

21.2.V. EXERCISE. Suppose Φ is the Frobenius morphism of Exercise 7.3.S. Show that $\Omega_\Phi = 0$.

21.3 Smoothness of varieties revisited

Suppose k is a field. Since §12.2.5, we have used an awkward definition of k-smoothness, and we now make our definition of k-smoothness more robust.

21.3.1. *Redefinition.* A k-scheme X is **k-smooth of dimension n** or **smooth of dimension n over** k if it is locally of finite type, of pure dimension n, and $\Omega_{X/k}$ is locally free of rank n. The dimension n is often omitted, and one might possibly want to call something smooth if it is the (scheme-theoretic) disjoint union of things smooth of various dimensions.

21.3.A. EXERCISE. Verify that this definition is equivalent to the one given in Definition 12.2.6.

21.3.B. $\star\star$ EXERCISE (FOR THOSE WITH BACKGROUND IN COMPLEX GEOMETRY). Suppose X is a complex algebraic variety. Show that the analytification X^{an} of X (defined in Exercise 5.3.F) is smooth (in the differential-geometric sense) if and only if X is smooth (in the algebro-geometric sense, over $\mathbb{C}$). In this case, show that complex dimension of the complex manifold X^{an} (half the real dimension) is $\dim X$. Hint: the Jacobian criterion applies in both settings.

21.3.2. *Important observation.* As a consequence of our better definition of smoothness, we see that it can be checked on *any* affine cover by using the Jacobian criterion on each affine open set in the cover, as hinted in §12.2.7.

21.3.C. EXERCISE (SMOOTHNESS IS INSENSITIVE TO EXTENSION OF BASE FIELD). Prove the converse to Exercise 12.2.F.

21.3.D. IMPORTANT EXERCISE. Suppose k is perfect, and X is a finite type k-scheme. Show that X is smooth if and only if X is regular at all closed points. In particular, Theorem 12.2.10(a) holds. Hint: let $X_{\overline{k}} := X \times_k \overline{k}$. Explain why X is smooth if and only if $X_{\overline{k}}$ is smooth if and only if $X_{\overline{k}}$ is regular at all closed points if and only if X is regular at all closed points. See Exercise 12.2.N for the last step.

21.3.3. *Remark: A short definition of smoothness of varieties.* Exercise 21.3.D gives another pithy characterization of smoothness of varieties: *a finite type k-scheme is smooth if it is geometrically regular*.

21.3.4. Generic smoothness.

We can now verify something you may already have intuited. In positive characteristic, this is a hard theorem, in that it uses a result from commutative algebra that we have not proved.

21.3.5. Theorem (generic smoothness of varieties). — *If X is an irreducible variety over a perfect field k of dimension n, there is a dense open subset U of X such that U is smooth of dimension n.*

Hence, by Fact 12.8.2, U is regular. Theorem 25.3.1 will generalize this to smooth *morphisms*, at the expense of restricting to characteristic 0.

Proof. The $n = 0$ case is immediate, so we assume $n > 0$.

We will show that the rank at the generic point is n. Then by upper semicontinuity of the rank of a coherent sheaf (Exercise 13.7.J), it must be n in an open neighborhood of the generic point, and we are done.

We thus have to check that if K is the fraction field of a dimension n integral finite type k-scheme, i.e., (by Theorem 11.2.1) if K/k is a transcendence degree n extension, then $\Omega_{K/k}$ is an n-dimensional vector space. But every extension of transcendence degree $n > 1$ is *separably generated* (see Exercise 21.2.M(b)): we can find n algebraically independent elements of K over k, say $x_1, \dots, x_n$, such that $K/k(x_1, \dots, x_n)$ is separable. (In characteristic 0, this is automatic from transcendence theory, see Exercise 11.2.A, as all finite extensions are separable. But it also holds for perfect fields in positive characteristic, see [**E**, Cor. A1.7] or [**Mat2**, Thm. 26.2].) Then $\Omega_{K/k}$ is generated by $dx_1, \dots, dx_n$ (by Exercise 21.2.M(b)). $\square$

Theorem 21.3.5 allows us to prove Theorem 12.8.3 (an important case of Fact 12.8.2), that the localization of regular local rings are regular in the case of varieties over perfect fields.

21.3.6. *Proof of Theorem 12.8.3.* Suppose Y is a variety over a perfect field k that is regular at its closed points (so Y is smooth, by Exercise 21.3.D), and let η be a point of Y. We will show that Y is regular at η. Let $X = \overline{\eta}$. By Theorem 21.3.5, X contains a dense (=nonempty) open subset of smooth points. By shrinking Y by discarding the points of X outside that open subset, we may assume X is smooth.

Then Theorem 21.3.8 (exactness of the conormal exact sequence for smooth varieties) implies that $\mathscr{I}/\mathscr{I}^2$ is a locally free sheaf of rank $\operatorname{codim}_{X/Y} = \dim \mathscr{O}_{Y,\eta}$.

21.3.E. EXERCISE. Let $\mathfrak{m}$ be the maximal ideal of $\mathscr{O}_{Y,\eta}$. Identify the stalk of $\mathscr{I}/\mathscr{I}^2$ at the generic point η of X with $\mathfrak{m}/\mathfrak{m}^2$. Conclude the proof of Theorem 12.8.3. $\square$

21.3.F. EXERCISE (LOCALIZATION OF REGULAR LOCAL RINGS OF VARIETIES ARE REGULAR, PROMISED JUST AFTER THEOREM 12.8.3). Suppose $(A, \mathfrak{m})$ is a regular local ring that is the localization of a finitely generated k-algebra, where k is perfect. Show that the localization of A at a prime is also a regular local ring.

21.3.7. Left-exactness of the conormal exact sequence for embeddings of smooth varieties.

As described in §21.2.13, we expect the conormal exact sequence to be exact on the left in appropriately "smooth" situations.

21.3.8. Theorem (conormal exact sequence for smooth varieties). — *Suppose $i : X \hookrightarrow Y$ is a closed embedding of smooth varieties over a field k, with conormal sheaf $\mathscr{N}^\vee_{X/Y}$. Then the conormal exact sequence (Theorem 21.2.25(b)) is exact on the left:*

$$0 \longrightarrow \mathscr{N}^\vee_{X/Y} \xrightarrow{\delta} i^*\Omega_{Y/k} \longrightarrow \Omega_{X/k} \longrightarrow 0$$

is exact.

By dualizing, i.e., applying $\mathscr{H}om(\cdot, \mathscr{O}_X)$, we obtain the **normal exact sequence**

$$0 \to \mathscr{T}_{X/k} \to \mathscr{T}_{Y/k}|_X \to \mathscr{N}_{X/Y} \to 0$$

which is geometrically more intuitive (see Figure 21.3 and the discussion after the proof of Theorem 21.2.12).

Proof. We use the fact that smooth k-varieties are regular at their closed points, which we have only proved for perfect fields, but which we shall later prove fully (see §12.2.9).

Let $\mathscr{I}$ be the ideal sheaf of $X \hookrightarrow Y$. By Exercise 12.2.K, i is a regular embedding in the neighborhood of all closed points, and hence everywhere. Thus by Proposition 21.2.16(b), $\mathscr{I}/\mathscr{I}^2$ is locally free of rank r. By Exercise 5.5.L, the associated points of $\ker \delta$ are a subset of the associated points of $\mathscr{I}/\mathscr{I}^2$, so to show that $\ker \delta = 0$, it suffices to check this at the associated points of $\mathscr{I}/\mathscr{I}^2$, which are precisely the associated points of X (as $\mathscr{I}/\mathscr{I}^2$ is locally free).

As stated above, X is regular at closed points, hence (by Theorem 12.2.13) reduced (the nonreduced locus is closed, §8.3.10, and hence if nonempty would necessarily contain a closed point). Thus we need only check at the generic components of X (as X has no embedded points, Exercise 5.5.C). But this involves checking the left-exactness of a right-exact sequence of fields, and as the dimension of the left (the codimension of X in Y by Proposition 21.2.16(b)) is precisely the difference of the other two, we are done. □

21.3.9. Important Corollary. — *Suppose Y is a smooth k-variety, and $q \in Y$ is a closed point whose residue field is separable over k. (This is automatic if $\operatorname{char} k = 0$ or if k is a finite field.) Then the conormal exact sequence for $q \hookrightarrow Y$ yields an isomorphism of the Zariski cotangent space of Y at q with the fiber of $\Omega_{Y/k}$ at q.*

Proof. Apply Theorem 21.3.8 with $X = q$. Note that q is indeed a smooth k-variety (!), and that $\Omega_{q/k} = 0$. □

21.3.10. *Remark.* This result generalizes Proposition 21.2.18 to separable closed points. As described after the statement of Proposition 21.2.18, this has a number of consequences. For example, it extends the Jacobian description of the Zariski tangent space to separable closed points.

21.3.G. EXERCISE (CF. §12.2.16). Suppose p is a closed point of a k-variety X, with residue field k' that is separable over k of degree d. Define $\pi : X_{\overline{k}} := X \times_k \overline{k} \to X$ by base change from $\operatorname{Spec} \overline{k} \to \operatorname{Spec} k$. Suppose $q \in \pi^{-1}(p)$. Show that $X_{\overline{k}}$ is regular at q if and only if X is regular at p.

21.3.11. *Remark.* The conormal sequence is exact on the left in even more general circumstances. Essentially all that is required is appropriate smoothness of $i \circ g : X \to Z$, see [**Stacks**, tag 06B7], and [**Stacks**, tag 06BB] for related facts.

21.4 Examples

The examples below are organized by topic, not by difficulty.

21.4.1. The geometric genus of a curve. A regular projective curve C (over a field k) has **geometric genus** $h^0(C, \Omega_{C/k})$. (This will be generalized to higher dimension in §21.5.3.) This is always finite, as $\Omega_{C/k}$ is coherent (Exercise 21.2.P), and coherent sheaves on projective k-schemes have finite-dimensional spaces of sections (Theorem 18.1.4(a)). (The geometric genus is also called the *first algebraic de Rham cohomology group*, in analogy with de Rham cohomology in the differentiable setting.) Sadly, this isn't really a new invariant. We will see in Exercise 21.4.C that this agrees with our earlier definition of genus, i.e., $h^0(C, \Omega_{C/k}) = h^1(C, \mathscr{O}_C)$.

21.4.2. The projective line. As an important first example, consider $\mathbb{P}^1_k$, with the usual projective coordinates x_0 and x_1. As usual, the first patch corresponds to $x_0 \neq 0$, and is of the form $\operatorname{Spec} k[x_{1/0}]$ where $x_{1/0} = x_1/x_0$. The second patch corresponds to $x_1 \neq 0$, and is of the form $\operatorname{Spec} k[x_{0/1}]$ where $x_{0/1} = x_0/x_1$.

Both patches are isomorphic to $\mathbb{A}^1_k$, and $\Omega_{\mathbb{A}^1_k} = \mathscr{O}_{\mathbb{A}^1_k}$. (More precisely, $\Omega_{k[x]/k} = k[x]\ dx$.) Thus $\Omega_{\mathbb{P}^1_k}$ is an invertible sheaf (a line bundle). The invertible sheaves on $\mathbb{P}^1_k$ are of the form $\mathscr{O}(m)$. So which invertible sheaf is $\Omega_{\mathbb{P}^1/k}$?

Let's take a section, $dx_{1/0}$ on the first patch. It has no zeros or poles there, so let's check what happens on the other patch. As $x_{1/0} = 1/x_{0/1}$, we have $dx_{1/0} = -(1/x_{0/1}^2)\ dx_{0/1}$. Thus this section has a double pole where $x_{0/1} = 0$. Hence $\Omega_{\mathbb{P}^1_k/k} \cong \mathscr{O}(-2)$.

Note that the above argument works equally well if k were replaced by $\mathbb{Z}$: our theory of Weil divisors and line bundles of Chapter 14 applies ($\mathbb{P}^1_{\mathbb{Z}}$ is factorial), so the previous argument essentially without change shows that $\Omega_{\mathbb{P}^1_{\mathbb{Z}}/\mathbb{Z}} \cong \mathscr{O}(-2)$. And because Ω behaves well with respect to base change (Exercise 21.2.27(b)), and any scheme maps to $\operatorname{Spec}\mathbb{Z}$, this implies that $\Omega_{\mathbb{P}^1_B/B} \cong \mathscr{O}_{\mathbb{P}^1_B}(-2)$ for *any* base scheme B.

(Also, as suggested by §18.5.2, this shows that $\Omega_{\mathbb{P}^1/k}$ is the dualizing sheaf for $\mathbb{P}^1_k$; see also Example 18.5.4. But given that we haven't yet proved Serre duality, this isn't so meaningful.)

Side Remark: the fact that the degree of the tangent bundle is 2 is related to the "Hairy Ball Theorem" (the dimension 2 case of [**Hat**, Thm. 2.28].

21.4.3. Hyperelliptic curves. Throughout this discussion of hyperelliptic curves, we suppose that $k = \overline{k}$ and $\operatorname{char} k \neq 2$, so we may apply the discussion of §19.5. Consider a double cover $\pi : C \to \mathbb{P}^1_k$ by a regular curve C, branched over $2g+2$ distinct points. We will use the explicit coordinate description of hyperelliptic curves of (19.5.2.1). By Exercise 19.5.1, C has genus g.

21.4.A. EXERCISE: DIFFERENTIALS ON HYPERELLIPTIC CURVES. What is the degree of the invertible sheaf $\Omega_{C/k}$? (Hint: let x be a coordinate on one of the coordinate patches of $\mathbb{P}^1_k$. Consider $\pi^* dx$ on C, and count poles and zeros. Use the explicit coordinates of §19.5. You should find that $\pi^* dx$ has $2g+2$ zeros and 4 poles (counted with multiplicity), for a total of $2g-2$.) Doing this exercise will set you up well for the Riemann-Hurwitz formula, in §21.7.

21.4.B. EXERCISE ("THE FIRST ALGEBRAIC DE RHAM COHOMOLOGY GROUP OF A HYPERELLIPTIC CURVE"). Show that $h^0(C, \Omega_{C/k}) = g$ as follows.
(a) Show that $\frac{dx}{y}$ is a (regular) differential on $\operatorname{Spec} k[x]/(y^2 - f(x))$ (i.e., an element of $\Omega_{(k[x]/(y^2-f(x)))/k}$).
(b) Show that $x^i(dx)/y$ extends to a global differential ω_i on C (i.e., with no poles).
(c) Show that the ω_i $(0 \leq i < g)$ are linearly independent differentials. (Hint: Show

that the valuation of ω_i at the origin is i. If $\omega := \sum_{j=i}^{g-1} a_j \omega_j$ is a nontrivial linear combination, with $a_j \in k$, and $a_i \neq 0$, show that the valuation of ω at the origin is i, and hence $\omega \neq 0$.)
$\star$(d) Show that the ω_i form a basis for the differentials. (Hint: consider the order of poles of the ω_i at $\pi^{-1}(\infty)$.)

21.4.C. $\star$ EXERCISE (TOWARD SERRE DUALITY).
(a) Show that $h^1(C, \Omega_{C/k}) = 1$. (In the course of doing this, you might interpret a generator of $H^1(C, \Omega_{C/k})$ as $x^{-1}\,dx$. In particular, the pullback map $H^1(\mathbb{P}^1, \Omega_{\mathbb{P}^1/k}) \to H^1(C, \Omega_{C/k})$ is an isomorphism.)
(b) Describe a natural perfect pairing

$$H^0(C, \Omega_{C/k}) \times H^1(C, \mathscr{O}_C) \to H^1(C, \Omega_{C/k}).$$

In terms of our explicit coordinates, you might interpret it as follows. Recall from the proof of the hyperelliptic Riemann-Hurwitz formula (Theorem 19.5.1) that $H^1(C, \mathscr{O}_C)$ can be interpreted as

$$\left\langle \frac{y}{x}, \frac{y}{x^2}, \dots, \frac{y}{x^g} \right\rangle.$$

Then the pairing

$$\left\langle \frac{dx}{y}, \dots, x^{g-1}\frac{dx}{y} \right\rangle \times \left\langle \frac{y}{x}, \dots, \frac{y}{x^g} \right\rangle \to \left\langle x^{-1}\,dx \right\rangle$$

is basically "multiply and read off the $x^{-1}\,dx$ term". Or in fancier terms: "multiply and take the residue". (You may want to compare this to Example 18.5.4.)

21.4.4. Discrete valuation rings. The following exercise is used in the proof of the Riemann-Hurwitz formula, §21.7.

21.4.D. EXERCISE. Suppose that the discrete valuation ring $(A, \mathfrak{m}, k)$ is a localization of a finitely generated k-algebra. Let t be a uniformizer of A. Show that the differentials are free of rank one and generated by dt, i.e., $\Omega_{A/k} = A\ dt$, as follows. (It is also possible to show this using the ideas from §21.5.)
(a) Show that $\Omega_{A/k}$ is a finitely generated A-module.
(b) Show that $\langle dt \rangle = \Omega_{A/k}$, i.e., that $\times dt : A \to \Omega_{A/k}$ is a surjection, as follows. Let π be the projection $A \to A/\mathfrak{m} = k$, so for $a \in A$, $a - \pi(a) \in \mathfrak{m}$. Define $\sigma(a) = (a - \pi(a))/t$. Show that $\Omega_{A/k} = \langle dt \rangle + \mathfrak{m}\Omega_{A/k}$, using the fact that for every $a \in A$,

$$da = \sigma(a)\ dt + t\ d\sigma(a).$$

Apply Nakayama's Lemma version 3 (Exercise 7.2.G) to $\langle dt \rangle \subset \Omega_{A/k}$. (This argument, with essentially no change, can be used to show that if $(A, \mathfrak{m}, k)$ is a localization of a finitely generated algebra over k, and $t_1, \dots, t_n$ generate $\mathfrak{m}$, then $dt_1, \dots, dt_n$ generate $\Omega_{A/k}$.)
(c) By part (b), $\Omega_{A/k}$ is a principal A-module. Show that $\times dt$ is an injection as follows. By the classification of finite generated modules over discrete valuation rings (Remark 12.5.15), it suffices to show that $t^m\ dt \neq 0$ for all m. The surjection $A \to A/(t^N)$ induces a map $\Omega_{A/k} \to \Omega_{(A/(t^N))/k}$, so it suffices to show that $t^m\ dt$ is nonzero in $\Omega_{(A/(t^N))/k}$. Show that $A/(t^N) \cong k[t]/(t^N)$. The usual differentiation rule for polynomials gives a map $k[t]/(t^N) \to k[t]/(t^{N-1})$ which is a

derivation of $k[t]/(t^N)$ over k, and $t^m\ dt$ will not map to 0 so long as N is sufficiently large. Put the pieces together and complete the proof. (An extension of these ideas can show that if $(A, \mathfrak{m}, k)$ is a localization of a finitely generated algebra over k that is a regular local ring, then $\Omega_{A/k}$ is free of rank $\dim A$.)

21.4.5. Projective space and the Euler exact sequence. We next examine the differentials of projective space $\mathbb{P}^n_k$, or more generally $\mathbb{P}^n_A$ where A is an arbitrary ring. As projective space is covered by affine open sets of the form $\mathbb{A}^n$, on which the differentials form a rank n locally free sheaf, $\Omega_{\mathbb{P}^n_A/A}$ is also a rank n locally free sheaf.

21.4.6. Theorem (the Euler exact sequence). — *The sheaf of differentials $\Omega_{\mathbb{P}^n_A/A}$ satisfies the following exact sequence*

$$0 \to \Omega_{\mathbb{P}^n_A/A} \to \mathscr{O}_{\mathbb{P}^n_A}(-1)^{\oplus(n+1)} \to \mathscr{O}_{\mathbb{P}^n_A} \to 0.$$

This is handy, because you can get a hold of $\Omega_{\mathbb{P}^n_A/A}$ in a concrete way. See Exercise 21.5.Q for an application. By dualizing this exact sequence, we have an exact sequence $0 \to \mathscr{O}_{\mathbb{P}^n_A} \to \mathscr{O}_{\mathbb{P}^n_A}(1)^{\oplus(n+1)} \to \mathscr{T}_{\mathbb{P}^n_A/A} \to 0$.

21.4.7. $\star$ *Proof of Theorem 21.4.6.* (What is really going on in this proof is that we consider those differentials on $\mathbb{A}^{n+1}_A \setminus \{0\}$ that are pullbacks of differentials on $\mathbb{P}^n_A$. For a different explanation, in terms of the Koszul complex, see [**E**, §17.5].)

We first describe a map $\phi : \mathscr{O}(-1)^{\oplus(n+1)} \to \mathscr{O}$, and later identify the kernel with $\Omega_{X/Y}$. The map is given by

$$\phi : (s_0, s_1, \dots, s_n) \mapsto x_0 s_0 + x_1 s_1 + \cdots + x_n s_n.$$

You should think of this as a "degree 1" map, as each x_i has degree 1.

21.4.8. *Remark.* The dual $\phi^\vee : \mathscr{O} \to \mathscr{O}(1)^{\oplus(n+1)}$ of ϕ gives a map $\mathbb{P}^n \to \mathbb{P}^n$ via Important Theorem 16.4.1 on maps to projective space. This map is the identity, and this is one way of describing ϕ in a "natural" (coordinate-free) manner.

21.4.E. Easy exercise. Show that ϕ is surjective, by checking on the open set $D(x_i)$. (There is a one-line solution.)

Now we must identify the kernel of ϕ with the differentials, and we can do this on each $D(x_i)$, so long as we do it in a way that works simultaneously for each open set. So we consider the open set U_0, where $x_0 \neq 0$, and we have the usual coordinates $x_{j/0} = x_j/x_0$ ($1 \leq j \leq n$). Given a differential

$$f_1(x_{1/0}, \dots, x_{n/0})\ dx_{1/0} + \cdots + f_n(x_{1/0}, \dots, x_{n/0})\ dx_{n/0}$$

we must produce $n+1$ sections of $\mathscr{O}(-1)$. As motivation, we just look at the first term, and pretend that the projective coordinates are actual coordinates.

$$\begin{aligned} f_1\ dx_{1/0} &= f_1\ d(x_1/x_0) \\ &= f_1 \frac{x_0\ dx_1 - x_1\ dx_0}{x_0^2} \\ &= -\frac{x_1}{x_0^2} f_1\ dx_0 + \frac{f_1}{x_0}\ dx_1 \end{aligned}$$

Note that x_0 times the "coefficient of dx_0" plus x_1 times the "coefficient of dx_1" is 0, and also both coefficients are of homogeneous degree -1. Motivated by this, we take:

(21.4.8.1) $$f_1\ dx_{1/0} + \cdots + f_n\ dx_{n/0} \mapsto \left(-\frac{x_1}{x_0^2} f_1 - \cdots - \frac{x_n}{x_0^2} f_n, \frac{f_1}{x_0}, \frac{f_2}{x_0}, \dots, \frac{f_n}{x_0} \right)$$

Note that over U_0, this indeed gives an injection of $\Omega_{\mathbb{P}^n_A}$ to $\mathscr{O}(-1)^{\oplus(n+1)}$ that surjects onto the kernel of $\mathscr{O}(-1)^{\oplus(n+1)} \to \mathscr{O}_X$ (if $(g_0, \dots, g_n)$ is in the kernel, take $f_i = x_0 g_i$ for $i > 0$).

Let's make sure this construction, applied to two different coordinate patches (say U_0 and U_1) gives the same answer. (This verification is best ignored on a first reading.) Note that

$$\begin{aligned} f_1\ dx_{1/0} + f_2\ dx_{2/0} + \cdots &= f_1\ d\frac{1}{x_{0/1}} + f_2\ d\frac{x_{2/1}}{x_{0/1}} + \cdots \\ &= -\frac{f_1}{x_{0/1}^2} dx_{0/1} + \frac{f_2}{x_{0/1}} dx_{2/1} - \frac{f_2 x_{2/1}}{x_{0/1}^2} dx_{0/1} + \cdots \\ &= -\frac{f_1 + f_2 x_{2/1} + \cdots}{x_{0/1}^2} dx_{0/1} + \frac{f_2 x_1}{x_0} dx_{2/1} + \cdots . \end{aligned}$$

Under this map, the $dx_{2/1}$ term goes to the second factor (where the factors are indexed 0 through n) in $\mathscr{O}(-1)^{\oplus(n+1)}$, and yields f_2/x_0 as desired (and similarly for $dx_{j/1}$ for $j > 2$). Also, the $dx_{0/1}$ term goes to the "zero" factor, and yields

$$\left(\sum_{j=1}^{n} f_i (x_i/x_1)/(x_0/x_1)^2 \right) /x_1 = f_i x_i / x_0^2$$

as desired. Finally, the "first" factor must be correct because the sum over i of x_i times the ith factor is 0.

21.4.F. EXERCISE. Finish the proof of Theorem 21.4.6, by verifying that this map $\Omega_{\mathbb{P}^n_A/A} \to \mathscr{O}_{\mathbb{P}^n_A}(-1)^{\oplus(n+1)}$ identifies $\Omega_{\mathbb{P}^n_A/A}$ with $\ker \phi$.

$\square$

21.4.G. EXERCISE (PROMISED IN §18.5.7). Show that $h^1(\mathbb{P}^2_k, \Omega_{\mathbb{P}^2_k/k}) > 0$. Show that $\Omega_{\mathbb{P}^2_k/k}$ is not the direct sum of line bundles. Show that Theorem 18.5.6 cannot be extended to rank 2 vector bundles on $\mathbb{P}^2_k$.

21.4.9. *Generalizations of the Euler exact sequence.* Generalizations of the Euler exact sequence are quite useful. We won't use them later, so no proofs will be given. First, the argument applies without change if $\operatorname{Spec} A$ is replaced by an arbitrary base scheme. The Euler exact sequence further generalizes in a number of ways. As a first step, suppose $\mathscr{V}$ is a rank $n+1$ locally free sheaf (or vector bundle) on a scheme X. Then $\Omega_{\mathbb{P}\mathscr{V}/X}$ sits in an Euler exact sequence:

$$0 \to \Omega_{\mathbb{P}\mathscr{V}/X} \to \mathscr{O}(-1) \otimes \mathscr{V}^\vee \to \mathscr{O}_X \to 0$$

If $\pi : \mathbb{P}\mathscr{V} \to X$, the map $\mathscr{O}(-1) \otimes \mathscr{V}^\vee \to \mathscr{O}_X$ is induced by $\mathscr{V}^\vee \otimes \pi_* \mathscr{O}(1) \cong (\mathscr{V}^\vee \otimes \mathscr{V}) \otimes \mathscr{O}_X \to \mathscr{O}_X$, where $\mathscr{V}^\vee \otimes \mathscr{V} \to \mathscr{O}_X$ is the trace map (§13.7.1). (You may wish to compare this to Remark 21.4.8.)

It is not obvious that this is useful, but we have already implicitly seen it in the case of $\mathbb{P}^1$-bundles over curves, in Exercise 20.2.J, where the normal bundle to a section was identified in this way.

21.4.10. $\star\star$ *Generalization to the Grassmannian.* For another generalization, fix a base field k, and let $G(m, n+1)$ be the space of sub-vector spaces of dimension m in an $(n+1)$-dimensional vector space V (the Grassmannian, §16.7). Over $G(m, n+1)$ we have a short exact sequence of locally free sheaves

$$0 \to \mathscr{S} \to \mathscr{O}_{G(m,n+1)} \otimes V^\vee \to \mathscr{Q} \to 0$$

where $\mathscr{O}_{G(m,n+1)} \otimes V^\vee$ is the "trivial bundle whose fibers are $V^\vee$" (do you understand what that means?), and $\mathscr{S}$ is the "universal subbundle". Then there is a canonical isomorphism

$$\Omega_{G(m,n+1)/k} \cong \mathscr{H}om(\mathscr{Q}, \mathscr{S}). \tag{21.4.10.1}$$

21.4.H. EXERCISE. Recall that in the case of projective space, i.e., $m = 1$, $\mathscr{S} = \mathscr{O}(-1)$ (Exercise 17.1.H). Verify (21.4.10.1) in this case using the Euler exact sequence (Theorem 21.4.6).

21.4.I. EXERCISE. Prove (21.4.10.1), and explain how it generalizes 20.2.I. (The hint to Exercise 20.2.I may help.)

This Grassmannian fact generalizes further to Grassmannian bundles, and to flag varieties, and to flag bundles.

21.5 Studying smooth varieties using their cotangent bundles

In this section, we construct birational invariants of varieties over algebraically closed fields (such as the geometric genus), motivate the notion of an unramified morphism, show that varieties are "smooth almost everywhere", and get a first glimpse of Hodge theory.

21.5.1. The geometric genus, and other birational invariants from i-forms $\Omega^i_{X/Y}$.

Suppose X is a projective scheme over k. Then for each i, $h^i(X, \Omega_{X/k})$ is an invariant of X, which can be useful. The first useful fact is that it, and related invariants, are *birational invariants* if X is smooth, as shown in the following exercise. We first define the **sheaf of (relative) i-forms** $\Omega^i_{X/Y} := \wedge^i \Omega_{X/Y}$. Sections of $\Omega^i_{X/Y}$ (over some open set) are called **(relative) i-forms** (over that open set).

21.5.2. *Joke.* Old Macdonald had a form; $e_i \wedge e_i = 0$.

21.5.A. EXERCISE ($h^0(X, \Omega^i_{X/k})$ ARE BIRATIONAL INVARIANTS). Suppose X and X' are birational smooth projective k-varieties. Show (for each i) that $H^0(X, \Omega^i_{X/k}) \cong H^0(X', \Omega^i_{X'/k})$. Hint: fix a birational map $\phi : X \dashrightarrow X'$. By Exercise 16.5.B, the complement of the domain of definition U of ϕ is codimension at least 2. By pulling back i-forms from X' to U, we get a map $\phi^* : H^0(X', \Omega^i_{X'/k}) \to H^0(U, \Omega^i_{X/k})$. Use Algebraic Hartogs's Lemma 11.3.10 and the fact that Ω^i is locally free to show the map extends to a map $\phi^* : H^0(X', \Omega^i_{X'/k}) \to H^0(X, \Omega^i_{X/k})$. If $\psi : X' \dashrightarrow X$ is the inverse rational map, we similarly get a map $\psi^* : H^0(X, \Omega^i_{X/k}) \to H^0(X', \Omega^i_{X'/k})$. Show that ϕ^* and ψ^* are inverse by showing that each composition is the identity on a dense open subset of X or X'.

21.5.3. *The canonical bundle $\mathscr{K}_X$ and the geometric genus $p_g(X)$.* If X is a dimension n smooth k-variety, the invertible sheaf (or line bundle) $\det \Omega_{X/k} = \Omega^n_{X/k}$ (the sheaf of "algebraic volume forms") has particular importance, and is called the **canonical (invertible) sheaf**, or the **canonical (line) bundle**. It is denoted $\mathscr{K}_X$ (or $\mathscr{K}_{X/k}$). As mentioned in §18.5.2, if X is projective, then $\mathscr{K}_X$ is the dualizing sheaf ω_X appearing in the statement of Serre duality, something we will establish in §30.4 (see Desideratum 30.1.1).

21.5.B. EXERCISE (THE ADJUNCTION FORMULA FOR $\mathscr{K}_X$). Suppose X is a smooth variety, and and Z is a smooth subvariety of X. Show that

$$\mathscr{K}_Z \cong \mathscr{K}_X|_Z \otimes \det \mathscr{N}_{Z/X}.$$

(Hint: apply Exercise 13.5.H to Theorem 21.3.8.) In particular, by Exercise 21.2.H, if Z is codimension 1, then

$$\mathscr{K}_Z \cong (\mathscr{K}_X \otimes \mathscr{O}_X(Z))\,|_Z.$$

This is often used inductively, for complete intersections. (See Exercise 20.2.B(a) and §30.4 for other versions of the adjunction formula.)

21.5.4. *Definition.* If X is a projective (or even proper) smooth k-variety, the birational invariant $h^0(X, \mathscr{K}_X) = h^0(X, \Omega^n_{X/k})$ has particular importance. It is called the **geometric genus**, and is denoted $p_g(X)$. We saw this in the case of curves in §21.4.1. If X is an irreducible variety that is *not* smooth or projective, the phrase geometric genus refers to $h^0(X', \mathscr{K}_{X'})$ for some smooth projective X' *birational* to X. (By Exercise 21.5.A, this is independent of the choice of X'.) For example, if X is an irreducible reduced projective curve over k, the geometric genus is the geometric genus of the normalization of X. (But in higher dimension, it is not clear if there exists such an X'. It is a nontrivial fact that this is true in characteristic 0 — Hironaka's resolution of singularities — and it is not yet known in positive characteristic; see Remark 22.4.6.)

It is a miracle that for a complex curve the geometric genus is the same as the topological genus and the arithmetic genus. We will connect the geometric genus to the topological genus in our discussion of the Riemann-Hurwitz formula soon (Exercise 21.7.I). We connect the geometric genus to the arithmetic genus in the following exercise.

21.5.C. EASY EXERCISE. Assuming Miracle 18.5.2 (that the canonical bundle is Serre-dualizing), show that the geometric genus of a smooth projective curve over $k = \overline{k}$ equals its arithmetic genus.

21.5.D. EXERCISE. Suppose Z is a regular degree d surface in $\mathbb{P}^3_k$. Compute the geometric genus $p_g(Z)$ of Z. Show that no regular quartic surface in $\mathbb{P}^3_k$ is rational (i.e., birational to $\mathbb{P}^2_k$, Definition 6.5.4). (Such quartic surfaces are examples of *K3 surfaces*, see Exercise 21.5.I.)

21.5.5. Important classes of varieties: Fano, Calabi-Yau, general type.

Suppose X is a smooth projective k-variety. Then X is said to be **Fano** if $\mathscr{K}_X^\vee$ is ample, and **Calabi-Yau** if $\mathscr{K}_X \cong \mathscr{O}_X$. (Caution: there are other definitions of Calabi-Yau.)

21.5.E. EXERCISE. The j**th plurigenus** of a smooth projective k-variety is $h^0(X, \mathscr{K}_X^{\otimes j})$. Show that the jth plurigenus is a birational invariant.

The **Kodaira dimension** of X, denoted $\kappa(X)$, tracks the rate of growth of the plurigenera. It is the smallest k such that $h^0(X, \mathscr{K}_X^{\otimes j})/j^k$ is bounded (as j varies through the positive integers), except that if all the plurigenera $h^0(X, \mathscr{K}_X^{\otimes j})$ are 0, we say $\kappa(X) = -1$. It is a nontrivial fact that the Kodaira dimension always exists, and that it is the maximum of the dimensions of images of the jth "pluricanonical rational maps". The latter fact implies that the Kodaira dimension is an integer between -1 and $\dim X$ inclusive. Exercise 21.5.E shows that the Kodaira dimension is a birational invariant. If $\kappa(X) = \dim X$, we say that X is of **general type**. (For more information on the Kodaira dimension, see [**Ii**, §10.5].)

21.5.F. EXERCISE. Show that if $\mathscr{K}_X$ is ample then X is of general type.

21.5.G. EXERCISE. Show that a smooth geometrically irreducible projective curve (over a field k) is Fano (resp. Calabi-Yau, general type) if its genus is 0 (resp. 1, greater than 1).

The "trichotomy" of Exercise 21.5.G is morally the reason that curves behave differently depending on which of the three classes they lie in (see §19.6.1). In some sense, important aspects of this trichotomy extend to higher dimension, which is part of the reason for making these definitions. We now explore this trichotomy in the case of complete intersections.

21.5.H. EXERCISE. Suppose X is a smooth complete intersection in $\mathbb{P}_k^N$ of hypersurfaces of degree $d_1, \dots, d_n$, where k is algebraically closed.
(a) Show that $\mathscr{K}_X \cong \mathscr{O}(-N-1+d_1+\cdots+d_n)|_X$.
(b) Show that $-N-1+d_1+\cdots+d_n$ is negative (resp. zero, positive) then X is Fano (resp. Calabi-Yau, general type).
(c) Find all possible values of N and $d_1, \dots, d_n$ where X is Calabi-Yau of dimension at most 3. Notice how small this list is.

21.5.I. EXERCISE. A **K3 surface** over a field k is a proper smooth geometrically connected Calabi-Yau surface X over k such that $H^1(X, \mathscr{O}_X) = 0$. (Weil has written that K3 surfaces were named in honor of Kummer, Kähler, and Kodaira, and the mountain K2, [**BPV**, p. 288].) Prove that the dimension 2 (smooth) Calabi-Yau complete intersections of Exercise 21.5.H(c) are all K3 surfaces.

21.5.6. *Tantalizing side remark.* If you compare the degrees of the hypersurfaces cutting out complete intersection K3 surfaces (in Exercise 21.5.H(c)), with the degrees of hypersurfaces cutting out complete intersection canonical curves (see §19.8.1 and Exercise 19.8.G), you will notice a remarkable coincidence. Of course this is not a coincidence at all.

21.5.J. EXERCISE (NOT EVERYTHING FITS INTO THIS TRICHOTOMY). Suppose C is a smooth projective irreducible complex curve of genus greater than 1. Show that $C \times_{\mathbb{C}} \mathbb{P}^1_{\mathbb{C}}$ is neither Fano, nor Calabi-Yau, nor general type. Hint: Exercise 21.2.U.

21.5.7. ⋆ **Kodaira vanishing.**

The Kodaira vanishing theorem is an important result that is an important tool in a number of different areas of algebraic geometry. We state it for the sake of culture, but do not prove it (and hence will not use it).

21.5.8. The Kodaira Vanishing Theorem. — *Suppose k is a field of characteristic 0, and X is a smooth projective k-variety. Then for any ample invertible sheaf $\mathscr{L}$, $H^i(X, \mathscr{K} \otimes \mathscr{L}) = 0$ for $i > 0$.*

21.5.9. The restriction on characteristic is necessary; Raynaud gave an example where Kodaira vanishing fails in positive characteristic, [**Ra**]. The original proof is by proving the complex-analytic version, and then transferring it into a complex algebraic statement using Serre's GAGA Theorem [**Se2**]. The general characteristic 0 case can be reduced to $\mathbb{C}$ — any reduction of this sort is often called (somewhat vaguely) an application of the *Lefschetz principle*. (For a precise formulation of the Lefschetz principle, with a short proof which applies in any characteristic see [**Ek**]. See [**FR**] for a generalization. See [**MO90551**] for more context.) Raynaud later gave a dramatic algebraic proof of Kodaira vanishing *using* positive characteristic (!), see [**DeI**] or [**Il**].

21.5.K. EXERCISE. Prove the Kodaira Vanishing Theorem in the case $\dim X = 1$, and for $\mathbb{P}^n_k$. (Neither case requires the characteristic 0 hypothesis.)

21.5.10. ⋆ **A first glimpse of Hodge theory.**

The invariant $h^j(X, \Omega^i_{X/k})$ is called the **Hodge number** $h^{i,j}(X)$. By Exercise 21.5.A, $h^{i,0}$ are birational invariants. We will soon see (in Exercise 21.5.O) that this isn't true for all $h^{i,j}$.

21.5.L. EXERCISE. Suppose X is a smooth projective variety over $k = \overline{k}$. Assuming Miracle 18.5.2 (that the canonical bundle is Serre-dualizing), show that Hodge numbers satisfy the symmetry $h^{p,q} = h^{n-p,n-q}$.

21.5.M. EXERCISE (THE HODGE NUMBERS OF PROJECTIVE SPACE). Show that $h^{p,q}(\mathbb{P}^n_k) = 1$ if $0 \le p = q \le n$ and $h^{p,q}(\mathbb{P}^n_k) = 0$ otherwise. Hint: use the Euler exact sequence (Theorem 21.4.6) and apply Exercise 13.5.G.

21.5.11. *Remark: The Hodge diamond.* Over $k = \mathbb{C}$, further miracles occur. If X is an irreducible smooth projective complex variety, then it turns out that there is a direct sum decomposition

$$H^m(X, \mathbb{C}) = \oplus_{i+j=m} H^j(X, \Omega^i_{X/\mathbb{C}}), \tag{21.5.11.1}$$

from which $h^m(X, \mathbb{C}) = \sum_{i+j=m} h^{i,j}$, so the Hodge numbers (purely algebraic objects) yield the Betti numbers (a priori topological information). Moreover, complex conjugation interchanges $H^j(X, \Omega^i_{X/\mathbb{C}})$ with $H^i(X, \Omega^j_{X/\mathbb{C}})$, from which

$$h^{i,j} = h^{j,i}. \tag{21.5.11.2}$$

(Aside: This additional symmetry holds in characteristic 0 in general, but can fail in positive characteristic, see for example [**Ig**, p. 966], [**Mu1**, §II], [**Se3**, Prop. 16].) This is the beginning of the vast and rich subject of Hodge theory (see [**GH1**, §0.6] for more, or [**Vo**] for much more).

If we write the Hodge numbers in a diamond, with $h^{i,j}$ the ith entry in the $(i+j)$th row, then the diamond has the two symmetries coming from Serre duality and complex conjugation. For example, the Hodge diamond of an irreducible smooth projective complex surface will be of the following form:

$$\begin{array}{ccccc} & & 1 & & \\ & q & & q & \\ p_g & & h^{1,1} & & p_g \\ & q & & q & \\ & & 1 & & \end{array}$$

where p_g is the geometric genus of the surface, and $q = h^{0,1} = h^{1,0} = h^{2,1} = h^{1,2}$ is called the **irregularity** of the surface. As another example, by Exercise 21.5.M, the Hodge diamond of $\mathbb{P}^n$ is all 0 except for 1's down the vertical axis of symmetry.

You won't need the unproved statements (21.5.11.1) or (21.5.11.2) to solve the following problems.

21.5.N. EXERCISE. Assuming Miracle 18.5.2 (that the canonical bundle is Serre-dualizing), show that the Hodge diamond of a smooth projective geometrically irreducible genus g curve over a field k is the following.

$$\begin{array}{ccc} & 1 & \\ g & & g \\ & 1 & \end{array}$$

21.5.O. EXERCISE. Show that the Hodge diamond of $\mathbb{P}^1_k \times \mathbb{P}^1_k$ is the following.

$$\begin{array}{ccccc} & & 1 & & \\ & 0 & & 0 & \\ 0 & & 2 & & 0 \\ & 0 & & 0 & \\ & & 1 & & \end{array}$$

By comparing your answer to the Hodge diamond of $\mathbb{P}^2_k$ (Exercise 21.5.M), show that $h^{1,1}$ is not a birational invariant.

Notice that in both cases, $h^{1,1}$ is the Picard number ρ (defined in §18.4.10). In general, $\rho \le h^{1,1}$ (see [**GH1**, §3.5, p. 456-7]).

21.5.12. $\star$ **Aside: Infinitesimal deformations and automorphisms.**

It is beyond the scope of this book to make this precise, but if X is a variety, $H^0(X, \mathscr{T}_X)$ parametrizes infinitesimal automorphisms of X, and $H^1(X, \mathscr{T}_X)$ parametrizes infinitesimal deformations. As an example if $X = \mathbb{P}^1$ (over a field), $\mathscr{T}_{\mathbb{P}^1} \cong \mathscr{O}(2)$ (§21.4.2), so $h^0(\mathbb{P}^1, \mathscr{T}_{\mathbb{P}^1}) = 3$, which is precisely the dimension of the automorphism group of $\mathbb{P}^1$ (Exercise 16.4.B).

21.5.P. EXERCISE. Compute $h^0(\mathbb{P}^n_k, \mathscr{T}_{\mathbb{P}^n_k})$ using the Euler exact sequence (Theorem 21.4.6). Compare this to the dimension of the automorphism group of $\mathbb{P}^n_k$ (Exercise 16.4.B).

21.5.Q. EXERCISE. Show that $H^1(\mathbb{P}^n_A, \mathscr{T}_{\mathbb{P}^n_A}) = 0$. (Thus projective space can't deform, and is "rigid".)

21.5.R. EXERCISE. Assuming Miracle 18.5.2 (that the canonical bundle is Serre-dualizing), compute $h^i(C, \mathscr{T}_C)$ for a genus g smooth projective geometrically irreducible curve over k, for $i = 0$ and 1. You should notice that $h^1(C, \mathscr{T}_C)$ for genus 0, 1, and $g > 1$ is 0, 1, and $3g - 3$ respectively; after doing this, re-read §19.8.2.

21.6 Unramified morphisms

Suppose $\pi : X \to Y$ is a morphism of schemes. The support of the quasicoherent sheaf $\Omega_\pi = \Omega_{X/Y}$ is called the **ramification locus**, and the image of its support, $\pi(\operatorname{Supp} \Omega_{X/Y})$, is called the **branch locus**. If $\Omega_\pi = 0$, we say that π is **formally unramified**, and if π is also furthermore locally of finite type, we say π is **unramified**. (Caution: there is some lack of consensus in the definition of "unramified"; "locally of finite type" is sometimes replaced by "locally of finite presentation", which was the definition originally used in [**Gr-EGA**].)

21.6.A. EASY EXERCISE (EXAMPLES OF UNRAMIFIED MORPHISMS).
(a) Show that locally closed embeddings are unramified.
(b) Show that if S is a multiplicative subset of the ring B, then $\operatorname{Spec} S^{-1}B \to \operatorname{Spec} B$ is formally unramified. (Thus if η is the generic point of an integral scheme Y, then $\operatorname{Spec} \mathscr{O}_{Y,\eta} \to Y$ is formally unramified.)
(c) Show that finite separable field extensions (or more correctly, the corresponding maps of schemes) are unramified.

21.6.B. EXERCISE (PRACTICE WITH THE CONCEPT).
(a) Show that the normalization of the node in Exercise 9.7.E (see Figure 7.4) is unramified.
(b) Show that the normalization of the cusp in Exercise 9.7.F (see Figure 9.4) is not unramified.

21.6.C. EASY EXERCISE (UNRAMIFIED MORPHISMS ARE PRESERVED BY COMPOSITION). Suppose $\pi : X \to Y$ and $\rho : Y \to Z$ are unramified. Show that $\rho \circ \pi$ is unramified.

21.6.D. EXERCISE (CHARACTERIZATIONS OF UNRAMIFIED MORPHISMS BY THEIR FIBERS). Suppose $\pi : X \to Y$ is locally of finite type.
(a) Show that π is unramified if and only if for each $q \in Y$, $\pi^{-1}(q)$ is the (scheme-theoretic) disjoint union of schemes of the form $\operatorname{Spec} K$, where K is a finite separable extension of $\kappa(q)$.
(b) Show that π is unramified if and only if for each *geometric* point $\overline{q}$, $\pi^{-1}(\overline{q}) := \overline{q} \times_Y X$ is the (scheme-theoretic) disjoint union of copies of $\overline{q}$.

21.6.E. UNIMPORTANT EXERCISE (FOR NUMBER THEORISTS). Suppose $\phi : (A, \mathfrak{m}) \to (B, \mathfrak{n})$ is a local homomorphism of local rings. In algebraic number theory, such a ring morphism is said to be *unramified* if $B/\phi(\mathfrak{m})B$ is a finite separable extension of $A/\mathfrak{m}$. Show that if ϕ is finite type, this agrees with our definition.

21.6.F. EXERCISE (FOR USE IN EXERCISE 21.6.G(B) AND EXERCISE 25.2.G).
(a) Suppose $\pi : X \to Y$ is a locally finite type morphism of locally Noetherian schemes. Show that π is unramified if and only if $\delta_\pi : X \to X \times_Y X$ is an open embedding. Hint: Show the following. If $\phi : X \to Z$ is a closed embedding of Noetherian schemes, and the ideal sheaf $\mathscr{I}$ of ϕ satisfies $\mathscr{I} = \mathscr{I}^2$, then ϕ is also an open embedding. For that, show that if $(A, \mathfrak{m})$ is a Noetherian local ring, and I is a proper ideal of A satisfying $I = I^2$, then $I = 0$. For that in turn, use Nakayama version 2 (Lemma 7.2.9). Also use the fact that $\operatorname{Supp} \mathscr{I}$ is closed (using Exercise 13.7.D), so its complement is open.
$\star$(b) Adapt your proof of (a) to drop the locally Noetherian hypothesis. Hint: show that if π is locally of finite type, then δ_π is locally finitely presented.

21.6.G. EASY EXERCISE. Suppose $\pi : X \to Y$ and $\rho : Y \to Z$ are locally of finite type. Let $\tau = \rho \circ \pi$:

$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "\tau"'] & & Y \arrow[dl, "\rho"] \\ & Z & \end{tikzcd}$$

(a) Show that if τ is unramified, then so is π.
(b) Show that if ρ is unramified, and τ is smooth of relative dimension n (e.g. étale if $n = 0$), then π is smooth of relative dimension n.
(Does this agree with your geometric intuition?)

21.6.H. UNIMPORTANT EXERCISE. Suppose $\pi : X \to Y$ is locally of finite type. Show that the locus in X where π is unramified is open.

21.6.1. *Arithmetic side remark: the different and discriminant.* If B is the ring of integers in a number field (§9.7.1), the **different ideal** of B is the annihilator of $\Omega_{B/\mathbb{Z}}$. It measures the failure of $\operatorname{Spec} B \to \operatorname{Spec} \mathbb{Z}$ to be unramified, and is a scheme-theoretic version of the ramification locus. The **discriminant ideal** can be interpreted as the ideal of $\mathbb{Z}$ corresponding to effective divisor on $\operatorname{Spec} \mathbb{Z}$ that is the "push forward" (not defined here, but defined as you might expect) of the divisor corresponding to the different. It is a scheme-theoretic version of the branch locus. If B/A is an extension of rings of integers of number fields, the **relative different ideal** (of B) and **relative discriminant ideal** (of A) are defined similarly. (We won't use these ideas.)

21.7 The Riemann-Hurwitz Formula

The Riemann-Hurwitz formula generalizes our calculation of the genus g of a double cover of $\mathbb{P}^1$ branched at $2g + 2$ points, Theorem 19.5.1, to higher degree covers, and to higher genus target curves.

21.7.1. *Definition.* A finite morphism between integral schemes $X \to Y$ is said to be **separable** if it is dominant, and the induced extension of function fields $K(X)/K(Y)$ is a separable extension. Similarly, a generically finite morphism is **generically**

separable if it is dominant, and the induced extension of function fields is a separable extension. Note that finite morphisms of integral schemes are automatically separable in characteristic 0.

21.7.2. Proposition. — *If $\pi : X \to Y$ is a generically separable morphism of irreducible smooth varieties of the same dimension n, then the relative cotangent sequence (Theorem 21.2.25) is exact on the left as well:*

$$0 \longrightarrow \pi^*\Omega_{Y/k} \xrightarrow{\phi} \Omega_{X/k} \longrightarrow \Omega_{X/Y} \longrightarrow 0. \tag{21.7.2.1}$$

This is an example of left-exactness of the relative cotangent sequence in the presence of appropriate "smoothness", see Remark 21.2.10.

Proof. We must check that ϕ is injective. Now $\Omega_{Y/k}$ is a rank n locally free sheaf on Y, so $\pi^*\Omega_{Y/k}$ is a rank n locally free sheaf on X. A locally free sheaf on an integral scheme (such as $\pi^*\Omega_{Y/k}$) is torsion-free (any section over any open set is nonzero at the generic point, see §13.5.4), so if a subsheaf of it (such as $\ker\phi$) is nonzero, it is nonzero at the generic point. Thus to show the injectivity of ϕ, we need only check that ϕ is an inclusion at the generic point. We thus tensor with $\mathscr{O}_\eta$ where η is the generic point of X. Tensoring with $\mathscr{O}_\eta$ is an exact functor (localization is exact, Exercise 1.6.F), and $\mathscr{O}_\eta \otimes \Omega_{X/Y} = 0$ (as $K(X)/K(Y)$ is a separable extension by hypothesis, and Ω for separable field extensions is 0 by Exercise 21.2.F(a)). Also, $\mathscr{O}_\eta \otimes \pi^*\Omega_{Y/k}$ and $\mathscr{O}_\eta \otimes \Omega_{X/k}$ are both n-dimensional $\mathscr{O}_\eta$-vector spaces (they are the stalks of rank n locally free sheaves at the generic point). Thus by considering

$$\mathscr{O}_\eta \otimes \pi^*\Omega_{Y/k} \to \mathscr{O}_\eta \otimes \Omega_{X/k} \to \mathscr{O}_\eta \otimes \Omega_{X/Y} \to 0$$

(which is $\mathscr{O}_\eta^{\oplus n} \to \mathscr{O}_\eta^{\oplus n} \to 0 \to 0$) we see that $\mathscr{O}_\eta \otimes \pi^*\Omega_{Y/k} \to \mathscr{O}_\eta \otimes \Omega_{X/k}$ is injective, and thus that $\pi^*\Omega_{Y/k} \to \Omega_{X/k}$ is injective. $\square$

People not confined to characteristic 0 should note what goes wrong for non-separable morphisms. For example, suppose k is a field of characteristic p, and consider the map $\pi : \mathbb{A}^1_k = \operatorname{Spec} k[t] \to \mathbb{A}^1_k = \operatorname{Spec} k[u]$ given by $u = t^p$. Then Ω_π is the trivial invertible sheaf generated by dt. As another (similar but different) example, if $K = k(x)$ and $K' = K(x^p)$, then the inclusion $K' \hookrightarrow K$ induces $\pi : \operatorname{Spec} K[t] \to \operatorname{Spec} K'[t]$. Once again, Ω_π is an invertible sheaf, generated by dx (which in this case is pulled back from $\Omega_{K/K'}$ on $\operatorname{Spec} K$). In both of these cases, we have maps from one affine line to another, and there are vertical tangent vectors.

21.7.A. EXERCISE. If X and Y are smooth varieties of dimension n, and $\pi : X \to Y$ is generically separable, show that the ramification locus is pure codimension 1, and has a natural interpretation as an effective divisor, as follows. Interpret ϕ as an $n \times n$ Jacobian matrix (12.1.6.1) in appropriate local coordinates, and hence interpret the locus where ϕ is not an isomorphism as (locally) the vanishing scheme of the determinant of an $n \times n$ matrix. Hence the branch locus is also pure codimension 1. (This is a special case of Zariski's theorem on *purity of (dimension of) the branch locus*.) Hence we use the terms **ramification divisor** and **branch divisor**.

Before getting down to our case of interest, dimension 1, we begin with something (literally) small but fun. Suppose $\pi : X \to Y$ is a surjective k-morphism from a smooth k-scheme that contracts a subset of codimension greater than 1. More precisely, suppose π is an isomorphism over an open subset of Y, from an open subset U of X whose complement has codimension greater than 1. Then by Exercise 21.7.A, Y *cannot* be smooth. (*Small resolutions*, to be defined in Exercise 22.4.N, are examples of such π. In particular, you can find an example there.)

Suppose now that X and Y are dimension 1. Then the ramification locus is a finite set (ramification *points*) of X, and the branch locus is a finite set (branch *points*) of Y. (Figure 21.4 shows a morphism with two ramification points and one branch point.) *Now assume that* $k = \overline{k}$. We examine $\Omega_{X/Y}$ near a point $p \in X$.

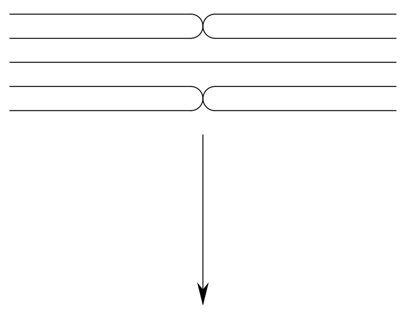


FIGURE 21.4. An example where the branch divisor appears with multiplicity 2 (see Exercise 21.7.C)

As motivation for what we will see, we note that in complex geometry, nonconstant maps from (complex) curves to curves may be written in appropriate local coordinates as $x \mapsto x^m = y$, from which we see that dy pulls back to $mx^{m-1}\,dx$, so $\Omega_{X/Y}$ locally looks like functions times dx modulo multiples of $mx^{m-1}\,dx$.

Consider now our map $\pi : X \to Y$, and fix $p \in X$, and $q = \pi(p)$. Because the construction of Ω behaves well under base change (Theorem 21.2.27(b)), we may replace Y with Spec of the local ring $\mathscr{O}_{Y,q}$ at q, i.e., we may assume $Y = \operatorname{Spec} B$, where B is a discrete valuation ring (as Y is a regular curve), with residue field $k = \overline{k}$ corresponding to q. Then as π is finite, X is affine too. Similarly, as the construction of Ω behaves well with respect to localization (Exercise 21.2.8), we may replace X by $\operatorname{Spec} \mathscr{O}_{X,p}$, and thus assume $X = \operatorname{Spec} A$, where A is a discrete valuation ring, and π corresponds to $B \to A$, inducing an isomorphism of residue fields (with k).

Suppose their uniformizers are s and t respectively, with $t \mapsto us^n$ where u is an invertible element of A.

$$\begin{tikzcd} X \arrow[d] & A & us^n \\ Y & B \arrow[u] & t \arrow[u, mapsto] \end{tikzcd}$$

Recall that the differentials of a discrete valuation ring over k are generated by d of the uniformizer (Exercise 21.4.D). Then

$$dt = d(us^n) = uns^{n-1}\,ds + s^n\,du.$$

This differential on $\operatorname{Spec} A$ vanishes to order at least $n-1$, and precisely $n-1$ if n does not divide the characteristic. The former case is called **tame** ramification, and the latter is called **wild** ramification. We call this order the **ramification order** at this point of X.

21.7.B. EXERCISE. Show that the degree of $\Omega_{X/Y}$ at $p \in X$ is precisely the ramification order of π at p. (The degree of a coherent sheaf on a curve was defined in §18.4.4. To do this exercise, you will have to explain why a coherent sheaf $\mathscr{F}$ supported on a finite set of points has a "degree" at each of these points, which sum to the total degree of $\mathscr{F}$.)

21.7.C. EXERCISE: INTERPRETING THE RAMIFICATION DIVISOR IN TERMS OF NUMBER OF PREIMAGES. Suppose all the ramification above $q \in Y$ is tame (which is always true in characteristic 0). Show that the degree of the branch divisor at q is $\deg \pi - |\pi^{-1}(q)|$. Thus the multiplicity of the branch divisor counts the extent to which the number of preimages is less than the degree (see Figure 21.4).

21.7.3. Theorem (the Riemann-Hurwitz formula). — *Suppose $\pi : X \to Y$ is a finite separable morphism of projective regular curves. Let $n = \deg \pi$, and let R be the ramification divisor. Then*

$$2g(X) - 2 = n(2g(Y) - 2) + \deg R.$$

21.7.D. EXERCISE. Prove the Riemann-Hurwitz formula. Hint: Apply the fact that degree is additive in exact sequences (Exercise 18.4.I) to (21.7.2.1). Recall that degrees of line bundles pull back well under finite morphisms of integral projective curves, Exercise 18.4.F. A torsion coherent sheaf on a reduced curve (such as Ω_π) is supported in dimension 0 (Exercise 13.7.G(b)), so $\chi(\Omega_\pi) = h^0(\Omega_\pi)$. Show that the degree of R as a divisor is the same as its degree in the sense of h^0.

Here are some applications of the Riemann-Hurwitz formula.

21.7.4. *Example.* The degree of R is always even: any cover of a curve must be branched over an even number of points (counted with appropriate multiplicity).

21.7.E. EASY EXERCISE. Show that there is no nonconstant map from a smooth projective irreducible genus 2 curve to a smooth projective irreducible genus 3 curve. (Hint: $\deg R \geq 0$.)

21.7.5. *Example.* If $k = \overline{k}$, the only connected unbranched finite separable cover of $\mathbb{P}^1_k$ is the isomorphism, for the following reason. Suppose X is connected and $X \to \mathbb{P}^1_k$ is unramified. Then X is a curve, and regular by Exercise 25.2.E(a). Applying the Riemann-Hurwitz theorem, using that the ramification divisor is 0, we have $2 - 2g_C = 2d$ with $d \geq 1$ and $g_c \geq 0$, from which $d = 1$ and $g_C = 0$.

21.7.F. EXERCISE. Show that if $k = \overline{k}$ has characteristic 0, the only connected unbranched cover of $\mathbb{A}^1_k$ is itself. (Aside: in characteristic p, this needn't hold; $\operatorname{Spec} k[x,y]/(y^p - y - x) \to \operatorname{Spec} k[x]$ is such a map, as you can show yourself. Once the theory of the algebraic fundamental group is developed, this translates to: "$\mathbb{A}^1$ is not simply connected in characteristic p." This cover is an example of an *Artin-Schreier cover*. Fun fact: the group $\mathbb{Z}/p$ acts on this cover via the map $y \mapsto y+1$. This is an example of a *Galois* cover; you can check that the extension of function fields is Galois.)

21.7.G. UNIMPORTANT EXERCISE. Extend Example 21.7.5 and Exercise 21.7.F, by removing the $k = \overline{k}$ hypothesis, and changing "connected" to "geometrically connected".

21.7.6. *Example: Lüroth's theorem.* Continuing the notation of Theorem 21.7.3, suppose $g(X) = 0$. Then from the Riemann-Hurwitz formula (21.7.2.1), $g(Y) = 0$. (Otherwise, if $g(Y)$ were at least 1, then the right side of the Riemann-Hurwitz formula would be non-negative, and thus couldn't be -2, which is the left side.) Informally: the only maps from a genus 0 curve to a curve of positive genus are the constant maps. This has a nonobvious algebraic consequence, by our identification of covers of curves with field extensions (Theorem 17.4.3): all subfields of $k(x)$ containing k are of the form $k(y)$ where $y = f(x)$ for some $f \in k(x)$.

$$\begin{tikzcd} k(x) & \mathbb{P}^1 \arrow[d] & \\ K(C) \arrow[u, hook] & C & = \mathbb{P}^1 \end{tikzcd}$$

(It turns out that the hypotheses $\operatorname{char} k = 0$ and $k = \overline{k}$ are not necessary.) This is Lüroth's theorem.

21.7.H. EXERCISE. Use Lüroth's Theorem to give new geometric solutions to Exercises 6.5.J and 6.5.L. (These arguments will be less ad hoc, and more suitable for generalization, than the algebraic solutions suggested in the hints to those exercises.)

21.7.I. $\star$ EXERCISE (GEOMETRIC GENUS EQUALS TOPOLOGICAL GENUS). This exercise is intended for those with some complex background, who know that the Riemann-Hurwitz formula holds in the complex analytic category. Suppose C is an irreducible regular projective complex curve. Show that there is an algebraic nonconstant map $\pi : C \to \mathbb{P}^1_{\mathbb{C}}$. Describe the corresponding map of Riemann surfaces. Use the previous exercise to show that the algebraic notion of genus (as computed using the branched cover π) agrees with the topological notion of genus (using the same branched cover). (Recall that assuming Miracle 18.5.2 — that the canonical bundle is Serre-dualizing — we know that the geometric genus equals the arithmetic genus, Exercise 21.5.C.)

21.7.J. UNIMPORTANT EXERCISE. Suppose $\pi : X \to Y$ is a dominant morphism of regular curves, and R is the ramification divisor of π. Show that $\Omega_X(-R) \cong \pi^*\Omega_Y$. (This exercise is geometrically pleasant, but we won't use it.) Hint: This says that we can interpret the invertible sheaf $\pi^*\Omega_Y$ over an open set U of X as precisely those differentials on U vanishing along the ramification divisor.

21.7.7. *Informal example: The degree of the discriminant of degree* d *polynomials in one variable.* You may be aware that there is a degree $2d-2$ polynomial in the coefficients $a_d, \dots, a_0$ of the degree d polynomial

$$f(x) = a_d x^d + a_{d-1}x^{d-1} + \cdots + a_0 = 0$$

that vanishes precisely when $f(x)$ has a multiple root. For example, when $d = 2$, the discriminant is $a_1^2 - 4a_0a_2$. We can "compute" this degree $2d-2$ using the Riemann-Hurwitz formula as follows. (You should try to make sense of the following informal and imprecise discussion.) We work over an algebraically closed field k of characteristic 0 for the sake of simplicity. If we take two general degree d polynomials, $g(x)$ and $h(x)$, the degree of the discriminant "should be" the number of $\lambda \in k$ for which $g(x) - \lambda h(x)$ has a double root. Consider the morphism $\mathbb{P}^1 \to \mathbb{P}^1$ given by $x \mapsto [g(x), h(x)]$. (Here we use "affine coordinates" on the source $\mathbb{P}^1$: by x we mean $[x, 1]$.) Then this morphism has degree d. A ramification point a mapping to the branch point $[\lambda, 1]$ in the target $\mathbb{P}^1$ corresponds to a being a double root of $g(x) - \lambda h(x)$. Thus the number of branch points should be the desired degree of the discriminant. By the Riemann-Hurwitz formula there are $2d-2$ branch points (admittedly, with multiplicity). It is possible to turn this into a proof, and it is interesting to do so.

21.7.8. Bounds on automorphism groups of curves.

It is a nontrivial fact that irreducible smooth projective curves of genus $g \geq 2$ have finite automorphism groups. (See [**FK**, §5.1] for a proof over $\mathbb{C}$. See [**Schm**] for the first proof in arbitrary characteristic, although better approaches are now available.) Granting this fact, we can show that in characteristic 0, the automorphism group has order at most $84(g-1)$ (*Hurwitz's automorphisms theorem*), as follows.

Suppose C is an irreducible smooth projective curve over an algebraically closed field $k = \overline{k}$ of characteristic 0, of genus $g \geq 2$. Suppose that G is a *finite* group of automorphisms of C. We now show that $|G| \leq 84(g-1)$. (The case where k is not algebraically closed is quickly dispatched by base-changing to $\overline{k}$.)

21.7.K. EXERCISE.
(a) Let C' be the smooth projective curve corresponding to the field extension $K(C)^G$ of k (via Theorem 17.4.3). ($K(C)^G$ means the G-invariants of $K(C)$.) Describe a morphism $\pi : C \to C'$ of degree $|G|$, as well as a faithful G-action on C that commutes with π.
(b) Show that above each branch point of π, the preimages are all ramified to the same order (as G acts transitively on them). Suppose there are n branch points and the ith one has ramification r_i (each $|G|/r_i$ times).
(c) Use the Riemann-Hurwitz formula to show that

$$(2g-2) = |G|\left(2g(C') - 2 + \sum_{i=1}^{n} \frac{r_i - 1}{r_i}\right)$$

To maximize $|G|$, we wish to minimize

$$2g(C') - 2 + \sum_{i=1}^{n} \frac{r_i - 1}{r_i} \qquad (21.7.8.1)$$

subject to (21.7.8.1) being positive. Note that $1/42$ is possible: take $g(C') = 0$, $n = 3$, and $(r_1, r_2, r_3) = (2, 3, 7)$.

21.7.L. EXERCISE. Show that you can't do better than $1/42$ by considering the following cases separately:

(a) $g(C') > 1$,
(b) $g(C') = 1$,
(c) $g(C') = 0$ and $n \geq 5$,
(d) $g(C') = 0$ and $n = 4$, and
(e) $g(C') = 0$ and $n = 3$.

21.7.M. EXERCISE. Use the fact that (21.7.8.1) is at least $1/42$ to prove the result.

21.7.9. *Remark.* In positive characteristic, there can be many more automorphisms, see for example [**Stic**].

CHAPTER 22

$\star$ Blowing up

We next discuss an important construction in algebraic geometry, the blow-up of a scheme along a closed subscheme (cut out by a finite type ideal sheaf). The theory could mostly be developed immediately after Chapter 17, but the interpretation in terms of the conormal cone/bundle/sheaf of many classical examples makes it natural to discuss blowing up after differentials.

We won't use blowing up much in later chapters, so feel free to skip this topic for now. But it is an important tool. For example, one can use it to resolve singularities, and more generally, indeterminacy of rational maps. In particular, blow-ups can be used to relate birational varieties to each other.

We will start with a motivational example that will give you a picture of the construction in a particularly important (and the historically earliest) case, in §22.1. We will then see a formal definition, in terms of a universal property, §22.2. The definition won't immediately have a clear connection to the motivational example. We will deduce some consequences of the definition (assuming that the blow-up actually exists). We then prove that the blow-up exists, by describing it quite explicitly, in §22.3. As a consequence, we will find that the blow-up morphism is projective, and we will deduce more consequences from this. In §22.4, we will do a number of explicit computations, to see various sorts of applications, and to see that many things can be computed by hand.

22.1 Motivating example: blowing up the origin in the plane

We will generalize the following notion, which will correspond to "blowing up" the origin of $\mathbb{A}^2_k$ (Exercise 9.3.F). Our discussion will be informal. Consider the subset of $\mathbb{A}^2 \times \mathbb{P}^1$ corresponding to the following. We interpret $\mathbb{P}^1$ as parametrizing the lines through the origin. Consider the subvariety $\mathrm{Bl}_{(0,0)} \mathbb{A}^2 := \{(p \in \mathbb{A}^2, [\ell] \in \mathbb{P}^1) \ : \ p \in \ell)\}$, which is the data of a point p in the plane, and a line ℓ containing both p and the origin. Algebraically: let x and y be coordinates on $\mathbb{A}^2$, and X and Y be projective coordinates on $\mathbb{P}^1$ ("corresponding" to x and y); we will consider the subset $\mathrm{Bl}_{(0,0)} \mathbb{A}^2$ of $\mathbb{A}^2 \times \mathbb{P}^1$ corresponding to $xY - yX = 0$. We have the useful diagram

$$\begin{tikzcd} \mathrm{Bl}_{(0,0)} \mathbb{A}^2 \arrow[r, hook] \arrow[dr, "\beta"'] & \mathbb{A}^2 \times \mathbb{P}^1 \arrow[r] \arrow[d] & \mathbb{P}^1 \\ & \mathbb{A}^2 & \end{tikzcd}$$

You can verify that it is smooth over k (Definition 12.2.6 or 21.3.1) directly (you can now make the paragraph after Exercise 9.3.F precise), but here is an informal argument, using the projection $\mathrm{Bl}_{(0,0)} \mathbb{A}^2 \to \mathbb{P}^1$. The projective line $\mathbb{P}^1$ is smooth, and for each point $[\ell]$ in $\mathbb{P}^1$, we have a smooth choice of points on the line ℓ. Thus we are verifying smoothness by way of a fibration over $\mathbb{P}^1$.

We next consider the projection to $\mathbb{A}^2$, $\beta : \mathrm{Bl}_{(0,0)} \mathbb{A}^2 \to \mathbb{A}^2$. This is an isomorphism away from the origin. Loosely speaking, if p is not the origin, there is precisely one line containing p and the origin. On the other hand, if p *is* the origin, then there is a full $\mathbb{P}^1$ of lines containing p and the origin. Thus the preimage of $(0,0)$ is a curve, and hence a divisor (an effective Cartier divisor, as the blown-up surface is regular). This is called the *exceptional divisor* of the blow-up.

If we have some curve $C \subset \mathbb{A}^2$ singular at the origin, it can be potentially partially desingularized, using the blow-up, by taking the closure of $C \setminus \{(0,0)\}$ in $\mathrm{Bl}_{(0,0)} \mathbb{A}^2$. (A **desingularization** or a **resolution of singularities** of a variety X is a proper birational morphism $\widetilde{X} \to X$ from a regular scheme.) For example, consider the curve $y^2 = x^3 + x^2$, which is regular except for a node at the origin. We can take the preimage of the curve minus the origin, and take the closure of this locus in the blow-up, and we will obtain a regular curve; the two branches of the node downstairs are separated upstairs. (You can check this in Exercise 22.4.B once we have defined things properly. The result will be called the *proper transform* (or *strict transform*) of the curve.) We are interested in desingularizations for many reasons. Because we understand regular curves quite well, we could hope to understand other curves through their desingularizations. This philosophy holds true in higher dimension as well.

More generally, we can blow up $\mathbb{A}^n$ at the origin (or more informally, "blow up the origin"), getting a subvariety of $\mathbb{A}^n \times \mathbb{P}^{n-1}$. Algebraically, If $x_1, \dots, x_n$ are coordinates on $\mathbb{A}^n$, and $X_1, \dots, X_n$ are projective coordinates on $\mathbb{P}^{n-1}$, then the blow-up $\mathrm{Bl}_{\vec{0}} \mathbb{A}^n$ is given by the equations $x_i X_j - x_j X_i = 0$. Once again, this is smooth: $\mathbb{P}^{n-1}$ is smooth, and for each point $[\ell] \in \mathbb{P}^{n-1}$, we have a smooth choice of $p \in \ell$.

We can extend this further, by blowing up $\mathbb{A}^{n+m}$ along a coordinate m-plane $\mathbb{A}^n$ by adding m more variables $x_{n+1}, \dots, x_{n+m}$ to the previous example; we get a subset of $\mathbb{A}^{n+m} \times \mathbb{P}^{n-1}$.

Because in complex geometry, submanifolds of manifolds locally "look like" coordinate m-planes in n-space, you might imagine that we could extend this to blowing up a regular subvariety of a regular variety. In the course of making this precise, we will accidentally generalize this notion greatly, defining the blow-up of any finite type sheaf of ideals in a scheme. In general, blowing up may not have such an intuitive description as in the case of blowing up something regular inside something regular — it can do great violence to the scheme — but even then, it is very useful.

Our description will depend only on the closed subscheme being blown up, and not on coordinates. That remedies a defect was already present in the first example, of blowing up the plane at the origin. It is not obvious that if we picked different coordinates for the plane (preserving the origin as a closed subscheme) that we wouldn't have two different resulting blow-ups.

As is often the case, there are two ways of understanding the notion of blowing up, and each is useful in different circumstances. The first is by universal property, which lets you show some things without any work. The second is an explicit construction, which lets you get your hands dirty and compute things (and implies for example that the blow-up morphism is projective).

The motivating example here may seem like a very special case, but if you understand the blow-up of the origin in n-space well enough, you will understand blowing up in general.

22.2 Blowing up, by universal property

We now define the blow-up by a universal property. The disadvantage of starting here is that this definition won't obviously be the same as (or even related to) the examples of §22.1.

Suppose $X \hookrightarrow Y$ is a closed subscheme corresponding to a finite type sheaf of ideals. (If Y is locally Noetherian, the "finite type" hypothesis is automatic, so Noetherian readers can ignore it.)

The blow-up of $X \hookrightarrow Y$ is a fiber diagram

(22.2.0.1)
$$\begin{tikzcd} E_X Y \arrow[r, hook] \arrow[d] & \mathrm{Bl}_X Y \arrow[d, "\beta"] \\ X \arrow[r, hook] & Y \end{tikzcd}$$

such that $E_X Y$ (the scheme-theoretical pullback of X by β) is an effective Cartier divisor (defined in §8.4.1) on $\mathrm{Bl}_X Y$, such that any other such fiber diagram

(22.2.0.2)
$$\begin{tikzcd} D \arrow[r, hook] \arrow[d] & W \arrow[d] \\ X \arrow[r, hook] & Y, \end{tikzcd}$$

where D is an effective Cartier divisor on W, factors uniquely through it:

$$\begin{tikzcd} D \arrow[r, hook] \arrow[d] & W \arrow[d] \\ E_X Y \arrow[r, hook] \arrow[d] & \mathrm{Bl}_X Y \arrow[d] \\ X \arrow[r, hook] & Y. \end{tikzcd}$$

We call $\mathrm{Bl}_X Y$ the **blow-up** (of Y along X, or of Y with **center** X). (Other somewhat archaic terms for this are *monoidal transformation*, *σ-process*, and *quadratic transformation*, and *dilation*.) We call $E_X Y$ the **exceptional divisor** of the blow-up. (Bl and β stand for "blow-up", and E stands for "exceptional".)

By a typical universal property argument, if the blow-up exists, it is unique up to unique isomorphism. (We can even recast this more explicitly in the language of Yoneda's lemma: consider the category of diagrams of the form (22.2.0.2), where morphisms are diagrams of the form

$$\begin{tikzcd}
D \arrow[r, hook] \arrow[drr] \arrow[ddr] & W \arrow[drr] \arrow[ddr] & & \\
& & D' \arrow[r, hook] \arrow[dl] & W' \arrow[dl] \\
& X \arrow[r, hook] & Y. &
\end{tikzcd}$$

Then the blow-up is a final object in this category, if one exists.)

If $Z \hookrightarrow Y$ is any closed subscheme of Y, then the (scheme-theoretic) pullback $\beta^{-1}Z$ is called the **total transform** of Z. We will soon see that β is an isomorphism away from X (Observation 22.2.2). $\overline{\beta^{-1}(Z-X)}$ is called the **proper transform** or **strict transform** of Z. (We will use the first terminology. We will also define it in a more general situation.) We will soon see (in the Blow-up Closure Lemma 22.2.6) that the proper transform is naturally isomorphic to $\mathrm{Bl}_{Z\cap X} Z$, where $Z \cap X$ is the scheme-theoretic intersection.

We will soon show that the blow-up always exists, and describe it explicitly. We first make a series of observations, *assuming that the blow up exists*.

22.2.1. *Observation.* If X is the empty set, then $\mathrm{Bl}_X Y = Y$. More generally, if X is an effective Cartier divisor, then the blow-up is an isomorphism. (Reason: $\mathrm{id}_Y : Y \to Y$ satisfies the universal property.)

22.2.A. EXERCISE. If U is an open subset of Y, then $\mathrm{Bl}_{U\cap X} U \cong \beta^{-1}(U)$, where $\beta : \mathrm{Bl}_X Y \to Y$ is the blow-up.

Thus "we can compute the blow-up locally."

22.2.B. EXERCISE. Show that if Y_α is an open cover of Y (as α runs over some index set), and the blow-up of Y_α along $X \cap Y_\alpha$ exists, then the blow-up of Y along X exists.

22.2.2. *Observation.* Combining Observation 22.2.1 and Exercise 22.2.A, we see that the blow-up is an isomorphism away from the locus you are blowing up:

$$\beta|_{\mathrm{Bl}_X Y - E_X Y} : \mathrm{Bl}_X Y - E_X Y \to Y - X$$

is an isomorphism.

22.2.3. *Observation.* If $X = Y$, then the blow-up is the empty set: the only map $W \to Y$ such that the pullback of X is a Cartier divisor is $\varnothing \hookrightarrow Y$. In this case we have "blown Y out of existence"!

22.2.C. EXERCISE (BLOW-UP PRESERVES IRREDUCIBILITY AND REDUCEDNESS). Show that if Y is irreducible, and X doesn't contain the generic point of Y, then $\mathrm{Bl}_X Y$ is irreducible. Show that if Y is reduced, then $\mathrm{Bl}_X Y$ is reduced.

22.2.4. Existence in a first nontrivial case: blowing up a locally principal closed subscheme.

We next see why $\mathrm{Bl}_X Y$ exists if $X \hookrightarrow Y$ is locally cut out by one equation. As the question is local on Y (Exercise 22.2.B), we reduce to the affine case $\operatorname{Spec} A/(t) \hookrightarrow$

Spec A. (A good example to think through is $A = k[x,y]/(xy)$ and $t = x$.) Let

$$I = \ker(A \to A_t) = \{a \in A \ : \ t^n a = 0 \text{ for some } n > 0\},$$

and let $\phi : A \to A/I$ be the projection.

22.2.D. EXERCISE. Show that $\phi(t)$ is not a zerodivisor in A/I.

22.2.E. EXERCISE. Show that $\beta : \operatorname{Spec} A/I \to \operatorname{Spec} A$ is the blow up of $\operatorname{Spec} A$ along $\operatorname{Spec} A/t$. In other words, show that

$$\begin{tikzcd} \operatorname{Spec} A/(t,I) \arrow[r] \arrow[d] & \operatorname{Spec} A/I \arrow[d, "\beta"] \\ \operatorname{Spec} A/t \arrow[r] & \operatorname{Spec} A \end{tikzcd}$$

is a "blow up diagram" (22.2.0.1). Hint: In checking the universal property reduce to the case where W (in (22.2.0.2)) is affine. Then solve the resulting problem about rings. Depending on how you proceed, you might find Exercise 10.2.G, about the uniqueness of extension of maps over effective Cartier divisors, helpful.

22.2.F. EXERCISE. Show that $\operatorname{Spec} A/I$ is the scheme-theoretic closure of $D(t)$ in $\operatorname{Spec} A$.

Thus you might geometrically interpret $\operatorname{Spec} A/I \to \operatorname{Spec} A$ as "shaving off any fuzz supported in $V(t)$". In the Noetherian case, this can be interpreted as removing those associated points lying in $V(t)$. This is intended to be vague, and you should think about how to make it precise only if you want to.

22.2.5. The Blow-up Closure Lemma.
Suppose we have a fibered diagram

$$\begin{tikzcd} W \arrow[r, hook, "\text{cl. emb.}"] \arrow[d] & Z \arrow[d] \\ X \arrow[r, hook, "\text{cl. emb.}"] & Y \end{tikzcd}$$

where the bottom closed embedding corresponds to a finite type ideal sheaf (and hence the upper closed embedding does too). The first time you read this, it may be helpful to consider only the special case where $Z \to Y$ is a closed embedding.

Then take the fibered product of this square by the blow-up $\beta : \mathrm{Bl}_X Y \to Y$, to obtain

$$\begin{tikzcd} W \times_Y E_X Y \arrow[r, hook] \arrow[d] & Z \times_Y \mathrm{Bl}_X Y \arrow[d] \\ E_X Y \arrow[r, hook, "\text{Cartier}"] & \mathrm{Bl}_X Y. \end{tikzcd}$$

The bottom closed embedding is locally cut out by one equation, and thus the same is true of the top closed embedding as well. However, the local equation on $Z \times_Y \mathrm{Bl}_X Y$ need not be a non-zerodivisor, and thus the top closed embedding is not necessarily an effective Cartier divisor.

Let $\overline{Z}$ be the scheme-theoretic closure of $Z \times_Y \mathrm{Bl}_X Y \setminus W \times_Y \mathrm{Bl}_X Y$ in $Z \times_Y \mathrm{Bl}_X Y$. (As $W \times_Y \mathrm{Bl}_X Y$ is locally principal, we are in precisely the situation of §22.2.4, so the scheme-theoretic closure is not mysterious.) Note that in the special case where $Z \to Y$ is a closed embedding, $\overline{Z}$ is the proper transform, as defined in §22.2. For this reason, it is reasonable to call $\overline{Z}$ the *proper transform of* Z even if Z *isn't* a closed embedding. Similarly, it is reasonable to call $Z \times_Y \mathrm{Bl}_X Y$ the *total transform of* Z even if Z isn't a closed embedding.

Define $E_{\overline{Z}} \hookrightarrow \overline{Z}$ as the pullback of $E_X Y$ to $\overline{Z}$, i.e., by the fibered diagram

$$\begin{tikzcd}
E_{\overline{Z}} \arrow[r, hook] \arrow[d, hook, "\text{cl. emb.}"] & \overline{Z} \arrow[d, hook, "\text{cl. emb.}"] & \text{proper transform} \\
W \times_Y E_X Y \arrow[r, hook, "\text{loc. prin.}"] \arrow[d] & Z \times_Y \mathrm{Bl}_X Y \arrow[d] & \text{total transform} \\
E_X Y \arrow[r, hook, "\text{Cartier}"] & \mathrm{Bl}_X Y. &
\end{tikzcd}$$

Note that $E_{\overline{Z}}$ is an effective Cartier divisor on $\overline{Z}$. (It is locally cut out by one equation, pulled back from a local equation of $E_X Y$ on $\mathrm{Bl}_X Y$. Can you see why this is locally not a zerodivisor?) It can be helpful to note that the top square of the diagram above is a blow-up square, by Exercises 22.2.E and 22.2.F (and the fact that blow-ups can be computed affine-locally).

22.2.6. Blow-up Closure Lemma. — $(\mathrm{Bl}_W Z, E_W Z)$ *is canonically isomorphic to* $(\overline{Z}, E_{\overline{Z}})$. *More precisely: if the blow-up* $\mathrm{Bl}_X Y$ *exists, then* $(\overline{Z}, E_{\overline{Z}})$ *is the blow-up of* Z *along* W.

This will be very useful. We make a few initial comments. The first three apply to the special case where $Z \to W$ is a closed embedding, and the fourth comment basically tells us we shouldn't have concentrated on this special case.

(1) First, note that if $Z \to Y$ is a closed embedding, then the Blow-Up Closure Lemma states that the proper transform (as defined in §22.2) is the blow-up of Z along the scheme-theoretic intersection $W = X \cap Z$.

(2) In particular, the Blow-Up Closure Lemma lets you actually compute blow-ups, and we will do lots of examples soon. For example, suppose C is a plane curve, singular at a point p, and we want to blow up C at p. Then we could instead blow up the plane at p (which we have already described how to do, even if we haven't yet proved that it satisfies the universal property of blowing up), and then take the scheme-theoretic closure of $C \setminus \{p\}$ in the blow-up.

(3) More generally, if W is some nasty subscheme of Z that we wanted to blow-up, and Z were a finite type k-scheme, then the same trick would work. We could work locally (Exercise 22.2.A), so we may assume that Z is affine. If W is cut out by r equations $f_1, \dots, f_r \in \Gamma(\mathscr{O}_Z)$, then complete the f's to a generating set $f_1, \dots, f_n$ of $\Gamma(\mathscr{O}_Z)$. This gives a closed embedding $Y \hookrightarrow \mathbb{A}^n$ such that W is the scheme-theoretic intersection of Y with a coordinate linear space $\mathbb{A}^r$.

22.2.7. (4) Most generally still, this reduces the existence of the blow-up to a specific special case. (If you prefer to work over a fixed field k, feel free to replace $\mathbb{Z}$ by k in this discussion.) Suppose that for each n, $\mathrm{Bl}_{(x_1,\dots,x_n)} \operatorname{Spec} \mathbb{Z}[x_1, \dots, x_n]$ exists. Then I claim that the blow-up always exists. Here's why. We may assume that Y is affine, say $\operatorname{Spec} B$, and $X = \operatorname{Spec} B/(f_1, \dots, f_n)$. Then we have a morphism $Y \to \mathbb{A}^n_{\mathbb{Z}}$

given by $x_i \mapsto f_i$, such that X is the scheme-theoretic pullback of the origin. Hence by the blow-up closure lemma, $\mathrm{Bl}_X Y$ exists.

22.2.G. $\star$ TRICKY EXERCISE. Prove the Blow-up Closure Lemma 22.2.6. Hint: obviously, construct maps in both directions, using the universal property. Constructing the following diagram may or may not help.

$$\begin{tikzcd}
 & & E_{\overline{Z}} \arrow[rr, hook, "\text{Cartier}"] \arrow[dll] \arrow[ddr, hook] & & \overline{Z} \arrow[dll] \arrow[ddr, hook] & & \\
E_W Z \arrow[rr, hook, "\text{Cartier}"] \arrow[drrr] \arrow[ddr] & & \mathrm{Bl}_W Z \arrow[drrr] \arrow[ddr] & & & & \\
 & & & Z \times_Y E_X Y \arrow[rr, hook, "\text{loc. prin.}"] \arrow[dll] \arrow[ddr] & & Z \times_Y \mathrm{Bl}_X Y \arrow[dll] \arrow[ddr] & \\
 & W \arrow[rr, hook] \arrow[ddr] & & Z \arrow[ddr] & & & \\
 & & & & E_X Y \arrow[rr, hook, "\text{Cartier}"] \arrow[dll] & & \mathrm{Bl}_X Y \arrow[dll] \\
 & & X \arrow[rr, hook] & & Y & &
\end{tikzcd}$$

Hooked arrows indicate closed embeddings; and when morphisms are furthermore locally principal or even effective Cartier, they are so indicated. Exercise 10.2.G, on the uniqueness of extension of maps over effective Cartier divisors, may or may not help as well. Note that if $Z \to Y$ is actually a closed embedding, then so is $Z \times_Y \mathrm{Bl}_X Y \to \mathrm{Bl}_X Y$ and hence also $\overline{Z} \to \mathrm{Bl}_X Y$.

22.3 The blow-up exists, and is projective

22.3.1. It is now time to show that the blow up always exists. We will see two arguments, which are enlightening in different ways. Both will imply that the blow-up morphism is projective, and hence quasicompact, proper, finite type, and separated. In particular, if $Y \to Z$ is quasicompact (resp. proper, finite type, separated), so is $\mathrm{Bl}_X Y \to Z$. (And if $Y \to Z$ is projective, and Z is quasicompact, then $\mathrm{Bl}_X Y \to Z$ is projective. See the solution to Exercise 17.3.B for the reason for this annoying extra hypothesis.) The blow-up of a k-variety is a k-variety (using the fact that reducedness is preserved, Exercise 22.2.C), and the blow-up of a irreducible k-variety is a irreducible k-variety (using the fact that irreducibility is preserved, also Exercise 22.2.C),

Approach 1. As explained in §22.2.7, it suffices to show that $\mathrm{Bl}_{V(x_1,\dots,x_n)} \operatorname{Spec} \mathbb{Z}[x_1,\dots,x_n]$ exists. But we know what it is supposed to be: the locus in $\operatorname{Spec} \mathbb{Z}[x_1,\dots,x_n] \times \operatorname{Proj} \mathbb{Z}[X_1,\dots,X_n]$ cut out by the equations $x_i X_j - x_j X_i = 0$. We will show this by the end of the section.

Approach 2. We can describe the blow-up all at once as a $\mathcal{P}roj$.

22.3.2. Theorem ($\mathcal{P}roj$ description of the blow-up). — *Suppose $X \hookrightarrow Y$ is a closed subscheme cut out by a finite type quasicoherent sheaf of ideals $\mathscr{I} \hookrightarrow \mathscr{O}_Y$. Then*

$$\mathcal{P}roj_Y \left(\mathscr{O}_Y \oplus \mathscr{I} \oplus \mathscr{I}^2 \oplus \mathscr{I}^3 \oplus \cdots \right) \to Y$$

satisfies the universal property of blowing up.

(We made sense of products of ideal sheaves, and hence $\mathscr{I}^n$, in Exercise 14.3.D.)

We will prove Theorem 22.3.2 soon (§22.3.3), after seeing what it tells us. Because I is finite type, the graded sheaf of algebras has degree 1 piece that is finite type. The graded sheaf of algebras is also clearly generated in degree 1. Thus the sheaf of algebras satisfy Hypotheses 17.2.1 ("finite generation in degree 1").

But first, we should make sure that the preimage of X is indeed an effective Cartier divisor. We can work affine-locally (Exercise 22.2.A), so we may assume that $Y = \operatorname{Spec} B$, and X is cut out by the finitely generated ideal I. Then

$$\operatorname{Bl}_X Y = \operatorname{Proj}_B \left(B \oplus I \oplus I^2 \oplus \cdots \right).$$

(You may recall that the ring $B \oplus I \oplus \cdots$ is called the *Rees algebra* of the ideal I in B, §12.9.1.) We are slightly abusing notation by using the notation $\operatorname{Bl}_X Y$, as we haven't yet shown that this satisfies the universal property.

The preimage of X isn't just any effective Cartier divisor; it corresponds to the invertible sheaf $\mathscr{O}(1)$ on this $\mathcal{P}roj$. Indeed, $\mathscr{O}(1)$ corresponds to taking our graded ring, chopping off the bottom piece, and sliding all the graded pieces to the left by 1 (§15.2); it is the invertible sheaf corresponding to the graded module

$$I \oplus I^2 \oplus I^3 \oplus \cdots$$

(where that first summand I has grading 0). But this can be interpreted as the scheme-theoretic pullback of X, which corresponds to the ideal I of B:

$$I \left(B \oplus I \oplus I^2 \oplus \cdots \right) \hookrightarrow B \oplus I \oplus I^2 \oplus \cdots.$$

Thus the scheme-theoretic pullback of $X \hookrightarrow Y$ to $\mathcal{P}roj(\mathscr{O}_Y \oplus \mathscr{I} \oplus \mathscr{I}^2 \oplus \cdots)$, the invertible sheaf corresponding to $\mathscr{I} \oplus \mathscr{I}^2 \oplus \mathscr{I}^3 \oplus \cdots$, is an effective Cartier divisor in class $\mathscr{O}(1)$. Once we have verified that this construction is indeed the blow-up, this divisor will be our exceptional divisor $E_X Y$.

Moreover, we see that the exceptional divisor can be described beautifully as a $\mathcal{P}roj$ over X:

$$E_X Y = \mathcal{P}roj_X \left(\mathscr{O}_Y/\mathscr{I} \oplus \mathscr{I}/\mathscr{I}^2 \oplus \mathscr{I}^2/\mathscr{I}^3 \oplus \cdots \right). \tag{22.3.2.1}$$

We will later see that in good circumstances (if X is a regular embedding in Y), this is a projectivization of a vector bundle (the "projectivized normal bundle"), see Exercise 22.3.D(a).

22.3.3. Proof of the universal property, Theorem 22.3.2. Let's prove that this $\mathcal{P}roj$ construction satisfies the universal property. Then Approach 1 will also follow, as a special case of Approach 2.

22.3.4. *Aside: why approach 1?.* Before we begin, you may be wondering why we bothered with Approach 1. One reason is that you may find it more comfortable to work with this one nice ring, and the picture may be geometrically clearer to you (in the same way that thinking about the Blow-up Closure Lemma 22.2.6

in the case where $Z \to Y$ is a closed embedding is more intuitive). Another reason is that, as you will find in the exercises, you will see some facts more easily in this explicit example, and you can then pull them back to more general examples. Perhaps most important, Approach 1 lets you actually compute blow-ups by working affine-locally: if $f_1, \dots, f_n$ are elements of a ring A, cutting a subscheme $X = \operatorname{Spec} A/(f_1, \dots, f_n)$ of $Y = \operatorname{Spec} A$, then $\mathrm{Bl}_X Y$ can be interpreted as a closed subscheme of $\mathbb{P}^{n-1}_A$, by pulling back from $\mathrm{Bl}_{V(x_1,\dots,x_n)} \operatorname{Spec} \mathbb{Z}[x_1, \dots, x_n]$, and taking the closure of the locus "above X" as dictated by the Blow-up Closure Lemma 22.2.6.

Proof. Reduce to the case of affine target $\operatorname{Spec} R$ with ideal $I \subset R$. Reduce to the case of affine source, with principal effective Cartier divisor t. (A principal effective Cartier divisor is locally cut out by a single non-zerodivisor.) Thus we have reduced to the case $\operatorname{Spec} S \to \operatorname{Spec} R$, corresponding to $f : R \to S$. Say $(x_1, \dots, x_n) = I$, with $(f(x_1), \dots, f(x_n)) = (t)$. We will describe *one* map $\operatorname{Spec} S \to \operatorname{Proj} R[I]$ that will extend the map on the open set $\operatorname{Spec} S_t \to \operatorname{Spec} R$. It is then unique, by Exercise 10.2.G. We map $R[I]$ to S as follows: the degree one part is $f : R \to S$, and $f(X_i)$ (where X_i corresponds to x_i, except it is in degree 1) goes to $f(x_i)/t$. Hence an element X of degree d goes to $X/(t^d)$. On the open set $D_+(X_1)$, we get the map $R[X_2/X_1, \dots, X_n/X_1]/(x_2 - X_2/X_1 x_1, \dots, x_i X_j - x_j X_i, \dots) \to S$ (where there may be many relations) which agrees with f away from $D(t)$. Thus this map does extend away from $V(I)$. $\square$

Here are some applications and observations arising from this construction of the blow-up. First, we can verify that our initial motivational examples are indeed blow-ups. For example, blowing up $\mathbb{A}^2$ (with coordinates x and y) at the origin yields: $B = k[x, y]$, $I = (x, y)$, and $\operatorname{Proj}(B \oplus I \oplus I^2 \oplus \cdots) = \operatorname{Proj} B[X, Y]$ where the elements of B have degree 0, and X and Y are degree 1 and "correspond to" x and y respectively.

22.3.5. *Normal bundles to exceptional divisors.* We will soon see that the normal bundle to a Cartier divisor D is the (space associated to the) invertible sheaf $\mathscr{O}(D)|_D$, the invertible sheaf corresponding to the D on the total space, then restricted to D (Exercise 21.2.H). Thus in the case of the blow-up of a point in the plane, the exceptional divisor has normal bundle $\mathscr{O}(-1)$. (As an aside: Castelnuovo's criterion, Theorem 29.7.1, states that conversely given a smooth surface containing $E \cong \mathbb{P}^1$ with normal bundle $\mathscr{O}(-1)$, E can be blown-down to a point on another smooth surface.) In the case of the blow-up of a regular subvariety of a regular variety, the blow up turns out to be regular (see Theorem 22.3.10), the exceptional divisor is a projective bundle over X, and the normal bundle to the exceptional divisor restricts to $\mathscr{O}(-1)$ (Exercise 22.3.D).

22.3.A. HARDER BUT ENLIGHTENING EXERCISE. If $X \hookrightarrow \mathbb{P}^n$ is a projective scheme, identify the exceptional divisor of the blow up of the affine cone over X (§8.2.12) at the origin with X itself, and that its normal bundle (§22.3.5) is isomorphic to $\mathscr{O}_X(-1)$. (In the case $X = \mathbb{P}^1$, we recover the blow-up of the plane at a point. In particular, we recover the important fact that the normal bundle to the exceptional divisor is $\mathscr{O}(-1)$.)

22.3.6. *The normal cone.* Motivated by (22.3.2.1), as well as Exercise 22.3.D below, we make the following definition. If X is a closed subscheme of Y cut out by $\mathscr{I}$, then the **normal cone** $N_X Y$ of X in Y is defined as

$$N_X Y := \mathcal{S}pec_X \left(\mathscr{O}_Y/\mathscr{I} \oplus \mathscr{I}/\mathscr{I}^2 \oplus \mathscr{I}^2/\mathscr{I}^3 \oplus \cdots \right). \tag{22.3.6.1}$$

This can profitably be thought of as an algebro-geometric version of a "tubular neighborhood". But some cautions are in order. If Y is smooth, $N_X Y$ may not be smooth. (You can work out the example of $Y = \mathbb{A}^2_k$ and $X = V(xy)$.) And even if X and Y are smooth, then although $N_X Y$ is smooth (as we will see shortly, Exercise 22.3.D), it doesn't "embed" in any way in Y (see Remark 22.3.9).

If X is a closed point p, then the normal cone is called the **tangent cone** to Y at p. The **projectivized tangent cone** is the exceptional divisor $E_X Y$ (the $\mathcal{P}roj$ of the same graded sheaf of algebras). Following §8.2.13, the tangent cone and the projectivized tangent cone can be put together in the projective completion of the tangent cone, which contains the tangent cone as an open subset, and the projectivized tangent cone as a complementary effective Cartier divisor.

In Exercise 22.3.D, we will see that at a regular point of Y, the tangent cone may be identified with the tangent space, and the normal cone may often be identified with the total space of the normal bundle.

22.3.B. EXERCISE. Suppose $Y = \operatorname{Spec} k[x,y]/(y^2 - x^2 - x^3)$ (the bottom of Figure 7.4). Assume (to avoid distraction) that $\operatorname{char} k \neq 2$. Show that the tangent cone to Y at the origin is isomorphic to $\operatorname{Spec} k[x,y]/(y^2 - x^2)$. Thus, informally, the tangent cone "looks like" the original variety "infinitely magnified".

22.3.C. EXERCISE. Suppose $S_\bullet$ is a finitely generated graded algebra over a field k. Exercise 22.3.A gives an isomorphism of $\operatorname{Proj} S_\bullet$ with the exceptional divisor to the blow-up of $\operatorname{Spec} S_\bullet$ at the origin. Show that the tangent cone to $\operatorname{Spec} S_\bullet$ at the origin is isomorphic to $\operatorname{Spec} S_\bullet$ itself. (Your geometric intuition should lead you to find these facts believable.)

22.3.7. Blowing up regular embeddings.

The case of blow-ups of regular embeddings $X \subset Y$ is particularly pleasant. For example, the exceptional divisor is a projective bundle over X. The central reason is the following result.

22.3.8. Theorem. — *If $I \subset A$ is generated by a regular sequence $a_1, \ldots, a_d$, then the natural map $\operatorname{Sym}^n_A(I/I^2) \to I^n/I^{n+1}$ is an isomorphism.*

Hence if a closed embedding $i : X \hookrightarrow Y$ is a regular embedding with ideal sheaf $\mathscr{I} \subset \mathscr{O}_Y$, then the natural map $\operatorname{Sym}^n(\mathscr{I}/\mathscr{I}^2) \to \mathscr{I}^n/\mathscr{I}^{n+1}$ is an isomorphism. Furthermore, in combination with Proposition 21.2.16, we see that $\mathscr{I}^n/\mathscr{I}^{n+1}$ is a locally free sheaf.

Before starting the proof of Theorem 22.3.8 in §22.3.11, we show its utility.

22.3.D. EXERCISE (ASSUMING THEOREM 22.3.8).
(a) Suppose $X \to Y$ is a regular embedding with ideal sheaf $\mathscr{I}$, identify the total space (§17.1.4) of the normal sheaf (the "normal bundle") with the normal cone $N_X Y$ (22.3.6.1), and show that the exceptional divisor $E_X Y$ is a projective bundle (the "projectivized normal bundle") over X.

(b) Show that the normal bundle to $E_X Y$ in Y is $\mathscr{O}(-1)$ (for the projective bundle over X).
(c) Assume further that X is a reduced closed point p. Show that p is a regular point of X. Identify the total space of the tangent space to p with the tangent cone to Y at p.

22.3.9. *Remark.* We can now make sense of a comment made in §22.3.6, that even if X and Y are smooth, then although $N_X Y$ is smooth, it needn't admit an open embedding in Y. To do this, start with a smooth complex quartic surface Y containing a line X. (Most smooth quartic surfaces don't contain a line, by Exercise 11.2.J, so you will have to construct one by hand.) Then $N_X Y$ is a line bundle over X, and thus rational (i.e., birational to $\mathbb{A}^2_{\mathbb{C}}$, Definition 6.5.4). But $N_X Y$ cannot admit an open embedding into Y, as otherwise Y would be rational, contradicting Exercise 21.5.D.

22.3.10. Theorem. — *Suppose $X \hookrightarrow Y$ is a closed embedding of smooth varieties over k. Then $\mathrm{Bl}_X Y$ is also smooth.*

Proof. (We use the fact that smooth varieties are regular, the Smoothness-Regularity Comparison Theorem 12.2.10(b), whose proof we still have to complete.)

We need only check smoothness of $\mathrm{Bl}_X Y$ at the points of $E_X Y$ (by Observation 22.2.2). By Exercise 12.2.K, $X \hookrightarrow Y$ is a regular embedding. Then by Exercise 22.3.D(a), $E_X Y$ is a projective bundle over X, and thus smooth, and hence regular at its closed points. But $E_X Y$ is an effective Cartier divisor on $\mathrm{Bl}_X Y$. By the slicing criterion for regularity (Exercise 12.2.B), it follows that $\mathrm{Bl}_X Y$ is regular at the closed points of $E_X Y$, hence smooth at all points of $E_X Y$. □

22.3.11. ⋆ *Proving Theorem 22.3.8.*

The proof of Theorem 22.3.8 may reasonably be skipped on a first reading. We prove Theorem 22.3.8 following [**F**, A.6.1], which in turn follows [**Da**]. The proof will be completed in §22.3.13. To begin, let α be the map of graded rings

$$\alpha : (A/I)[X_1, \dots, X_d] \to \oplus_{n=0}^{\infty} I^n/I^{n+1}$$

which takes X_i to the image of a_i in I/I^2. Clearly α is surjective.

22.3.E. Exercise. Show that Theorem 22.3.8 would follow from the statement that α is an isomorphism.

Because a_1 is a non-zerodivisor, we can interpret $A[a_2/a_1, \dots, a_d/a_1]$ as a subring of the total fraction ring (defined in §5.5.7). In particular, as A is a subring of its total fraction ring, the map $A \to A[a_2/a_1, \dots, a_d/a_1]$ is an injection. Define

$$\beta : A[T_2, \dots, T_d] \to A[a_2/a_1, \dots, a_d/a_1]$$

by $T_i \mapsto a_i/a_1$. Clearly, the map β is surjective, and $L_i := a_1 T_i - a_i$ lies in $\ker(\beta)$.

22.3.12. Lemma. — *The kernel of β is $(L_2, \dots, L_d)$.*

Proof. We prove the result by induction on d. We consider first the base case $d = 2$. Suppose $F[T_2] \in \ker\beta$, so $F(a_2/a_1) = 0$. Then applying the algorithm for the remainder theorem, dividing $a_1^{\deg F} F(T_2)$ by $a_1 T_2 - a_2 = L_2$,

$$a_1^{\deg F} F(T_2) = G(T_2)(a_1 T_2 - a_2) + R \tag{22.3.12.1}$$

where $G(T_2) \in A[T_2]$, and $R \in A$ is the remainder. Substituting $T_2 = a_2/a_1$ (and using the fact that $A \to A[a_2/a_1]$ is injective), we have that $R = 0$. Then $(a_1 T_2 - a_2)G(T_2) \equiv 0 \pmod{(a_1^{\deg F})}$. Using the fact that a_2 is a non-zerodivisor modulo $(a_1^{\deg F})$ (as $a_1^{\deg F}$, a_2 is a regular sequence by Exercise 8.4.E), a short induction shows that the coefficients of $G(T_2)$ must all be divisible by $a_1^{\deg F}$. Thus $F(T_2)$ is divisible by $a_1 T_2 - a_2 = L_2$, so the case $d = 2$ is proved.

We now consider the general case $d > 2$, assuming the result for all smaller d. Let $A' = A[a_2/a_1]$. Then $a_1, a_3, a_4, \dots, a_d$ is a regular sequence in A'. Reason: a_1 is a non-zerodivisor in A'. $A'/(a_1) = (A[T_2]/(a_1 T_2 - a_2))/(a_1) = A[T_2]/(a_1 T_2 - a_2, a_1) = A[T_2]/(a_1, a_2)$. Then a_3 is a non-zerodivisor in this ring because it is a nonzero divisor in $A/(a_1, a_2)$, a_4 is a non-zerodivisor in $A[T_2]/(a_1, a_2, a_3)$ because it is a non-zerodivisor in $A/(a_1, a_2, a_3)$, and so forth.

Consider the composition

$$A[T_2, \dots, T_d] \to A'[T_3, \dots, T_d] \to A'[a_3/a_1, \dots, a_d/a_1] = A[a_2/a_1, \dots, a_d/a_1].$$

By the case $d = 2$, the kernel of the first map is L_2. By the inductive hypothesis, the kernel of the second map is $(L_3, \dots, L_d)$. The result follows. □

22.3.13. *Proof of Theorem 22.3.8.* By Exercise 22.3.E, it suffices to prove that the surjection α is an isomorphism. Suppose $F \in \ker(\alpha)$; we wish to show that $F = 0$. We may assume that F is homogeneous, say of degree n. Consider the map $\alpha' : A[X_1, \dots, X_d] \to \oplus_{n=0}^{\infty} I^n/I^{n+1}$ lifting α. Lift F to $A[X_1, \dots, X_d]$, so $F \in \ker(\alpha')$. We wish to show that $F \in IA[X_1, \dots, X_d]$. Suppose $F(a_1, \dots, a_d) = x \in I^{n+1}$. Then we can write x as $F'(a_1, \dots, a_d)$, where F' is a homogeneous polynomial of the same degree as F, with coefficients in I. Then by replacing F by $F - F'$, we are reduced to the following problem: suppose $F \in A[X_1, \dots, X_d]$ is homgeneous of degree n, and $F(a_1, \dots, a_d) = 0$, we wish to show that $F \in IA[X_1, \dots, X_n]$. But if $F(a_1, \dots, a_d) = 0$, then $F(1, a_2/a_1, \dots, a_d/a_1) = 0$ in $A[a_2/a_1, \dots, a_d/a_1]$. Since β is an isomorphism, $F(1, T_2, T_3, \dots, T_d) \in (a_1 T_2 - a_2, a_1 T_3 - a_3, \dots, a_1 T_d - a_d)$. Thus the coefficients of F are in $(a_1, \dots, a_d) = I$ as desired. □

22.4 Examples and computations

In this section we will work through a number of explicit of examples, to get a sense of how blow-ups behave, how they are useful, and how one can work with them explicitly. **Throughout we work over a field k, and we assume throughout that** $\operatorname{char} k = 0$ **to avoid distraction.** The examples and exercises are loosely arranged by topic, but not in order of importance.

22.4.1. Example: Blowing up the plane along the origin.

Let's first blow up the plane $\mathbb{A}^2$ along the origin, and see that the result agrees with our discussion in §22.1. Let x and y be the coordinates on $\mathbb{A}^2$. The blow-up is $\operatorname{Proj} k[x, y, X, Y]$, where $xY - yX = 0$. (Here x and y have degree 0 and X and Y have degree 1.) This is naturally a closed subscheme of $\mathbb{A}^2 \times \mathbb{P}^1$, cut out (in terms of the projective coordinates X and Y on $\mathbb{P}^1$) by $xY - yX = 0$. We consider the two usual patches on $\mathbb{P}^1$: $[X, Y] = [s, 1]$ and $[1, t]$. The first patch yields $\operatorname{Spec} k[x, y, s]/(sy - x)$, and the second gives $\operatorname{Spec} k[x, y, t]/(y - xt)$. Notice that both are smooth: the first is $\operatorname{Spec} k[y, s] \cong \mathbb{A}^2$, and the second is $\operatorname{Spec} k[x, t] \cong \mathbb{A}^2$.

We now describe the exceptional divisor. We first consider the first (s) patch. The ideal is generated by (x, y), which in our ys-coordinates is $(ys, y) = (y)$, which is indeed principal. Thus on this patch the exceptional divisor is generated by y. Similarly, in the second patch, the exceptional divisor is cut out by x. (This can be a little confusing, but there is no contradiction!) This explicit description will be useful in working through some of the examples below.

22.4.A. EXERCISE. Let p be a k-valued point of $\mathbb{P}^2$. Exhibit an isomorphism between $\operatorname{Bl}_p \mathbb{P}^2$ and the Hirzebruch surface $\mathbb{F}_1 = \mathbb{P}_{\mathbb{P}^1}(\mathscr{O}_{\mathbb{P}^1} \oplus \mathscr{O}_{\mathbb{P}^1}(1))$ (Example 17.2.4). (The map $\operatorname{Bl}_p \mathbb{P}^2 \to \mathbb{P}^1$ informally corresponds to taking a point to the line connecting it to the origin. Do not be afraid: You can do this by explicitly working with coordinates.)

22.4.2. Resolving singularities.

22.4.3. *The proper transform of a nodal curve (Figure 22.1).* (You may wish to flip to Figure 7.4 while thinking through this exercise.) Consider next the curve $y^2 = x^3 + x^2$ inside the plane $\mathbb{A}^2$. Let's blow up the origin, and compute the total and proper transform of the curve. (By the Blow-up Closure Lemma 22.2.6, the latter is the blow-up of the nodal curve at the origin.) In the first patch, we get $y^2 - s^2y^2 - s^3y^3 = 0$. This factors: we get the exceptional divisor y with multiplicity two, and the curve $1 - s^2 - s^3y = 0$. You can easily check that the proper transform is regular. Also, notice that the proper transform $\widetilde{C}$ meets the exceptional divisor at two points, $s = \pm 1$. This corresponds to the two tangent directions at the origin (as $s = x/y$).

22.4.B. EXERCISE (FIGURE 22.1). Describe both the total and proper transform of the curve C given by $y = x^2 - x$ in $\operatorname{Bl}_{(0,0)} \mathbb{A}^2$. Show that the proper transform of C is isomorphic to C. Interpret the intersection of the proper transform of C with the exceptional divisor E as the slope of C at the origin.

22.4.C. EXERCISE: BLOWING UP A CUSPIDAL PLANE CURVE (CF. EXERCISE 9.7.F). Describe the proper transform of the cuspidal curve C given by $y^2 = x^3$ in the plane $\mathbb{A}^2$. Show that it is regular. Show that the proper transform of C meets the exceptional divisor E at one point, and is tangent to E there.

The previous two exercises are the first in an important sequence of singularities, which we now discuss.

22.4.D. EXERCISE: RESOLVING A_n CURVE SINGULARITIES. Resolve the singularity $y^2 = x^{n+1}$ in $\mathbb{A}^2$, by first blowing up its singular point, then considering its proper transform and deciding what to do next.

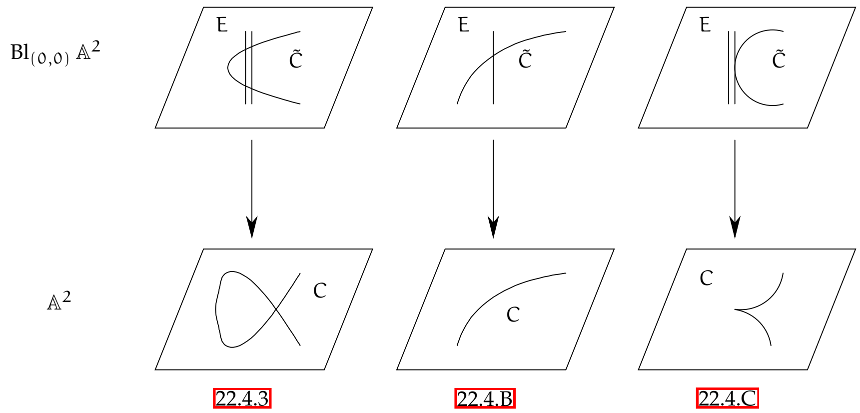


FIGURE 22.1. Resolving curve singularities (§22.4.3, Exercise 22.4.B, and Exercise 22.4.C)

22.4.4. *Toward a definition of* A_n *curve singularities.* You will notice that your solution to Exercise 22.4.D depends only on the "power series expansion" of the singularity at the origin, and not on the precise equation. For example, if you compare your solution to Exercise 22.4.B with the $n = 1$ case of Exercise 22.4.D, you will see that they are "basically the same". We will make this precise in Definition 29.3.C.

22.4.E. EXERCISE (WARM-UP TO EXERCISE 22.4.F). Blow up the cone point $z^2 = x^2 + y^2$ (Figure 3.4) at the origin. Show that the resulting surface is regular. Show that the exceptional divisor is isomorphic to $\mathbb{P}^1$. (Remark: You can check that the normal bundle to this $\mathbb{P}^1$ is not $\mathscr{O}(-1)$, as is the case when you blow up a point on a smooth surface, see §22.3.5; it is $\mathscr{O}(-2)$.)

22.4.F. EXERCISE (RESOLVING A_n SURFACE SINGULARITIES). Resolve the singularity $z^2 = y^2 + x^{n+1}$ in $\mathbb{A}^3$ by first blowing up its singular point, then considering its proper transform, and deciding what to do next. (A k-surface singularity analytically isomorphic this is called an A_n **surface singularity**. For example, the cone shown in Figure 3.4 is an A_1 surface singularity. We make this precise in Exercise 29.3.C.) This exercise is a bit time consuming, but is rewarding in that it shows that you can really resolve singularities by hand.

22.4.5. *Remark:* ADE*-surface singularities and Dynkin diagrams (see Figure 22.2).* A k-singularity analytically isomorphic to $z^2 = x^2 + y^{n+1}$ (resp. $z^2 = x^3 + y^4$, $z^2 = x^3 + xy^3$, $z^2 = x^3 + y^5$) is a called a D_n surface singularity (resp. E_6, E_7, E_8 surface singularity). We will make this precise in Exercise 29.3.C, and you will then be able to guess the definition of the corresponding curve singularity. If you (minimally) desingularize each of these surfaces by sequentially blowing up singular points as in Exercise 22.4.F, and look at the arrangement of exceptional divisors (the various exceptional divisors and how they meet), you will discover the corresponding Dynkin diagram. More precisely, if you create a graph, where the vertices correspond to exceptional divisors, and two vertices are joined by an edge if the two divisors meet, you will find the underlying graph of the corresponding Dynkin diagram. This is the start of several very beautiful stories; see Remark 27.4.4 for a first glimpse of one of them.

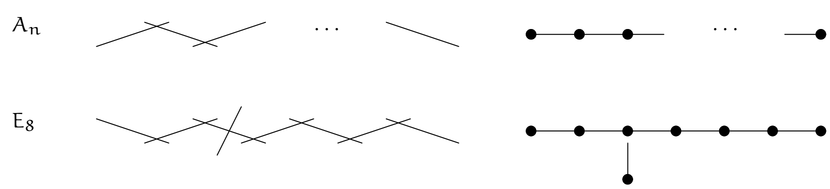


FIGURE 22.2. The exceptional divisors for resolutions of some ADE surface singularities, and their corresponding dual graphs (see Remark 22.4.5)

22.4.6. *Remark: Resolution of singularities.* Hironaka's theorem on resolution of singularities implies that this idea of trying to resolve singularities by blowing up singular loci in general can succeed in characteristic 0. More precisely, if X is a variety over a field of characteristic 0, then X can be resolved by a sequence of blow-ups, where the nth blow-up is along a regular subvariety that lies in the singular locus of the variety produced after the $(n-1)$st stage (see [**Hir**], and [**Ko2**]). (The case of dimension 1 will be shown in §29.5.4, and the case of dimension 2 will be discussed in §29.7.4.) It is not known if an analogous statement is true in positive characteristic (except in dimension at most 2, by Lipman, see [**Ar3**]), but de Jong's Alteration Theorem [**dJ**] gives a result which is good enough for most applications. Rather than producing a birational proper map $\widetilde{X} \to X$ from something regular, it produces a proper map from something regular that is generically finite (and the corresponding extension of function fields is separable).

Here are some other exercises related to resolution of singularities.

22.4.G. EXERCISE. Blowing up a nonreduced subscheme of a regular scheme can give you something singular, as shown in this example. Describe the blow up of the closed subscheme $V(y, x^2)$ in $\operatorname{Spec} k[x,y] = \mathbb{A}^2$. Show that you get an A_1 surface singularity.

22.4.H. EXERCISE. Desingularize the tacnode $y^2 = x^4$, not in two steps (as in Exercise 22.4.D), but in a single step by blowing up (y, x^2).

22.4.I. EXERCISE (RESOLVING A SINGULARITY BY AN UNEXPECTED BLOW-UP). Suppose Y is the cone $x^2 + y^2 = z^2$, and X is the line cut out by $x = 0$, $y = z$ on Y. Show that $\mathrm{Bl}_X Y$ is regular. (In this case we are blowing up a codimension 1 locus that is not an effective Cartier divisor, see Problem 12.1.3. But it *is* an effective Cartier divisor away from the cone point, so you should expect your answer to be an isomorphism away from the cone point.)

22.4.7. *Multiplicity of an function at a point of a regular scheme.* In order to pose Exercise 22.4.J, we introduce a useful concept. If f is a function on a locally Noetherian scheme X, its **multiplicity at a regular point** p is the largest m such that f lies in the mth power of the maximal ideal in the local ring $\mathscr{O}_{X,p}$. For example, if $f \neq 0$, $V(f)$ is singular at p if and only if $m > 1$. (Do you see why?)

22.4.J. EXERCISE. Show that the multiplicity of the exceptional divisor in the total transform of a subscheme Z of $\mathbb{A}^n$ when you blow up the origin is the smallest multiplicity (at the origin) of a defining equation of Z. (For example, in the case of the nodal and cuspidal curves above, Example 22.4.3 and Exercise 22.4.C respectively, the exceptional divisor appears with multiplicity 2.)

22.4.8. Resolving rational maps.

22.4.K. EXERCISE (UNDERSTANDING THE BIRATIONAL MAP $\mathbb{P}^2 \dashrightarrow \mathbb{P}^1 \times \mathbb{P}^1$ VIA BLOW-UPS). Let p and q be two distinct k-points of $\mathbb{P}^2$, and let r be a k-point of $\mathbb{P}^1 \times \mathbb{P}^1$. Describe an isomorphism $\mathrm{Bl}_{\{p,q\}} \mathbb{P}^2 \leftrightarrow \mathrm{Bl}_r \mathbb{P}^1 \times \mathbb{P}^1$. (Possible hint: Consider lines ℓ through p and m through q; the choice of such a pair corresponds to the parametrized by $\mathbb{P}^1 \times \mathbb{P}^1$. A point s of $\mathbb{P}^2$ not on line pq yields a pair of lines $(\overline{ps}, \overline{qs})$ of $\mathbb{P}^1 \times \mathbb{P}^1$. Conversely, a choice of lines (ℓ, m) such that neither ℓ and m is line $\overline{pq}$ yields a point $s = \ell \cap m \in \mathbb{P}^2$. This describes a birational map $\mathbb{P}^2 \dashrightarrow \mathbb{P}^1 \times \mathbb{P}^1$. Exercise 22.4.A is related.)

Exercise 22.4.K is an example of the general phenomenon explored in the next two exercises.

22.4.L. HARDER BUT USEFUL EXERCISE (BLOW-UPS RESOLVE BASE LOCI OF RATIONAL MAPS TO PROJECTIVE SPACE). Suppose we have a scheme Y, an invertible sheaf $\mathscr{L}$, and a number of sections $s_0, \dots, s_n$ of $\mathscr{L}$ (a *linear series*, Definition 15.3.6). Then away from the closed subscheme X cut out by $s_0 = \cdots = s_n = 0$ (the base locus of the linear series), these sections give a morphism to $\mathbb{P}^n$. Show that this morphism extends uniquely to a morphism $\mathrm{Bl}_X Y \to \mathbb{P}^n$, where this morphism corresponds to the invertible sheaf $(\beta^*\mathscr{L})(-E_X Y)$, where $\beta : \mathrm{Bl}_X Y \to Y$ is the blow-up morphism. In other words, "blowing up the base scheme resolves this rational map". Hint: it suffices to consider an affine open subset of Y where $\mathscr{L}$ is trivial. Uniqueness might use Exercise 10.2.G.

22.4.9. *Remarks.* (i) Exercise 22.4.L immediately implies that blow-ups can be used to resolve rational maps to projective schemes $Y \dashrightarrow Z \hookrightarrow \mathbb{P}^n$.

(ii) The following interpretation is enlightening. The linear series on Y pulls back to a linear series on $\mathrm{Bl}_X Y$, and the base locus of the linear series on Y pulls back to the base locus on $\mathrm{Bl}_X Y$. The base locus on $\mathrm{Bl}_X Y$ is $E_X Y$, an effective Cartier divisor. Because $E_X Y$ is not just locally principal, but also locally a non-zerodivisor, it can be "divided out" from the $\beta^* s_i$ (yielding a section of $(\beta^*\mathscr{L})(-E_X Y)$, thereby removing the base locus, and leaving a base-point-free linear series. (In a sense that can be made precise through the universal property, this is the smallest "modification" of Y that can remove the base locus.) If X is already Cartier (as for example happens with any nontrivial linear system if Y is a regular pure-dimensional curve), then we can remove a base locus by just "dividing out X".

(iii) You may wish to revisit Exercise 19.7.B, and interpret it in terms of Exercise 22.4.L.

22.4.10. *Examples.* (i) The rational map $\mathbb{P}^n \dashrightarrow \mathbb{P}^{n-1}$ given by $[x_0, \cdots, x_n] \dashrightarrow [x_1, \cdots, x_n]$, defined away from $p = [1, 0, \cdots, 0]$, is resolved by blowing up p. Then by the Blow-up Closure Lemma 22.2.6, if Y is any locally closed subscheme of $\mathbb{P}^n$, we can project to $\mathbb{P}^{n-1}$ once we blow up p in Y, and the invertible sheaf giving the map to $\mathbb{P}^{n-1}$ is (somewhat informally speaking) $\beta^*(\mathscr{O}_{\mathbb{P}^n}(1)) \otimes \mathscr{O}(-E_p Y)$.

(ii) Consider two general cubic equations C_1 and C_2 in three variables, yielding two cubic curves in $\mathbb{P}^2$. They are smooth (by Bertini's Theorem 12.4.2, and meet in 9 points $p_1, \dots, p_9$ (using our standing assumption that we work over an algebraically closed field). Then $[C_1, C_2]$ gives a rational map $\mathbb{P}^2 \dashrightarrow \mathbb{P}^1$. To resolve the rational map, we blow up $p_1, \dots, p_9$. The result is (generically) an *elliptic fibration* $\mathrm{Bl}_{\{p_1,\dots,p_9\}} \mathbb{P}^2 \to \mathbb{P}^1$. (This is by no means a complete argument.)

(iii) Fix six general points $p_1, \dots, p_6$ in $\mathbb{P}^2$. There is a four-dimensional vector space of cubics vanishing at these points, and they vanish scheme-theoretically precisely at these points. This yields a rational map $\mathbb{P}^2 \dashrightarrow \mathbb{P}^3$, which is resolved by blowing up the six points. The resulting morphism turns out to be a closed embedding, and the image in $\mathbb{P}^3$ is a (smooth) cubic surface. This is the famous fact that the blow up of the plane at six general points may be represented as a (smooth) cubic in $\mathbb{P}^3$. (Again, this argument is not intended to be complete.) See 27.4.2 for a more precise and more complete discussion.

In reasonable circumstances, Exercise 22.4.L has an interpretation in terms of graphs of rational maps.

22.4.M. EXERCISE. Suppose $s_0, \dots, s_n$ are sections of an invertible sheaf $\mathscr{L}$ on an integral scheme X, not all 0. By Remark 16.4.3, these data give a rational map $\phi : X \dashrightarrow \mathbb{P}^n$. Give an isomorphism between the graph of ϕ (§10.2.4) and $\mathrm{Bl}_{V(s_0,\dots,s_n)} X$. (Your argument will not require working over a field k; it will work in general.)

You may enjoy exploring the previous idea by working out how the Cremona transformation $\mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ (Exercise 6.5.I) can be interpreted in terms of the graph of the rational map $[x, y, z] \dashrightarrow [1/x, 1/y, 1/z]$.

22.4.N. $\star$ EXERCISE. Resolve the rational map

$$\operatorname{Spec} k[w, x, y, z]/(wz - xy) \xrightarrow{[w,x]} \mathbb{P}^1$$

from the cone over the quadric surface to the projective line. Let X be the resulting variety, and $\pi : X \to \operatorname{Spec} k[w, x, y, z]/(wz - xy)$ the projection to the cone over the quadric surface. Show that π is an isomorphism away from the cone point, and that the preimage of the cone point is isomorphic to $\mathbb{P}^1$ (and thus has codimension 2, and therefore is different from the resolution obtained by simply blowing up the cone point). Possible hint: if Q is the quadric in $\mathbb{P}^3$ cut out by $wz - xy = 0$, then factor the rational map as $\operatorname{Spec} k[w, x, y, z]/(wz - xy) \setminus \{0\} \to Q$ (cf. Exercise 8.2.P), followed by the isomorphism $Q \to \mathbb{P}^1 \times \mathbb{P}^1$ (Example 9.6.2), followed by projection onto one of the factors.

This is an example of a small resolution. (A **small resolution** $X \to Y$ is a resolution where the space of points of Y where the fiber has dimension r is of codimension greater than $2r$. We will not use this notion again in any essential way.) Notice that this resolution of the morphism involves blowing up the base locus $w = x = 0$, which is a cone over one of the lines on the quadric surface $wz = xy$. We are blowing up an effective Weil divisor, which is necessarily not Cartier as the blow-up is not an isomorphism. In Exercise 12.1.E, we saw that (w, x) was not principal, while here we see that (w, x) is not even locally principal. Essentially by Exercise 14.2.R, $V(w, x)$ cannot even be the support of a locally principal divisor.

22.4.11. *Remark: Non-isomorphic small resolutions.* If you instead resolved the map $[w, y]$, you would obtain a similar looking small resolution $\pi' : X' \to \operatorname{Spec} k[w, x, y, z]/(wz-xy)$ (it is an isomorphism away from the origin, and the fiber over the origin is $\mathbb{P}^1$). But it is different! More precisely, there is no morphism $X \to X'$ making the following the diagram commute.

$$\begin{tikzcd} X \arrow[rr] \arrow[dr, "\pi"'] & & X' \arrow[dl, "\pi'"] \\ & \operatorname{Spec} k[w,x,y,z]/(wz-xy) & \end{tikzcd}$$

The birational map $X \dashrightarrow X'$ is called the **Atiyah flop**, [**At1**].

22.4.12. *Factorization of birational maps.* We end our discussion of resolution of rational maps by noting that just as Hironaka's theorem states that one may resolve all singularities of varieties in characteristic by a sequence of blow-ups along smooth centers, the weak factorization theorem (first proved by Włoldarczyk) states that any two birational varieties X and Y in characteristic 0 may be related by blow-ups and blow-downs along smooth centers. More precisely, there are varieties X_0, $\dots, X_n, X_{01}, \dots, X_{(n-1)n}$, with $X_0 = X$ and $X_n = Y$, with morphisms $X_{i(i+1)} \to X_i$ and $X_{i(i+1)} \to X_{i+1}$ $(0 \le i < n)$ which are blow-ups of smooth subvarieties.

22.4.13. Blow-ups and line bundles.

22.4.O. EXERCISE (GENERALIZING EXERCISE 20.2.D). Suppose X is a regular projective surface over k, and p is a k-valued point. Let $\beta : \operatorname{Bl}_p X \to X$ be the blow-up morphism, and let $E = E_p X$ be the exceptional divisor. Consider the exact sequence

$$\mathbb{Z} \xrightarrow{\gamma: 1 \mapsto [E]} \operatorname{Pic} \operatorname{Bl}_p X \xrightarrow{\alpha} \operatorname{Pic}(\operatorname{Bl}_p X \setminus E) \longrightarrow 0$$

(from (14.2.8.1)). Note that $\operatorname{Bl}_p X \setminus E = X \setminus p$. Show that $\operatorname{Pic}(X \setminus p) = \operatorname{Pic} X$. Show that $\beta^* : \operatorname{Pic} X \to \operatorname{Pic} \operatorname{Bl}_p X$ gives a section to α. Use §22.3.5 to show that $E^2 = -1$ (so E is a (-1)-curve, Definition 20.2.6), and from that show that γ is an injection. Conclude that $\operatorname{Pic} \operatorname{Bl}_p X \cong \operatorname{Pic} X \oplus \mathbb{Z}$. Describe how to find the intersection matrix on $N^1_{\mathbb{Q}}(\operatorname{Bl}_p X)$ from that of $N^1_{\mathbb{Q}}(X)$.

22.4.P. EXERCISE. Suppose D is an effective Cartier divisor (a curve) on X. Let $\operatorname{mult}_p D$ be the multiplicity of D at p (Exercise 22.4.J), and let D^{pr} be the proper transform of D. Show that $\pi^* D = D^{pr} + (\operatorname{mult}_p D) E$ as effective Cartier divisors. More precisely, show that the product of the local equation for D^{pr} and the $(\operatorname{mult}_p D)$th power of the local equation for E is the local equation for $\pi^* D$, and hence that (i) $\pi^* D$ is an effective Cartier divisor, and (ii) $\pi^* \mathscr{O}_X(D) \cong \mathscr{O}_{\operatorname{Bl}_p X}(D^{pr}) \otimes \mathscr{O}_{\operatorname{Bl}_p X}(E)^{\otimes(\operatorname{mult}_p D)}$. (A special case is the equation $\ell = e + m$ in Hint 20.2.5.)

22.4.14. *Change of the canonical line bundle under blow-ups.*

As motivation for how the canonical line bundle changes under blowing up, consider $\pi : \mathrm{Bl}_{(0,0)} \mathbb{A}^2 \to \mathbb{A}^2$. Let $X = \mathrm{Bl}_{(0,0)} \mathbb{A}^2$ and $Y = \mathbb{A}^2$ for convenience. We use Exercise 21.7.A to relate $\pi^* \mathscr{K}_Y$ to $\mathscr{K}_X$.

We pick a generator for $\mathscr{K}_Y$ near $(0,0)$: $dx \wedge dy$. (This is in fact a generator for $\mathscr{K}_Y$ everywhere on $\mathbb{A}^2$, but for the sake of generalization, we point out that all that matters is that is a generator at $(0,0)$, and hence *near* $(0,0)$ by geometric Nakayama, Exercise 13.7.E.) When we pull it back to X, we can interpret it as a section of $\mathscr{K}_X$, which will generate $\mathscr{K}_X$ away from the exceptional divisor E, but may contain E with some multiplicity μ. Recall that X can be interpreted as the data of a point in $\mathbb{A}^2$ as well as the choice of a line through the origin. We consider the open subset U where the line is not vertical, and thus can be written as $y = mx$. Here we have natural coordinates: $U = \operatorname{Spec} k[x, y, m]/(y - mx)$, which we can interpret as $\operatorname{Spec} k[x, m]$. The exceptional divisor E meets U, at $x = 0$ (in the coordinates on U), so we can calculate μ on this open set. Pulling back $dx \wedge dy$ to U, we get

$$dx \wedge dy = dx \wedge d(xm) = m(dx \wedge dx) + x(dx \wedge dm) = x(dx \wedge dm)$$

as $dx \wedge dx = 0$. Thus $\pi^*(dx \wedge dy)$ vanishes to order 1 along E.

22.4.Q. EXERCISE (CF. UNIMPORTANT EXERCISE 21.7.J). Explain how this determines an isomorphism $\mathscr{K}_X \cong (\pi^* \mathscr{K}_Y)(E)$.

22.4.R. EXERCISE. Repeat the above calculation in dimension n. Show that the exceptional divisor appears with multiplicity $(n-1)$.

22.4.S. $\star$ EXERCISE. Suppose k is perfect.
(a) Suppose Y is a surface over k, and p is a regular k-valued point, and let $\beta : X \to Y$ be the blow-up of Y at p. Show that $\mathscr{K}_X \cong (\beta^* \mathscr{K}_Y)(E)$. Hint: to find a generator of $\mathscr{K}_X$ near p, choose generators $\overline{x}$ and $\overline{y}$ of $\mathfrak{m}/\mathfrak{m}^2$ (where $\mathfrak{m}$ is the maximal ideal of $\mathscr{O}_{Y,p}$), and lift them to elements of $\mathscr{O}_{X,p}$. Why does $dx \wedge dy$ generate $\mathscr{K}_X$ at p?
(b) Repeat part (a) in arbitrary dimension (following Exercise 22.4.R).
(c) Suppose Z is a smooth m-dimensional (closed) subvariety of a smooth n-dimensional variety Y, and let $\beta : X \to Y$ be the blow-up of Y along Z. Show that $\mathscr{K}_X \cong (\beta^* \mathscr{K}_Y)((n - m - 1)E)$. (Recall from Theorem 22.3.10 that $X = \mathrm{Bl}_Z Y$ is smooth.)

22.4.15. $\star$ **Dimensional vanishing for quasiprojective schemes (promised in §18.2.7).**

Using the theory of blowing up, Theorem 18.2.6 (dimensional vanishing for quasicoherent sheaves on projective k-schemes) can be extended to quasiprojective k-schemes. Suppose X is a quasiprojective k-variety of dimension n. We show that X may be covered by $n+1$ affine open subsets. As X is quasiprojective, there is some projective variety Y with an open embedding $X \hookrightarrow Y$. By replacing Y with the closure of X in Y, we may assume that $\dim Y = n$. Put any subscheme structure Z on the complement of X in Y (for example the reduced subscheme structure, §8.3.9). Let $Y' = \mathrm{Bl}_Z Y$. Then Y' is a projective variety (§22.3.1), which can be covered by $n+1$ affine open subsets. The complement of X in Y' is an effective Cartier divisor $(E_Z Y)$, so the restriction to X of each of these affine open subsets of Y is also affine, by Exercise 7.3.F.

22.4.16. *Remarks.* (i) You might then hope that *any* dimension n variety can be covered by $n+1$ affine open subsets. This is not true. For each integer m, there is a threefold that requires at least m affine open sets to cover it, see [**RV**, Ex. 4.9]. By the discussion above, this example is necessarily not quasiprojective. (ii) Here is a fact useful in invariant theory, which can be proved in the same way. Suppose $p_1, \dots, \dots, p_n$ are closed points on a quasiprojective k-variety X. Then there is an affine open subset of X containing all of them.

22.4.T. EXERCISE (DIMENSIONAL VANISHING FOR QUASIPROJECTIVE VARIETIES). Suppose X is a quasiprojective k-scheme of dimension d. Show that for any quasicoherent sheaf $\mathscr{F}$ on X, $H^i(X, \mathscr{F}) = 0$ for $i > d$.

22.4.17. ⋆⋆ **Deformation to the normal cone.**

The following construction is key to the modern understanding of intersection theory in algebraic geometry, as developed by Fulton and MacPherson, [**F**].

22.4.U. EXERCISE: DEFORMATION TO THE NORMAL CONE. Suppose Y is a k-variety, and $X \hookrightarrow Y$ is a closed subscheme.
(a) Show that the exceptional divisor of $\beta : \mathrm{Bl}_{X\times 0}(Y \times \mathbb{P}^1) \to Y \times \mathbb{P}^1$ is isomorphic to the projective completion of the normal cone to X in Y.
(b) Let $\pi : \mathrm{Bl}_{X\times 0}(Y \times \mathbb{P}^1) \to \mathbb{P}^1$ be the composition of β with the projection to $\mathbb{P}^1$. Show that $\pi^*(0)$ is the scheme-theoretic union of $\mathrm{Bl}_X Y$ with the projective completion of the normal cone to X and Y, and the intersection of these two subschemes may be identified with $E_X Y$, which is a closed subscheme of $\mathrm{Bl}_X Y$ in the usual way (as the exceptional divisor of the blow-up $\mathrm{Bl}_X Y \to Y$), and a closed subscheme of the projective completion of the normal cone as described in Exercise 8.2.Q.

The map

$$\mathrm{Bl}_{X\times 0}(Y \times \mathbb{P}^1) \setminus \mathrm{Bl}_X Y \to \mathbb{P}^1$$

is called the **deformation to the normal cone** (short for *deformation of* Y *to the normal cone of* X *in* Y). Notice that the fiber above every k-point away from $0 \in \mathbb{P}^1$ is canonically isomorphic to Y, and the fiber over 0 is the normal cone. Because this family is "nice" (more precisely, *flat*, the topic of Chapter 24), we can prove things about general Y (near X) by way of this degeneration. (We will see in §24.4.9 that the deformation to the normal cone is a *flat* morphism, which is useful in intersection theory.)

Part VI

More

CHAPTER 23

Derived functors

Ça me semble extrêmement plaisant de ficher comme ça beaucoup de choses, pas drôles quand on les prend séparément, sous le grand chapeau des foncteurs dérivés.

I find it very agreeable to stick all sorts of things, which are not much fun when taken individually, together under the heading of derived functors.

— *A. Grothendieck, letter to J.-P. Serre* [**GrS**, p. 6]

In this chapter, we discuss derived functors, introduced by Grothendieck in his celebrated "Tôhoku article" [**Gr1**], and their applications to sheaves. For quasicoherent sheaves on quasicompact separated schemes, derived functor cohomology will agree with Čech cohomology (§23.5). Čech cohomology will suffice for most of our purposes, and is quite down to earth and computable, but derived functor cohomology is worth seeing. First, it will apply much more generally in algebraic geometry (e.g. étale cohomology) and elsewhere, although this is beyond the scope of this book. Second, it will easily provide us with some useful notions, such as the Ext functors and the Leray spectral sequence. But derived functors can be intimidating the first time you see them, so feel free to just skim the main results, and to return to them later.

23.1 The Tor functors

We begin with a warm-up: the case of Tor. This is a hands-on example, but if you understand it well, you will understand derived functors in general. Tor will be useful to prove facts about flatness, which we will discuss in §24.3. Tor is short for "torsion" (see Remark 24.3.1).

If you have never seen this notion before, you may want to just remember its properties. But I will prove everything anyway — it is surprisingly easy.

The idea behind Tor is as follows. Whenever we see a right-exact functor, we always hope that it is the end of a long exact sequence. Informally, given a short exact sequence

$$0 \longrightarrow N' \longrightarrow N \longrightarrow N'' \longrightarrow 0, \tag{23.1.0.1}$$

we hope $M \otimes_A N' \to M \otimes_A N \to M \otimes_A N'' \to 0$ will extend to a long exact sequence
(23.1.0.2)

$$\cdots \longrightarrow \mathrm{Tor}_i^A(M, N') \longrightarrow \mathrm{Tor}_i^A(M, N) \longrightarrow \mathrm{Tor}_i^A(M, N'') \longrightarrow \cdots$$

$$\longrightarrow \mathrm{Tor}_1^A(M, N') \longrightarrow \mathrm{Tor}_1^A(M, N) \longrightarrow \mathrm{Tor}_1^A(M, N'')$$

$$\longrightarrow M \otimes_A N' \longrightarrow M \otimes_A N \longrightarrow M \otimes_A N'' \longrightarrow 0.$$

More precisely, we are hoping for *covariant functors* $\mathrm{Tor}_i^A(M, \cdot)$ from A-modules to A-modules (functoriality giving 2/3 of the morphisms in (23.1.0.2)), with $\mathrm{Tor}_0^A(M, N) \equiv M \otimes_A N$, and natural "connecting" homomorphisms $\delta : \mathrm{Tor}_{i+1}^A(M, N'') \to \mathrm{Tor}_i^A(M, N')$ for every short exact sequence (23.1.0.1) giving the long exact sequence (23.1.0.2). ("Natural" means: given a morphism of short exact sequences, the natural square you would write down involving the δ-morphism must commute.)

It turns out to be not too hard to make this work, and this will also motivate derived functors. Let's now define $\mathrm{Tor}_i^A(M, N)$.

Take any resolution $\mathscr{R}$ of N by free modules:

$$\cdots \longrightarrow A^{\oplus n_2} \longrightarrow A^{\oplus n_1} \longrightarrow A^{\oplus n_0} \longrightarrow N \longrightarrow 0.$$

More precisely, build this resolution from right to left. Start by choosing generators of N as an A-module, giving us $A^{\oplus n_0} \to N \to 0$. Then choose generators of the kernel, and so on. Note that we are not requiring the n_i to be finite (although we could, if N is a finitely generated module and A is Noetherian). Truncate the resolution, by stripping off the last term N (replacing $\to N \to 0$ with $\to 0$). Then tensor with M (which does not preserve exactness). Note that $M \otimes (A^{\oplus n_i}) = M^{\otimes n_i}$, as tensoring with M commutes with arbitrary direct sums (Exercise 1.3.M). Let $\mathrm{Tor}_i^A(M, N)_{\mathscr{R}}$ be the homology of this complex at the ith stage ($i \geq 0$). The subscript $\mathscr{R}$ reminds us that our construction depends on the resolution, although we will soon see that it is independent of $\mathscr{R}$.

We make some quick observations.

- $\mathrm{Tor}_0^A(M, N)_{\mathscr{R}} \cong M \otimes_A N$, canonically. Reason: as tensoring is right-exact, and $A^{\oplus n_1} \to A^{\oplus n_0} \to N \to 0$ is exact, we have that $M^{\oplus n_1} \to M^{\oplus n_0} \to M \otimes_A N \to 0$ is exact, and hence that the homology of the truncated complex $M^{\oplus n_1} \to M^{\oplus n_0} \to 0$ is $M \otimes_A N$.
- If $M \otimes \cdot$ is exact (i.e., M is *flat*, §1.6.11), then $\mathrm{Tor}_i^A(M, N)_{\mathscr{R}} = 0$ for all $i > 0$. (This characterizes flatness, see Exercise 23.1.D.)

Now given two modules N and N' and resolutions $\mathscr{R}$ and $\mathscr{R}'$ of N and N', we can "lift" any morphism $N \to N'$ to a morphism of the two resolutions:

$$\begin{tikzcd}
\cdots \arrow[r] & A^{\oplus n_i} \arrow[r] \arrow[d] & \cdots \arrow[r] & A^{\oplus n_1} \arrow[r] \arrow[d] & A^{\oplus n_0} \arrow[r] \arrow[d] & N \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & A^{\oplus n_i'} \arrow[r] & \cdots \arrow[r] & A^{\oplus n_1'} \arrow[r] & A^{\oplus n_0'} \arrow[r] & N' \arrow[r] & 0
\end{tikzcd}$$

We do this inductively on i. Here we use the freeness of $A^{\oplus n_i}$: if $a_1, \ldots, a_{n_i}$ are generators of $A^{\oplus n_i}$, to lift the map $b : A^{\oplus n_i} \to A^{\oplus n'_{i-1}}$ (the composition of the differential $A^{\oplus n_i} \to A^{\oplus n_{i-1}}$ with the previously constructed $A^{\oplus n_{i-1}} \to A^{\oplus n'_{i-1}}$) to $c : A^{\oplus n_i} \to A^{\oplus n'_i}$, we arbitrarily lift $b(a_j)$ from $A^{\oplus n'_{i-1}}$ to $A^{\oplus n'_i}$, and declare this to be $c(a_j)$. (Warning for people who care about such things: we are using the axiom of choice here.)

Denote the choice of lifts by $\mathscr{R} \to \mathscr{R}'$. Now truncate both complexes (remove column $N \to N'$) and tensor with M. Maps of complexes induce maps of homology (Exercise 1.6.D), so we have described maps (a priori depending on $\mathscr{R} \to \mathscr{R}'$)

$$\mathrm{Tor}_i^A(M, N)_{\mathscr{R}} \to \mathrm{Tor}_i^A(M, N')_{\mathscr{R}'}.$$

We say two maps of complexes $f, g : C_\bullet \to C'_\bullet$ are **homotopic** if there is a sequence of maps $w : C_i \to C'_{i+1}$ such that $f - g = dw + wd$.

23.1.A. EXERCISE. Show that two homotopic maps give the same map on homology.

23.1.B. CRUCIAL EXERCISE. Show that any two lifts $\mathscr{R} \to \mathscr{R}'$ are homotopic.

We now pull these observations together. (Be sure to digest these completely!)

(1) We get a map of A-modules $\mathrm{Tor}_i^A(M, N)_{\mathscr{R}} \to \mathrm{Tor}_i^A(M, N')_{\mathscr{R}'}$, independent of the lift $\mathscr{R} \to \mathscr{R}'$.

(2) Hence for any two resolutions $\mathscr{R}$ and $\mathscr{R}'$ of an A-module N, we get a canonical isomorphism $\mathrm{Tor}_i^A(M, N)_{\mathscr{R}} \cong \mathrm{Tor}_i^A(M, N)_{\mathscr{R}'}$. Here's why. Choose lifts $\mathscr{R} \to \mathscr{R}'$ and $\mathscr{R}' \to \mathscr{R}$. The composition $\mathscr{R} \to \mathscr{R}' \to \mathscr{R}$ is homotopic to the identity (as it is a lift of the identity map $N \to N$). Thus if $f_{\mathscr{R} \to \mathscr{R}'} : \mathrm{Tor}_i^A(M, N)_{\mathscr{R}} \to \mathrm{Tor}_i^A(M, N)_{\mathscr{R}'}$ is the map induced by $\mathscr{R} \to \mathscr{R}'$, and similarly $f_{\mathscr{R}' \to \mathscr{R}}$ is the map induced by $\mathscr{R}' \to \mathscr{R}$, then $f_{\mathscr{R}' \to \mathscr{R}} \circ f_{\mathscr{R} \to \mathscr{R}'}$ is the identity, and similarly $f_{\mathscr{R} \to \mathscr{R}'} \circ f_{\mathscr{R}' \to \mathscr{R}}$ is the identity.

(3) Hence $\mathrm{Tor}_i^A(M, \cdot)$ doesn't depend on the choice of resolution. It is a covariant functor $Mod_A \to Mod_A$.

23.1.1. *Remark.* Note that if N is a free module, then $\mathrm{Tor}_i^A(M, N) = 0$ for all M and all $i > 0$, as N has the trivial resolution $0 \to N \to N \to 0$ (it is "its own resolution").

Finally, we get long exact sequences:

23.1.2. Proposition. — *For any short exact sequence* (23.1.0.1) *we get a long exact sequence of* Tor*'s* (23.1.0.2).

Proof. Given a short exact sequence (23.1.0.1), choose resolutions of N' and N''. Then use these to get a resolution for N as follows.

(23.1.2.1)
$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
\cdots \arrow[r] & A^{\oplus n'_1} \arrow[r] \arrow[d] & A^{\oplus n'_0} \arrow[r] \arrow[d] & N' \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & A^{\oplus(n'_1+n''_1)} \arrow[r] \arrow[d] & A^{\oplus(n'_0+n''_0)} \arrow[r] \arrow[d] & N \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & A^{\oplus n''_1} \arrow[r] \arrow[d] & A^{\oplus n''_0} \arrow[r] \arrow[d] & N'' \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

The map $A^{\oplus(n'_{i+1}+n''_{i+1})} \to A^{\oplus(n'_i+n''_i)}$ is the composition $A^{\oplus n'_{i+1}} \to A^{\oplus n'_i} \hookrightarrow A^{\oplus(n'_i+n''_i)}$ along with a lift of $A^{\oplus n''_{i+1}} \to A^{\oplus n''_i}$ to $A^{\oplus(n'_i+n''_i)}$ ensuring that the middle row is a *complex*.

23.1.C. EXERCISE. Verify that it is possible to choose such a lift of $A^{\oplus n''_{i+1}} \to A^{\oplus n''_i}$ to $A^{\oplus(n'_i+n''_i)}$.

Hence (23.1.2.1) is *exact* (not just a complex), using the long exact sequence in cohomology (Theorem 1.6.6), and the fact that the top and bottom rows are exact. Thus the middle row is a resolution, and (23.1.2.1) is a short exact sequence of resolutions. (This is sometimes called the *horseshoe construction*, as the filling in of the middle row looks like filling in the middle of a horseshoe.) It may be helpful to notice that the columns other than the "N-column" are all "direct sum exact sequences", and the horizontal maps in the middle row are "block upper triangular".

Then truncate (removing the right column $0 \to N' \to N \to N'' \to 0$), tensor with M (obtaining a short exact sequence of complexes) and take cohomology, yielding the desired long exact sequence. $\square$

23.1.D. EXERCISE. Show that the following are equivalent conditions on an A-module M.

(i) M is flat.
(ii) $\mathrm{Tor}_i^A(M, N) = 0$ for all $i > 0$ and all A-modules N.
(iii) $\mathrm{Tor}_1^A(M, N) = 0$ for all A-modules N.

23.1.3. *Caution.* Given that free modules are immediately seen to be flat, you might think that Exercise 23.1.D implies Remark 23.1.1. This would follow if we knew that $\mathrm{Tor}_i^A(M, N) \cong \mathrm{Tor}_i^A(N, M)$, which is clear for $i = 0$ (as $\otimes$ is symmetric), but we won't know this about Tor_i when $i > 0$ until Exercise 23.3.A.

23.1.E. EXERCISE. Show that the connecting homomorphism δ constructed above is independent of all of choices (of resolutions, etc.). Try to do this with as little annoyance as possible. (Possible hint: given two sets of choices used to build (23.1.2.1), build a map — a three-dimensional diagram — from one version of (23.1.2.1) to the other version.)

23.1.F. UNIMPORTANT EXERCISE. Show that $\mathrm{Tor}_i^A(M, \cdot)$ is an *additive* functor (Definition 1.6.1). (We won't use this later, so feel free to skip it.)

We have thus established the foundations of Tor.

23.2 Derived functors in general

23.2.1. Projective resolutions. We used very little about free modules in the above construction of Tor — in fact we used only that free modules are **projective**, i.e., those modules P such that for any surjection $M \twoheadrightarrow N$, it is possible to lift any morphism $P \to N$ to $P \to M$:

(23.2.1.1)
$$\begin{tikzcd} P \arrow[d, dashed, "\text{exists}"'] \arrow[dr] & \\ M \arrow[r, two heads] & N \end{tikzcd}$$

(As noted in §23.1, this needs the axiom of choice.) Equivalently, $\mathrm{Hom}(P, \cdot)$ is an exact functor (recall that $\mathrm{Hom}(Q, \cdot)$ is always left-exact for any Q). More generally, the same idea yields the definition of a **projective object in any abelian category**. Hence by following through our entire argument with projective modules replacing free modules throughout, (i) we can compute $\mathrm{Tor}_i^A(M, N)$ by taking any projective resolution of N, and (ii) $\mathrm{Tor}_i^A(M, N) = 0$ for any projective A-module N.

23.2.A. EXERCISE. Show that an object P is projective if and only if every short exact sequence $0 \to A \to B \to P \to 0$ splits. Hence show that an A-module M is projective if and only if M is a direct summand of a free module.

23.2.B. EXERCISE. Show that projective modules are flat. (Hint: Exercise 23.2.A. Be careful if you want to use Exercise 23.1.D; see Caution 23.1.3.)

23.2.2. Definition: Derived functors.

The above description was low-tech, but immediately generalizes drastically. All we are using is that $M \otimes_A \cdot$ is a right-exact functor, and that for any A-module N, we can find a surjection $P \twoheadrightarrow N$ from a projective module. In general, if F is *any* right-exact covariant functor from the category of A-modules to any abelian category, this construction will define a sequence of functors $L_i F$ such that $L_0 F = F$ and the $L_i F$'s give a long exact sequence. We can make this more general still. We say that an abelian category **has enough projectives** if for any object N there is a surjection onto it from a projective object. Then if F is any right-exact covariant functor from an abelian category with enough projectives to any abelian category, then we can define the *left derived functors* to F, denoted $L_i F$ ($i \geq 0$). You should reread §23.1 and see that throughout we only use the fact we have a projective resolution (repeatedly lifting maps as in (23.2.1.1)), as well as the fact that F sends products to products (a consequence of additivity of the functor, see Remark 1.6.2) to show that F applied to (23.1.2.1) preserves the exactness of the columns.

23.2.C. EXERCISE. The notion of an **injective object** in an abelian category is dual to the notion of a projective object.
(a) State precisely the definition of an injective object.
(b) Define derived functors for (i) covariant left-exact functors (these are called **right derived functors**), (ii) contravariant left-exact functors (also called **right derived functors**), and (iii) contravariant right-exact functors (these are called **left derived functors**), making explicit the necessary assumptions of the category having enough injectives or projectives.

23.2.3. *Notation.* If F is a right-exact functor, its (left-)derived functors are denoted $L_i F$ ($i \geq 0$, with $L_0 F = F$). If F is a left-exact functor, its (right-) derived functors are denoted $R^i F$. The i is a superscript, to indicate that the long exact sequence is "ascending in i".

23.2.4. The Ext **functors.**

23.2.D. EASY EXERCISE (AND DEFINITION): Ext FUNCTORS FOR A-MODULES, FIRST VERSION. As $\operatorname{Hom}(\cdot, N)$ is a contravariant left-exact functor in Mod_A, which has enough projectives, define $\operatorname{Ext}^i_A(M, N)$ as the ith right derived functor of $\operatorname{Hom}(\cdot, N)$, applied to M. State the corresponding long exact sequence for Ext-modules.

23.2.E. EASY EXERCISE (AND DEFINITION): Ext FUNCTORS FOR A-MODULES, SECOND VERSION. The category Mod_A has enough injectives (see §23.2.5). As $\operatorname{Hom}(M, \cdot)$ is a covariant left-exact functor in Mod_A, define $\operatorname{Ext}^i_A(M, N)$ as the ith right derived functor of $\operatorname{Hom}(M, \cdot)$, applied to N. State the corresponding long exact sequence for Ext-modules.

We seem to have a problem with the previous two exercises: we have defined $\operatorname{Ext}^i(M, N)$ twice, and we have two different long exact sequences! Fortunately, these two definitions agree (see Exercise 23.3.B).

23.2.F. EASY EXERCISE. Use the definition of Ext in Exercise 23.2.D to show that if A is a Noetherian ring, and M and N are finitely generated A-modules, then $\operatorname{Ext}^i_A(M, N)$ is a finitely generated A-module.

Ext-functors (for sheaves) will play a key role in Serre duality, see §30.2.

23.2.5. $\star$ **The category of A-modules has enough injectives.** We will need the fact that Mod_A has enough injectives, but the details of the proof won't come up again, so feel free to skip this discussion.

23.2.G. EXERCISE. Suppose Q is an A-module, such that for every ideal $I \subset A$, every homomorphism $I \to Q$ extends to $A \to Q$. Show that Q is an injective A-module. Hint: suppose $N \subset M$ is an inclusion of A-modules, and we are given $\beta : N \to Q$. We wish to show that β extends to $M \to Q$. Use the axiom of choice to show that among those A-modules N' with $N \subset N' \subset M$, such that β extends to

N', there is a maximal one. If this N' is not M, give an extension of β to $N' + Am$, where $m \in M \setminus N'$, obtaining a contradiction.

23.2.H. EASY EXERCISE (USING THE AXIOM OF CHOICE, IN THE GUISE OF ZORN'S LEMMA). Show that a $\mathbb{Z}$-module (i.e., abelian group) Q is injective if and only if it is **divisible** (i.e., for every $q \in Q$ and $n \in \mathbb{Z}^{\neq 0}$, there is $q' \in Q$ with $nq' = q$). Hence show that any quotient of an injective $\mathbb{Z}$-module is also injective.

23.2.I. EXERCISE. Show that the category of $\mathbb{Z}$-modules $Mod_{\mathbb{Z}} = Ab$ has enough injectives. (Hint: if M is a $\mathbb{Z}$-module, then write it as the quotient of a free $\mathbb{Z}$-module F by some K. Show that M is contained in the divisible group $(F \otimes_{\mathbb{Z}} \mathbb{Q})/K$.)

23.2.J. EXERCISE. Suppose Q is an injective $\mathbb{Z}$-module, and A is a ring. Show that $\mathrm{Hom}_{\mathbb{Z}}(A, Q)$ is an injective A-module. Hint: First describe the A-module structure on $\mathrm{Hom}_{\mathbb{Z}}(A, Q)$. You will only use the fact that $\mathbb{Z}$ is a ring, and that A is an algebra over that ring.

23.2.K. EXERCISE. Show that Mod_A has enough injectives. Hint: suppose M is an A-module. By Exercise 23.2.I, we can find an inclusion *of $\mathbb{Z}$-modules* $M \hookrightarrow Q$ where Q is an injective *$\mathbb{Z}$-module*. Describe a sequence of inclusions of A-modules

$$M \hookrightarrow \mathrm{Hom}_{\mathbb{Z}}(A, M) \hookrightarrow \mathrm{Hom}_{\mathbb{Z}}(A, Q).$$

(The A-module structure on $\mathrm{Hom}_{\mathbb{Z}}(A, M)$ is via the A-action on the left argument A, not via the A-action on the right argument M.) The right term is injective by the previous Exercise 23.2.J.

23.2.6. ⋆ **Universal δ-functors.**

(This discussion is best skipped on a first reading; you should move directly to §23.3.) We now describe a more general variant of derived functors, as you may use them in the discussion of Serre duality in Chapter 30. The advantage of the notion of universal δ-functor is that we can apply it even in cases where $\mathscr{A}$ does not have enough injectives.

Abstracting key properties of derived functors, we define the data of a (cohomological) δ-functor, from an abelian category $\mathscr{A}$ to another abelian category $\mathscr{B}$. A **δ-functor** is a collection of additive functors $T^i : \mathscr{A} \to \mathscr{B}$ (where T^i is taken to be 0 if $i < 0$), along with morphisms $\delta^i : T^i(A'') \to T^{i+1}(A')$ for each short exact sequence

$$0 \to A' \to A \to A'' \to 0 \tag{23.2.6.1}$$

in $\mathscr{A}$, satisfying two properties:

(i) *(short exact sequences yield long exact sequences)* For each short exact sequence (23.2.6.1), the sequence

$$\cdots \longrightarrow T^{i-1}(A'') \xrightarrow{\delta^{i-1}} T^i(A') \longrightarrow T^i(A) \longrightarrow T^i(A'') \xrightarrow{\delta^i} T^{i+1}(A') \longrightarrow \cdots$$

(where the unlabeled maps come from the covariance of the T^i) is exact. In particular, T^0 is left-exact.

(ii) *(functoriality of (i))* For each morphism of short exact sequences in $\mathscr{A}$

$$\begin{tikzcd} 0 \arrow[r] & a' \arrow[r] \arrow[d] & a \arrow[r] \arrow[d] & a'' \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & A' \arrow[r] & A \arrow[r] & A'' \arrow[r] & 0 \end{tikzcd}$$

(where the squares commute), the δ^i's give a commutative diagram

$$\begin{tikzcd} T^i(a'') \arrow[r, "\delta^i"] \arrow[d] & T^{i+1}(a') \arrow[d] \\ T^i(A'') \arrow[r, "\delta^i"] & T^{i+1}(A') \end{tikzcd}$$

(where the vertical arrows come from the covariance of T^i and T^{i+1}).

Derived functor cohomology is clearly an example of a δ-functor; Čech cohomology of sheaves on quasicompact separated schemes is another. (You can make these statements precise if you wish.)

23.2.L. EXERCISE. Figure out the right definition of **morphism of δ-functors** $\mathscr{A} \to \mathscr{B}$. (It should then be clear that the δ-functors from $\mathscr{A}$ to $\mathscr{B}$ form a category.)

23.2.7. *Definition.* A (cohomological) δ-functor (T^i, δ^i) is **universal** if for any other δ-functor (T'^i, δ'^i), and any natural transformation of functors $\alpha : T^0 \to T'^0$, there is a unique morphism of δ-functors $(T^i, \delta^i) \to (T'^i, \delta'^i)$ extending α. By universal property nonsense (and Exercise 23.2.L), given any covariant left-exact functor $F : \mathscr{A} \to \mathscr{B}$, there is at most one universal δ-functor (T^i, δ^i) extending F (i.e., with a natural isomorphism $T^0 \cong F$). The key fact about universal δ-functors is the following.

23.2.8. Theorem. — *Suppose (T^i, δ^i) is a covariant δ-functor from $\mathscr{A}$ to $\mathscr{B}$, and for all $A \in \mathscr{A}$, there exists a monomorphism $A \to J$ with $T^i J = 0$ for all $i > 0$. Then (T^i, δ^i) is universal.*

23.2.M. ⋆⋆ EXERCISE. Prove Theorem 23.2.8. Partial hint: motivated by Corollary 23.2.10 below, follow our discussion of derived functors. Better hint (because this exercise is hard): follow the hints in [**Wei**, Exercise 2.4.5], or follow the proof of [**Lan**, Thm. 7.1].

23.2.9. *Remark.* An additive functor $F : \mathscr{A} \to \mathscr{B}$ is said to be **effaceable** if for every $A \in \mathscr{A}$, there is a monomorphism $A \to J$ with $F(J) = 0$. The hypotheses of Theorem 23.2.8 can be weakened to require only that T^i is effaceable for each $i > 0$, and you are welcome to prove that instead. (Indeed, [**Wei**, Exercise 2.4.5], [**Lan**, Thm. 7.1], and the original source [**Gr1**, II.2.2.1] deal with this case.) We give the statement of Theorem 23.2.8 for simplicity, as we will only use this version.

23.2.10. Corollary. — *If $\mathscr{A}$ has enough injectives, and F is a left-exact covariant functor $\mathscr{A} \to \mathscr{B}$, then the R^iF (with the δ^i that accompany them) form a universal δ-functor.*

Proof. Each element of $\mathscr{A}$ admits a monomorphism into an injective element; this is just the definition of "enough injectives" (Exercise 23.2.C). Higher derived functors of an injective elements I are always 0: just compute the higher derived functor by taking the injective resolution of I "by itself". □

23.3 Derived functors and spectral sequences

A number of useful facts can be easily proved using spectral sequences. By doing these exercises, you will lose any fear of spectral sequence arguments in similar situations, as you will realize they are all the same.

Before you read this section, you should read §1.7 on spectral sequences.

23.3.1. Symmetry of Tor.

23.3.A. EXERCISE (SYMMETRY OF Tor). Show that there is an isomorphism $\mathrm{Tor}_i^A(M, N) \cong \mathrm{Tor}_i^A(N, M)$. (Hint: take a free resolution of M and a free resolution of N. Take their "product" to somehow produce a double complex. Use both orientations of the obvious spectral sequence and see what you get.)

On a related note:

23.3.B. EXERCISE. Show that the two definitions of $\mathrm{Ext}^i(M, N)$ given in Exercises 23.2.D and 23.2.E agree.

23.3.2. Derived functors can be computed using acyclic resolutions. Suppose $F : \mathscr{A} \to \mathscr{B}$ is a right-exact additive functor of abelian categories, and that $\mathscr{A}$ has enough projectives. (In other words, the hypotheses ensure the existence of left derived functors of F. Analogous facts will hold with the other types of derived functors, Exercise 23.2.C(b).) We say that $A \in \mathscr{A}$ is F**-acyclic** (or just **acyclic** if the F is clear from context) if $L_i F\, A = 0$ for $i > 0$.

The following exercise is a good opportunity to learn a useful trick (Hint 23.3.3).

23.3.C. EXERCISE. Show that you can also compute the derived functors of an objects B of $\mathscr{A}$ using **acyclic resolutions**, i.e., by taking a resolution

$$\cdots \longrightarrow A_2 \longrightarrow A_1 \longrightarrow A_0 \longrightarrow B \longrightarrow 0$$

by F-acyclic objects A_i, truncating, applying F, and taking homology. Hence $\mathrm{Tor}_i(M, N)$ can be computed with a flat resolution of M or N.

23.3.3. *Hint for Exercise 23.3.C (and a useful trick: building a "projective resolution of a complex").* Show that you can construct a double complex

(23.3.3.1)
$$\begin{tikzcd}
& \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
\cdots \arrow[r] & P_{2,1} \arrow[r] \arrow[d] & P_{1,1} \arrow[r] \arrow[d] & P_{0,1} \arrow[r] \arrow[d] & P_1 \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & P_{2,0} \arrow[r] \arrow[d] & P_{1,0} \arrow[r] \arrow[d] & P_{0,0} \arrow[r] \arrow[d] & P_0 \arrow[r] \arrow[d] & 0 \\
\cdots \arrow[r] & A_2 \arrow[r] \arrow[d] & A_1 \arrow[r] \arrow[d] & A_0 \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & 0 \\
& 0 & 0 & 0 & 0 &
\end{tikzcd}$$

where the rows and columns are exact and the $P_?$'s are projective. Do this by constructing the $P_?$'s inductively from the bottom right. Remove the bottom row, and the right-most nonzero column, and then apply F, to obtain a new double complex. Use a spectral sequence argument to show that (i) the double complex has homology equal to $L_i F\, B$, and (ii) the homology of the double complex agrees with the construction given in the statement of the exercise. If this is too confusing, read more about the Cartan-Eilenberg resolution below.

23.3.4. The Grothendieck composition-of-functors spectral sequence.

Suppose $\mathscr{A}$, $\mathscr{B}$, and $\mathscr{C}$ are abelian categories, $F : \mathscr{A} \to \mathscr{B}$ and $G : \mathscr{B} \to \mathscr{C}$ are a left-exact additive covariant functors, and $\mathscr{A}$ and $\mathscr{B}$ have enough injectives. Thus right derived functors of F, G, and $G \circ F$ exist. A reasonable question is: how are they related?

23.3.5. Theorem (Grothendieck composition-of-functors spectral sequence). — *Suppose $F : \mathscr{A} \to \mathscr{B}$ and $G : \mathscr{B} \to \mathscr{C}$ are left-exact additive covariant functors, and $\mathscr{A}$ and $\mathscr{B}$ have enough injectives. Suppose further that F sends injective elements of $\mathscr{A}$ to G-acyclic elements of $\mathscr{B}$. Then for each $X \in \mathscr{A}$, there is a spectral sequence with $_{\rightarrow}E_2^{p,q} = R^q G(R^p F(X))$ converging to $R^{p+q}(G \circ F)(X)$.*

We will soon see the Leray spectral sequence as an application (Theorem 23.4.4).

There is more one might want to extract from the proof of Theorem 23.3.5. For example, although E_0 page of the spectral sequence will depend on some choices (of injective resolutions), the E_2 page will be independent of choice. For our applications, we won't need this refinement.

We will have to work to establish Theorem 23.3.5, so the proof is possibly best skipped on a first reading.

23.3.6. ⋆ **Proving Theorem 23.3.5.**

Before we give the proof (in §23.3.8), we begin with some preliminaries to motivate it. In order to discuss derived functors applied to X, we choose an injective resolution of X:

$$0 \longrightarrow X \longrightarrow I^0 \longrightarrow I^1 \longrightarrow \cdots .$$

To compute the derived functors $R^p F(X)$, we apply F to the injective resolution $I^\bullet$:

$$0 \longrightarrow F(I^0) \longrightarrow F(I^1) \longrightarrow F(I^2) \longrightarrow \cdots .$$

Note that $F(I^p)$ is G-acyclic, by hypothesis of Theorem 23.3.5. If we were to follow our nose, we might take simultaneous injective resolutions $I^{\bullet,\bullet}$ of the terms in the above complex $F(I^\bullet)$ (the "dual" of Hint 23.3.3 — note that only the columns are required to be exact), and apply G, and consider the resulting double complex:

(23.3.6.1)
$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
0 \arrow[r] & G(I^{0,2}) \arrow[r] \arrow[u] & G(I^{1,2}) \arrow[r] \arrow[u] & G(I^{2,2}) \arrow[r] \arrow[u] & \cdots \\
0 \arrow[r] & G(I^{0,1}) \arrow[r] \arrow[u] & G(I^{1,1}) \arrow[r] \arrow[u] & G(I^{2,1}) \arrow[r] \arrow[u] & \cdots \\
0 \arrow[r] & G(I^{0,0}) \arrow[r] \arrow[u] & G(I^{1,0}) \arrow[r] \arrow[u] & G(I^{2,0}) \arrow[r] \arrow[u] & \cdots \\
 & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] &
\end{tikzcd}$$

23.3.D. EXERCISE. Consider the spectral sequence with upward orientation, starting with (23.3.6.1) as page E_0. Show that $E_2^{p,q}$ is $R^p(G \circ F)(X)$ if $q = 0$, and 0 otherwise.

We now see half of the terms in the conclusion of Theorem 23.3.5; we are halfway there. To complete the proof, we would want to consider another spectral sequence, with rightward orientation, but we need to know more about (23.3.6.1); we will build it more carefully.

23.3.7. *Cartan-Eilenberg resolutions.*

Suppose $\cdots \to C^{p-1} \to C^p \to C^{p+1} \to \cdots$ is a complex in an abelian category $\mathscr{B}$. We will build an injective resolution of $C^\bullet$

(23.3.7.1)
$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
0 \arrow[r] & I^{0,2} \arrow[r] \arrow[u] & I^{1,2} \arrow[r] \arrow[u] & I^{2,2} \arrow[r] \arrow[u] & \cdots \\
0 \arrow[r] & I^{0,1} \arrow[r] \arrow[u] & I^{1,1} \arrow[r] \arrow[u] & I^{2,1} \arrow[r] \arrow[u] & \cdots \\
0 \arrow[r] & I^{0,0} \arrow[r] \arrow[u] & I^{1,0} \arrow[r] \arrow[u] & I^{2,0} \arrow[r] \arrow[u] & \cdots \\
0 \arrow[r] & C^{0} \arrow[r] \arrow[u] & C^{1} \arrow[r] \arrow[u] & C^{2} \arrow[r] \arrow[u] & \cdots \\
 & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] &
\end{tikzcd}$$

satisfying some further properties.

We first define some notation for functions on a complex.

- Let $Z^p(K^\bullet)$ be the kernel of the pth differential of a complex $K^\bullet$.
- Let $B^{p+1}(K^\bullet)$ be the image of the pth differential of a complex $K^\bullet$. (The superscript is chosen so that $B^{p+1}(K^\bullet) \subset K^{p+1}$.)
- As usual, let $H^p(K^\bullet)$ be the homology at the pth step of a complex $K^\bullet$.

For each p, we have complexes

(23.3.7.2)
$$\begin{tikzcd}
\vdots & \vdots & \vdots \\
Z^p(I^{\bullet,1}) \arrow[u] & B^p(I^{\bullet,1}) \arrow[u] & H^p(I^{\bullet,1}) \arrow[u] \\
Z^p(I^{\bullet,0}) \arrow[u] & B^p(I^{\bullet,0}) \arrow[u] & H^p(I^{\bullet,0}) \arrow[u] \\
Z^p(C^\bullet) \arrow[u] & B^p(C^\bullet) \arrow[u] & H^p(C^\bullet) \arrow[u] \\
0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u]
\end{tikzcd}$$

We will construct (23.3.7.1) so that the three complexes (23.3.7.2) are all injective resolutions (of their first nonzero terms). We begin by choosing injective resolutions $B^{p,*}$ of $B^p(C^\bullet)$ and $H^{p,*}$ of $H^p(C^\bullet)$; these will eventually be the last two lines of (23.3.7.2).

23.3.E. EXERCISE. Describe an injective resolution $Z^{p,*}$ of $Z^p(C^\bullet)$ (the first line of (23.3.7.2)) making the following diagram a short exact sequence of complexes.

(23.3.7.3)
$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
0 \arrow[r] & B^{p,1} \arrow[r] \arrow[u] & Z^{p,1} \arrow[r] \arrow[u] & H^{p,1} \arrow[r] \arrow[u] & 0 \\
0 \arrow[r] & B^{p,0} \arrow[r] \arrow[u] & Z^{p,0} \arrow[r] \arrow[u] & H^{p,0} \arrow[r] \arrow[u] & 0 \\
0 \arrow[r] & B^p(C^\bullet) \arrow[r] \arrow[u] & Z^p(C^\bullet) \arrow[r] \arrow[u] & H^p(C^\bullet) \arrow[r] \arrow[u] & 0 \\
 & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] &
\end{tikzcd}$$

Hint: the "dual" problem was solved in (23.1.2.1), by a "horseshoe construction".

23.3.F. EXERCISE. Describe an injective resolution $I^{p,*}$ of C^p making the following diagram a short exact sequence of complexes.

(23.3.7.4)
$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
0 \arrow[r] & Z^{p,1} \arrow[r] \arrow[u] & I^{p,1} \arrow[r] \arrow[u] & B^{p+1,1} \arrow[r] \arrow[u] & 0 \\
0 \arrow[r] & Z^{p,0} \arrow[r] \arrow[u] & I^{p,0} \arrow[r] \arrow[u] & B^{p+1,0} \arrow[r] \arrow[u] & 0 \\
0 \arrow[r] & Z^p(C^\bullet) \arrow[r] \arrow[u] & C^p \arrow[r] \arrow[u] & B^{p+1}(C^\bullet) \arrow[r] \arrow[u] & 0 \\
 & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] &
\end{tikzcd}$$

(The hint for the previous problem applies again. We remark that the first nonzero columns of (23.3.7.3) and (23.3.7.4) appeared in (1.6.5.3).)

23.3.G. EXERCISE/DEFINITION. Build an injective resolution (23.3.7.1) of $C^\bullet$ such that $Z^{p,*} = Z^p(I^{\bullet,*})$, $B^{p,*} = B^p(I^{\bullet,*})$, $H^{p,*} = H^p(I^{\bullet,*})$, so the three complexes (23.3.7.2) are injective resolutions. This is called a **Cartan-Eilenberg resolution** of $C^\bullet$.

23.3.8. *Proof of the Grothendieck spectral sequence, Theorem 23.3.5.* We pick up where we left off before our digression on Cartan-Eilenberg resolutions. Choose an injective resolution $I^\bullet$ of X. Apply the functor F, then take a Cartan-Eilenberg resolution $I^{\bullet,\bullet}$ of $FI^\bullet$, and then apply G, to obtain (23.3.6.1).

Exercise 23.3.D describes what happens when we take (23.3.6.1) as E_0 in a spectral sequence with upward orientation. So we now consider the rightward orientation.

From our construction of the Cartan-Eilenberg resolution, we have injective resolutions (23.3.7.2), and short exact sequences

(23.3.8.1) $$0 \longrightarrow B^p(I^{\bullet,q}) \longrightarrow Z^p(I^{\bullet,q}) \longrightarrow H^p(I^{\bullet,q}) \longrightarrow 0$$

(23.3.8.2) $$0 \longrightarrow Z^p(I^{\bullet,q}) \longrightarrow I^{p,q} \longrightarrow B^{p+1}(I^{\bullet,q}) \longrightarrow 0$$

of *injective* objects (from the columns of (23.3.7.3) and (23.3.7.4)). This means that both are *split* exact sequences (the central term can be expressed as a direct sum of the outer two terms), so upon application of G, both exact sequences remain exact.

Applying the left-exact functor G to

$$0 \longrightarrow Z^p(I^{\bullet,q}) \longrightarrow I^{p,q} \longrightarrow I^{p+1,q},$$

we find that $GZ^p(I^{\bullet,q}) = \ker(GI^{p,q} \to GI^{p+1,q})$. But this kernel is the *definition* of $Z^p(GI^{\bullet,q})$, so we have an induced isomorphism $GZ^p(I^{\bullet,q}) = Z^p(GI^{\bullet,q})$ ("G and Z^p commute"). From the exactness of (23.3.8.2) upon application of G, we see that $GB^{p+1}(I^{\bullet,q}) = B^{p+1}(GI^{\bullet,q})$ (both are $\operatorname{coker}(GZ^p(I^{\bullet,q}) \to GI^{p,q})$). From the exactness of (23.3.8.1) upon application of G, we see that $GH^p(I^{\bullet,q}) = H^p(GI^{\bullet,q})$ (both are $\operatorname{coker}(GB^p(I^{\bullet,q}) \to GZ^p(I^{\bullet,q}))$ — so "G and H^p commute").

We return to considering the rightward-oriented spectral sequence with (23.3.6.1) as E_0. Taking cohomology in the rightward direction, we find $E_1^{p,q} = H^p(GI^{\bullet,q}) = GH^p(I^{\bullet,q})$ (as G and H^p commute). Now $H^p(I^{\bullet,q})$ is an injective resolution of $(R^pF)(X)$ (the last resolution of (23.3.7.2)). Thus when we compute E_2 by using the vertical arrows, we find $E_2^{p,q} = R^qG(R^pF(X))$.

You should now verify yourself that this (combined with Exercise 23.3.D) concludes the proof of Theorem 23.3.5. □

23.4 Derived functor cohomology of $\mathscr{O}$-modules

We wish to apply the machinery of derived functors to define cohomology of quasicoherent sheaves on a scheme X. Rather than working in the category $QCoh_X$, for a number of reasons it is simpler to work in the larger category $Mod_{\mathscr{O}_X}$ (see Unimportant Remark 23.5.7).

23.4.1. Theorem. — *Suppose $(X, \mathscr{O}_X)$ is a ringed space. Then the category of $\mathscr{O}_X$-modules $Mod_{\mathscr{O}_X}$ has enough injectives.*

As a side benefit (of use to others more than us), taking $\mathscr{O}_X = \underline{\mathbb{Z}}$, we see that the category of sheaves of abelian groups on a fixed topological space have enough injectives.

We prove Theorem 23.4.1 in a series of exercises, following Godement, [**GrS**, p.27-28]. Suppose $\mathscr{F}$ is an $\mathscr{O}_X$-module. We will exhibit an injection $\mathscr{F} \hookrightarrow \mathscr{Q}'$ into an injective $\mathscr{O}_X$-module. For each $p \in X$, choose an inclusion $\mathscr{F}_p \hookrightarrow Q_p$ into an injective $\mathscr{O}_{X,p}$-module (possible as the category of $\mathscr{O}_{X,p}$-modules has enough injectives, Exercise 23.2.K).

23.4.A. EXERCISE. Show that the skyscraper sheaf $\mathscr{Q}_p := i_{p,*}Q_p$, with module Q_p at point $p \in X$, is an injective $\mathscr{O}_X$-module. (You can cleverly do this by abstract nonsense, using Exercise 23.5.B, but it is just as quick to do it by hand.)

23.4.B. EASY EXERCISE. Show that the direct product (possibly infinite) of injective objects in an abelian category is also injective.

By the previous two exercises, $\mathscr{Q}' := \prod_{p \in X} \mathscr{Q}_p$ is an injective $\mathscr{O}_X$-module.

23.4.C. EASY EXERCISE. By considering stalks, show that the natural map $\mathscr{F} \to \mathscr{Q}'$ is an injection.

This completes the proof of Theorem 23.4.1. $\square$

We can now make a number of definitions.

23.4.2. Definition. If $(X, \mathscr{O}_X)$ is a ringed space, and $\mathscr{F}$ is an $\mathscr{O}_X$-module, define $H^i(X, \mathscr{F})$ as $R^i\Gamma(X, \mathscr{F})$. If furthermore $\pi : (X, \mathscr{O}_X) \to (Y, \mathscr{O}_Y)$ is a map of ringed spaces, we have derived pushforwards $R^i\pi_* : Mod_{\mathscr{O}_X} \to Mod_{\mathscr{O}_Y}$.

We have defined these notions earlier in special cases, for quasicoherent sheaves on quasicompact separated schemes (for H^i), or for quasicompact separated morphisms of schemes (for $R^i\pi_*$), in Chapter 18. We will soon (§23.5) show that these older definitions agree with Definition 23.4.2. Thus the derived functor definition applies much more generally than our Čech definition. But it is worthwhile to note that almost everything we use will come out of the Čech definition. A notable exception is the Leray spectral sequence, which we now discuss.

23.4.3. The Leray spectral sequence.

23.4.4. Theorem (the Leray spectral sequence). — *Suppose $\pi : (X, \mathscr{O}_X) \to (Y, \mathscr{O}_Y)$ is a morphism of ringed spaces. Show that for any $\mathscr{O}_X$-module $\mathscr{F}$, there is a spectral sequence with E_2 term given by $H^q(Y, R^p\pi_*\mathscr{F})$ abutting to $H^{p+q}(X, \mathscr{F})$.*

This is an immediate consequence of the Grothendieck composition-of-functors spectral sequence (Theorem 23.3.5) once we prove that the pushforward of an injective $\mathscr{O}$-module is an acyclic $\mathscr{O}$-module. We do this now.

23.4.5. *Definition.* We make an intermediate definition that is independently important. A sheaf $\mathscr{F}$ on a topological space is **flasque** (also sometimes called *flabby*) if all restriction maps are surjective, i.e., if $\mathrm{res}_{U \subset V} : \mathscr{F}(V) \to \mathscr{F}(U)$ is surjective for all $U \subset V$.

23.4.D. EXERCISE. Suppose $(X, \mathscr{O}_X)$ is a ringed space.
(a) Show that if

$$0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0 \tag{23.4.5.1}$$

is an exact sequence of $\mathscr{O}_X$-modules, and $\mathscr{F}'$ is flasque, then (23.4.5.1) is exact on sections over any open set U, i.e. $0 \to \mathscr{F}'(U) \to \mathscr{F}(U) \to \mathscr{F}''(U) \to 0$ is exact.
(b) Given an exact sequence (23.4.5.1), if $\mathscr{F}'$ is flasque, show that $\mathscr{F}$ is flasque if and only if $\mathscr{F}''$ is flasque.

23.4.E. EASY EXERCISE (PUSHFORWARD OF FLASQUES ARE FLASQUE).
(a) Suppose $\pi : X \to Y$ is a continuous map of topological spaces, and $\mathscr{F}$ is a

flasque sheaf of sets on X. Show that $\pi_*\mathscr{F}$ is a flasque sheaf on Y.
(b) Suppose $\pi : (X, \mathscr{O}_X) \to (Y, \mathscr{O}_Y)$ is a morphism of ringed spaces, and $\mathscr{F}$ is a flasque $\mathscr{O}_X$-module. Show that $\pi_*\mathscr{F}$ is a flasque $\mathscr{O}_Y$-module.

23.4.F. EXERCISE (INJECTIVE IMPLIES FLASQUE). Suppose $(X, \mathscr{O}_X)$ is a ringed space, and $\mathscr{Q}$ is an injective $\mathscr{O}_X$-module. Show that $\mathscr{Q}$ is flasque. Hint: If $U \subset V \subset X$, then describe an injection of $\mathscr{O}_X$-modules $0 \to (i_V)_!\mathscr{O}_V \to (i_U)_!\mathscr{O}_U$, where $i_U : U \hookrightarrow X$ and $i_V : V \hookrightarrow X$ are the obvious open embeddings. Apply the exact contravariant functor $\operatorname{Hom}(\cdot, \mathscr{Q})$. (The morphisms $(i_V)_!$ and $(i_U)_!$ are extensions by zero, see Exercise 2.6.G. Now might be a good time to do that Exercise.)

23.4.G. EXERCISE (FLASQUE IMPLIES Γ-ACYCLIC). Suppose $\mathscr{F}$ is a flasque $\mathscr{O}_X$-module. Show that $\mathscr{F}$ is Γ-acyclic (that $H^i(X, \mathscr{F}) = 0$ for $i > 0$, §23.3.2) as follows. As $Mod_{\mathscr{O}_X}$ has enough injectives, choose an inclusion of $\mathscr{F}$ into some injective $\mathscr{I}$, and call its cokernel $\mathscr{G}$:

$$0 \to \mathscr{F} \to \mathscr{I} \to \mathscr{G} \to 0.$$

Then $\mathscr{I}$ is flasque by Exercise 23.4.F, so $\mathscr{G}$ is flasque by Exercise 23.4.D(b). Take the long exact sequence in (derived functor) cohomology, and show that $H^1(X, \mathscr{F}) = 0$. Your argument works for *any* flasque sheaf $\mathscr{F}$, so $H^1(X, \mathscr{G}) = 0$ as well. Show that $H^2(X, \mathscr{F}) = 0$. Turn this into an induction.

Thus if $\pi : X \to Y$ is a morphism of ringed spaces, and $\mathscr{I}$ is an injective $\mathscr{O}_X$-module, then $\mathscr{I}$ is flasque (Exercise 23.4.F), so $\pi_*\mathscr{I}$ is flasque (Exercise 23.4.E(b)), so $\pi_*\mathscr{I}$ is acyclic for the functor Γ (Exercise 23.4.G), so this completes the proof of the Leray spectral sequence (Theorem 23.4.4). $\square$

23.4.H. EXERCISE. Extend the Leray spectral sequence (Theorem 23.4.4) to deal with a composition of derived pushforwards for

$$(X, \mathscr{O}_X) \xrightarrow{\pi} (Y, \mathscr{O}_Y) \xrightarrow{\rho} (Z, \mathscr{O}_Z).$$

23.4.6. $\star\star$ **The category of $\mathscr{O}_X$-modules need not have enough projectives.** In contrast to Theorem 23.4.1, the category of $\mathscr{O}_X$-modules need not have enough projectives. For example, let X be $\mathbb{P}^1_k$ with the Zariski-topology (in fact we will need very little about X — only that it is not an *Alexandrov space*), but take $\mathscr{O}_X$ to be the constant sheaf $\underline{\mathbb{Z}}$. We will see that $Mod_{\mathscr{O}_X}$ — i.e., the category of sheaves of abelian groups on X — does not have enough projectives. If $Mod_{\mathscr{O}_X}$ had enough projectives, then there would be a surjection $\psi : P \to \underline{\mathbb{Z}}$ from a projective sheaf. Fix a closed point $q \in X$. We will show that the map on stalks $\psi_q : P_q \to \underline{\mathbb{Z}}_q$ is the zero map, contradicting the surjectivity of ψ. For each open subset U of X, denote by $\underline{\mathbb{Z}}_U$ the extension to X of the constant sheaf associated to $\mathbb{Z}$ on U by 0 (Exercise 2.6.G — $\underline{\mathbb{Z}}_U(V) = \mathbb{Z}$ if $V \subset U$, and $\underline{\mathbb{Z}}_U(V) = 0$ otherwise). For each open neighborhood V of q, let W be a strictly smaller open neighborhood. Consider the surjection $\underline{\mathbb{Z}}_{X-q} \oplus \underline{\mathbb{Z}}_W \to \underline{\mathbb{Z}}$. By projectivity of P, the surjection ψ lifts to $P \to \underline{\mathbb{Z}}_{X-q} \oplus \underline{\mathbb{Z}}_W$. The map $P(V) \to \underline{\mathbb{Z}}(V)$ factors through $\underline{\mathbb{Z}}_{X-q}(V) \oplus \underline{\mathbb{Z}}_W(V) = 0$, and hence must be the zero map. Thus the map $\psi_q : P_q \to \underline{\mathbb{Z}}_q$ map is zero as well (do you see why?) as desired.

23.5 Čech cohomology and derived functor cohomology agree

We next prove that Čech cohomology and derived functor cohomology agree, where the former is defined.

23.5.1. Theorem. — *Suppose X is a quasicompact separated scheme, and $\mathscr{F}$ is a quasicoherent sheaf. Then the Čech cohomology of $\mathscr{F}$ agrees with the derived functor cohomology of $\mathscr{F}$.*

This statement is not as precise as it should be. We would want to know that this isomorphism is functorial in $\mathscr{F}$, and that it respects long exact sequences (so the connecting homomorphism defined for Čech cohomology agrees with that for derived functor cohomology). There is also an important extension to higher pushforwards. We leave these issues for the end of this section, §23.5.5.

In case you are curious: so long as it is defined appropriately, Čech cohomology agrees with derived functor cohomology in a wide variety of circumstances outside of scheme theory (if the underlying topological space is paracompact), but not always (see [**Gr1**, §3.8] for a counterexample).

The central idea in the proof (albeit with a twist) is a spectral sequence argument in the same style as those of §23.3, and uses two "cohomology-vanishing" ingredients, one for each orientation of the spectral sequence.

(A) If $(X, \mathscr{O}_X)$ is a ringed space, $\mathscr{Q}$ is an injective $\mathscr{O}_X$-module, and $X = \cup_i U_i$ is a finite open cover, then $\mathscr{Q}$ has no ith Čech cohomology with respect to this cover for $i > 0$.

(B) If X is an affine scheme, and $\mathscr{F}$ is a quasicoherent sheaf on X, then $R^i\Gamma\mathscr{F} = 0$ for $i > 0$.

Translation: **(A)** says that building blocks of derived functor cohomology have no Čech cohomology, and **(B)** says that building blocks of Čech cohomology have no derived functor cohomology.

23.5.A. PRELIMINARY EXERCISE. Suppose $(X, \mathscr{O}_X)$ is a ringed space, $\mathscr{Q}$ is an injective $\mathscr{O}$-module, and $i : U \hookrightarrow X$ is an open subset. Show that $\mathscr{Q}|_U$ is injective on U. Hint: use the fact that i^{-1} has an *exact* left adjoint $i_!$ (extension by zero), see Exercise 2.6.G, and the following diagrams.

$$\begin{tikzcd} 0 \arrow[r] & \mathscr{A} \arrow[r] \arrow[d] & \mathscr{B} \arrow[dl, dotted, "?"] \\ & \mathscr{Q}|_U & \end{tikzcd} \qquad \begin{tikzcd} 0 \arrow[r] & i_!\mathscr{A} \arrow[r] \arrow[d] & i_!\mathscr{B} \arrow[dl, dotted, "?"] \\ & \mathscr{Q} & \end{tikzcd}$$

In the course of Exercise 23.5.A, you will have proved the following fact, which we shall use again in Exercise 30.3.C. (You can also use it to solve Exercise 23.4.A.)

23.5.B. EXERCISE. Show that if (F, G) is an adjoint pair of additive functors between abelian categories, and F is exact, then G sends injective elements to injective elements.

23.5.2. *Proof of Theorem 23.5.1, assuming* **(A)** *and* **(B)**. As with the facts proved in §23.3, we take the only approach that is reasonable: we choose an injective resolution $0 \to \mathscr{F} \to \mathscr{Q}_\bullet$ of $\mathscr{F}$ and a Čech cover of X, mix these two types of information in a double complex, and toss it into our spectral sequence machine (§1.7). More precisely, choose a finite affine open cover $X = \cup_i U_i$ and an injective resolution

$$0 \to \mathscr{F} \to \mathscr{Q}_0 \to \mathscr{Q}_1 \to \cdots .$$

Consider the double complex

(23.5.2.1)
$$\begin{tikzcd}
\vdots & \vdots & \vdots & \vdots & \\
0 \arrow[u] \arrow[r] & \oplus_i \mathscr{Q}_2(U_i) \arrow[u] \arrow[r] & \oplus_{i,j} \mathscr{Q}_2(U_{ij}) \arrow[u] \arrow[r] & \oplus_{i,j,k} \mathscr{Q}_2(U_{ijk}) \arrow[u] \arrow[r] & \cdots \\
0 \arrow[u] \arrow[r] & \oplus_i \mathscr{Q}_1(U_i) \arrow[u] \arrow[r] & \oplus_{i,j} \mathscr{Q}_1(U_{ij}) \arrow[u] \arrow[r] & \oplus_{i,j,k} \mathscr{Q}_1(U_{ijk}) \arrow[u] \arrow[r] & \cdots \\
0 \arrow[u] \arrow[r] & \oplus_i \mathscr{Q}_0(U_i) \arrow[u] \arrow[r] & \oplus_{i,j} \mathscr{Q}_0(U_{ij}) \arrow[u] \arrow[r] & \oplus_{i,j,k} \mathscr{Q}_0(U_{ijk}) \arrow[u] \arrow[r] & \cdots \\
0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] &
\end{tikzcd}$$

We take this as the E_0 term in a spectral sequence. First, we use the rightward filtration. As higher Čech cohomology of injective $\mathscr{O}$-modules is 0 (assumption **(A)**), we get 0's everywhere except in "column 0", where we get $\mathscr{Q}_i(X)$ in row i:

$$\begin{tikzcd}
\vdots & \vdots & \vdots & \vdots & \\
0 \arrow[u] & \mathscr{Q}_2(X) \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & \cdots \\
0 \arrow[u] & \mathscr{Q}_1(X) \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & \cdots \\
0 \arrow[u] & \mathscr{Q}_0(X) \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & \cdots \\
0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] &
\end{tikzcd}$$

Then we take cohomology in the vertical direction, and we get derived functor cohomology of $\mathscr{F}$ on X on the E_2 page:

$$\begin{tikzcd}
\vdots & \vdots & \vdots & \vdots & \\
0 & R^2\Gamma(X,\mathscr{F}) & 0 & 0 & \cdots \\
0 & R^1\Gamma(X,\mathscr{F}) & 0 & 0 & \cdots \\
0 & \Gamma(X,\mathscr{F}) \arrow[uul] & 0 \arrow[uul] & 0 \arrow[uul] & \cdots \\
 & 0 \arrow[uul] & 0 \arrow[uul] & 0 \arrow[uul] &
\end{tikzcd}$$

We then start over on the E_0 page, and this time use the filtration corresponding to choosing the upward arrow first. By Proposition 23.5.A, $\mathscr{Q}_i|_{U_J}$ is injective on U_J, so we are computing the derived functor cohomology of $\mathscr{F}$ on U_J. Then the higher derived functor cohomology is 0 (assumption **(B)**), so all entries are 0 except possibly on row 0. Thus the E_1 term is:

(23.5.2.2)

$$\begin{tikzcd}
0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & \cdots \\
0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & \cdots \\
0 \arrow[r] & \oplus_i\Gamma(U_i,\mathscr{F}) \arrow[r] & \oplus_{i,j}\Gamma(U_{ij},\mathscr{F}) \arrow[r] & \oplus_{i,j,k}\Gamma(U_{ijk},\mathscr{F}) \arrow[r] & \cdots \\
0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & 0 \arrow[r] & \cdots
\end{tikzcd}$$

Row 0 is precisely the Čech complex of $\mathscr{F}$, so the spectral sequence converges at the E_2 term, yielding the Čech cohomology. Since one orientation yields derived functor cohomology and one yields Čech cohomology, we are done. $\square$

So it remains to show **(A)** and **(B)**.

23.5.3. Ingredient (A): injectives have no Čech cohomology.

23.5.C. EXERCISE. Suppose $X = \cup_j U_j$ is a finite cover of X by open sets, and $\mathscr{F}$ is a flasque sheaf (Definition 23.4.5) on X. Show that the Čech complex for $\mathscr{F}$ with respect to $\cup_j U_j$ has no cohomology in positive degree, i.e., that it is exact except in degree 0 (where it has cohomology $\mathscr{F}(X)$, by the sheaf axioms). Hint: use induction on j. Consider the short exact sequence of complexes (18.2.4.2) (see also (18.2.3.1)). The corresponding long exact sequence will immediately give the desired result for $i > 1$, and flasqueness will be used for $i = 1$.

Thus flasque sheaves have no Čech cohomology, so injective $\mathscr{O}$-modules in particular (Exercise 23.4.F) have none. This is all we need for our algebro-geometric applications, but to show you how general this machinery is, we give an entertaining application.

23.5.D. Unimportant Exercise (perverse proof of inclusion-exclusion through cohomology of sheaves). The inclusion-exclusion principle is (equivalent to) the following: suppose that X is a finite set, and U_i ($1 \leq i \leq n$) are finite sets covering X. As usual, define $U_I = \cap_{i \in I} U_i$ for $I \subset \{1, \ldots, n\}$. Then

$$|X| = \sum |U_i| - \sum_{|I|=2} |U_{|I|}| + \sum_{|I|=3} |U_{|I|}| - \sum_{|I|=4} |U_{|I|}| + \cdots .$$

Prove this by endowing X with the discrete topology, showing that the constant sheaf $\underline{\mathbb{Q}}$ is flasque, considering the Čech complex computing $H^i(X, \underline{\mathbb{Q}})$ using the cover $\{U_i\}$, and using Exercise 1.6.B.

23.5.4. $\star$ **Ingredient (B): quasicoherent sheaves on affine schemes have no derived functor cohomology (M. Olsson).**

We show the following statement by induction on k. Suppose X is an affine scheme, and $\mathscr{F}$ is a quasicoherent sheaf on X. Then $R^i\Gamma(X, \mathscr{F}) = 0$ for $0 < i \leq k$. The result is vacuously true for $k = 0$; so suppose we know the result for all $0 < k' < k$ (with X replaced by *any* affine scheme). Suppose $\alpha \in R^k\Gamma(X, \mathscr{F})$. We wish to show that $\alpha = 0$. Choose an injective resolution *by* $\mathscr{O}_X$*-modules*

$$0 \longrightarrow \mathscr{F} \longrightarrow \mathscr{Q}_0 \xrightarrow{d_0} \mathscr{Q}_1 \xrightarrow{d_1} \cdots .$$

Then α has a representative α' in $\mathscr{Q}_k(X) = \Gamma(X, \mathscr{Q}_k)$, such that $d\alpha' = 0$. Because the injective resolution is exact, α' is locally a boundary. In other words, in the neighborhood of any point $p \in X$, there is an open set V_p such that $\alpha|_{V_p} = d\alpha'$ for some $\alpha' \in \mathscr{Q}_{k-1}(V_p)$. (Be sure you see why this is true! Recall that taking cokernel of a map of sheaves requires sheafification, see Proposition 2.5.1.) By shrinking V_p if necessary, we can assume V_p is affine. By the quasicompactness of X, we can choose a finite number of the V_p's that cover X. Rename these U_i, so we have an affine cover X. Consider the Čech cover of X with respect to *this* affine cover (*not* the affine cover you might have thought we would use — that of X by itself — but instead an affine cover tailored to our particular α). Consider the double complex (23.5.2.1), as the E_0 term in a spectral sequence.

First consider the rightward orientation. As in the argument in §23.5.2, the spectral sequence converges at E_2, where we get 0 everywhere, except that the derived functor cohomology appears in the 0th column.

Next, start over again, choosing the upward filtration. On the E_1 page, row 0 is the Čech complex, as in (23.5.2.2). All the rows between 1 and $k-1$ are 0 by our inductive hypothesis, but we don't yet know anything about the higher rows. Because we are interested in the kth derived functor, we focus on the kth antidiagonal ($E_\bullet^{p,k-p}$). The only possibly nonzero terms in this antidiagonal are $E_1^{k,0}$ and $E_1^{0,k}$. We look first at the term on the bottom row $E_1^{k,0} = \prod_{|I|=k} \Gamma(U_I, \mathscr{F})$, which is part of the Čech complex:

$$\cdots \to \prod_{|I|=k-1} \Gamma(U_I, \mathscr{F}) \to \prod_{|I|=k} \Gamma(U_I, \mathscr{F}) \to \prod_{|I|=k+1} \Gamma(U_I, \mathscr{F}) \to \cdots.$$

But we have already verified that the Čech cohomology of a quasicoherent sheaf on an affine scheme vanishes — this is the one spot where we use the quasicoherence of $\mathscr{F}$. Thus this term vanishes by the E_2 page (i.e., $E_i^{k,0} = 0$ for $i \geq 2$).

So the only term of interest in the kth antidiagonal of E_1 is $E_1^{0,k}$, which is the homology of

$$\prod_i \mathscr{Q}_{k-1}(U_i) \to \prod_i \mathscr{Q}_k(U_i) \to \prod_i \mathscr{Q}_{k+1}(U_i), \tag{23.5.4.1}$$

which is $\prod_i R^k\Gamma(U_i, \mathscr{F})$ (using Preliminary Exercise 23.5.A which stated that the $\mathscr{Q}_j|_{U_i}$ are injective on U_i, and they can be used to compute $R^k(\Gamma(U_i, \mathscr{F}))$. So $E_2^{0,k}$ is the homology of

$$0 \to \prod_i R^k\Gamma(U_i, \mathscr{F}) \to \prod_{i,j} R^k\Gamma(U_{ij}, \mathscr{F})$$

and thereafter all differentials to and from the $E_\bullet^{0,k}$ terms will be 0, as the sources and targets of those arrows will be 0. Consider now our lift of α' of our original class $\alpha \in R^k\Gamma(X, \mathscr{F})$ to $\prod_i R^k\Gamma(U_i, \mathscr{F})$. Its image in the homology of (23.5.4.1) is *zero* — this was how we chose our cover U_i to begin with! Thus $\alpha = 0$ as desired, completing our proof. $\square$

23.5.E. ⋆⋆ EXERCISE. The proof is not quite complete. We have a class $\alpha \in R^k\Gamma(X, \mathscr{F})$, and we have interpreted $R^k\Gamma(X, \mathscr{F})$ as

$$\ker\left(\prod_i R^k\Gamma(U_i, \mathscr{F}) \to \prod_{i,j} R^k\Gamma(U_{ij}, \mathscr{F})\right).$$

We have two maps $R^k\Gamma(X, \mathscr{F}) \to R^k\Gamma(U_i, \mathscr{F})$, one coming from the natural restriction (under which we can see that the image of α is zero), and one coming from the actual spectral sequence machinery. Verify that they are the same map. (Possible hint: with the filtration used, the $E_\infty^{0,k}$ term is indeed the quotient of the homology of the double complex, so the map goes the right way.)

23.5.5. ⋆ **Tying up loose ends.**

23.5.F. IMPORTANT EXERCISE. State and prove the generalization of Theorem 23.5.1 to higher pushforwards $R^i\pi_*$, where $\pi : X \to Y$ is a quasicompact separated morphism of schemes.

23.5.G. EXERCISE. Show that the isomorphism of Theorem 23.5.1 is functorial in $\mathscr{F}$, i.e., given a morphism $\mathscr{F} \to \mathscr{G}$, the diagram

$$\begin{tikzcd} H^i(X, \mathscr{F}) \arrow[r, leftrightarrow] \arrow[d] & R^i\Gamma(X, \mathscr{F}) \arrow[d] \\ H^i(X, \mathscr{G}) \arrow[r, leftrightarrow] & R^i\Gamma(X, \mathscr{G}) \end{tikzcd}$$

commutes, where the horizontal arrows are the isomorphisms of Theorem 23.5.1, and the vertical arrows come from functoriality of H^i and $R^i\Gamma$. (Hint: "spectral sequences are functorial in E_0", which can be easily seen from the construction, although we haven't said it explicitly.)

23.5.H. EXERCISE. Show that the isomorphisms of Theorem 23.5.1 induce isomorphisms of long exact sequences.

23.5.6. *Remark.* If you wish, you can use the above argument to prove the following theorem of Leray. Suppose we have a sheaf of abelian groups $\mathscr{F}$ on a topological space X, and some covering $\{U_i\}$ of X such that the (derived functor) cohomology of $\mathscr{F}$ in positive degree vanishes on every finite intersection of the U_i. Then the cohomology of $\mathscr{F}$ can be calculated by the Čech cohomology of the cover $\{U_i\}$; there is no need to pass to the inductive limit of all covers, as is the case for Čech cohomology in general.

23.5.7. *Unimportant Remark: Working in $QCoh_X$ rather than $Mod_{\mathscr{O}_X}$.* In our definition of derived functors of quasicoherent sheaves on X, we could have tried to work in the category of quasicoherent sheaves $QCoh_X$ itself, rather than in the larger category $Mod_{\mathscr{O}_X}$. There are several reasons why this would require more effort. It is not hard to show that $QCoh_X$ has enough injectives if X is Noetherian (see for example [**Ha1**, Exer. III.3.6(a)]). Because we don't have "extension by zero" (Exercise 2.6.G) in *QCoh*, the proofs that injective quasicoherent sheaves on an open set U restrict to injective quasicoherent sheaves on smaller open subsets V (the analog of Exercise 23.5.A) and that injective quasicoherent sheaves are flasque (the analog of Exercise 23.4.F) are harder. You can use this to show that H^i and $R^i\pi_*$ computed in *QCoh* are the same as those computed in $Mod_{\mathscr{O}}$ (once you make the statements precise). It is true that injective elements of $QCoh_X$ (X Noetherian) are injective in $Mod_{\mathscr{O}_X}$, but this requires work (see [**Mur**, Prop. 68]).

It is true that $QCoh_X$ has enough injectives for *any* scheme X, but this is much harder, see [**EE**]. And as is clear from the previous paragraph, "enough injectives" is only the beginning of what we want.

23.5.8. *Unimportant Remark.* Theorem 23.5.1 implies that if $\pi : X \to Y$ is quasicompact and separated, then $R^i\pi_*$ sends $QCoh_X$ to $QCoh_Y$ (by showing an isomorphism with Čech cohomology). If X and Y are Noetherian, the hypothesis "separated" can be relaxed to "quasiseparated", which can be shown using the ideas of Unimportant Remark 23.5.7. This is not nearly as useful as the separated case, because without Čech cohomology, it is hard to compute anything.

CHAPTER 24

Flatness

The concept of flatness is a riddle that comes out of algebra, but which technically is the answer to many prayers.
— *D. Mumford* [**Mu7**, III.10]

It is a riddle, wrapped in a mystery, inside an enigma; but perhaps there is a key.
— *W. Churchill*

24.1 Introduction

We come next to the important concept of flatness (first introduced in §16.3.8). We could have discussed flatness at length as soon as we had discussed quasicoherent sheaves and morphisms. But it is an unexpected idea, and the algebra and geometry are not obviously connected, so we have left it for relatively late. The translation of the french word "plat" that best describes this notion is "phat", but unfortunately that word had not yet been coined when flatness first made its appearance.

Serre has stated that he introduced flatness purely for reasons of algebra in his landmark "GAGA" paper [**Se2**], and that it was Grothendieck who recognized its geometric significance.

A flat morphism $\pi : X \to Y$ is the right notion of a "nice", or "nicely varying" family over Y. For example, if π is a projective flat family over a connected Noetherian base (translation: $\pi : X \to Y$ is a projective flat morphism, with Y connected and Noetherian), we will see that various numerical invariants of fibers are constant, including the dimension (§24.5.4), and numbers interpretable in terms of an Euler characteristic (see §24.7):

(a) the Hilbert polynomial (Corollary 24.7.2),
(b) the degree (in projective space) (Exercise 24.7.B(a)),
(b') (as a special case of (b)) if π is finite, the degree of π (recovering and extending the fact that the degree of a projective map between regular curves is constant, §17.4.4, see Exercise 24.4.H and §24.4.11),
(c) the arithmetic genus (Exercise 24.7.B(b)),
(d) the degree of a line bundle if the fiber is a curve (Corollary 24.7.3), and
(e) intersections of divisors and line bundles (Exercise 24.7.D).

One might think that the right hypothesis might be smoothness, or more generally some sort of equisingularity, but we only need something weaker. And this is a good thing: branched covers are not fibrations in any traditional sense, yet they still behave well — the double cover $\mathbb{A}^1 \to \mathbb{A}^1$ given by $y \mapsto x^2$ has constant degree 2 (§9.3.3, revisited in §17.4.8). Another key example is that of a family of smooth curves degenerating to a nodal curve (Figure 24.1) — the topology of the (underlying analytic) curve changes, but the arithmetic genus remains constant. One can prove things about regular curves by first proving them about a nodal degeneration, and then showing that the result behaves well in flat families. Degeneration techniques such as this are ubiquitous in algebraic geometry.

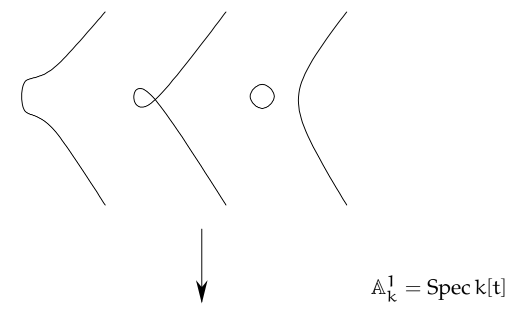


FIGURE 24.1. A flat family of smooth curves degenerating to a nodal curve: $y^2 = x^3 - tx^2$.

Given the cohomological nature of the constancy of Euler characteristic result, you should not be surprised that the hypothesis needed (flatness) is cohomological in nature — it can be characterized by vanishing of Tor (Exercise 23.1.D), which we use to great effect in §24.3.

But flatness is important for other reasons too. As a start: as this is the right notion of a "nice family", it allows us to correctly define the notion of moduli space. For example, the *Hilbert scheme* of $\mathbb{P}^n$ "parametrizes closed subschemes of $\mathbb{P}^n$". Maps from a scheme B to the Hilbert scheme correspond to (finitely presented) closed subschemes of $\mathbb{P}^n_B$ *flat* over B. By universal property nonsense, this defines the Hilbert scheme up to unique isomorphism (although we of course must show that it exists, which takes some effort — [**Mu2**] gives an excellent exposition). The moduli space of smooth projective curves is defined by the universal property that maps to the moduli space correspond to projective flat (finitely presented) families whose geometric fibers are smooth curves. (Sadly, this moduli space does not exist...) On a related note, flatness is central in deformation theory: it is key to understanding how schemes (and other geometric objects, such as vector bundles) can deform (cf. §21.5.12). Finally, the notion of Galois descent generalizes to (faithfully) "flat descent", which allows us to "glue" in more exotic Grothendieck topologies in the same way we do in the Zariski topology (or more classical topologies); but this is beyond the scope of our current discussion.

24.1.1. Structure of the chapter.

Flatness has many aspects of different flavors, and it is easy to lose sight of the forest for the trees. Because the algebra of flatness seems so unrelated to the geometry, it can be nonintuitive. We will necessarily begin with algebraic foundations, but you should focus on the following points: methods of showing things are flat (both general criteria and explicit examples), and classification of flat modules over particular kinds of rings. You should try every exercise dealing with explicit examples such as these.

Here is an outline of the chapter, to help focus your attention.

- In §24.2, we discuss some of the easier facts, which are algebraic in nature.
- §24.3, §24.4, and §24.6 give ideal-theoretic criteria for flatness. §24.3 and §24.4 should be read together. The first uses Tor to understand flatness, and the second uses these insights to develop ideal-theoretic criteria for flatness. §24.6, on local criteria for flatness, is harder.
- §24.5 is relatively free-standing, and could be read immediately after §24.2. It deals with topological aspects of flatness, such as the fact that flat morphisms are open in good situations.
- In §24.7, we explain fact that "the Euler characteristic of quasicoherent sheaves is constant in flat families" (with appropriate hypotheses), and its many happy consequences. This section is surprisingly easy given its utility.

You should focus on what flatness implies and how to "picture" it, but also on explicit criteria for flatness in different situations, such as for integral domains (Observation 24.2.2), principal ideal domains (Exercise 24.4.B), discrete valuation rings (Exercise 24.4.C), the dual numbers (Exercise 24.4.D), and local rings (Theorem 24.4.5).

24.2 Easier facts

Many facts about flatness are easy or immediate, although a number are tricky. As always, I will try to make clear which is which, to help you remember the easy facts and the key ideas of proofs of the harder facts. We will pick the low-hanging fruit first.

We recall the definition of a *flat* A*-module* (§1.6.11). If $M \in Mod_A$, $M \otimes_A \cdot$ is always right-exact (Exercise 1.3.H). We say that M **is a flat** A**-module** (or *flat over* A or A-*flat*) if $M \otimes_A \cdot$ is an exact functor. We say that a *ring morphism* $B \to A$ *is* **flat** if A is flat as a B-module. (In particular, the algebra structure of A is irrelevant.)

24.2.1. *Two key examples.*

(i) Free A-modules (even of infinite rank) are clearly flat. More generally, projective modules are flat (Exercise 23.2.B).

(ii) Localizations are flat: Suppose S is a multiplicative subset of B. Then $B \to S^{-1}B$ is a flat ring morphism (Exercise 1.6.F(a)).

24.2.A. EASY EXERCISE: FIRST EXAMPLES.
(a) *(trick question)* Classify flat modules over a field k.
(b) Show that $A[x_1, \dots, x_n]$ is a flat A-module.
(c) Show that the ring morphism $\mathbb{Q}[x] \to \mathbb{Q}[y]$, with $x \mapsto y^2$, is flat. (This will help us understand Example 9.3.3 better, see §24.4.11.)

We make some quick but important observations.

24.2.2. *Important Observation.* If x is a non-zerodivisor of A, and M is a flat A-module, then $M \xrightarrow{\times x} M$ is injective. (Reason: apply the exact functor $M \otimes_A$ to the exact sequence $0 \longrightarrow A \xrightarrow{\times x} A$.) In particular, *flat modules are torsion-free*. (Torsion-freeness was defined in §13.5.4.) This observation gives an easy way of recognizing when a module is *not* flat. We will use it many times.

24.2.B. EXERCISE: ANOTHER EXAMPLE. Show that a finitely generated module over a discrete valuation ring is flat if and only if it is torsion-free if and only if it is free. Hint: Remark 12.5.15 classifies finitely generated modules over a discrete valuation ring. (Exercise 24.4.B sheds more light on flatness over a discrete valuation ring. Proposition 13.7.3 is also related.)

24.2.C. EXERCISE (FLATNESS IS PRESERVED BY CHANGE OF BASE RING). Show that if M is a flat B-module, $B \to A$ is a homomorphism, then $M \otimes_B A$ is a flat A-module.

24.2.D. EXERCISE (TRANSITIVITY OF FLATNESS). Show that if A is a flat B-algebra, and M is A-flat, then M is also B-flat.

24.2.3. Proposition (flatness is a stalk/prime-local property). — *An A-module M is flat if and only if $M_{\mathfrak{p}}$ is a flat $A_{\mathfrak{p}}$-module for all primes $\mathfrak{p}$.*

Proof. Suppose first that M is a flat A-module. Given any exact sequence of $A_{\mathfrak{p}}$-modules

$$0 \to N' \to N \to N'' \to 0, \tag{24.2.3.1}$$

$$0 \to M \otimes_A N' \to M \otimes_A N \to M \otimes_A N'' \to 0$$

is exact too. But $M \otimes_A N$ is canonically isomorphic to $M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N$ (do you see why?), so $M_{\mathfrak{p}}$ is a flat $A_{\mathfrak{p}}$-module.

Suppose next that $M_{\mathfrak{p}}$ is a flat $A_{\mathfrak{p}}$-module for all $\mathfrak{p}$. Given any short exact sequence (24.2.3.1), tensoring with M yields

$$0 \longrightarrow K \longrightarrow M \otimes_A N' \longrightarrow M \otimes_A N \longrightarrow M \otimes_A N'' \longrightarrow 0 \tag{24.2.3.2}$$

(using right-exactness of $\otimes$, Exercise 1.3.H) where K is the kernel of $M \otimes_A N' \to M \otimes_A N$. We wish to show that $K = 0$. It suffices to show that $K_{\mathfrak{p}} = 0$ for every prime $\mathfrak{p} \subset A$ (see the comment after Exercise 4.3.F). Given any $\mathfrak{p}$, localizing (24.2.3.1) at $\mathfrak{p}$ and tensoring with the exact $A_{\mathfrak{p}}$-module $M_{\mathfrak{p}}$ yields

$$0 \longrightarrow M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N'_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N''_{\mathfrak{p}} \longrightarrow 0. \tag{24.2.3.3}$$

But localizing (24.2.3.2) at $\mathfrak{p}$ and using the isomorphisms $M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N_{\mathfrak{p}} \cong (M \otimes_A N')_{A_{\mathfrak{p}}}$, we obtain the exact sequence

$$0 \longrightarrow K_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N'_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N_{\mathfrak{p}} \longrightarrow M_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} N''_{\mathfrak{p}} \longrightarrow 0,$$

which is the same as the exact sequence (24.2.3.3) except for the $K_{\mathfrak{p}}$. Hence $K_{\mathfrak{p}} = 0$ as desired. ∎

24.2.4. Flatness for schemes.

Motivated by Proposition 24.2.3, the extension of the notion of flatness to schemes is straightforward.

24.2.5. *Definition: flat quasicoherent sheaves.* We say that a quasicoherent sheaf $\mathscr{F}$ on a scheme X is **flat at** $p \in X$ if $\mathscr{F}_p$ is a flat $\mathscr{O}_{X,p}$-module. We say that a quasicoherent sheaf $\mathscr{F}$ on a scheme X is **flat** (over X) if it is flat at all $p \in X$. In light of Proposition 24.2.3, we can check this notion on affine open cover of X.

24.2.6. *Definition: flat morphism.* Similarly, we say that a morphism of schemes $\pi : X \to Y$ is **flat at** $p \in X$ if $\mathscr{O}_{X,p}$ is a flat $\mathscr{O}_{Y,\pi(p)}$-module. We say that a morphism of schemes $\pi : X \to Y$ is **flat** if it is flat at all $p \in X$. We can check flatness locally on the source and target.

We can combine these two definitions into a single fancy definition.

24.2.7. *Definition: flat quasicoherent sheaf over a base.* Suppose $\pi : X \to Y$ is a morphism of schemes, and $\mathscr{F}$ is a quasicoherent sheaf on X. We say that $\mathscr{F}$ **is flat (over** Y**) at** $p \in X$ if $\mathscr{F}_p$ is a flat $\mathscr{O}_{Y,\pi(p)}$-module. We say that $\mathscr{F}$ **is flat (over** Y**)** if it is flat at all $p \in X$.

Definitions 24.2.5 and 24.2.6 correspond to the cases $X = Y$ and $\mathscr{F} = \mathscr{O}_X$ respectively. (Definition 24.2.7 applies without change to the category of ringed spaces, but we won't use this.)

24.2.E. EASY EXERCISE (REALITY CHECK). Show that open embeddings are flat.

Our results about flatness over rings above carry over easily to schemes.

24.2.F. EXERCISE. Show that a map of rings $B \to A$ is flat if and only if the corresponding morphism of schemes $\operatorname{Spec} A \to \operatorname{Spec} B$ is flat. More generally, if $B \to A$ is a map of rings, and M is an A-module, show that M is B-flat if and only if $\widetilde{M}$ is flat over $\operatorname{Spec} B$.

24.2.G. EASY EXERCISE (EXAMPLES AND REALITY CHECKS).
(a) If X is a scheme, and $p \in X$, show that the natural morphism $\operatorname{Spec} \mathscr{O}_{X,p} \to X$ is flat. (Hint: localization is flat, §24.2.1.)
(b) Show that $\mathbb{A}^n_A \to \operatorname{Spec} A$ is flat.
(c) If $\mathscr{F}$ is a locally free sheaf on a scheme X, show that $\mathbb{P}\mathscr{F} \to X$ (Definition 17.2.3) is flat.
(d) Show that $\operatorname{Spec} k \to \operatorname{Spec} k[t]/(t^2)$ is not flat. (Draw a picture to try to see what is not "nice" about this morphism. Some more insight about flatness of the dual numbers will be given in the criterion of Exercise 24.4.D.)

24.2.H. EXERCISE (TRANSITIVITY OF FLATNESS). Suppose $\pi : X \to Y$ and $\mathscr{F}$ is a quasicoherent sheaf on X, flat over Y. Suppose also that $\psi : Y \to Z$ is a flat morphism. Show that $\mathscr{F}$ is flat over Z.

24.2.I. EXERCISE (FLATNESS IS PRESERVED BY BASE CHANGE). Suppose $\pi : X \to Y$ is a morphism, and $\mathscr{F}$ is a quasicoherent sheaf on X, flat over Y. If $\rho : Y' \to Y$ is any morphism, and $\rho' : X \times_Y Y' \to X$ is the induced morphism, show that $(\rho')^*\mathscr{F}$

is flat over Y'.

$$\begin{tikzcd} (\rho')^*\mathscr{F} \arrow[d, dash] & \mathscr{F} \arrow[d, dash] \\ X \times_Y Y' \arrow[r, "\rho'"] \arrow[d] & X \arrow[d, "\pi"] \\ Y' \arrow[r, "\rho"] & Y \end{tikzcd}$$

In particular, using Exercise 24.2.A(a), if X and Y' are k-schemes, and $\mathscr{F}$ is *any* quasicoherent sheaf on X, then $(\rho')^*\mathscr{F}$ is flat over Y'. For example, $X \times_k Y'$ is *always* flat over Y' — "products over a field are flat over their factors". (Feel free to immediately generalize this further; for example, $\mathscr{F}$ can be a quasicoherent sheaf on a scheme Z over X, flat over Y.)

The following exercise is very useful for visualizing flatness and non-flatness (see for example Figure 24.2).

24.2.J. Exercise (flat maps send associated points to associated points). Suppose $\pi : X \to Y$ is a flat morphism of locally Noetherian schemes. Show that any associated point of X must map to an associated point of Y. (Feel free to immediately generalize this to a coherent sheaf $\mathscr{F}$ on X, flat over Y, without π itself needing to be flat.) Hint: suppose $\pi^\sharp : (B, \mathfrak{n}) \to (A, \mathfrak{m})$ is a local morphism of Noetherian local rings (i.e., $\pi^\sharp(\mathfrak{n}) \subset \mathfrak{m}$, §6.3.1). Suppose $\mathfrak{n}$ is not an associated prime of B. Show that there is an element $f \in \mathfrak{n}$ not in any associated prime of B (perhaps using prime avoidance, Proposition 11.2.13), and hence is a non-zerodivisor. Show that $\pi^\sharp f \in \mathfrak{m}$ is a non-zerodivisor of A using Observation 24.2.2, and thus show that $\mathfrak{m}$ is not an associated prime of A.

24.2.K. Exercise. Use Exercise 24.2.J to show that the following morphisms are not flat (see Figure 24.2):

(a) $\operatorname{Spec} k[x,y]/(xy) \to \operatorname{Spec} k[x]$,
(b) $\operatorname{Spec} k[x,y]/(y^2, xy) \to \operatorname{Spec} k[x]$,
(c) $\operatorname{Bl}_{(0,0)} \mathbb{A}^2_k \to \mathbb{A}^2_k$.

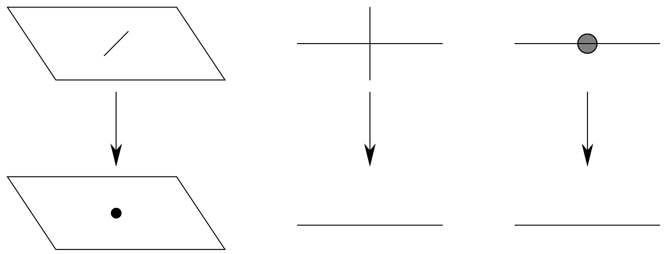
Hint for (c): first pull back to a line through the origin to obtain a something akin to (a). (This foreshadows the statement and proof Proposition 24.5.5, which says that for flat morphisms "there is no jumping of fiber dimension".)

24.2.L. Exercise. Show that the earlier definition of flat morphism in Remark 16.3.8 agrees with Definition 24.2.6. (We will not use the earlier definition, so in this sense the exercise is unimportant.)

24.2.8. Theorem (cohomology commutes with flat base change). — *Suppose*

$$\begin{tikzcd} X' \arrow[r, "g'"] \arrow[d, "f'"'] & X \arrow[d, "f"] \\ Y' \arrow[r, "g"] & Y \end{tikzcd}$$

is a fiber diagram, and f (and thus f') is quasicompact and separated (so higher pushforwards of quasicoherent sheaves by f and f' exist, as described in §18.8). Suppose also



(a) (b) (c)

[Figure to be updated to reflect ordering in Exercise 24.2.K later]

FIGURE 24.2. Morphisms that are not flat (Exercise 24.2.K)

that g *is flat, and* $\mathscr{F}$ *is a quasicoherent sheaf on* X*. Then the natural morphisms (Exercise 18.8.B(a))* $g^*(R^i f_* \mathscr{F}) \to R^i f'_*(g'^* \mathscr{F})$ *are isomorphisms.*

24.2.M. EXERCISE. Prove Theorem 24.2.8. Hint: Exercise 18.8.B(b) is the special case where f is affine. Extend it to the quasicompact and separated case using the same idea as the proof of Theorem 16.2.1 (which was actually proved in Exercise 13.3.F, using Exercise 13.3.E). Your proof of the case $i = 0$ will only need a quasiseparated hypothesis in place of the separated hypothesis.

A useful special case is where Y' is the generic point of a reduced component of Y. In other words, in light of Exercise 24.2.G(a), the stalk of the higher pushforward of $\mathscr{F}$ at the generic point is the cohomology of $\mathscr{F}$ on the fiber over the generic point. This is a first example of something important: understanding cohomology of (quasicoherent sheaves on) fibers in terms of higher pushforwards. (We would certainly hope that higher pushforwards would tell us something about higher cohomology of fibers, but this is certainly not a priori clear!) In comparison to this result, which shows that cohomology of *any* quasicoherent sheaf commutes with *flat* base change, §24.7 and Chapter 28 deal with when and how cohomology of a *flat* quasicoherent sheaf commutes with *any* base change.

24.2.9. Pulling back closed subschemes (and ideal sheaves) by flat morphisms. Closed subschemes pull back particularly well under flat morphisms, and this can be helpful to keep in mind. As pointed out in Remarks 16.3.8 and 16.3.9, in the case of flat morphisms, pullback of ideal sheaves *as quasicoherent sheaves* agrees with pullback in terms of the pullback of the corresponding closed subschemes. In other words, closed subscheme exact sequences pull back (remain exact) under flat pullbacks. This is in fact not just a necessary condition for flatness; it is also sufficient, which can be shown using the ideal-theoretic criterion for flatness (Theorem 24.4.1). There is an analogous fact about pulling ideal sheaves of *flat* subschemes by *arbitrary* pullbacks, see §24.3.2.

24.2.N. EXERCISE.
(a) Suppose D is an effective Cartier divisor on Y and $\pi : X \to Y$ is a flat morphism.

Show that the pullback of D to X (by π) is also an effective Cartier divisor.
(b) Use part (a) to show that under a flat morphism, regular embeddings pull back to regular embeddings.

24.2.O. UNIMPORTANT EXERCISE.
(a) Suppose $\pi : X \to Y$ is a morphism, and $Z \hookrightarrow Y$ is a closed embedding cut out by an ideal sheaf $\mathscr{I} \subset \mathscr{O}_Y$. Show that $(\pi^*\mathscr{I})^n = \pi^*(\mathscr{I}^n)$.
(b) Suppose further that π is flat, $Y = \mathbb{A}^n_k$, and Z is the origin. Let $\mathscr{J} = \pi^*\mathscr{I}$ be the quasicoherent sheaf of algebras on X cutting out the pullback W of Z. Prove that the graded sheaf of algebras $\oplus_{n \geq 0} \mathscr{J}^n/\mathscr{J}^{n+1}$ (do you understand the multiplication?) is isomorphic to $\mathscr{O}_W[x_1, \dots, x_n]$ (interpreted as a graded sheaf of algebras). (Hint: first show that $\mathscr{J}^n/\mathscr{J}^{n+1} \cong \mathrm{Sym}^n(\mathscr{J}/\mathscr{J}^2)$.)

24.2.P. UNIMPORTANT EXERCISE.
(a) Show that blowing up commutes with flat base change. More precisely, if $\pi : X \to Y$ is any morphism, and $Z \hookrightarrow Y$ is any closed embedding, give a canonical isomorphism $(\mathrm{Bl}_Z Y) \times_Y X \cong \mathrm{Bl}_{Z \times_Y X} X$. (You can proceed by universal property, using Exercise 24.2.N(a), or by using the Proj construction of the blow up and Exercise 24.2.O.)
(b) Give an example to show that blowing up does not commute with base change in general.

24.3 Flatness through Tor

We defined the Tor (bi-)functor in §23.1: $\mathrm{Tor}_i^A(M, N)$ is obtained by taking a free resolution of N, removing the N, tensoring it with M, and taking homology. Exercise 23.1.D characterized flatness in terms of Tor: M is A-flat if $\mathrm{Tor}_1^A(M, N) = 0$ for all N. In this section, we reap the easier benefits of this characterization, recalling key properties of Tor when needed. In §24.4, we work harder to extract more from Tor.

It is sometimes possible to compute Tor from its definition, as shown in the following exercise that we will use repeatedly.

24.3.A. EXERCISE. Define $(M : x)$ as $\{m \in M \ : \ xm = 0\} \subset M$ — it consists of the elements of M annihilated by x. If x is not a zerodivisor, show that

$$\mathrm{Tor}_i^A(M, A/(x)) = \begin{cases} M/xM & \text{if } i = 0; \\ (M : x) & \text{if } i = 1; \\ 0 & \text{if } i > 1. \end{cases}$$

Hint: use the resolution

$$0 \longrightarrow A \xrightarrow{\times x} A \longrightarrow A/(x) \longrightarrow 0$$

of $A/(x)$.

24.3.1. *Remark.* As a corollary of Exercise 24.3.A, we see again that flat modules over an integral domain are torsion-free (and more generally, Observation 24.2.2). Also, Exercise 24.3.A gives the reason for the notation Tor — it is short for *torsion*.

24.3.B. EXERCISE. If B is A-flat, use the FHHF theorem (Exercise 1.6.H(c)) to give an isomorphism $B \otimes_A \operatorname{Tor}_i^A(M, N) \cong \operatorname{Tor}_i^B(B \otimes M, B \otimes N)$.

Recall that the Tor functor is symmetric in its entries (there is an isomorphism $\operatorname{Tor}_i^A(M, N) \leftrightarrow \operatorname{Tor}_i^A(N, M)$, Exercise 23.3.A). This gives us a quick but very useful result.

24.3.C. EASY EXERCISE. If $0 \to N' \to N \to N'' \to 0$ is an exact sequence of A-modules, and N'' is flat (e.g. free), show that $0 \to M \otimes_A N' \to M \otimes_A N \to M \otimes_A N'' \to 0$ is exact for *any* A-module M.

We would have cared about this result long before learning about Tor, so it gives some motivation for learning about Tor. (Unimportant side question: Can you prove this without Tor, using a diagram chase?)

24.3.D. EXERCISE (IMPORTANT CONSEQUENCE OF EXERCISE 24.3.C). Suppose $0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0$ is a short exact sequence of quasicoherent sheaves on a scheme Y, and $\mathscr{F}''$ is flat (e.g. locally free). Show that if $\pi : X \to Y$ is any morphism of schemes, the pulled back sequence $0 \to \pi^*\mathscr{F}' \to \pi^*\mathscr{F} \to \pi^*\mathscr{F}'' \to 0$ remains exact.

24.3.E. EXERCISE (CF. EXERCISE 13.5.B FOR THE ANALOGOUS FACTS ABOUT VECTOR BUNDLES). Suppose $0 \to M' \to M \to M'' \to 0$ is an exact sequence of A-modules.
(a) If M and M'' are both flat, show that M' is too. (Hint: Recall the long exact sequence for Tor, Proposition 23.1.2. Also, use that N is flat if and only if $\operatorname{Tor}_i(N, N') = 0$ for all $i > 0$ and all N', Exercise 23.1.D.)
(b) If M' and M'' are both flat, show that M is too. (Same hint.)
(c) If M' and M are both flat, show that M'' need not be flat.

24.3.F. EXERCISE. If $0 \to M_0 \to M_1 \to \cdots \to M_n \to 0$ is an exact sequence of flat A-modules, show that it remains exact upon tensoring with any other A-module. (Hint: as always, break the exact sequence into short exact sequences.)

24.3.G. EASY EXERCISE. If $0 \to M_0 \to M_1 \to \cdots \to M_n \to 0$ is an exact sequence, and M_i is flat for $i > 0$, show that M_0 is flat too. (Hint: as always, break the exact sequence into short exact sequences.)

We will use the Exercises 24.3.F and 24.3.G later in this chapter.

24.3.2. Pulling back quasicoherent ideal sheaves of flat closed subschemes by arbitrary morphisms (promised in §24.2.9). Suppose

$$\begin{tikzcd} W \arrow[r, "\alpha"] \arrow[d] & X \arrow[d] \\ Y \arrow[r, "\beta"] & Z \end{tikzcd}$$

is a fibered product, and $V \hookrightarrow X$ is a closed subscheme. Then $Y \times_Z V$ is a closed subscheme of W (§9.2.1). There are two possible senses in which $\mathscr{I}_{V/X}$ can be "pulled back" to W: as a quasicoherent sheaf $\alpha^*\mathscr{I}_{V/X}$, and as the ideal of the "pulled back" closed subscheme $\mathscr{I}_{Y\times_Z V/W}$. As pointed out in Remark 16.3.9, these are not necessarily the same, but they *are* the same if β is flat. We now give another important case in which they are the same.

24.3.H. EXERCISE. If V is flat over Z (with no hypotheses on β), show that $\alpha^*\mathscr{I}_{V/X} \cong \mathscr{I}_{Y\times_Z V/W}$. Hint: Easy Exercise 24.3.C.

24.4 Ideal-theoretic criteria for flatness

The following theorem will allow us to classify flat modules over a number of rings. It is a refined version of Exercise 23.1.D, that M is a flat A-module if and only if $\operatorname{Tor}_1^A(M,N) = 0$ for all A-modules N.

24.4.1. Theorem (ideal-theoretic criterion for flatness). — *M is flat if and only if $\operatorname{Tor}_1^A(M, A/I) = 0$ for every ideal I.*

24.4.2. *Remarks.* Before getting to the proof, we make some side remarks that may give some insight into how to think about flatness. Theorem 24.4.1 is profitably stated without the theory of Tor. It is equivalent to the statement that M is flat if and only if for all ideals $I \subset A$, $I \otimes_A M \to M$ is an injection, and you can reinterpret the proof in this guise. Perhaps better, M is flat if and only if $I \otimes_A M \to IM$ is an isomorphism for every ideal I.

Flatness is often informally described as "continuously varying fibers", and this can be made more precise as follows. An A-module M is flat if and only if it restricts nicely to closed subschemes of $\operatorname{Spec} A$. More precisely, what we lose in this restriction, the submodule IM of elements which "vanish on Z", is easy to understand: it consists of formal linear combinations of elements $i \otimes m$, with no surprise relations among them — i.e., the tensor product $I \otimes_A M$. This is the content of the following exercise.

24.4.A. $\star$ EXERCISE (THE EQUATIONAL CRITERION FOR FLATNESS). Show that an A-module M is flat if and only if for every relation $\sum a_i m_i = 0$ with $a_i \in A$ and $m_i \in M$, there exist $m'_j \in M$ and $a_{ij} \in A$ such that $\sum_j a_{ij} m'_j = m_i$ for all i and $\sum_i a_i a_{ij} = 0$ in A for all j. (Translation: whenever elements of M satisfy an A-linear relation, this is "because" of linear equations holding in A.)

24.4.3. *Unimportant remark.* In the statement of Theorem 24.4.1, it suffices to check only finitely generated ideals. This is essentially the content of the following statement, which you can prove if you wish: Show that an A-module M is flat if and only if for all *finitely generated ideals* I, the natural map $I \otimes_A M \to M$ is an injection. Hint: if there is a counterexample for an ideal J that is not finitely generated, use it to find another counterexample for an ideal I that *is* finitely generated.

24.4.4. *Proof of the ideal-theoretic criterion for flatness, Theorem 24.4.1.* By Exercise 23.1.D, we need only show that $\operatorname{Tor}_1^A(M, A/I) = 0$ for all I implies $\operatorname{Tor}_1^A(M,N) = 0$ for all A-modules N, and hence that M is flat.

We first prove that $\operatorname{Tor}_1^A(M,N) = 0$ for all *finitely generated* modules N, by induction on the number n of generators $a_1, \dots, a_n$ of N. The base case (if $n = 1$, so $N \cong A/\operatorname{Ann}(a_1)$) is our assumption. If $n > 1$, then $Aa_n \cong A/\operatorname{Ann}(a_n)$ is a submodule of N, and the quotient Q is generated by the images of $a_1, \dots, a_{n-1}$, so the result follows by considering the Tor_1 portion of the Tor long exact sequence for

$$0 \to A/\operatorname{Ann}(a_n) \to N \to Q \to 0.$$

We deal with the case of general N by abstract nonsense. Notice that N is the union of its finitely generated submodules $\{N_\alpha\}$. In fancy language, this union is a filtered colimit — any two finitely generated submodules are contained in a finitely generated submodule (specifically, the submodule they generate). Filtered colimits of modules commute with cohomology (Exercise 1.6.L), so $\operatorname{Tor}_1(M, N)$ is the colimit over α of $\operatorname{Tor}_1(M, N_\alpha) = 0$, and is thus 0. □

We now use Theorem 24.4.1 to get explicit characterizations of flat modules over three (types of) rings: principal ideal domains, dual numbers, and Noetherian local rings.

Recall Observation 24.2.2, that flatness implies torsion-free. The converse is true for principal ideal domains:

24.4.B. EXERCISE (FLAT = TORSION-FREE FOR A PID). Show that a module over a principal ideal domain is flat if and only if it is torsion-free.

24.4.C. EXERCISE (FLATNESS OVER A DVR). Suppose M is a module over a discrete valuation ring A with uniformizer t. Show that M is flat if and only if t is not a zerodivisor on M, i.e., (using the notation defined in Exercise 24.3.A) $(M : t) = 0$. (See Exercise 24.2.B for the case of finitely generated modules.) This yields a simple geometric interpretation of flatness over a regular curve, which we discuss in §24.4.8.

24.4.D. EXERCISE (FLATNESS OVER THE DUAL NUMBERS). Show that M is flat over $k[t]/(t^2)$ if and only if the "multiplication by t" map $M/tM \to tM$ is an isomorphism. (This fact is important in deformation theory and elsewhere.) Hint: $k[t]/(t^2)$ has only three ideals.

24.4.5. Important Theorem (flat = free = projective for finitely presented modules over local rings). — *Suppose $(A, \mathfrak{m})$ is a local ring (not necessarily Noetherian), and M is a finitely presented A-module. Then M is flat if and only if it is free if and only if it is projective.*

24.4.6. *Remarks.* Warning: modules over local rings can be flat without being free: $\mathbb{Q}$ is a flat $(\mathbb{Z})_{(p)}$-algebra ($(\mathbb{Z})_{(p)}$ is the localization of $\mathbb{Z}$ at p, not the p-adics), as all localizations are flat §24.2.1, but it is not free (do you see why?).

Also, non-Noetherian people may be pleased to know that with a little work, "finitely presented" can be weakened to "finitely generated": use [**Mat2**, Thm. 7.10] in the proof below, where finite presentation comes up.

Proof. For any ring, free modules are projective (§23.2.1), and projective modules are flat (Exercise 23.2.B), so we need only show that flat modules are free for a local ring.

(At this point, you should see Nakayama coming from a mile away.) Now $M/\mathfrak{m}M$ is a finite-dimensional vector space over the field $A/\mathfrak{m}$. Choose a basis of

$M/\mathfrak{m}M$, and lift it to elements $m_1, \dots, m_n \in M$. Consider $A^{\oplus n} \to M$ given by $e_i \mapsto m_i$. We will show this is an isomorphism. It is surjective by Nakayama's Lemma (see Exercise 7.2.H): the image is all of M modulo the maximal ideal, hence is everything. As M is finitely presented, by Exercise 13.6.A ("finitely presented implies always finitely presented"), the kernel K is finitely generated. Tensor $0 \to K \to A^{\oplus n} \to M \to 0$ with $A/\mathfrak{m}$. As M is flat, the result is still exact (Exercise 24.3.C):

$$0 \to K/\mathfrak{m}K \to (A/\mathfrak{m})^{\oplus n} \to M/\mathfrak{m}M \to 0.$$

But $(A/\mathfrak{m})^{\oplus n} \to M/\mathfrak{m}M$ is an isomorphism by construction, so $K/\mathfrak{m}K = 0$. As K is finitely generated, $K = 0$ by Nakayama's Lemma 7.2.9. □

Here is an immediate and useful corollary — really just a geometric interpretation.

24.4.7. Corollary (flat = locally free for coherent sheaves). — *A coherent sheaf $\mathscr{F}$ on X is flat (over X) if and only if it is locally free.*

Proof. Local freeness of a coherent sheaf can be checked at the stalks, Exercise 13.7.F. □

24.4.E. EXERCISE. Suppose $\pi : X \to Y$ is a finite flat morphism of locally Noetherian schemes, and $\mathscr{F}$ is a finite rank locally free sheaf on X. Show that $\pi_* \mathscr{F}$ is a finite rank locally free sheaf on Y. If Y is irreducible with generic point η, the degree of π above η is n, and $\mathscr{F}$ is locally free of rank r, show that $\pi_*\mathscr{F}$ is locally free of rank nr. (Hint: transitivity of flatness, Exercise 24.2.H.)

24.4.F. ⋆ EXERCISE (INTERESTING VARIANT OF THEOREM 24.4.5, BUT UNIMPORTANT FOR US). Suppose A is a ring (not necessarily local), and M is a finitely presented A module. Show that M is flat if and only if it is projective. Hint: show that M is projective if and only if $M_\mathfrak{m}$ is free for every maximal ideal $\mathfrak{m}$. The harder direction of this implication uses the fact that $\operatorname{Hom}_{A_\mathfrak{m}}(M_\mathfrak{m}, N_\mathfrak{m}) = \operatorname{Hom}_A(M, N)_\mathfrak{m}$, which follows from Exercise 1.6.G. (Remark: There exist finitely *generated* flat modules that are not projective. They are necessarily not finitely presented. Example without proof: let $A = \prod_{i=1}^\infty \mathbb{F}_2$, interpreted as functions $\mathbb{Z}^{\geq 0} \to \mathbb{Z}/2$, and let M be the module of functions modulo those of proper support, i.e., those vanishing at almost all points of $\mathbb{Z}^{\geq 0}$.)

24.4.G. EXERCISE. Make precise and prove the following statement: "finite flat morphisms have locally constant degree". (You may want to glance at §17.4.4 — and in particular, Exercise 17.4.D(a) — to make this precise. We will revisit the example of §17.4.4 in §24.4.11.)

24.4.H. EXERCISE. Prove the following useful criterion for flatness: Suppose $\pi : X \to Y$ is a finite morphism, and Y is reduced and locally Noetherian. Then π is flat if and only if $\pi_* \mathscr{O}_X$ is locally free, if and only if the rank of $\pi_*\mathscr{O}_X$ is locally constant ($\dim_{\kappa(q)} (\pi_* \mathscr{O}_X)_q \otimes \kappa(q)$ is a locally constant function of $q \in Y$). Partial hint: Exercise 13.7.K.

24.4.I. EXERCISE. Show that the normalization of the node (see Figure 7.4) is not flat. (Hint: use Exercise 24.4.H.)

This exercise can be strengthened to show that nontrivial normalizations are *never* flat. The following exercise shows that our old friend, two planes attached at a point, yield an interesting example of this fact.

24.4.J. EXERCISE. In $\mathbb{A}^4_k = \operatorname{Spec} k[w,x,y,z]$, let X be the union of the wx-plane with the yz-plane:

$$X = \operatorname{Spec} k[w,x,y,z]/(wy, wz, xy, xz). \tag{24.4.7.1}$$

The projection $\mathbb{A}^4_k \to \mathbb{A}^2_k$ given by $k[a,b] \to k[w,x,y,z]$ with $a \mapsto w-y$, $b \mapsto x-z$ restricts to a morphism $X \to \mathbb{A}^2_k$. Show that this morphism is not flat.

24.4.8. Flat families over regular curves. Exercise 24.4.C gives an elegant geometric criterion for when morphisms to regular curves are flat.

24.4.K. EXERCISE (CRITERION FOR FLATNESS OVER A REGULAR CURVE). Suppose $\pi : X \to Y$ is a morphism from a locally Noetherian scheme to a regular (locally Noetherian) curve. (The local Noetherian hypothesis on X is so we can discuss its associated points.) Show that π is flat if and only if all associated points of X map to a generic point of Y. (This is a partial converse to Exercise 24.2.J, that flat maps always send associated points to associated points. As with Exercise 24.2.J, feel free to immediately generalize your argument to a coherent sheaf on $\mathscr{F}$ on X.)

24.4.9. For example, a nonconstant map from an integral (locally Noetherian) scheme to a regular curve must be flat. (As another example, the deformation to the normal cone, discussed in the double-starred section §22.4.17, is flat.) Exercise 24.4.I (and the comment after it) shows that the regular condition is necessary. The example of two planes meeting at a point in Exercise 24.4.J shows that the dimension 1 condition is necessary.

24.4.10. $\star$ *Remark: A valuative criterion for flatness.* Exercise 24.4.K shows that flatness over a regular curves is geometrically intuitive (and is "visualizable"). It gives a criterion for flatness in general: suppose $\pi : X \to Y$ is finitely presented morphism. If π is flat, then for every morphism $Y' \to Y$ where Y' is the Spec of a discrete valuation ring, $\pi' : X \times_Y Y' \to Y'$ is flat, so no associated points of $X \times_Y Y'$ map to the closed point of Y'. If Y is reduced and locally Noetherian, then this is a sufficient condition; this can reasonably be called a *valuative criterion for flatness*. (Reducedness is necessary: consider Exercise 24.2.G(d).) This gives an excellent way to visualize flatness, which you should try to put into words (perhaps after learning about flat limits below). See [**Gr-EGA**, IV_3.3.11.8] for a proof (and an extension without Noetherian hypothesis).

24.4.11. *Revisiting the degree of a projective morphism from a curve to a regular curve.* As hinted after the statement of Proposition 17.4.5, we can now better understand why nonconstant projective morphisms from a curve to a regular curve have a well-defined degree, which can be determined by taking the preimage of any point (§17.4.4). (Example 9.3.3 was particularly enlightening.) This is because such maps are flat by Exercise 24.4.K, and then the degree is constant by Exercise 24.4.G (see also Exercise 24.4.H). Also, Exercise 24.4.H yields a new proof of Proposition 17.4.5.

24.4.12. Flat limits. Here is an important consequence of Exercise 24.4.K, which we can informally state as: we can take flat limits over one-parameter families. More precisely: suppose A is a discrete valuation ring, and let 0 be the closed point of $\operatorname{Spec} A$ and η the generic point. Suppose X is a locally Noetherian scheme over A, and Y is a closed subscheme of $X|_\eta$. Let Y' be the scheme-theoretic closure of Y in X. Then Y' is flat over A. Similarly, suppose Z is a one-dimensional Noetherian scheme, 0 is a regular point of Z, and $\pi : X \to Z$ is a morphism from a locally Noetherian scheme to Z. If Y is a closed subscheme of $\pi^{-1}(Z - \{0\})$, and Y' is the scheme-theoretic closure of Y in X, then Y' is flat over Z. In both cases, the closure $Y'|_0$ is often called the **flat limit** of Y. (Feel free to weaken the Noetherian hypotheses on X.)

24.4.L. EXERCISE. Suppose (with the language of the previous paragraph) that A is a discrete valuation ring, X is a locally Noetherian A-scheme, and Y is a closed subscheme of the generic fiber $X|_\eta$. Show that there is only one closed subscheme Y' of X such that $Y'|_\eta = Y$, and Y' is flat over A.

24.4.M. HARDER EXERCISE (AN EXPLICIT FLAT LIMIT). Let $X = \mathbb{A}^3 \times \mathbb{A}^1 \to Y = \mathbb{A}^1$ over a field k, where the coordinates on $\mathbb{A}^3$ are x, y, and z, and the coordinates on $\mathbb{A}^1$ are t. Define X away from $t = 0$ as the union of the two lines $y = z = 0$ (the x-axis) and $x = z - t = 0$ (the y-axis translated by t). Find the flat limit at $t = 0$. (Hints: (i) it is *not* the union of the two axes, although it includes this union. The flat limit is nonreduced at the node, and the "fuzz" points out of the plane they are contained in. (ii) $(y, z)(x, z) \neq (xy, z)$. (iii) Once you have a candidate flat limit, be sure to check that it *is* the flat limit. (iv) If you get stuck, read Example 24.4.13 below.)

Consider a projective version of the previous example, where two lines in $\mathbb{P}^3$ degenerate to meet. The limit consists of two lines meeting at a node, with some nonreduced structure at the node. Before the two lines come together, their space of global sections is two-dimensional. When they come together, it is not immediately obvious that their flat limit also has two-dimensional space of global sections as well. The reduced version (the union of the two lines meeting at a point) has a one-dimensional space of global sections, but the effect of the nonreduced structure on the space of global sections may not be immediately clear. However, we will see that "cohomology groups can only jump up in flat limits", as a consequence (indeed the main moral) of the Semicontinuity Theorem 28.1.1.

24.4.13. ⋆⋆ **Example of variation of cohomology groups in flat families.** We can use a variant of Exercise 24.4.M to see an example of a cohomology group actually jumping. We work over an algebraically closed field to avoid distractions. Before we get down to explicit algebra, here is the general idea. Consider a twisted cubic C in $\mathbb{P}^3$. A projection pr_p from a random point $p \in \mathbb{P}^3$ will take C to a nodal plane cubic. Picture this projection "dynamically", by choosing coordinates so p is at $[1, 0, 0, 0]$, and considering the map $\phi_t : [w, x, y, z] \mapsto [w, tx, ty, tz]$; ϕ_1 is the identity on $\mathbb{P}^3$, ϕ_t is an automorphism of $\mathbb{P}^3$ for $t \neq 0$, and ϕ_0 is the projection. The limit of $\phi_t(C)$ as $t \to 0$ will be a nodal cubic, with nonreduced structure at the node "analytically the same" as what we saw when two lines came together (Exercise 24.4.M). (The phrase "analytically the same" can be made precise once we define completions in §29.)

Let's now see this in practice. Rather than working directly with the twisted cubic, we use another example where we saw a similar picture. Consider the nodal (affine) plane cubic $y^2 = x^3 + x^2$. Its normalization (see Figure 7.4, Example (3) of §7.3.8, Exercise 9.7.E, ...) was obtained by adding an extra variable m corresponding to y/x (which can be interpreted as blowing up the origin, see §22.4.3). We use the variable m rather than t (used in §7.3.8) in order to reserve t for the parameter for the flat family.

We picture the nodal cubic C as lying in the xy-plane in 3-space $\mathbb{A}^3 = \operatorname{Spec} k[x, y, m]$, and the normalization $\widetilde{C}$ projecting to it, with $m = y/x$. What are the equations for $\widetilde{C}$? Clearly, they include the equations $y^2 = x^3 + x^2$ and $y = mx$, but these are not enough — the m-axis (i.e., $x = y = 0$) is also in $V(y^2 - x^3 - x^2, y - mx)$. A little thought (and the algebra we have seen earlier in this example) will make clear that we have a third equation $m^2 = (x+1)$, which along with $y = mx$ implies $y^2 = x^2+x^3$. *Now* we have enough equations: $k[x, y, m]/(m^2-(x+1), y-mx)$ is an integral domain, as it is clearly isomorphic to $k[m]$. Indeed, you should recognize this as the algebra appearing in Exercise 9.7.E.

Next, we want to formalize our intuition of the dynamic projection to the xy-plane of $\widetilde{C} \subset \mathbb{A}^3$. We picture it as follows. Given a point (x, y, m) at time 1, at time t we want it to be at (x, y, mt). At time $t = 1$, we "start with" $\widetilde{C}$, and at time $t = 0$ we have (set-theoretically) C. Thus *at time* $t \neq 0$, the curve $\widetilde{C}$ is sent to the curve cut out by equations

$$k[x, y, m]/(m^2 - t(x + 1), ty - mx).$$

The family over $\operatorname{Spec} k[t, t^{-1}]$ is thus

$$k[x, y, m, t, t^{-1}]/(m^2 - t(x + 1), ty - mx).$$

Notice that we have inverted t because we are so far dealing only with nonzero t. For $t \neq 0$, this is certainly a "nice" family, and so surely flat. Let's make sure this is true.

24.4.N. EXERCISE. Check this, as painlessly as possible! Hint: by a clever change of coordinates, show that the family is constant "over $\operatorname{Spec} k[t, t^{-1}]$", and hence pulled back (in some way you must figure out) via $k[t, t^{-1}] \to k$ from

$$\operatorname{Spec} k[X, Y, M]/(M^2 - (X + 1), Y - MX) \to \operatorname{Spec} k,$$

which is flat by Trick Question 24.2.A(a).

We now figure out the flat limit of this family over $t = 0$, in $\operatorname{Spec} k[x, y, m, t] \to \mathbb{A}^1 = \operatorname{Spec} k[t]$. We first hope that our flat family is given by the equations we have already written down:

$$\operatorname{Spec} k[x, y, m, t]/(m^2 - t(x + 1), ty - mx).$$

But this is *not* flat over $\mathbb{A}^1 = \operatorname{Spec} k[t]$, as the fiber dimension jumps (§24.5.4): substituting $t = 0$ into the equations (obtaining the fiber over $0 \in \mathbb{A}^1$), we find $\operatorname{Spec} k[x, y, m]/(m^2, mx)$. This is set-theoretically the xy-plane ($m = 0$), which of course has dimension 2. Notice for later reference that this "false limit" is scheme-theoretically the xy-plane, *with some nonreduced structure along the* y*-axis*. (This may remind you of Figure 4.4.)

So we are missing at least one equation. One clue as to what equation is missing: the equation $y^2 = x^3 + x^2$ clearly holds for $t \neq 0$, and does *not* hold for our naive attempt at a limit scheme $m^2 = mx = 0$. So we put this equation back in, and have a second hope for describing the flat family over $\mathbb{A}^1$:

$$\operatorname{Spec} k[x, y, m, t]/(m^2 - t(x+1), ty - mx, y^2 - x^3 - x^2) \to \operatorname{Spec} k[t].$$

Let $A = k[x, y, m, t]/(m^2 - t(x+1), ty - mx, y^2 - x^3 - x^2)$ for convenience. The morphism $\operatorname{Spec} A \to \mathbb{A}^1$ *is* flat at $t = 0$. How can we show it? We could hope to show that A is an integral domain, and thus invoke Exercise 24.4.K. Instead we use Exercise 24.4.B, and show that t is not a zerodivisor on A. We do this by giving a "normal form" for elements of A.

24.4.O. EXERCISE. Show that each element of A can be written uniquely as a polynomial in x, y, m, and t such that no monomial in it is divisible by m^2, mx, or y^2. Then show that t is not a zerodivisor on A, and conclude that $\operatorname{Spec} A \to \mathbb{A}^1$ is indeed flat.

Thus the flat limit when $t = 0$ is given by

$$\operatorname{Spec} k[x, y, m]/(m^2, mx, y^2 - x^2 - x^3).$$

24.4.P. EXERCISE. Show that the flat limit is nonreduced, and the "nonreducedness has length 1 and supported at the origin". More precisely, if $X = \operatorname{Spec} A/(t)$, show that $\mathscr{I}_{X^{\mathrm{red}}}$ is a skyscraper sheaf, with value k, supported at the origin. Sketch this flat limit X.

24.4.14. Note that we have a nonzero global function on X, given by m, which is supported at the origin (i.e., 0 away from the origin).

We now use this example to get a projective example with interesting behavior. We take the projective completion of this example, to get a family of cubic curves in $\mathbb{P}^3$ degenerating to a nodal cubic C with a nonreduced point.

24.4.Q. EXERCISE. Do this: describe this family (in $\mathbb{P}^3 \times \mathbb{A}^1$) precisely.

Take the long exact sequence corresponding to

$$0 \longrightarrow \mathscr{I}_{C^{\mathrm{red}}} \longrightarrow \mathscr{O}_C \longrightarrow \mathscr{O}_{C^{\mathrm{red}}} \longrightarrow 0,$$

to get

$$\begin{tikzcd} H^1(C, \mathscr{I}_{C^{\mathrm{red}}}) \arrow[r] & H^1(C, \mathscr{O}_C) \arrow[r] & H^1(C, \mathscr{O}_{C^{\mathrm{red}}}) \arrow[r] & {} \\ H^0(C, \mathscr{I}_{C^{\mathrm{red}}}) \arrow[r, "\alpha"] & H^0(C, \mathscr{O}_C) \arrow[r] & H^0(C, \mathscr{O}_{C^{\mathrm{red}}}) \arrow[r] & 0 \end{tikzcd}$$

We have $H^1(C, \mathscr{I}_{C^{\mathrm{red}}}) = 0$ as $\mathscr{I}_{C^{\mathrm{red}}}$ is supported in dimension 0 (by dimensional vanishing, Theorem 18.2.6). Also, $H^i(C^{\mathrm{red}}, \mathscr{O}_{C^{\mathrm{red}}}) = H^i(C, \mathscr{O}_{C^{\mathrm{red}}})$ (property **(v)** of cohomology, see §18.1). The (reduced) nodal cubic C^{red} has $h^0(\mathscr{O}) = 1$ (§10.3.7) and $h^1(\mathscr{O}) = 1$ (cubic plane curves have genus 1, (18.6.6.1)). Also, $h^0(C, \mathscr{I}_{C^{\mathrm{red}}}) = 1$ as observed above. Finally, α is not 0, as there exists a nonzero function on C vanishing on C^{red} (§24.4.14 — convince yourself that this function extends from the affine patch $\operatorname{Spec} A$ to the projective completion).

Using the long exact sequence, we conclude $h^0(C, \mathscr{O}_C) = 2$ and $h^1(C, \mathscr{O}_C) = 1$. Thus in this example we see that $(h^0(\mathscr{O}), h^1(\mathscr{O})) = (1, 0)$ for the general member of the family (twisted cubics are isomorphic to $\mathbb{P}^1$), and the special member (the flat limit) has $(h^0(\mathscr{O}), h^1(\mathscr{O})) = (2, 1)$. Notice that both cohomology groups have jumped, yet the Euler characteristic has remained the same. The first behavior, as stated after Exercise 24.4.M, is an example of the Semicontinuity Theorem 28.1.1. The second, constancy of Euler characteristics in flat families, is what we turn to next. (It is no coincidence that the example had a singular limit, see §28.1.3.)

24.5 Topological aspects of flatness

We now discuss some topological aspects and consequences of flatness, that boil down to the Going-Down theorem for flat morphisms (§24.5.2), which in turn comes from faithful flatness. Because dimension in algebraic geometry is a topological notion, we will show that dimensions of fibers behave well in flat families (§24.5.4).

24.5.1. Faithful flatness. The notion of faithful flatness is handy for many reasons, and we describe only a few. A B-module M is **faithfully flat** if for all complexes of B-modules

(24.5.1.1) $$N' \to N \to N'',$$

(24.5.1.1) is exact if and only if (24.5.1.1)$\otimes_B M$ is exact. A B**-algebra** A **is faithfully flat** if it is faithfully flat as a B-module.

24.5.A. Exercise. Show that a flat B-module M is faithfully flat if and only if for all B-modules N, $M \otimes_B N = 0$ implies that $N = 0$.

24.5.B. Exercise. Suppose M is a flat B-module. Show that the following are equivalent.

(a) M is faithfully flat;
(b) for all prime ideals $\mathfrak{p} \subset B$, $M \otimes_B \kappa(\mathfrak{p})$ is nonzero (i.e., $\operatorname{Supp} M = \operatorname{Spec} B$);
(c) for all maximal ideals $\mathfrak{m} \subset B$, $M \otimes_B \kappa(\mathfrak{m}) = M/\mathfrak{m}M$ is nonzero.

Suppose $\pi : X \to Y$ is a morphism of schemes. We say that π is **faithfully flat** if it is flat and surjective. (Unlike flatness, faithful flatness is not that useful a notion for quasicoherent sheaves, so we do not define faithfully flat quasicoherent sheaves over a base.)

24.5.C. Exercise (cf. 24.5.B). Suppose $B \to A$ is a ring morphism. Show that A is faithfully flat over B if and only if $\operatorname{Spec} A \to \operatorname{Spec} B$ is faithfully flat.

Faithful flatness is preserved by base change, as both surjectivity and flatness are (Exercises 9.4.D and 24.2.I respectively).

24.5.D. Exercise. Suppose $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$ is flat.
(a) Show that π is faithfully flat if and only if every *closed* point $x \in \operatorname{Spec} B$ is in the image of π. (Hint: Exercise 24.5.B(c).)
(b) Hence show that *every* flat (local) morphism of local rings (Definition 6.3.1) is

faithfully flat. (Morphisms of local rings are assumed to be local, i.e., the maximal ideal pulls back to the maximal ideal.)

24.5.2. Going-Down for flat morphisms. A consequence of Exercise 24.5.D is the following useful result, whose statement makes no mention of faithful flatness. (The statement is not coincidentally reminiscent of the Going-Down Theorem for finite extensions of integrally closed domains, Theorem 11.2.12.)

24.5.E. EXERCISE (GOING-DOWN THEOREM FOR FLAT MORPHISMS).
(a) Suppose that $B \to A$ is a flat morphism of rings, corresponding to a map $\pi : \operatorname{Spec} A \to \operatorname{Spec} B$. Suppose $\mathfrak{q} \subset \mathfrak{q}'$ are prime ideals of B, and $\mathfrak{p}'$ is a prime ideal of A with $\pi([\mathfrak{p}']) = [\mathfrak{q}']$. Show that there exists a prime $\mathfrak{p} \subset \mathfrak{p}'$ of A with $\pi([\mathfrak{p}]) = [\mathfrak{q}]$. Hint: show that $B_{\mathfrak{q}'} \to A_{\mathfrak{p}'}$ is a flat local ring homomorphism, and hence faithfully flat by the Exercise 24.5.D(b).
(b) Part (a) gives a geometric consequence of flatness. Draw a picture illustrating this.
(c) Recall the Going-Up Theorem, described in §7.2.4. State the Going-Down Theorem for flat morphisms in a way parallel to Exercise 7.2.F, and prove it.

24.5.F. EXERCISE. Suppose $\pi : X \to Y$ is an integral (e.g. finite) flat morphism, and Y has pure dimension n. Show that X has pure dimension n. (This generalizes Exercise 11.1.G(a).) Hint: π satisfies both Going-Up (see Exercise 7.2.F) and Going-Down.

24.5.G. IMPORTANT EXERCISE: FLAT MORPHISMS ARE OPEN (IN REASONABLE SITUATIONS). Suppose $\pi : X \to Y$ is locally of finite type and flat, and Y (and hence X) is locally Noetherian. Show that π is an open map (i.e., sends open sets to open sets). Hint: reduce to showing that $\pi(X)$ is open for all such π. Reduce to the case where X is affine. Use Chevalley's Theorem 7.4.2 to show that $\pi(X)$ is constructible. Use the Going-Down Theorem for flat morphisms, Exercise 24.5.E, to show that $\pi(X)$ is closed under generization. Conclude using Exercise 7.4.C.

24.5.H. EASY EXERCISE. Suppose A and B are finite type k-algebras. Show that $\operatorname{Spec} A \times_{\operatorname{Spec} k} \operatorname{Spec} B \to \operatorname{Spec} B$ is an open map. (This is the long-promised proof to Proposition 9.5.4.)

24.5.3. *Follow-ups to Exercise 24.5.G.*
(i) Of course, not all open morphisms are flat: witness $\operatorname{Spec} k[t]/(t) \to \operatorname{Spec} k[t]/(t^2)$.
(ii) Also, in quite reasonable circumstances, flat morphisms are *not* open: witness $\operatorname{Spec} k(t) \to \operatorname{Spec} k[t]$ (flat by Example 24.2.1(b)).
(iii) On the other hand, you can weaken the hypotheses of "locally of finite type" and "locally Noetherian" to just "locally finitely presented" [**Gr-EGA**, IV$_2$.2.4.6] — as with the similar generalization in Exercise 9.3.I of Chevalley's Theorem 7.4.2, use the fact that any such morphisms is "locally" pulled back from a Noetherian situation. We won't use this, and hence omit the details.

24.5.4. Dimensions of fibers are well-behaved for flat morphisms.

24.5.5. Proposition. — *Suppose $\pi : X \to Y$ is a flat morphism of locally Noetherian schemes, with $p \in X$ and $q \in Y$ such that $\pi(p) = q$. Then*

$$\operatorname{codim}_X p = \operatorname{codim}_Y q + \operatorname{codim}_{\pi^{-1}(q)} p$$

(see Figure 11.3).

Informal translation: the dimension of the fibers is the difference of the dimensions of X and Y (at least locally). Compare this to Exercise 11.4.A, which stated that without the flatness hypothesis, we would only have inequality ($\leq$).

24.5.I. EXERCISE. Prove Proposition 24.5.5 as follows. As just mentioned, Exercise 11.4.A gives one inequality, so show the other. Given a chain of irreducible closed subsets in Y containing $\overline{q}$, and a chain of irreducible closed subsets in $\pi^{-1}(q) \subset X$ containing $\overline{p}$, construct a chain of irreducible closed subsets in X containing $\overline{p}$, using the Going-Down Theorem for flat morphisms (Exercise 24.5.E).

As a consequence of Proposition 24.5.5, if $\pi : X \to Y$ is a flat map of irreducible varieties, then the fibers of π all have pure dimension $\dim X - \dim Y$. (Warning: $\operatorname{Spec} k[t]/(t) \to \operatorname{Spec} k[t]/(t^2)$ does not exhibit dimensional jumping of fibers, is open, and sends associated points to associated points, cf. Exercise 24.2.J, but is not flat. If you prefer a reduced example, the normalization $\operatorname{Spec} k[t] \to \operatorname{Spec} k[x,y]/(y^2 - x^3)$, shown in Figure 9.4, also has these properties.) This leads us to the following useful definition.

24.5.6. *Definition.* If a morphism $\pi : X \to Y$ is flat morphism that is locally of finite type, and all fibers of π have pure dimension n, we say that π is **flat of relative dimension** n. (In particular: when one says a morphism is flat of relative dimension n, the locally finite type hypotheses are implied. Remark 24.5.7 motivates this hypothesis.)

24.5.J. EXERCISE. Suppose $\pi : X \to Y$ is a flat morphism of finite type k-schemes, and Y is pure dimensional (so "codimension is the difference of dimensions", cf. Theorem 11.2.9). Show that the following are equivalent.

(i) The scheme X has pure dimension $\dim Y + n$.
(ii) The morphism π is flat of relative dimension n.

24.5.K. EXERCISE. Suppose $\pi : X \to Y$ and $\rho : Y \to Z$ are morphisms of locally Noetherian schemes, flat of relative dimension m and n respectively (hence locally of finite type). Show that $\rho \circ \pi$ is flat of relative dimension $m + n$. Hint: Exercise 24.5.J.

24.5.L. EXERCISE. Show that the notion of a morphism being "flat of relative dimension n" is preserved by arbitrary base change. Hint to show that fiber dimension is preserved: Exercise 11.2.I.

24.5.M. EXERCISE (DIMENSION IS ADDITIVE FOR PRODUCTS OF VARIETIES). If X and Y are k-varieties of pure dimension m and n respectively, show that $X \times_k Y$ has pure dimension $m + n$. (We have waited egregiously long to prove this basic fact!)

24.5.7. *Remark.* The reason for the "locally finite type" assumption in the definition of "flat of relative dimension n" is that we want any class of morphism to be behave "reasonably" (in the sense of §7.1.1). In particular, we want our notion of "flatness of relative dimension n" to be preserved by base change. Consider the fibered diagram

$$\begin{tikzcd} \operatorname{Spec} k(x) \otimes_k k(y) \arrow[r] \arrow[d, "\pi'"'] & \operatorname{Spec} k(y) \arrow[d, "\pi"] \\ \operatorname{Spec} k(x) \arrow[r] & \operatorname{Spec} k. \end{tikzcd}$$

Both π and π' are trivially flat, because they are morphisms to Spec's of fields (Exercise 24.2.A(a)). But the dimension of the fiber of π is 0, while (as described in Remark 11.2.16) the dimension of the fiber of π' is 1.

24.5.8. Generic flatness. (This would be better called "general flatness".)

24.5.N. EASY EXERCISE (GENERIC FLATNESS). Suppose $\pi : X \to Y$ is a finite type morphism to a Noetherian integral scheme, and $\mathscr{F}$ is a coherent sheaf on X. Show that there is a dense open subset $U \subset Y$ over which $\mathscr{F}$ is flat. (An important special case is if $\mathscr{F} = \mathscr{O}_X$, in which case this shows there is a dense open subset U over which π is flat.) Hint: Grothendieck's Generic Freeness Lemma 7.4.4.

24.5.9. *Interpretation of the degree of a generically finite morphism.* We use this to interpret the degree of a rational map of varieties in terms of counting preimages, fulfilling a promise made in §11.2.2. Suppose $\pi : X \dashrightarrow Y$ is a rational map of k-varieties of degree d. For simplicity we assume that X and Y are irreducible. (But feel free to relax this.) Replace X by an open subset on which π is a morphism. By Proposition 11.4.3 ("generically finite implies generally finite"), there is a dense open subset V of Y over which π is finite. By "generic flatness" (Exercise 24.5.N, there is a dense open subset V' of V over which π is *finite and flat*. Then by Exercise 24.4.G ("finite flat morphisms have locally constant degree"), over this open subset V', π has "locally constant degree".

24.5.10. Exercise 24.5.N can be improved:

24.5.11. Theorem (generic flatness, improved version, [Stacks, tag 052B]). — *If $\pi : X \to Y$ is a morphism of schemes, and $\mathscr{F}$ is finite type quasicoherent on X, Y is reduced, π is finite type, then there is an open dense subset $U \subset Y$ over which π is flat and finite presentation, and such that $\mathscr{F}$ is flat and of finite presentation over Y.*

We won't use this result, so we omit the proof.

24.5.12. *Flatness is an open condition.* Generic flatness can be used to show that in reasonable circumstances, the locus where a quasicoherent sheaf is flat over a base is an open subset. More precisely:

24.5.13. Theorem (flatness is an open condition). — *Suppose $\pi : X \to Y$ is a locally finite type morphism of locally Noetherian schemes, and $\mathscr{F}$ is a finite type quasicoherent sheaf on X.*

(a) *The locus of points of X at which $\mathscr{F}$ is Y-flat is an open subset of X.*

(b) *If π is closed (e.g. proper), then the locus of points of Y over which $\mathscr{F}$ is flat is an open subset of Y.*

Part (b) follows immediately from part (a). Part (a) reduces to a nontrivial statement in commutative algebra, see for example [**Mat2**, Thm. 24.3] or [**Gr-EGA**, IV_3.11.1.1]. As is often the case, Noetherian hypotheses can be dropped in exchange for local finite presentation hypotheses on the morphism π, see [**Gr-EGA**, IV_3.11.3.1] or [**Stacks**, tag 00RC].

24.6 Local criteria for flatness

(This is the hardest section on ideal-theoretic criteria for flatness, and could profitably be postponed to a second reading.)

In the case of a Noetherian local ring, there is a greatly improved version of the ideal-theoretic criterion of Theorem 24.4.1: we need check only *one* ideal — the maximal ideal. The price we pay for the simplicity of this "local criterion for flatness" is that it is harder to prove.

24.6.1. Theorem (local criterion for flatness). — *Suppose $(A, \mathfrak{m})$ is a Noetherian local ring, and M is a finitely generated A-module. Then M is flat if and only if $\operatorname{Tor}_1^A(M, A/\mathfrak{m}) = 0$.*

This is a miracle: flatness over all of $\operatorname{Spec} A$ is determined by what happens over the closed point. (Caution: the finite generation is necessary. Let $A = k[x,y]_{(x,y)}$ and $M = k(x)$, with y acting as 0. Then M is not flat by Observation 24.2.2, but it turns out that it satisfies the local criterion otherwise.)

Theorem 24.6.1 is an immediate consequence of the following more general statement.

24.6.2. Theorem (local criterion for flatness, more general version). — *Suppose $(B, \mathfrak{n}) \to (A, \mathfrak{m})$ is a local morphism of Noetherian local rings (§6.3.1), and that M is a finitely generated A-module. Then M is B-flat if and only if $\operatorname{Tor}_1^B(M, B/\mathfrak{n}) = 0$.*

24.6.3. ⋆ *Proof of Theorem 24.6.2.* A sign of the difficulty of this result is that the Artin-Rees Lemma 12.9.3 (or a consequence thereof) is used twice — once for the local ring $(A, \mathfrak{m})$ (in the guise of the Krull Intersection Theorem), and once for the local ring $(B, \mathfrak{n})$.

Recall from Exercise 18.4.R that a B-module N has finite length if there is a finite sequence $0 = N_0 \subset N_1 \subset \cdots \subset N_n = N$ with $N_m/N_{m-1} \cong B/\mathfrak{n}$ for $1 \leq m \leq n$.

24.6.A. Easy Preliminary Exercise. With the same hypotheses as Theorem 24.6.2, suppose that $\operatorname{Tor}_1^B(M, B/\mathfrak{n}) = 0$. Show that $\operatorname{Tor}_1^B(M, N) = 0$ for all B-modules N of finite length, by induction on the length of N.

By Exercise 23.1.D, if M is B-flat, then $\operatorname{Tor}_1^B(M, B/\mathfrak{n}) = 0$, so it remains to assume that $\operatorname{Tor}_1^B(M, B/\mathfrak{n}) = 0$ and show that M is B-flat.

By the ideal theoretic criterion for flatness (Theorem 24.4.1, see §24.4.2), we wish to show that $\phi : I \otimes_B M \to M$ is an injection for all ideals I of B, i.e., that $\ker \phi = 0$.

Note that $I \otimes_B M$ inherits an A-module structure (as M is an A-module). It is furthermore a *finitely generated* A-module (do you see why?), so by the Krull Intersection Theorem (Exercise 12.9.A), $\cap_t \mathfrak{m}^t(I \otimes_B M) = 0$. Thus it suffices to show that $\ker \phi \subset \mathfrak{m}^t(I \otimes_B M)$ for all t.

As $\mathfrak{n} \subset \mathfrak{m}$ (or more correctly, the image of $\mathfrak{n}$ is contained in $\mathfrak{m}$), it suffices to show that $\ker \phi \subset \mathfrak{n}^t(I \otimes_B M)$ for all t. Notice that $\mathfrak{n}^t(I \otimes_B M)$ is (the image in $I \otimes_B M$ of) $(\mathfrak{n}^t I) \otimes_B M$.

24.6.B. EXERCISE. Show that for each s, $\mathfrak{n}^t \cap I \subset \mathfrak{n}^s I$ for $t \gg 0$, so it suffices to show that $\ker \phi \subset (\mathfrak{n}^t \cap I) \otimes_B M$ for all t. Hint: Use the Artin-Rees Lemma 12.9.3, taking A there to be B here, $M_t = \mathfrak{n}^t$, $I = \mathfrak{n}$, and $L = I$.

Consider the short exact sequence

$$0 \to \mathfrak{n}^t \cap I \to I \to I/(\mathfrak{n}^t \cap I) \to 0.$$

Applying $(\cdot) \otimes_B M$, and using the fact that $I/(\mathfrak{n}^t \cap I)$ is finite length, we have that

$$0 \to (\mathfrak{n}^t \cap I) \otimes_B M \to I \otimes_B M \to \left(I/(\mathfrak{n}^t \cap I)\right) \otimes_B M \to 0$$

is exact using Exercise 24.6.A. Our goal is thus to show that $\ker \phi$ maps to 0 in

$$\left(I/(\mathfrak{n}^t \cap I)\right) \otimes_B M = \left((I + \mathfrak{n}^t)/\mathfrak{n}^t\right) \otimes_B M. \tag{24.6.3.1}$$

Applying $(\cdot) \otimes_B M$ to the short exact sequence

$$0 \to (I + \mathfrak{n}^t)/\mathfrak{n}^t \to B/\mathfrak{n}^t \to B/(I + \mathfrak{n}^t) \to 0, \tag{24.6.3.2}$$

and using Exercise 24.6.A (as $B/(I+\mathfrak{n}^t)$ is finite length), the top row of the diagram
(24.6.3.3)

$$\begin{tikzcd} 0 \arrow[r] & ((I + \mathfrak{n}^t)/\mathfrak{n}^t) \otimes_B M \arrow[r, "\alpha"] & (B/\mathfrak{n}^t) \otimes_B M \arrow[r] & (B/(I + \mathfrak{n}^t)) \otimes_B M \arrow[r] & 0 \\ & I \otimes_B M \arrow[r, "\phi"] \arrow[u] & B \otimes_B M \arrow[u] & & \end{tikzcd}$$

is exact, and the square clearly commutes. But then any element of $I \otimes_B M$ mapping to 0 in $B \otimes_B M = M$ must map to 0 (under the right vertical arrow) in $(B/\mathfrak{n}^t) \otimes_B M$, and hence must have mapped to 0 in $((I + \mathfrak{n}^t)/\mathfrak{n}^t) \otimes_B M$ by the injectivity of α, as desired. □

This argument basically shows that flatness is an "infinitesimal" property, depending only on the completion of the scheme at the point in question. This is made precise as follows.

Suppose $(B, \mathfrak{n}) \to (A, \mathfrak{m})$ is a (local) homomorphism of Noetherian local rings, and M is an A-module. If M is flat over B, then for each $t \in \mathbb{Z}^{\geq 0}$, $M/(\mathfrak{n}^t M)$ is flat over $B/\mathfrak{n}^t$ (flatness is preserved by base change, 24.2.I). (You should of course restate this in your mind in the language of schemes and quasicoherent sheaves.) The *infinitesimal criterion for flatness* states that this necessary criterion for flatness is actually sufficient.

24.6.C. ⋆ EXERCISE (THE INFINITESIMAL CRITERION FOR FLATNESS). Suppose $(B, \mathfrak{n}) \to (A, \mathfrak{m})$ is a (local) homomorphism of local rings, and M is a finitely generated A-module. Suppose further that for each $t \in \mathbb{Z}^{\geq 0}$, $M/(\mathfrak{n}^t M)$ is flat over $B/\mathfrak{n}^t$. Show that M is flat over B. (In combination with Exercise 24.6.A, this gives another proof of the local criterion of flatness, Theorem 24.6.2.) Hint: follow the proof of Theorem 24.6.2. Given the hypothesis, then for each t, we wish to show that $\ker \phi$ maps to 0 in (24.6.3.1). We wish to apply $(\cdot) \otimes_B M$ to (24.6.3.2) and obtain the (exactness of the) top row of (24.6.3.3). To do this, show that applying $(\cdot) \otimes_B M$ to (24.6.3.2) is the same as applying $(\cdot) \otimes_{B/\mathfrak{n}^t} (M/\mathfrak{n}^t M)$. Then proceed as in the rest of the proof of Theorem 24.6.2.

Exercise 24.6.C implies the useful fact that if A is a Noetherian local ring, then its completion (at its maximal ideal) is flat over A.

24.6.4. The slicing criterion for flatness.

A useful variant of the local criterion is the following. Suppose t is a non-zerodivisor of B in $\mathfrak{m}$ (geometrically: an effective Cartier divisor on the target passing through the closed point). If M is flat over B, then t is not a zerodivisor of M (Observation 24.2.2). Also, M/tM is a flat B/tB-module (flatness commutes with base change, Exercise 24.2.I). The next result says that this is a characterization of flatness, at least when M is finitely generated, or somewhat more generally.

24.6.5. Theorem (slicing criterion for flatness). — *Suppose $(B, \mathfrak{n}) \to (A, \mathfrak{m})$ is a local morphism of Noetherian local rings, M is a finitely generated A-module, and t is a non-zerodivisor on B. Then M is B-flat if and only t is not a zerodivisor on M, and M/tM is flat over $B/(t)$.*

(For two other slicing criteria, see Exercise 12.2.B and Theorem 26.2.3.)

Proof. Assume that t is not a zerodivisor on M, and M/tM is flat over $B/(t)$. We will show that M is B-flat. (As stated at the start of §24.6.4, the other implication is a consequence of what we have already shown.)

By the local criterion, Theorem 24.6.2, we know $\mathrm{Tor}_1^{B/(t)}(M/tM, (B/(t))/\mathfrak{n}) = 0$, and we wish to show that $\mathrm{Tor}_1^B(M, B/\mathfrak{n}) = 0$. (Note that $(B/(t))/\mathfrak{n} = B/\mathfrak{n}$.) The result then follows from the following lemma. □

24.6.6. Lemma. — *Suppose M is a B-module, and $t \in B$ is not a zerodivisor on M. Then for any $B/(t)$-module N, we have*

$$\mathrm{Tor}_i^B(M, N) = \mathrm{Tor}_i^{B/(t)}(M/tM, N). \tag{24.6.6.1}$$

Proof. We calculate the left side of (24.6.6.1) by taking a free resolution of M:

$$\cdots \to F_2 \to F_1 \to F_0 \to M \to 0. \tag{24.6.6.2}$$

By Exercise 24.3.A, $\mathrm{Tor}_i^B(M, B/(t)) = 0$ for $i > 0$ (here we use that t is not a zerodivisor on M, to show that $\mathrm{Tor}_1^B(M, B/(t)) = 0$). But this Tor module is computed by tensoring the free resolution (24.6.6.2) of M with $B/(t)$. Thus the complex

$$\cdots \to F_2/tF_2 \to F_1/tF_1 \to F_0/tF_0 \to M/tM \to 0 \tag{24.6.6.3}$$

is exact (exactness except at the last term comes from the vanishing of Tor_i). This is a free resolution of M/tM over the ring $B/(t)$! The left side of (24.6.6.1) is obtained by tensoring (24.6.6.2) by N and truncating and taking homology, and the right side is obtained by tensoring (24.6.6.3) by N and truncating and taking homology. As $(\cdot) \otimes_B N = (\cdot \otimes_B (B/t)) \otimes_{B/t} N$, we have established (24.6.6.1) as desired. $\square$

24.6.D. EXERCISE. Show that $\operatorname{Spec} k[x,y,z]/(x^2+y^2+z^2) \to \operatorname{Spec} k[x,y]$ is flat using the local slicing criterion.

24.6.E. EXERCISE. Give a second (admittedly less direct) proof of the criterion for flatness over a discrete valuation ring of Exercise 24.4.K, using the slicing criterion for flatness (Theorem 24.6.5).

24.6.F. EXERCISE. Use the slicing criterion to give a second solution to Exercise 24.4.J, on two planes in $\mathbb{P}^4$ meeting at a point.

The following exercise gives a sort of slicing criterion for flatness in the *source*.

24.6.G. EXERCISE. Suppose A is an B-algebra, A and B are Noetherian, M is a finitely generated A-module, and $f \in A$ has the property that for all maximal ideals $\mathfrak{m} \subset B$, multiplication by f is injective on $M/\mathfrak{m}M$. Show that if M is B-flat, then M/fM is also B-flat. (Hint: Use the local criterion for flatness, Theorem 24.6.2. Notice that

$$0 \longrightarrow M \xrightarrow{\times f} M \longrightarrow M/fM \longrightarrow 0$$

is a flat resolution of M/fM.)

Exercise 24.6.G has an immediate geometric interpretation: "Suppose $\pi : X \to Y$ is a morphism of Noetherian schemes, $\mathscr{F}$ is a coherent sheaf on X, and $Z \hookrightarrow X$ is a locally principal subscheme ..." In the special case where $\mathscr{F} = \mathscr{O}_X$, this leads to the notion of a **relative effective Cartier divisor**: a locally principal subscheme of X that is an effective Cartier divisor on all the fibers of π. Exercise 24.6.G implies that if π is flat, then any relative Cartier divisor is also flat. (See Exercise 28.3.F for an important application.)

24.6.7. *Remark: Local slicing criterion for flatness in the source, without Noetherian assumptions.*

The Noetherian hypotheses in Exercise 24.6.G can be removed. Suppose $(B, \mathfrak{n}) \to (A, \mathfrak{m})$ is a flat (local) homomorphism of local rings (not necessarily Noetherian). Suppose further that A is the localization of a finitely presented B-algebra. (This means that A is *essentially of finite presentation* over B, but we won't need this language.) If $f \in A$ is a nonzero divisor in $B/\mathfrak{m}S$, then A/f is flat over B. See [**Stacks**, tag 046Z] for a proof.

24.6.8. ⋆⋆ **Fibral flatness.** We conclude by mentioning a criterion for flatness that is useful enough to be worth recognizing, but not so useful as to merit proof here.

24.6.H. EXERCISE. Suppose we have a commuting diagram

$$\begin{tikzcd} X \arrow[rr, "\pi"] \arrow[dr, "\tau"'] & & Y \arrow[dl, "\rho"] \\ & Z & \end{tikzcd} \tag{24.6.8.1}$$

and a quasicoherent sheaf $\mathscr{F}$ on X, and points $p \in X$, $q = \pi(p) \in Y$, $r = \tau(p) \in Z$. Suppose ρ is flat at q, and $\mathscr{F}$ is π-flat at p. Show that $\mathscr{F}$ is τ-flat at p, and $\mathscr{F}|_r$ is π_r-flat at p. Here $\mathscr{F}|_r$ is the restriction of $\mathscr{F}$ to the fiber above r, and $\pi_r : \tau^{-1}(r) \to \rho^{-1}(r)$ is the restriction of π above r.

The Fibral Flatness Theorem states that in good circumstances the converse is true.

24.6.9. The Fibral Flatness Theorem ([Gr-EGA, IV$_3$.11.3.10], see also [Stacks, tag 039A]). — *Suppose we have a commuting diagram (24.6.8.1) and a finitely presented quasicoherent sheaf $\mathscr{F}$ on X, and points $p \in X$, $q = \pi(p) \in Y$, $r = \tau(p) \in Z$, with $\mathscr{F}_p \neq 0$. Suppose either X and Y are locally Noetherian, or ρ and τ are locally of finite presentation. Then the following are equivalent.*

(a) *$\mathscr{F}$ is τ-flat at p, and $\mathscr{F}|_r$ is π_r-flat at p.*
(b) *ρ is flat at q, and $\mathscr{F}$ is π-flat at p.*

This is a useful way of showing that a $\mathscr{F}$ is π-flat. The architecture of the argument is as follows. First reduce to the case where X, Y, and Z are affine. Cleverly reduce to the Noetherian case (see [**Gr-EGA**, IV.11.2.7]), then prove the resulting nontrivial problem in commutative algebra (see [**Gr-EGA**, IV.11.3.10.1]).

24.7 Flatness implies constant Euler characteristic

We come to an important consequence of flatness promised in §24.1. We will see that this result implies many answers and examples to questions that we would have asked before we even knew about flatness.

24.7.1. Important Theorem (Euler characteristic is constant in flat families). — *Suppose $\pi : X \to Y$ is a projective morphism of locally Noetherian schemes, and $\mathscr{F}$ is a coherent sheaf on X, flat over Y. Then $\chi(X_q, \mathscr{F}|_{X_q}) = \sum_{i \geq 0} (-1)^i h^i(X_q, \mathscr{F}|_{X_q})$ is a locally constant function of $q \in Y$ (where $X_q = \pi^{-1}(q)$).*

This is first sign that "cohomology behaves well in flat families." (We will soon see a second: the Semicontinuity Theorem 28.1.1. A different proof, giving an extension to the proper case, will be given in §28.2.5.) The Noetherian conditions are used to ensure that $\pi_* \mathscr{F}(m)$ is a coherent sheaf.

Theorem 24.7.1 gives a necessary condition for flatness. Converses (yielding a sufficient condition) are given in Exercise 24.7.A(b)–(d).

Proof. We make three quick reductions. (i) The question is local on the target Y, so we may reduce to case Y is affine, say $Y = \operatorname{Spec} B$, so π factors through a closed embedding $X \hookrightarrow \mathbb{P}^n_B$ for some n. (ii) We may reduce to the case $X = \mathbb{P}^n_B$, by considering $\mathscr{F}$ as a sheaf on $\mathbb{P}^n_B$. (iii) We may reduce to showing that for $m \gg 0$,

$h^0(X_q, \mathscr{F}(m)|_{X_q})$ is a locally constant function of $q \in Y$ (by Serre vanishing for $m \gg 0$, Theorem 18.1.4(ii), h^0 agrees with the Euler characteristic).

Twist by $\mathscr{O}(m)$ for $m \gg 0$, so that all the higher pushforwards vanish. Now consider the Čech complex $\mathscr{C}^\bullet$ for $\mathscr{F}(m)$. Note that all the terms in the Čech complex $\mathscr{C}^\bullet$ are flat, because $\mathscr{F}$ is flat. (Do you see why?) As all higher cohomology groups (higher pushforwards) vanish, $\mathscr{C}^\bullet$ is exact except at the first term, where the cohomology is $\Gamma(\pi_*\mathscr{F}(m))$. We add the module $\Gamma(\pi_*\mathscr{F}(m))$ to the front of the complex, so it is once again exact:

(24.7.1.1)

$$0 \longrightarrow \Gamma(\pi_*\mathscr{F}(m)) \longrightarrow \mathscr{C}^1 \longrightarrow \mathscr{C}^2 \longrightarrow \cdots \longrightarrow \mathscr{C}^{n+1} \longrightarrow 0.$$

(We have done this trick of tacking on a module before, for example in (18.2.4.1).) Thus by Exercise 24.3.G, as we have an exact sequence in which all but the first terms are flat, the first term is flat as well. Thus $\pi_*\mathscr{F}(m)$ is a flat coherent sheaf on Y, and hence locally free (Corollary 24.4.7), and thus has locally constant rank.

Suppose $q \in Y$. We wish to show that the Hilbert function $h_{\mathscr{F}|_{X_q}}(m)$ is a locally constant function of q. To compute $h_{\mathscr{F}|_{X_q}}(m)$, we tensor the Čech resolution with $\kappa(q)$ and take cohomology. Now the extended Čech resolution (with $\Gamma(\pi_*\mathscr{F}(m))$ tacked on the front), (24.7.1.1), is an exact sequence of flat modules, and hence remains exact upon tensoring with $\kappa(q)$ (Exercise 24.3.F). Thus $\Gamma(\pi_*\mathscr{F}(m)) \otimes \kappa(q) \cong \Gamma(\pi_*\mathscr{F}(m)|_q)$, so the Hilbert function $h_{\mathscr{F}|_{X_q}}(m)$ is the rank at q of a locally free sheaf, which is a locally constant function of $q \in Y$. $\square$

Before we get to the interesting consequences of Theorem 24.7.1, we mention some converses.

24.7.A. UNIMPORTANT EXERCISE (CONVERSES TO THEOREM 24.7.1). (We won't use this exercise for anything.)

(a) Suppose A is a ring, and $S_\bullet$ is a finitely generated A-algebra that is flat over A. Show that $\operatorname{Proj} S_\bullet$ is flat over A.

(b) Suppose $\pi : X \to Y$ is a projective morphism of locally Noetherian schemes (which as always includes the data of an invertible sheaf $\mathscr{O}_X(1)$ on X), such that $\pi_*\mathscr{O}_X(m)$ is locally free for all $m \geq m_0$ for some m_0. Show that π is flat. Hint: describe X as

$$\operatorname{Proj}\left(\mathscr{O}_Y \bigoplus (\oplus_{m \geq m_0} \pi_*\mathscr{O}_X(m))\right).$$

(c) More generally, suppose $\pi : X \to Y$ is a projective morphism of locally Noetherian schemes, and $\mathscr{F}$ is a coherent sheaf on X, such that $\pi_*\mathscr{F}(m)$ is locally free for all $m \geq m_0$ for some m_0. Show that $\mathscr{F}$ is flat over Y.

(d) Suppose $\pi : X \to Y$ is a projective morphism of locally Noetherian schemes, and $\mathscr{F}$ is a coherent sheaf on X, such that $\sum(-1)^i h^i(X_q, \mathscr{F}|_q)$ is a locally constant function of $q \in Y$. If Y is reduced, show that $\mathscr{F}$ must be flat over Y. (Hint: Exercise 13.7.K shows that constant rank implies local freeness in particularly nice circumstances.)

We now give some ridiculously useful consequences of Theorem 24.7.1.

24.7.2. Corollary. — *Assume the same hypotheses and notation as in Theorem 24.7.1. Then the Hilbert polynomial of $\mathscr{F}|_{X_q}$ is locally constant as a function of $q \in Y$.*

24.7.B. CRUCIAL EXERCISE. Suppose $X \to Y$ is a projective flat morphism, where Y is connected. Show that the following functions of $q \in Y$ are constant: (a) the degree of the fiber, (b) the dimension of the fiber, (c) the arithmetic genus of the fiber.

24.7.C. EXERCISE. Use §24.4.8 and Exercise 24.7.B(a) to give another solution to Exercise 17.4.D(a) ("the degree of a finite map from a curve to a regular curve is constant").

Another consequence of Corollary 24.7.2 is something remarkably useful.

24.7.3. Corollary. — *An invertible sheaf on a flat projective family of curves has locally constant degree on the fibers.*

(Recall that the degree of a line bundle on a projective curve requires no hypotheses on the curve such as regularity, see (18.4.4.1).)

Proof. An invertible sheaf $\mathscr{L}$ on a flat family of curves is always flat (as locally it is isomorphic to the structure sheaf). Hence $\chi(X_q, \mathscr{L}_q)$ is a constant function of q. By the definition of degree given in (18.4.4.1), $\deg(X_q, \mathscr{L}_q) = \chi(X_q, \mathscr{L}_q) - \chi(X_q, \mathscr{O}_{X_q})$. The result follows from the local constancy of $\chi(X_q, \mathscr{L}_q)$ and $\chi(X_q, \mathscr{O}_{X_q})$ (Theorem 24.7.1). $\square$

The following exercise is a serious generalization of Corollary 24.7.3.

24.7.D. $\star$ EXERCISE FOR THOSE WHO HAVE READ STARRED CHAPTER 20: INTERSECTION NUMBERS ARE LOCALLY CONSTANT IN FLAT FAMILIES. Suppose $\pi : X \to Y$ is a proper morphism to a connected scheme; $\mathscr{L}_1, \dots, \mathscr{L}_n$ are line bundles on X; and $\mathscr{F}$ is a coherent sheaf on X, flat over Y, such that the support of $\mathscr{F}$ when restricted to any fiber of π has dimension at most n. If q is any point of Y, define (the temporary notation) $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})_q$ to be the intersection on the fiber X_q of $\mathscr{L}_1, \dots, \mathscr{L}_n$ with $\mathscr{F}|_{X_q}$ (Definition 20.1.1). Show that $(\mathscr{L}_1 \cdot \mathscr{L}_2 \cdots \mathscr{L}_n \cdot \mathscr{F})_q$ is independent of q.

Corollary 24.7.3 motivates the following definition.

24.7.4. *Definition.* Suppose $\mathscr{L}_1$ and $\mathscr{L}_2$ are line bundles on a k-variety X. We say that $\mathscr{L}_1$ and $\mathscr{L}_2$ are **algebraically equivalent** if there exists a connected k-variety Y with two k-valued points q_1 and q_2, and a line bundle $\mathscr{L}$ on $X \times Y$ such that the restriction of $\mathscr{L}$ to the fibers over q_1 and q_2 are isomorphic to $\mathscr{L}_1$ and $\mathscr{L}_2$ respectively.

24.7.E. EXERCISE. Show that "algebraic equivalence" really is an equivalence relation. Show that the line bundles algebraically equivalent to $\mathscr{O}$ form a subgroup of $\operatorname{Pic} X$. This subgroup is denoted $\operatorname{Pic}^0 X$.

24.7.5. Identify the group of line bundles $\operatorname{Pic} X$ modulo algebraic equivalence with $\operatorname{Pic} X / \operatorname{Pic}^0 X$. This quotient is called the **Néron-Severi group**. (This definition was promised in §18.4.10.) By Proposition 20.1.4, $\operatorname{Pic}^\tau X \subset \operatorname{Pic}^0 X$: algebraic equivalence implies numerical equivalence. (Side Remark: a line bundle on a proper k-scheme X is numerically trivial if and only if there exists an integer $m \neq 0$ with

$L^{\otimes m}$ algebraically trivial. Thus $\mathrm{Pic}^\tau X/\mathrm{Pic}^0 X$ is torsion. See [**SGA6**, XIII, Thm. 4.6] for a proof, or [**Laz**, Cor. 1.4.38] for the projective case.)

24.7.6. ⋆ **Hironaka's example of a proper nonprojective smooth threefold.**
In §16.4.10, we produced a proper nonprojective variety, but it was singular. We can use Corollary 24.7.3 to give a *smooth* example, due to Hironaka.

Inside $\mathbb{P}^3_k$, fix two conics C_1 and C_2, which meet in two (k-valued) points, p_1 and p_2. We construct a proper map $\pi : X \to \mathbb{P}^3_k$ as follows. Away from p_i, we blow up C_i and then blow up the proper transform of C_{3-i} (see Figure 24.3). This is well-defined, as away from p_1 and p_2, C_1 and C_2 are disjoint, blowing up one and then the other is the same as blowing up their union, and thus the order doesn't matter.

[picture to be made later]

FIGURE 24.3. Hironaka's example of a proper nonprojective smooth threefold

Note that π is proper, as it is proper away from p_1, and proper away from p_2, and the notion of properness is local on the base (Proposition 10.3.4(b)). As $\mathbb{P}^3_k$ is proper (over k, Theorem 7.4.7), and compositions of proper morphisms are proper (Proposition 10.3.4(c)), X is proper.

24.7.F. EXERCISE. Show that X is smooth. (Hint: Theorem 22.3.10.) Let E_i be the preimage of $C_i \setminus \{p_1, p_2\}$. Show that $\pi|_{E_i} : E_i \to C_i \setminus \{p_1, p_2\}$ is a $\mathbb{P}^1$-bundle (and flat).

24.7.G. EXERCISE. Let $\overline{E}_i$ be the closure of E_i in X. Show that $\overline{E}_i \to C_i$ is flat. (Hint: Exercise 24.4.K.)

24.7.H. EXERCISE. Show that $\pi^*(p_i)$ is the union of two $\mathbb{P}^1$'s, say Y_i and Z_i, meeting at a point, such that $Y_i, Y_{3-i}, Z_{3-i} \in \overline{E}_i$ but $Z_i \notin \overline{E}_i$.

24.7.I. EXERCISE. Show that X is not projective as follows. Suppose otherwise $\mathscr{L}$ is a very ample line bundle on X, so $\mathscr{L}$ has positive degree on every curve (including the Y_i and Z_i). Using flatness of $\overline{E}_i \to C_i$, and constancy of degree in flat families (Exercise 24.7.D), show that $\deg_{Y_i} \mathscr{L} = \deg_{Y_{3-i}} \mathscr{L} + \deg_{Z_{3-i}} \mathscr{L}$. Obtain a contradiction. (This argument will remind you of the argument of §16.4.10.)

24.7.7. *The notion of "projective morphism" is not local on the target.* Note that $\pi : X \to \mathbb{P}^3$ is not projective, as otherwise X would be projective (as the composition of projective morphisms is projective if the final target is quasicompact, Exercise 17.3.B). But away from each p_i, π *is* projective (as it is a composition of blow-ups, which are projective by construction, and the final target is quasicompact, so Exercise 17.3.B applies). Thus the notion of "projective morphism" is not local on the target.

CHAPTER 25

Smooth, étale, and unramified morphisms revisited

We have defined smooth and étale morphisms earlier (Definition 12.6.2), but we are now in a position to understand them much better. We will see that the notion of unramified morphism (§21.6) is a natural companion to them.

Our three algebro-geometric definitions won't be so obviously a natural triplet, but we will discuss the definitions given in [**Gr-EGA**] (§25.2.6), and in this context the three types of morphisms look very similar. (We briefly mention other approaches and definitions in §25.2.7.)

25.1 Some motivation

The three classes of morphisms we will discuss in this chapter are the analogs of the following types of maps of manifolds, in differential geometry.

- *Submersions* are maps inducing surjections of tangent spaces everywhere. They are useful in the notion of a fibration. (Perhaps a more relevant notion from differential geometry, allowing singularities, is: "locally on the source a smooth fibration".)
- *Isomorphisms locally on the source* (or *local isomorphisms*) are maps inducing isomorphisms of tangent spaces.
- *Immersions* are maps inducing injections of tangent spaces.

(Recall our warning from §8.1.2: "immersion" is often used in algebraic geometry with a different meaning.)

[picture to be made later]

FIGURE 25.1. Sketches of notions from differential geometry: (a) maps "locally on the source a smooth fibration", (b) local isomorphisms, and (c) immersions. (In algebraic geometry: (a) smooth morphisms, (b) étale morphisms, and (c) unramified morphisms.)

In order to better understand smooth and étale morphisms, we temporarily forget our earlier definitions, and consider some examples of things we want to be analogs of "local isomorphism" (or "locally on the source an isomorphism") and "locally on the source a smooth fibration", and see if they help us make good definitions.

25.1.1. "Local isomorphisms" (étale morphisms). Consider the parabola $x = y^2$ projecting to the x-axis, over the complex numbers. (This example has come up repeatedly, in one form or another, including in Exercise 12.6.G.) We might reasonably want this to be a local isomorphism (on the source) away from the origin. We might also want the notion of local isomorphism space to be an open condition: the locus where a morphism is a local isomorphism should be open on the source. This is true for the differential geometric definition. But then this morphism should be a local isomorphism over the generic point, and here we get a non-trivial residue field extension ($\mathbb{C}(y)/\mathbb{C}(y^2)$), not an isomorphism. Thus we are forced to consider (the Spec's of) certain finite extensions of fields to be "isomorphisms locally on the source" — in this case, even "covering spaces". (We will see in Exercise 25.2.C that we want precisely the finite separable extensions; you could have shown this earlier.)

Note also in this example there are no (nonempty) Zariski-open subsets $U \subset \operatorname{Spec} \mathbb{C}[x] \setminus \{0\}$ and $V \subset \operatorname{Spec} \mathbb{C}[x,y]/(x-y^2) \setminus \{(0,0)\}$ where the map sends U into V isomorphically (Exercise 12.6.G), so this is not a local isomorphism in a way you may have seen before. This leads to the notion of the étale topology, which is not even a topology in the usual sense, but a "Grothendieck topology" (§13.3.4). The étale topology is beyond the scope of this book.

25.1.2. Submersions (smooth morphisms).

25.1.3. *Fibers are smooth varieties.* As a first approximation of the algebro-geometric version of submersion, we will want the fibers to be smooth varieties (over the residue field). So the very first thing we need is to generalize the notion of "variety" over a base. It is reasonable to do this by having a locally finite type hypothesis. For somewhat subtle reasons, we will require the stronger condition that the morphism be locally of finite presentation. (If you really care, you can see where it comes up in our discussion. But of course there is no difference for Noetherian readers.)

The fibers are not just varieties; they should be *smooth* varieties (of dimension n, say). From the case of smoothness over a field, or from our intuition of what smooth varieties should look like, we expect the sheaf of differentials on the fibers to be locally free of rank n, or even better, that the sheaf of relative differentials be locally free of rank n.

25.1.4. *Flatness.* At this point, our first approximation of "smooth morphism" is some version of "locally finitely presented, and fibers are smooth varieties". But that isn't quite enough. For example, a horrible map from a scheme X to a curve Y that maps a different regular variety to a each point Y (X is the infinite disjoint union of these) should not be considered a smooth fibration in any reasonable sense. Also, we might not want to consider $\operatorname{Spec} k \to \operatorname{Spec} k[\epsilon]/(\epsilon^2)$ to be a submersion; for example, this isn't surjective on tangent spaces, and more generally the picture "doesn't look like a fibration".

[picture to be made later]

FIGURE 25.2. We don't want these to be smooth morphisms

Both problems are failures of $\pi : X \to Y$ to be a nice, "'continuous" family. Whenever we are looking for some vague notion of "niceness" we know that "flatness" will be in the definition.

For comparison, note that "unramified" has no flatness hypothesis, and indeed we didn't expect it, as we would want the inclusion (closed embedding) of the origin into $\mathbb{A}^1$ to be unramified. But then weird things may be unramified. For example, if $X = \coprod_{z \in \mathbb{C}} \operatorname{Spec} \mathbb{C}$, then the morphism $X \to \mathbb{A}^1_{\mathbb{C}}$ sending the point corresponding to z to the point $z \in \mathbb{A}^1_{\mathbb{C}}$ is unramified. Such is life.

25.1.5. *Desired Alternate Definitions.* We might hope that a morphism $\pi : X \to Y$ is **smooth of relative dimension** n if and only if

(i) π is locally of finite presentation,
(ii) π is flat of relative dimension n, and
(iii) $\Omega_{X/Y}$ is locally free of rank n.

We might similarly hope that a morphism $\pi : X \to Y$ is **étale** if and only if

(i) π is locally of finite presentation,
(ii) π is flat, and
(iii) $\Omega_{X/Y} = 0$.

We will shortly show (in Theorem 25.2.2 and Exercise 25.2.B respectively) that these desired definitions are equivalent to Definition 12.6.2.

25.2 Different characterizations of smooth and étale morphisms

The main result in this section is a description of equivalent characterizations of smooth morphisms, Theorem 25.2.2. We will state the theorem, and then give consequences, and then finally give a proof.

25.2.1. But first, we discuss a preliminary fact. It requires something from §26.2.7: if k is a field, and $f_1, \dots, f_i \in k[x_1, \dots, x_N]$ cut out a dimension $N - i$ subscheme X of $\mathbb{A}^N_k$, and $f_{i+1} \in k[x_1, \dots, x_N]$ does not vanish on any irreducible component of X, then f_{i+1} is not a zerodivisor on X. For now, take this fact for granted; you shouldn't read all of Chapter 26 in order to do this exercise. There are other ways around this issue, but it is not worth the effort.

25.2.A. EXERCISE. Suppose $\pi : X \to Y$ is smooth of relative dimension n. Either (i) assume Y (and hence X) is locally Noetherian, or (ii) accept Remark 24.6.7 (the slicing criterion for flatness in the source without Noetherian hypotheses, which had a reference rather than a proof). Show that π is flat. Hint: suppose π is the morphism $\operatorname{Spec} B[x_1, \dots x_{n+r}]/(f_1, \dots, f_r) \to \operatorname{Spec} B$. Consider $B[x_1, \dots, x_{n+r}]/(f_1, \dots, f_i)$ as i runs from 0 to r, and use induction on i and the slicing criterion for flatness in the source ((i) Exercise 24.6.G or (ii) Remark 24.6.7).

We now come to the central result of this section.

25.2.2. Theorem. — *Suppose $\pi : X \to Y$ is a morphism of schemes. Then the following are equivalent.*

(i) *The morphism π is smooth of relative dimension n (Definition 12.6.2).*
(ii) *The morphism π is locally finitely presented and flat of relative dimension n; and Ω_π is locally free of rank n (Desired Alternate Definition 25.1.5).*
(iii) *The morphism π is locally finitely presented and flat, and the fibers are smooth varieties of dimension n.*

(iv) *The morphism π is locally finitely presented and flat, and the geometric fibers are smooth varieties of dimension n.*

Before proving Theorem 25.2.2, we motivate it, by giving a number of exercises and results assuming it.

25.2.B. EXERCISE. Show that a morphism π is étale (Definition 12.6.2) if and only if π is locally of finite presentation, flat, and $\Omega_\pi = 0$ (Desired Alternate Definition 25.1.5).

25.2.C. EASY EXERCISE. Suppose $\pi : X \to Y$ is a morphism. Show that the following are equivalent.

(a) π is étale.
(b) π is smooth and unramified.
(c) π is locally finitely presented, flat, and unramified.
(d) π is locally finitely presented, flat, and for each $q \in Y$, $\pi^{-1}(q)$ is the disjoint union of schemes of the form $\operatorname{Spec} K$, where K is a finite separable extension of $\kappa(q)$.
(e) π is locally finitely presented, flat, and for each *geometric* point $\overline{q}$ of Y, $\pi^{-1}(\overline{q})$ is the (scheme-theoretic) disjoint union of copies of $\overline{q}$.

25.2.D. EXERCISE. If $\pi : X \to Y$ is étale, show that the preimage $p \in X$ of any regular point $q \in Y$ whose local ring has dimension n is also a regular point whose local ring has dimension n. Hint: Prove the result by induction on $\dim \mathscr{O}_{Y,q}$. "Slice" by an element of $\mathfrak{m}_{Y,q} \setminus \mathfrak{m}_{Y,q}^2$. Use the slicing criterion for regularity (Exercise 12.2.B).

25.2.3. *Proof of the Smoothness-Regularity Comparison Theorem 12.2.10(b) (every smooth k-scheme is regular).* By Exercise 12.6.E, any dimension n smooth k-variety X can locally be expressed as an étale cover of $\mathbb{A}^n_k$, which is regular by Exercise 12.3.O. Then by Exercise 25.2.D, X is regular. $\square$

25.2.E. EXERCISE. Suppose $\pi : X \to Y$ is a morphism of k-varieties, and Y is smooth. Use the conormal exact sequence (Theorem 21.2.12) to show the following.
(a) Suppose that $\dim X = \dim Y = n$, and π is unramified. Show that X is smooth.
$\star$(b) Suppose $\dim X = m > \dim Y = n$, and the (scheme-theoretic) fibers of π over closed points are smooth of dimension $m - n$. Show that X is smooth.

25.2.F. EXERCISE. Suppose $\pi : X \to Y$ is a morphism (over k) of smooth k-varieties, where $\dim X = m$ and $\dim Y = n$. (These two parts are clearly essentially the same geometric situation.)
(a) Show that if the fibers of π are all smooth of dimension $m-n$, then π is smooth of relative dimension $m - n$.
(b) Show that if $\pi^*\Omega_Y \to \Omega_X$ is injective ("T_π is surjective", or "the relative cotangent sequence is exact on the left"), then π is smooth of relative dimension $m - n$. Hint: for each point $q \in Y$, $\mathscr{O}_{Y,q}$ is a regular local ring. Work by induction on $\dim \mathscr{O}_{Y,q}$, and "slice" by an element of $\mathfrak{m}_{Y,q} \setminus \mathfrak{m}_{Y,q}^2$.

25.2.G. EXERCISE. Suppose $\pi : X \to Y$ is locally finitely presented, $\rho : Y \to Z$ is étale, and $\tau = \rho \circ \pi$. Show that π is smooth of dimension n (e.g. étale, taking

$n = 0$) if and only if τ is. (We showed earlier that π is unramified if and only if τ is. Do you see why and where?) Hint: use the Cancellation Theorem 10.1.19 for flat morphisms and Exercise 21.6.F. For the Ω part of the problem, use the relative cotangent sequence, Theorem 21.2.25. (If you only solved Exercise 21.6.F(a), then you will still be able to use it to prove this in the Noetherian case.)

25.2.4. $\star\star$ *Proof of Theorem 25.2.2.*

(i) implies (ii). As noted in §12.6.2, finite presentation is clear. Flatness follows from Exercise 25.2.A. Finally, the sheaf of relative differentials is locally free of rank n by Exercise 21.2.Q.

(ii) implies (iii). Suppose π satisfies *(ii)*; we wish to show that the fibers are smooth. The hypotheses of *(ii)* are preserved by base change, so we may assume that $Y = \operatorname{Spec} k$. Then X is indeed smooth (see the equivalent Definition 21.3.1).

(iii) implies (iv) by Exercise 12.2.F.

We now come to the main part of the argument: *(iv) implies (i)*. (Part of this argument will be reminiscent of the proof of Theorem 12.6.3, in which we also had to prove smoothness.)

Fix a point $p \in X$. Let $q = \pi(p) \in Y$, say $q = \operatorname{Spec} k$. We will show that there is a neighborhood U_i of p and V_i of q of the form stated in Definition 12.6.2. It suffices to deal with the case that p is a closed point in $\pi^{-1}(q)$, because closed points are dense in finite type schemes (Exercise 5.3.E). (We cannot and do not assume that q is a closed point of Y.)

Because π is finitely presented, there are affine open neighborhoods U of p and V of q, with $\pi(U) \subset V$, such that the morphism $\pi|_U : U \to V$ can be written as

$$\operatorname{Spec} B[x_1, \dots, x_N]/I \to \operatorname{Spec} B,$$

where I *is finitely generated*.

Fix a geometric point $\overline{q} = \operatorname{Spec} \overline{k}$ of Y mapping to q. Fix a point $\overline{p}$ of the geometric fiber over $\overline{q}$, which maps to $p \in X$. (Why is there such a $\overline{p}$?) We then have a diagram

$$\begin{tikzcd}
U_{\overline{k}} \arrow[r] \arrow[d, equal] & U_k \arrow[r, hook] \arrow[d, equal] & U \arrow[d, equal] \\
\operatorname{Spec} \overline{k}[x_1, \dots, x_N]/I_{\overline{k}} \arrow[r] \arrow[d] & \operatorname{Spec} k[x_1, \dots, x_N]/I_k \arrow[r, hook] \arrow[d] & \operatorname{Spec} B[x_1, \dots, x_N]/I \arrow[d, "\pi|_U"] \\
\operatorname{Spec} \overline{k} \arrow[r] \arrow[d, equal] & \operatorname{Spec} k \arrow[r, hook] \arrow[d, equal] & \operatorname{Spec} B \arrow[d, equal] \\
\overline{q} \arrow[r] & q \arrow[r, hook] & V
\end{tikzcd}$$

where $U_{\overline{k}}$ (resp. U_k) are the fibers of U over $\overline{q}$ (resp. q), and $I_{\overline{k}}$ (resp. I_k) is the ideal cutting out $U_{\overline{k}}$ (resp. U_k) in $\mathbb{A}^N_{\overline{k}}$ (resp. $\mathbb{A}^N_k$).

We make an important observation, using §24.3.2 (for a flat closed subscheme, the pullback of the defining ideal as a quasicoherent sheaf is the same as the ideal of the pulled back closed subscheme). Because $B[x_1, \dots, x_N]/I$ is flat over B, we have that $I_{\overline{k}} = I \otimes_B \overline{k}$ and $I_k = I \otimes_B k$.

By our hypothesis *(iv)*, the geometric fiber $U_{\overline{k}} \subset \mathbb{A}^N_{\overline{k}}$ is smooth, and hence regular at its closed points by Exercise 12.2.C. (Here we use the fact that $\overline{k}$ is algebraically closed.)

Let $r = N - n$ be the codimension of U_k in $\mathbb{A}^N_k$ (and the codimension of $U_{\overline{k}}$ in $\mathbb{A}^N_{\overline{k}}$). By the Jacobian criterion for smoothness over a field (§21.3.2), the cotangent space for $U_{\overline{k}}$ at $\overline{p}$ is cut out in $\Omega_{\mathbb{A}^N_{\overline{k}}}|_{\overline{p}}$ by r linear equations (over $\overline{k}$). By the conormal exact sequence for $U_{\overline{k}} \subset \mathbb{A}^N_{\overline{k}}$ (Theorem 21.2.25(b)), these equations are spanned by the elements of $I_{\overline{k}}$, which in turn are spanned by elements of I (as $I_{\overline{k}} = I \otimes_B \overline{k}$, so elements of $I_{\overline{k}}$ are finite sums of pure tensors $i \otimes a$ with $i \in I$ and $a \in \overline{k}$). Thus we can choose $f_1, \dots, f_r$ in I so that the images of $f_1, \dots, f_r$ in $\overline{k}[x_1, \dots, x_N]$ cut out the Zariski tangent space of $U_{\overline{k}}$ at $\overline{p}$ (as the Zariski tangent space at a closed point of a variety over an algebraically closed field is computed by the Jacobian matrix, Exercise 12.2.C).

25.2.H. EXERCISE. Show that $f_1, \dots, f_r$ generate the ideal $(I_{\overline{k}})_{\overline{p}}$. Hint: look at your argument for Exercise 12.2.K.

Rearrange the x_i's so that the Jacobian matrix of the f_i with respect to the *first* r of the x_i is invertible at $\overline{p}$, and hence at p, and hence in a neighborhood of p. Then (a neighborhood of p in) $\operatorname{Spec} B[x_1, \dots, x_N]/(f_1, \dots, f_r)$ is of the form we are looking for. We will show that U is the same as $\operatorname{Spec} B[x_1, \dots, x_N]/(f_1, \dots, f_r)$ (i.e., $I = (f_1, \dots, f_r)$) near p, thereby completing the proof.

For notational compactness (at the cost of having subscripts later on with confusingly different meanings), define

$$\begin{aligned} A &:= B[x_1, \dots, x_N]/(f_1, \dots, f_r), \\ A_k &:= A \otimes_B k = k[x_1, \dots, x_N]/(f_1, \dots, f_r), \\ A_{\overline{k}} &:= A \otimes_B \overline{k} = \overline{k}[x_1, \dots, x_N]/(f_1, \dots, f_r). \end{aligned}$$

We have the diagram

$$\begin{tikzcd} A/J & & A \arrow[ll] \\ & B \arrow[ul] \arrow[ur] & \end{tikzcd}$$

where J denotes the image of I in A. We wish to show that "J is 0 near p".

Let $\mathfrak{m}$ be the maximal ideal of A corresponding to p, and let $\overline{\mathfrak{m}}$ be the maximal ideal of $A_{\overline{k}}$ corresponding to $\overline{p}$.

Let $J_{\overline{k}}$ be the image of $I_{\overline{k}}$ in $A_{\overline{k}}$. We use §24.3.2 a second time in this proof. Because $A/J = B[x_1, \dots, x_N]/I$ is flat over B, $J_{\overline{k}}$ can be interpreted both as the pullback of J as a quasicoherent sheaf, or as the ideal sheaf of the "pulled back closed subscheme corresponding to J". More precisely,

$$J \otimes_B \overline{k} = J A_{\overline{k}}. \tag{25.2.4.1}$$

Similarly,

$$J \otimes_B k = J A_k. \tag{25.2.4.2}$$

Now Exercise 25.2.H means that

$$J \cdot \left(A_{\overline{k}}\right)_{\overline{\mathfrak{m}}} = (0),$$

so taking the quotient by the maximal ideal $\overline{\mathfrak{m}}$ of $(A_{\overline{k}})_{\overline{\mathfrak{m}}}$ yields

$$\left(J \cdot A_{\overline{k}}\right) \otimes_{A_{\overline{k}}} \left(A_{\overline{k}}/\overline{\mathfrak{m}}\right) = (0),$$

which by (25.2.4.1) can be rewritten as

$$\left(J \otimes_B \overline{k}\right) \otimes_{A \otimes_B \overline{k}} \left(A_{\overline{k}}/\overline{\mathfrak{m}}\right) = (0),$$

which in turn can be readily rewritten as

$$(J \otimes_B k) \otimes_{A \otimes_B k} \left(A_{\overline{k}}/\overline{\mathfrak{m}}\right) = (0)$$

and then as

$$((J \otimes_B k) \otimes_{A \otimes_B k} (A_k/\mathfrak{m})) \otimes_{A_k/\mathfrak{m}} \left(A_{\overline{k}}/\overline{\mathfrak{m}}\right) = (0)$$

But $A_{\overline{k}}/\overline{\mathfrak{m}} = \overline{k}$ is a field extension of $A_k/\mathfrak{m} = k$, and a k-vector space V is 0 if (and only if) $V \otimes_k \overline{k} = 0$, so we have shown that

$$((J \otimes_B k) \otimes_{A \otimes_B k} (A_k/\mathfrak{m})) = (0).$$

Using (25.2.4.2) (essentially reversing our steps, with $\overline{k}$ replaced by k), we have

$$(J \cdot A_k) \otimes_{A_k} (A_k/\mathfrak{m}) = 0,$$

which means that $J \otimes_A (A/\mathfrak{m}) = 0$.

25.2.I. EXERCISE. Show that $J_{\mathfrak{m}} = 0$, and from this conclude that there is an element $f \in A \setminus \mathfrak{m}$ such that $fJ = 0$. (Hint: Nakayama, and finite generation of J.)

Then $J = 0$ in the neighborhood $D(f) \subset \operatorname{Spec} A$ of $\mathfrak{p}$, implying that *(i)* holds. □

This proof can be made notably easier by working in characteristic 0, where we can take *(iii)* rather than *(iv)* as our starting point to prove *(i)*, using the fact that smooth schemes over perfect fields k are regular at closed points (see Exercise 21.3.D).

25.2.5. $\star\star$ **Formally unramified, smooth, and étale.**

[**Gr-EGA**] takes a different starting point for the definition of unramified, smooth, and étale. The definitions there make clear that these three definitions form a family.

The cost of these definitions are that they are perhaps less immediately motivated by geometry, and it is harder to show some basic properties. The benefit is that it is possible to show more (for example, left-exactness of the relative cotangent and conormal sequences, and good interpretations in terms of completions of local rings). But we simply introduce these ideas here, and do not explore them. See [**BLR**, §2.2] for an excellent discussion. (You should largely ignore what follows, unless you find later in life that you really care.)

25.2.6. *Definition.* We say that $\pi : X \to Y$ is **formally smooth** (resp. **formally étale**, **formally unramified**) if for all affine schemes Z, and every closed subscheme $Z_0 \subset Z$ defined by a nilpotent ideal, and every morphism $Z \to Y$, the canonical map $\operatorname{Hom}_Y(Z, X) \to \operatorname{Hom}_Y(Z_0, X)$ is surjective (resp. bijective, injective). This is summarized in the following diagram, which is reminiscent of the valuative criteria for separatedness and properness.

$$\begin{tikzcd} Z_0 \arrow[r] \arrow[d, hook, "\text{nilpotent ideal}"'] & X \arrow[d, "\pi"] \\ Z \arrow[r] \arrow[ur, dotted, "?"] & Y \end{tikzcd}$$

(You can check that this is the same as the definition we would get by replacing "nilpotent" by "square-zero". This is sometimes an easier formulation to work with.)

Then [**Gr-EGA**] defines smooth as morphisms that are formally smooth and locally of finite presentation, and similarly for unramified and étale.

One can show that [**Gr-EGA**]'s definitions of formally unramified, and smooth agree with the definitions we give. For "formally unramified" (where our definition given in §21.6 is not obviously the same as the definition of [**Gr-EGA**] given here), see [**Gr-EGA**, IV_4.17.2.1] or [**Stacks**, tag 00UO]. For "smooth", see [**Gr-EGA**, IV_4.17.5.1] or [**Stacks**, tag 00TN]. (Our characterization of étale as smooth of relative dimension 0 then agrees with [**Gr-EGA**]. Our definition of unramified as formally unramified plus locally of finite *type* disagrees with [**Gr-EGA**], as mentioned in §21.6.)

25.2.7. Other starting points. Unlike many other definitions in algebraic geometry, there are a number of quite different ways of defining smooth, étale, and unramified morphisms, and it is nontrivial to relate them. We have just described the approach of [**Gr-EGA**]. Another common approach is the characterization of smooth morphisms as locally finitely presented, flat, and with regular geometric fibers (see Theorem 25.2.2). Yet another definition is via a naive version of the cotangent complex; this is the approach taken by [**Stacks**, tag 00T2], and is less frightening than it sounds. Finally, the different characterizations of Exercise 25.2.C give a number of alternate initial definitions of étaleness.

25.3 Generic smoothness and the Kleiman-Bertini Theorem

We will now discuss a number of important results that fall under the rubric of "generic smoothness", an idea we first met in §21.3.4. All will require working over a field of characteristic 0 in an essential way.

25.3.1. Theorem (generic smoothness on the source). — *Let k be a field of characteristic 0, and let $\pi : X \to Y$ be a dominant morphism of integral k-varieties. Then there is a nonempty (=dense) open set $U \subset X$ such that $\pi|_U$ is smooth of dimension* $\dim X - \dim Y$.

The key idea already appeared in Theorem 21.3.5, when we showed that every variety has an open subset that is regular.

Proof. Define $n = \dim X - \dim Y$ (the "relative dimension"). Now $K(X)/K(Y)$ is a finitely generated field extension of transcendence degree n (from transcendence theory, see for example Exercise 11.2.A), so $\Omega_{X/Y}$ has rank n at the generic point, by Exercise 21.2.M(b). (Here we use the hypothesis $\operatorname{char} k = 0$, as we are using the fact that $K(X)/K(Y)$ is separably generated by any transcendence basis.) By upper semicontinuity of fiber rank of a coherent sheaf (Exercise 13.7.J), it is rank n for every point in a dense open set. On a reduced scheme, constant rank implies locally free of that rank (Exercise 13.7.K), so $\Omega_{X/Y}$ is locally free of rank n on that set. Also, by generic flatness (Exercise 24.5.N), it is flat on a dense open set. Let U be the intersection of these two open sets. $\square$

25.3.2. *Example.* An examination of the proof yields an example showing that this result fails in positive characteristic: consider the purely inseparable extension $\mathbb{F}_p(t)/\mathbb{F}_p(t^p)$. The same problem can arise even over an *algebraically closed* field of characteristic p: consider $\mathbb{A}^1_k = \operatorname{Spec} k[t] \to \operatorname{Spec} k[u] = \mathbb{A}^1_k$, given by $u \mapsto t^p$. The field example is just the situation at the generic point of $\operatorname{Spec} k[u]$.

If furthermore X is smooth, the situation is even better.

25.3.3. Theorem (generic smoothness in the target). — *Suppose $\pi : X \to Y$ is a morphism of k-varieties, where $\operatorname{char} k = 0$, and X is smooth (over k). Then there is a dense open subset U of Y such that $\pi|_{\pi^{-1}(U)}$ is a smooth morphism.*

Note that $\pi^{-1}(U)$ may be empty! Indeed, if π is not dominant, we will have to take such a U.

To prove Theorem 25.3.3, we use a neat trick.

25.3.4. Lemma. — *Suppose $\pi : X \to Y$ is a morphism of schemes that are finite type over k, where $\operatorname{char} k = 0$. Define*

$$X_r = \left\{ p \in X \ : \ \operatorname{rank} \left(\pi^*\Omega_{Y/k} \to \Omega_{X/k}\right)\Big|_p \leq r \right\}.$$

Then $\dim \pi(X_r) \leq r$.

It is profitable to write $\pi^*\Omega_{Y/k} \to \Omega_{X/k}$ as Ω_π. It is intuitively helpful to interpret $\operatorname{rank} \Omega_\pi|_p$ as $\operatorname{rank} T_\pi|_p$.

In order to even make sense of $\dim \pi(X_r)$, you will need to do the following exercise, or else extract it from the proof of Lemma 25.3.4.

25.3.A. EXERCISE. Using the notation of Lemma 25.3.4, show that $\pi(X_r)$ is a constructible subset of Y, so we can take its dimension. (Possible hint: if X and Y are both smooth, show that the rank condition implies that X_r is cut out by "determinantal equations".)

25.3.5. *Proof of Lemma 25.3.4.* In this proof, we make repeated use of the identification of Zariski cotangent spaces at closed points with the fibers of cotangent sheaves, using Corollary 21.3.9. We can replace X by an irreducible component of X_r, and Y by the closure of that component's image in Y (with reduced subscheme structure, see §8.3.9). (The resulting map will have all of X contained in X_r. This boils down to the following linear algebra observation: if a linear map $\rho : V_1 \to V_2$ has rank at most r, and V_i' is a subspace of V_i, with ρ sending V_1' to V_2', then the restriction of ρ to V_1' has rank at most that of ρ itself.) Thus we have a dominant morphism $\pi : X \to Y$, and we wish to show that $\dim Y \leq r$. Using generic smoothness on the source (Theorem 25.3.1) for $Y \to \operatorname{Spec} k$, we can shrink Y further so as to assume that it is smooth. By generic smoothness on the source for $\pi : X \to Y$, there is a nonempty open subset $U \subset X$ such that $\pi : U \to Y$ is smooth. But then for any point $p \in U$, the tangent map $\Omega_{\pi(p),Y} \to \Omega_{p,X}$ is injective (Exercise 21.2.S), and has rank at most r. Taking p to be a *closed* point, we have $\dim Y = \dim_{\pi(p)} Y \leq \dim \Omega_{\pi(p),Y} \leq r$. $\square$

There is not much left to do to prove the theorem.

25.3.6. *Proof of Theorem 25.3.3.* Reduce to the case Y smooth over k (by restricting to a smaller open set, using generic smoothness of Y, Theorem 25.3.1). Say $n = \dim Y$. Now $\dim \overline{\pi(X_{n-1})} \leq n-1$ by Lemma 25.3.4, so remove $\overline{\pi(X_{n-1})}$ from Y as well. Then the rank of Ω_π is at least n for each closed point of X. But as Y is regular of dimension n, we have that $\Omega_\pi = (\pi^*\Omega_{Y/k} \to \Omega_{X/k})$ is injective for every closed point of X, and hence $\pi^*\Omega_{Y/k} \to \Omega_{X/k}$ is an injective map of sheaves (do you see why?). Thus π is smooth by Exercise 25.2.F(b). $\square$

25.3.7. ⋆⋆ **The Kleiman-Bertini Theorem.** The same idea of bounding the dimension of some "bad locus" can be used to prove the Kleiman-Bertini Theorem 25.3.8 (due to Kleiman), which is useful in (for example) enumerative geometry. Throughout this discussion $k = \overline{k}$, although the definitions and results can be generalized. Suppose G is a group variety (over $k = \overline{k}$), and we have a G-action on a variety X. We say that the **action is transitive** if it is transitive on closed points. A better definition (that you can show is equivalent) is that the action morphism $G \times X \to X$ restricted to a fiber above any closed point of X is surjective. A variety X with a transitive G-action is said to be a **homogeneous space** for G. For example, G acts on itself by left-translation, and via this action G is a homogeneous space for itself.

25.3.B. EASY EXERCISE. Suppose G is a group variety over an algebraically closed field k of characteristic 0. Show that every homogeneous space X for G is smooth. (In particular, taking $X = G$, we see that G is smooth.) Hint: X has a dense open set U that is smooth by Theorem 25.3.1, and G acts transitively on the closed points of X, so we can cover X with translates of U.

25.3.8. The Kleiman-Bertini Theorem, [Kl2, Thm. 2]. — *Suppose X is homogeneous space for a group variety G (over a field $k = \overline{k}$ of characteristic 0). Suppose $\alpha : Y \to X$ and $\beta : Z \to X$ are morphisms from smooth k-varieties Y and Z.*
(a) Then there is a nonempty open subset $V \subset G$ such that for every $\sigma \in V(k)$, $Y \times_X Z$ defined by

$$\begin{tikzcd} Y \times_X Z \arrow[r] \arrow[d] & Z \arrow[d, "\beta"] \\ Y \arrow[r, "\sigma \circ \alpha"] & X \end{tikzcd}$$

(Y is "translated by σ") is smooth of dimension exactly $\dim Y + \dim Z - \dim X$.
(b) Furthermore, there is a nonempty open subset of $V \subset G$ such that

$$(G \times_k Y) \times_X Z \to G \tag{25.3.8.1}$$

is a smooth morphism of relative dimension $\dim Y + \dim Z - \dim X$ *over* V.

The first time you hear this, you should think of the special case where $Y \to X$ and $Z \to X$ are locally closed embeddings (Y and Z are smooth subvarieties of X).

In this case, the Kleiman-Bertini theorem says that the second subvariety will meet a "general translate" of the first "transversely".

Proof. It is more pleasant to describe this proof "backwards", by considering how we would prove it ourselves. We will use generic smoothness twice.

Clearly (b) implies (a), so we prove (b).

In order to show that the morphism (25.3.8.1) is generically smooth on the target, it would suffice to apply generic smoothness on the target (Theorem 25.3.3), so we wish to show that $(G \times_k Y) \times_X Z$ is a smooth k-variety. Now Z is smooth over k, so it suffices to show that $(G \times_k Y) \times_X Z \to Z$ is a smooth morphism (as the composition of two smooth morphisms is smooth, Exercise 12.6.D). But this is obtained by base change from $G \times_k Y \to X$, so it suffices to show that this latter morphism is smooth (as smoothness is preserved by base change).

Now $G \times_k Y \to X$ is a G-equivariant morphism. (By "G-equivariant", we mean that the G-action on both sides respects the morphism.) By generic smoothness of the target (Theorem 25.3.3), this is smooth over a dense open subset X. But then by transitivity of the G-action, this morphism is smooth everywhere. $\square$

25.3.C. EXERCISE (POOR MAN'S KLEIMAN-BERTINI). Prove Theorem 25.3.8(a) without the hypotheses on k (on algebraic closure or characteristic), and without the smoothness in the conclusion. Hint: This is a question about dimensions of fibers of morphisms, so you could have solved this after reading §11.4.

25.3.D. EXERCISE (IMPROVED CHARACTERISTIC 0 BERTINI). Suppose Z is a smooth k-variety, where $\operatorname{char} k = 0$ and $k = \overline{k}$. Let V be a finite-dimensional base-point-free linear series on D, i.e., a finite vector space of sections of some invertible sheaf $\mathscr{L}$ on D. Show that a general element of V, considered as a closed subscheme of Z, is regular. (More explicitly: each element $s \in V$ gives a closed subscheme of Z. Then for a general s, considered as a point of $\mathbb{P}V$, the corresponding closed subscheme is smooth over k.) Hint: figure out what this has to do with the Kleiman-Bertini Theorem 25.3.8. Let $n = \dim V$, $G = GL(V)$, $X = \mathbb{P}|V^\vee|$, take Z in Kleiman-Bertini to be the Z of the problem, and let Y be the "universal hyperplane" over $\mathbb{P}|V^\vee|$ (the variety $I \subset \mathbb{P}V \times \mathbb{P}V^\vee$ of Definition 12.4.1).

25.3.E. EASY EXERCISE. Interpret Bertini's Theorem 12.4.2 over a characteristic 0 field as a corollary of Exercise 25.3.D.

In characteristic 0, Exercise 25.3.D is a good improvement on Bertini's Theorem. For example, we don't need $\mathscr{L}$ to be very ample, or X to be projective. But unlike Bertini's Theorem, Exercise 25.3.D fails in positive characteristic, as shown by the one-dimensional linear series $\{pQ \ : \ Q \in \mathbb{P}^1\}$. This is essentially Example 25.3.2. (Do you see why this does not contradict Bertini's Theorem 12.4.2?)

CHAPTER 26

Depth and Cohen-Macaulayness

We now introduce the notion of depth. Depth is an algebraic rather than geometric concept, so we concentrate on developing some geometric sense of what it means. Most important is the geometric bound on depth by dimension of associated points (Theorem 26.1.2). A central tool to understanding depth is the Koszul complex, but we avoid this approach, as we can prove what we need directly.

When the depth of a local ring equals its dimension, the ring is said to be Cohen-Macaulay, and this is an important way in which schemes can be "nice". For example, regular local rings are Cohen-Macaulay (§26.2.5), as are regular embeddings in smooth varieties (Proposition 26.2.6). Cohen-Macaulayness will be key to the proof of Serre duality in Chapter 30, through the Miracle Flatness Theorem 26.2.11.

Another application of depth is Serre's $R_1 + S_2$ criterion for normality (Theorem 26.3.2), which we will use to prove that regular schemes are normal (§26.3.5) without having to show that regular local rings are factorial (Fact 12.8.5), and to prove that regular embeddings in smooth schemes are normal if they are regular in codimension 1 (§26.3.3).

26.1 Depth

Recall the theory of regular sequences from §8.4.3.

26.1.1. *Definition.* Suppose $(A, \mathfrak{m})$ is a Noetherian local ring, and M is a finitely generated A-module. The **depth of** M (denoted $\operatorname{depth} M$) is the length of the longest M-regular sequence with elements in $\mathfrak{m}$. (More generally, if R is a Noetherian ring, $I \subset R$ is an ideal, and M is a finitely generated R-module, then the I**-depth of** M, denoted $\operatorname{depth}_I M$, is the length of the longest M-regular sequence with elements in I. We won't need this.)

26.1.A. IMPORTANT EXERCISE. Suppose M is a finitely generated module over a Noetherian local ring $(A, \mathfrak{m})$. Show that $\operatorname{depth} M = 0$ if and only if every element of $\mathfrak{m}$ is a zerodivisor of M if and only if $\mathfrak{m}$ is an associated prime of M.

26.1.B. EXERCISE. Suppose M is a finitely generated module over a Noetherian local ring A. Show that $\operatorname{depth} M \leq \dim \operatorname{Supp} M$. In particular,

(26.1.1.1) $$\operatorname{depth} A \leq \dim A.$$

Hint: Krull's Principal Ideal Theorem 11.3.3. (We will improve this result in Theorem 26.1.6.)

At this point, it is hard to determine the depth of an A-module M. You can start trying to build an M-regular sequence by successively choosing $x_1, x_2, \dots$, but how do you know you have made the right choices to find the *longest* one? The happy answer is that you can't go wrong; this is the content of the next result. (We then describe how to find the depth of a module M in practice, in §26.1.4.)

26.1.2. Theorem. — *Suppose M is a finitely generated module over a Noetherian local ring $(A, \mathfrak{m})$. Then all maximal M-regular sequences contained in $\mathfrak{m}$ have the same length. Thus the depth of M is the length of any maximal M-regular sequence.*

We prove Theorem 26.1.2 by giving a cohomological criterion, Theorem 26.1.3, for a regular sequence to be maximal. (You will then prove Theorem 26.1.2 in Exercise 26.1.D.) We will also use this criterion to give a better bound on the depth in Theorem 26.1.6. Theorem 26.1.2 is the key technical result of this Chapter. An important moral is that depth should be understood as a "cohomological" property.

26.1.3. Theorem (cohomological criterion for existence of regular sequences). — *Suppose $(A, \mathfrak{m})$ is a Noetherian local ring, and M is a finitely generated A-module. The following are equivalent.*

(i) *For every finitely generated A-module N with $\operatorname{Supp} N = \{[\mathfrak{m}]\}$, $\operatorname{Ext}^i_A(N, M) = 0$ for all $i < n$.*
(ii) $\operatorname{Ext}^i_A(A/\mathfrak{m}, M) = 0$ *for all $i < n$.*
(iii) *There exists an M-regular sequence in $\mathfrak{m}$ of length n.*

This result can be extended in various ways, see for example [**Mat1**, Thm. 28].

Proof. Clearly *(i) implies (ii)*.

26.1.C. EXERCISE. Prove that *(ii) implies (i)*. Hint: apply Exercise 5.5.M to N, so that it admits a filtration with each subquotient isomorphic to $A/\mathfrak{m}$, or see Exercise 18.4.R. Use induction on the length of the filtration, and the long exact sequence for $\operatorname{Ext}^i_A(\cdot, M)$.

Proof that (iii) implies (ii). The case $n = 0$ is vacuous. We inductively prove the result for all n. Suppose (iii) is satisfied, where $n \geq 1$, and assume that we know (ii) for "all smaller n". Choose a regular sequence $x_1, \dots, x_n$ of length n. Then x_1 is a non-zerodivisor on M, so we have an exact sequence

$$0 \longrightarrow M \xrightarrow{\times x_1} M \longrightarrow M/x_1 M \longrightarrow 0. \tag{26.1.3.1}$$

Then M/x_1M has a regular sequence $x_2, \dots, x_{n-1}$ of length $n-1$, so by the inductive hypothesis, $\operatorname{Ext}^i_A(A/\mathfrak{m}, M/x_1M) = 0$ for $i < n-1$. Taking the Ext long exact sequence for $\operatorname{Ext}^i_A(A/\mathfrak{m}, \cdot)$ for (26.1.3.1), we find that

$$\operatorname{Ext}^i_A(A/\mathfrak{m}, M) \xrightarrow{\times x_1} \operatorname{Ext}^i_A(A/\mathfrak{m}, M)$$

is an injection for $i < n$. Now $\operatorname{Ext}^i_A(A/\mathfrak{m}, M)$ can be computed by taking an injective resolution of M, and applying $\operatorname{Hom}_A(A/\mathfrak{m}, \cdot)$. Hence as x_1 lies in $\mathfrak{m}$ (and thus annihilates $A/\mathfrak{m}$), multiplication by x_1 is the zero map. Thus (ii) holds for n as well.

Proof that (ii) implies (iii). The case $n = 0$ is vacuous.

We deal next with the case $n = 1$, by showing the contrapositive. Assume that there are no non-zerodivisors in $\mathfrak{m}$ on M, so by Exercise 26.1.A, $\mathfrak{m}$ is an associated prime of M. Thus from §5.5.9 we have an injection $A/\mathfrak{m} \hookrightarrow M$, yielding $\mathrm{Hom}_A(A/\mathfrak{m}, M) \neq 0$ as desired.

We now inductively prove the result for all $n > 1$. Suppose (ii) is satisfied, where $n \geq 2$, and assume that we know (iii) for "all smaller n". Then by the case $n = 1$, there exists a non-zerodivisor x_1 on M, so we have a short exact sequence (26.1.3.1). A portion of the Ext long exact sequence for $\mathrm{Ext}^i_A(A/\mathfrak{m}, \cdot)$ for (26.1.3.1) is

$$\mathrm{Ext}^i_A(A/\mathfrak{m}, M) \longrightarrow \mathrm{Ext}^i_A(A/\mathfrak{m}, M/x_1M) \longrightarrow \mathrm{Ext}^{i+1}_A(A/\mathfrak{m}, M).$$

By assumption, both $\mathrm{Ext}^i_A(A/\mathfrak{m}, M)$ and $\mathrm{Ext}^{i+1}_A(A/\mathfrak{m}, M)$ are 0 for $i < n-1$, so $\mathrm{Ext}^i_A(A/\mathfrak{m}, M/x_1M) = 0$ for $i < n-1$, so by the inductive hypothesis, we have an (M/x_1M)-regular sequence $x_2, \dots, x_n$ of length $n-1$ in $\mathfrak{m}$. Adding x_1 to the front of this sequence, we are done. $\square$

26.1.D. EXERCISE. Prove Theorem 26.1.2. Hint: by Theorem 26.1.3 (notably, the equivalence of (ii) and (iii)), you have control of how long an M-regular sequence in $\mathfrak{m}$ can be. Use the criterion, and the long exact sequence used in the proof of Theorem 26.1.3, to show that any M-regular sequence in $\mathfrak{m}$ can be extended to this length.

26.1.E. EXERCISE. Suppose M is a finitely generated module over a Noetherian local ring $(A, \mathfrak{m})$. If x is a non-zerodivisor in $\mathfrak{m}$, show that $\operatorname{depth}(M/xM) = \operatorname{depth} M - 1$. Hint: Theorem 26.1.2.

26.1.4. *Finding the depth of a module.* We can now compute the depth of M by successively finding non-zerodivisors, as follows. Is there a zerodivisor x on M in $\mathfrak{m}$?

(a) If not, then $\operatorname{depth} M = 0$ (Exercise 26.1.A).
(b) If so, then choose any such x, and (using the previous exercise) repeat the process with M/xM.

The process must terminate by Exercise 26.1.B.

26.1.F. IMPORTANT EXERCISE. Suppose $(A, \mathfrak{m})$ is a dimension d regular local ring. Show that $\operatorname{depth} A = d$. Hint: if $x \in \mathfrak{m} \setminus \mathfrak{m}^2$, consider the non-zerodivisor x.

26.1.G. EXERCISE (CF. EXERCISE 24.4.J). Suppose $X = \operatorname{Spec} R$, where

$$R = k[w, x, y, z]/(wy, wz, xy, xz),$$

the union of two co-ordinate two-planes in $\mathbb{A}^4_k$ meeting at the origin. Show that the depth of the local ring of X at the origin is 1. Hint: Show that $w - y$ is not a zerodivisor, and that $R/(w-y)$ has an embedded point at the origin.

26.1.5. Depth is bounded by the dimension of associated primes.

Theorem 26.1.3 can be used to give an important improvement of the bound (26.1.1.1) on depth by the dimension:

26.1.6. Theorem. — *The depth of a module* M *is at most the smallest* $\dim A/\mathfrak{p}$ *as* $\mathfrak{p}$ *runs over the associated primes of* M.

(The example of two planes meeting at a point in Exercise 26.1.G shows that this bound is not sharp.) The key step in the proof of Theorem 26.1.6 is the following result of Ischebeck.

26.1.7. Lemma. — *Suppose* $(A, \mathfrak{m})$ *is a Noetherian local ring, and* M *and* N *are nonzero finitely generated* A*-modules. Then* $\operatorname{Ext}_A^i(N, M) = 0$ *for* $i < \operatorname{depth} M - \dim \operatorname{Supp} N$.

Proof. Consider the following statements.

(*general*$_r$) Lemma 26.1.7 holds for $\dim \operatorname{Supp} N \leq r$.
(*prime*$_r$) Lemma 26.1.7 holds for $\dim \operatorname{Supp} N \leq r$ and $N = A/\mathfrak{p}$ for some prime $\mathfrak{p}$.

Note that (*general*$_0$) (and hence (*prime*$_0$)) is true, as in this case $\operatorname{Supp} N = \{[\mathfrak{m}]\}$, and the result follows from Theorem 26.1.3 (from "(iii) implies (i)").

26.1.H. EXERCISE. Use Exercise 5.5.M, and Exercise 5.5.L (or the easier fact that if $0 \to M' \to M \to M'' \to 0$ is a short exact sequence of A-modules, then $\operatorname{Supp} M = \operatorname{Supp} M' \cup \operatorname{Supp} M''$), to show that (*prime*$_r$) implies (*general*$_r$).

We conclude the proof by showing that (*general*$_{r-1}$) implies (*prime*$_r$) for $r \geq 1$. Fix a prime $\mathfrak{p} \subset A$ with $\dim A/\mathfrak{p} = r$. Since $\dim A/\mathfrak{p} > 0$, we can choose $x \in \mathfrak{m} \setminus \mathfrak{p}$. Consider the exact sequence

$$0 \longrightarrow A/\mathfrak{p} \xrightarrow{\times x} A/\mathfrak{p} \longrightarrow A/(\mathfrak{p} + (x)) \longrightarrow 0, \qquad (26.1.7.1)$$

noting that $\dim \operatorname{Supp} A/(\mathfrak{p}+(x)) \leq r-1$ (do you see why?). Then the Ext long exact sequence obtained by applying $\operatorname{Hom}_A(\cdot, M)$ to (26.1.7.1), along with the vanishing of $\operatorname{Ext}_A^i(A/(\mathfrak{p} + (x)), M)$ for $i < \operatorname{depth} M - r + 1$ (by (*general*$_{r-1}$)), implies that

$$\operatorname{Ext}_A^i(N, M) \xrightarrow{\times x} \operatorname{Ext}_A^i(N, M)$$

is an isomorphism for $i < \operatorname{depth} M - r$. But $\operatorname{Ext}_A^i(N, M)$ is a finitely generated A-module (Exercise 23.2.F), so by Nakayama's Lemma 7.2.8, $\operatorname{Ext}_A^i(N, M) = 0$ for $i < \operatorname{depth} M - r$. □

26.1.I. EASY EXERCISE. Prove Theorem 26.1.6. Hint: from §5.5.9, if $\mathfrak{p} \in \operatorname{Ass}(M)$, then $\operatorname{Hom}_A(A/\mathfrak{p}, M) \neq 0$.

26.2 Cohen-Macaulay rings and schemes

26.2.1. *Definition.* A Noetherian local ring $(A, \mathfrak{m})$ is **Cohen-Macaulay** (or often **CM** for short) if $\operatorname{depth} A = \dim A$, i.e., if equality holds in (26.1.1.1). (One may define **Cohen-Macaulay module** similarly, but we won't need this concept.) A locally Noetherian scheme is **Cohen-Macaulay** if all of its local rings are Cohen-Macaulay.

26.2.A. EXERCISE. Show that every locally Noetherian scheme of dimension 0 is Cohen-Macaulay. Show that a locally Noetherian scheme of dimension 1 is Cohen-Macaulay if and only if it has no embedded points.

26.2.2. *(Counter)example.* Let X be the example of Exercise 26.1.G — two planes meeting at a point. By Exercise 26.1.G, X is not Cohen-Macaulay.

26.2.B. EXERCISE. Suppose A is Cohen-Macaulay Noetherian local ring. Use Theorem 26.1.6 to show that $\operatorname{Spec} A$ is pure dimensional, and has no embedded points. (It is not true that Noetherian local rings of pure dimension having no embedded primes are Cohen-Macaulay, see Example 26.2.2.)

26.2.3. Theorem (slicing criterion for Cohen-Macaulayness). — *Suppose $(A, \mathfrak{m})$ is a Noetherian local ring, and $x \in \mathfrak{m}$ is a non-zerodivisor. Then $(A, \mathfrak{m})$ is Cohen-Macaulay if and only if $(A/x, \mathfrak{m})$ is Cohen-Macaulay.*

Compare this to the slicing criteria for regularity and flatness (Exercise 12.2.B and Theorem 24.6.5 respectively).

26.2.4. *Geometric interpretation of the slicing criterion.* Suppose X is a locally Noetherian scheme, and D is an effective Cartier divisor. If X is Cohen-Macaulay, then so is D. If D is Cohen-Macaulay, then X is Cohen-Macaulay at the points of D.

26.2.C. EXERCISE. Prove Theorem 26.2.3, using Theorem 26.1.2, the fact that maximal regular sequences (in $\mathfrak{m}$) all have the same length.

26.2.D. EXERCISE. Show that if $(A, \mathfrak{m})$ is Cohen-Macaulay, then a set of elements $x_1, \dots, x_r \in \mathfrak{m}$ is a regular sequence (for A) if and only if $\dim A/(x_1, \dots, x_r) = \dim A - r$.

26.2.5. By Exercise 26.1.F, regular local rings are Cohen-Macaulay. In particular, as smooth schemes over a field k are regular (Theorem 12.2.10(b)), we see that smooth k-schemes are Cohen-Macaulay. (In combination with Exercise 26.2.B, this explains why effective Cartier divisors on smooth varieties have no embedded points, which justifies a comment made in Aside 20.2.1.) Combining this with Exercise 26.2.D or §26.2.4, we have the following.

26.2.6. Proposition. — *Regular embeddings in smooth k-schemes are Cohen-Macaulay.*

26.2.7. As a consequence of Proposition 26.2.6 and the fact that Cohen-Macaulay schemes have no embedded points (Exercise 26.2.B), we see that regular embeddings in smooth k-schemes (in $\mathbb{A}^n_k$ or $\mathbb{P}^n_k$ for example) have no embedded points, generalizing Exercise 5.5.I (the hypersurface in $\mathbb{A}^n_k$ case). This is not clear without the theory of Cohen-Macaulayness! (This fact was used in Exercise 25.2.A, see §25.2.1.)

26.2.8. *Alternate definition of Cohen-Macaulayness.* The slicing criterion (Theorem 26.2.3) gives an enlightening alternative inductive definition of Cohen-Macaulayness in terms of effective Cartier divisors, in the spirit of the method of §26.1.4 for computing depth. Suppose as before that $(A, \mathfrak{m})$ is a Noetherian local ring.

(i) If $\dim A = 0$, then A is Cohen-Macaulay (Exercise 26.2.A).
(iia) If $\dim A > 0$, and every element of $\mathfrak{m}$ is a zerodivisor, then A is *not* Cohen-Macaulay (by Exercise 26.1.A).

(iib) Otherwise, choose *any* non-zerodivisor x in $\mathfrak{m}$. Then A is Cohen-Macaulay if and only if $A/(x)$ (necessarily of dimension $\dim A - 1$ by Krull's Principal Ideal Theorem 11.3.3) is Cohen-Macaulay.

The following example could have been stated (but not proved) before we knew any algebraic geometry at all. (We work over $\mathbb{C}$ rather than over an arbitrary field simply to ensure that the statement requires as little background as possible.)

26.2.E. FUN EXERCISE (MAX NOETHER'S $AF + BG$ THEOREM). Suppose $f, g \in \mathbb{C}[x_0, x_1, x_2]$ are two homogeneous polynomials, cutting out two curves in $\mathbb{P}^2_{\mathbb{C}}$ that meet "transversely", i.e., at a finite number of reduced points. Suppose $h \in \mathbb{C}[x_0, x_1, x_2]$ is a homogeneous polynomial vanishing at these points. Show that $h \in (f, g)$. Hint: show that the intersection of the affine cones $V(f)$ and $V(g)$ in $\mathbb{A}^3$ has no embedded points. (This problem is quite nontrivial to do without the theory developed in this chapter! As a sign that this is subtle: you can easily construct *three* quadratics $e, f, g \in \mathbb{C}[x_0, x_1, x_2]$ cutting out precisely the two points $[1, 0, 0]$ and $[0, 1, 0] \in \mathbb{P}^2_{\mathbb{C}}$, yet the line $z = 0$ is not in the ideal (e, f, g) for degree reasons.)

26.2.9. Miracle flatness.

We conclude with a remarkably simple and useful criterion for flatness, which we shall use in the proof of Serre duality. The main content is the following algebraic result.

26.2.10. Miracle Flatness Theorem (algebraic version). — *Suppose $\phi : (B, \mathfrak{n}) \to (A, \mathfrak{m})$ is a (local) homomorphism of Noetherian local rings, such that A is Cohen-Macaulay, and B is regular, and $A/\mathfrak{n}A = A \otimes_B (B/\mathfrak{n})$ (the ring corresponding to the fiber) has dimension $\dim A - \dim B$. Then ϕ is flat.*

Proof. We prove Theorem 26.2.10 by induction on $\dim B$. If $\dim B = 0$, then B is a field (Exercise 12.2.A), so the result is immediate, as everything is flat over a field (Exercise 24.2.A(a)). Assume next that $\dim B > 0$, and we have proved the result for all "B of smaller dimension". Choose $x \in \mathfrak{n} - \mathfrak{n}^2$, so B/x is a regular local ring of dimension $\dim B - 1$. Then

$$\begin{aligned}
\dim A/xA &\leq \dim B/(x) + \dim A/\mathfrak{n}A && \text{(Key Exercise 11.4.A)} \\
&= \dim B - 1 + \dim A/\mathfrak{n}A && \text{(Krull's Principal Ideal Theorem 11.3.3)} \\
&= \dim A - 1 && \text{(by hypothesis of Theorem 26.2.10)}.
\end{aligned}$$

By Krull's Principal Ideal Theorem 11.3.3, $\dim A/xA \geq \dim A - 1$, so we have $\dim A/xA = \dim B/(x) + \dim A/\mathfrak{n}A = \dim A - 1$. By Exercise 26.2.D, A/xA is a Cohen-Macaulay ring, and x is a non-zerodivisor on A. The inductive hypothesis then applies to $(B/(x), \mathfrak{n}) \to (A/xA, \mathfrak{m})$, so A/xA is flat over $B/(x)$. Then by the local slicing criterion for flatness (Theorem 24.6.5), $B \to A$ is flat, as desired. $\square$

26.2.11. Miracle Flatness Theorem. — *Suppose $\pi : X \to Y$ is a morphism of equidimensional finite type k-schemes, where X is Cohen-Macaulay, Y is regular, and the fibers of π have dimension $\dim X - \dim Y$. Then π is flat.*

26.2.F. EXERCISE. Prove the Miracle Flatness Theorem 26.2.11. (Do not forget that schemes usually have non-closed points!)

The geometric situation in the Miracle Flatness Theorem 26.2.11 is part of the following pretty package.

26.2.G. EXERCISE. Suppose $\pi : X \to Y$ is a map of locally Noetherian schemes, where both X and Y are equidimensional, and Y is regular. Show that if any two of the following hold, then the third does as well:

(i) π is flat of relative dimension $\dim X - \dim Y$.
(ii) X is Cohen-Macaulay.
(iii) Every fiber X_y is Cohen-Macaulay of dimension $\dim X - \dim Y$.

Hint: if $\phi : B \to A$ is a flat ring map, then ϕ sends non-zerodivisors to non-zerodivisors (Observation 24.2.2).

The statement of Exercise 26.2.G can be improved, at the expense of killing the symmetry. In the implication (ii) and (iii) imply (i), the Cohen-Macaulay hypotheses in (iii) are not needed.

26.2.12. *Example.* As an example of Exercise 26.2.G in action, we consider the example of two planes meeting a point, continuing the notation of Exercise 26.1.G. Consider the morphism $\pi : X \to Y := \mathbb{A}^2$ given by $(w - y, x - z)$. Then Y is regular (by the Smoothness-Regularity Comparison Theorem 12.2.10(b)). But X is not Cohen-Macaulay (Exercise 26.1.G) and π is not flat (Exercise 24.4.J), so by Exercise 26.2.G, you can use either of these to prove the other.

26.2.13. Fancy properties of Cohen-Macaulayness.
We mention a few additional properties of Cohen-Macaulayness without proof. They are worth seeing, but we will not use them.

If X is a locally Noetherian scheme, the locus of Cohen-Macaulay points is open, see [**Mat2**, Thm 24.5] (and also [**Stacks**, tag 00RD]). In particular:

26.2.14. Fact (cf. Fact 12.8.5). — *Any localization of a Cohen-Macaulay Noetherian local ring at a prime ideal is also Cohen-Macaulay.*

(See [**Stacks**, tag 00NB], [**E**, Prop. 18.8], [**Mat2**, Thm. 17.3(iii)], or [**Mat1**, Thm. 30] for a direct proof.) In particular, for finite type k-schemes, Cohen-Macaulayness may be checked at closed points. The fact that Cohen-Macaulayness is preserved by localization can be used to quickly show that Cohen-Macaulay local rings are catenary (see [**E**, Cor. 18.10], [**Mat1**, Thm. 31(ii)], or [**Mat2**, Thm. 17.9]). A ring A is Cohen-Macaulay if and only if $A[x]$ is Cohen-Macaulay ([**E**, Prop 18.9], [**Mat2**, Thm 17.7], [**Mat1**, Thm. 33]), if and only if $A[[x]]$ is [**Mat2**, p. 137]. A local ring is Cohen-Macaulay if and only if its completion is Cohen-Macaulay ([**E**, Prop. 18.8], [**Mat2**, Thm 17.5]).

26.3 ⋆⋆ Serre's $R1 + S2$ criterion for normality

The notion of depth yields a useful criterion for normality, due to Serre.

26.3.1. *Definition.* Suppose A is a Noetherian ring, and $k \in \mathbb{Z}^{\geq 0}$. We say A has **property** R_k (A is **regular in codimension** $\leq k$, or more sloppily, A is **regular in codimension** k) if for every prime $\mathfrak{p} \subset A$ of codimension at most k, $A_\mathfrak{p}$ is regular.

(In light of Fact 12.8.2, a Noetherian local ring is regular and only if it has property R_k for all k.)

We say A has **property** S_k if for every prime $\mathfrak{p} \subset A$, the local ring $A_{\mathfrak{p}}$ has depth at least $\min(k, \dim A_{\mathfrak{p}})$ — "the local rings are Cohen-Macaulay up until codimension k, and have depth at least k thereafter". (In light of Fact 26.2.14, a Noetherian local ring is Cohen-Macaulay if and only if it has property S_k for all k.)

26.3.A. EASY EXERCISE. Note that a Noetherian ring A trivially always has property S_0. Show that a Noetherian ring A has property R_0 if $\operatorname{Spec} A$ is "generically reduced": it is reduced at the generic point of each of its irreducible components. Show that a Noetherian ring A has property S_1 if and only if $\operatorname{Spec} A$ has no embedded points. (Possible hint: Exercise 26.1.A.)

26.3.B. EXERCISE. Show that a Noetherian ring A has no nilpotents ($\operatorname{Spec} A$ is reduced) if and only if it has properties R_0 and S_1. (Hint: show that a Noetherian scheme is reduced if and only if each irreducible component is generically reduced, and it has no embedded points.)

Incrementing the subscripts in Exercise 26.3.B yields Serre's criterion.

26.3.2. Theorem (Serre's criterion for normality). — *A Noetherian ring A is normal if and only if it has properties R_1 and S_2.*

Thus failure of normality can have two possible causes: it can be failure of R_1 (something we already knew, from the equivalence of (a) and (g) in Theorem 12.5.8), or it can be the more subtle failure of S_2. Examples of varieties satisfying R_1 but not S_2 are given in Example 26.3.4 (two planes meeting at a point) and Exercise 26.3.D (the crumpled plane).

26.3.3. *Applications.* As usual, before giving a proof, we give some applications to motivate the result. First, it implies that Cohen-Macaulay schemes are normal if and only if they are regular in codimension 1. Thus checking normality of hypersurfaces (or more generally, regular embeddings) in $\mathbb{P}^n_k$ (or more generally, in any smooth variety), it suffices to check that their singular locus has codimension greater than one. (You should think through the details of why these statements are true.)

26.3.C. EXERCISE (PRACTICE WITH THE CONCEPT). Show that two-dimensional normal varieties are Cohen-Macaulay.

26.3.4. *Example: two planes meeting at a point.* The variety X of Exercise 26.1.G (two planes meeting at a point) is not normal (why?), but it is regular away from the origin. This implies that X does not have property S_2 (and hence is not Cohen-Macaulay), without the algebraic manipulations of Exercise 26.1.G.

We already knew that this example was not Cohen-Macaulay, but the same idea can show that the crumpled plane $\operatorname{Spec} k[x^3, x^2, xy, y]$ (appearing in Exercise 12.5.I) is not Cohen-Macaulay. Because of the "extrinsic" description of the ring, it is difficult to do this in another way.

26.3.D. EXERCISE: THE CRUMPLED PLANE IS NOT COHEN-MACAULAY. Let A be the subring $k[x^3, x^2, xy, y]$ of $k[x, y]$. Show that A is not Cohen-Macaulay. Hint: Exercise 12.5.I showed that A is not integrally closed.

26.3.5. *Regular local rings are integrally closed.* Serre's criterion can be used to show that regular local rings are integrally closed without going through the hard Fact 12.8.5 that they are unique factorization domains. Regular local rings are Cohen-Macaulay (§26.2.5), and (by Exercise 26.3.B) regular in codimension 0 as they are integral domains (Theorem 12.2.13), so we need only show that regular local rings are regular in codimension 1. We can invoke a different hard Fact 12.8.2 that localizations of regular local rings are again regular, but at least we have shown this for localizations of finitely generated algebras over a perfect field (see Theorem 12.8.3 and Exercise 21.3.F).

26.3.6. *Caution.* As is made clear by the following exercise, the condition S_2 is a condition on *all* primes, not just those of codimension at most 2.

26.3.E. EXERCISE. Give an example of a variety satisfying R_1, and Cohen-Macaulay at all points of codimension at most 2, which is not normal.

26.3.7. Proof of Theorem 26.3.2. The proof of Serre's criterion will take us until the end of this section.

26.3.8. *Normal implies* $R_1 + S_2$.

Suppose first that A is normal (and thus that all of its localizations are normal, by Exercise 5.4.A). It is reduced, and thus satisfies (R_0 and) S_1 by Exercise 26.3.B. It satisfies R_1 from the equivalence of (a) and (g) in Theorem 12.5.8. All that is left is to verify property S_2. We thus must show that if $(A, \mathfrak{m})$ is a normal Noetherian local ring of dimension greater than 1, then A has depth at least 2.

Choose any nonzero $x \in \mathfrak{m}$. (Don't forget that normality implies that the local ring A is an integral domain.) We wish to show that $\operatorname{depth} A/(x) > 0$. Assume otherwise that $\operatorname{depth} A/(x) = 0$, so (by Exercise 26.1.A) $\mathfrak{m}$ is an associated prime of $A/(x)$. Thus there is some nonzero $y \in A/(x)$ (i.e., $y \in (x)$) with $y\mathfrak{m} = 0$ in $A/(x)$, i.e., *multiplying any element* $z \in \mathfrak{m}$ *by* y *gives you a multiple of* x. Let $w = y/x \in K(A)$. We will show that $w \in A$, and thus that $y \in (x)$, yielding a contradiction.

For each prime ideal $\mathfrak{p} \subset A$ of codimension 1, choose any $z \in \mathfrak{m} \setminus \mathfrak{p}$. Then $zw = zy/x \in (y/x)\mathfrak{m} \subset A$ by the italicized statement in the previous paragraph, so $w \in A_{\mathfrak{p}}$. Thus $w \in A_{\mathfrak{p}}$ for all codimension 1 prime ideals $\mathfrak{p} \subset A$, so by Algebraic Hartogs's Lemma 11.3.10 (using the hypothesis that A is integrally closed), $w \in A$ as desired.

26.3.9. $R_1 + S_2$ *implies normal: the integral domain case.* Normality is a stalk-local condition (Proposition 5.4.2), so it suffices to show:

(†) if $(A, \mathfrak{m})$ is a dimension d Noetherian local ring satisfying R_1 and S_2 then it is integrally closed.

(As A satisfies R_0 and S_1, A is reduced by Exercise 26.3.B.)

The case in which A is an integral domain is notably easier, and may help motivate the statement of the criterion, so we deal with this case first, and leave the general case as something to read when you have too much time on your hands.

Note that this case suffices for many of the consequences we discuss, notably the normality of regular schemes (§26.3.5).

Suppose now that A is an integral domain. Suppose $x \in K(A)$ is integral over A. We must prove that $x \in A$. Write $x = f/g$, where $f, g \in A$. We wish to show that $f \in (g)$, or equivalently that f restricts to 0 on the closed subscheme $W := V(g)$ of $\operatorname{Spec} A$.

26.3.F. EXERCISE. Show that $W := V(g)$ has no embedded points. Translation: $A/(g)$ satisfies S_1.

Let $\eta_1, \dots, \eta_m$ be the generic points of the irreducible components of W, corresponding to primes $\mathfrak{p}_1, \dots, \mathfrak{p}_m$ of A. By Exercise 26.3.F, $\eta_1, \dots, \eta_m$ are *all* the associated points of W, so the natural map

$$A/(g) \to \prod_{i=1}^{m} \mathscr{O}_{W,\eta_i} = \prod_{i=1}^{m} (A/(g))_{\mathfrak{p}_i} \tag{26.3.9.1}$$

is an injection, by Theorem 5.5.10(b).

The ring $A_{\mathfrak{p}_i}$ is regular in codimension 1, so by parts (a) and (g) of Theorem 12.5.8, $A_{\mathfrak{p}_i}$ is integrally closed. Now w is integral over A and thus over $A_{\mathfrak{p}_i}$, so by the integral closure of $A_{\mathfrak{p}_i}$, $w \in A_{\mathfrak{p}_i}$. Hence f is 0 in each $A_{\mathfrak{p}_i}/(g) = (A/(g))_{\mathfrak{p}_i}$. From the injectivity of (26.3.9.1), we have that f is 0 on W as desired.

26.3.10. $\star\star$ $R_1 + S_2$ *implies normal: the general case.* We prove (†) by induction on d.

26.3.G. EXERCISE. Prove the desired result if $d = 0$ and $d = 1$.

Our next step is to show that A is integrally closed in its total fraction ring, the localization of A at the multiplicative subset of non-zerodivisors (defined in §5.5.7). Suppose we have $f, g \in A$, with g a non-zerodivisor, such that f/g satisfies the monic equation

$$(f/g)^n + \sum_{i=1}^{n} a_i (f/g)^{n-i} = 0 \tag{26.3.10.1}$$

with $a_i \in A$. We prove $f \in (g)$ by induction on d. The cases $d = 0$ and 1 follow from Exercise 26.3.G. Assume now that $d > 1$, and that the result is known for "smaller d".

As $\operatorname{depth} A \geq 2$, $\operatorname{depth} A/(g) \geq 1$, so there is a non-zerodivisor t on $A/(g)$ with $t \in \mathfrak{m}$. Consider the localization of (26.3.10.1) at all the primes corresponding to points of $D(t)$ (those primes not containing t), each of which has codimension $< d$. By the inductive hypothesis, $f \in (g)$ in each of these localizations, so f is zero in $(A/(g))_t$. But as t is a non-zerodivisor on $A/(g)$, the map $A/(g) \to (A/(g))_t$ is an injection (Exercise 1.3.C), so f is zero in $A/(g)$, as desired. We have thus shown that A is integrally closed in its total fraction ring.

26.3.11. Lemma. — *Suppose R is a reduced ring with finitely many minimal primes $\mathfrak{p}_1, \dots, \mathfrak{p}_n$. If R is integrally closed in its total fraction ring, then R is a finite product of integrally closed integral domains.*

26.3.12. *Proof of Lemma 26.3.11.* The $\mathfrak{p}_i$ are the associated primes of R, by Exercise 5.5.C, so the natural map $\phi : R \to \prod_{i=1}^n R_{\mathfrak{p}_i}$ is an inclusion (by Theorem 5.5.10(b)).

26.3.H. EXERCISE. Show that the map ϕ identifies the total fraction ring of R with $\prod_{i=1}^n R_{\mathfrak{p}_i}$.

Suppose e_i is the ith idempotent of the total fraction ring $\prod R_{\mathfrak{p}_i}$ (i.e., $(0, \dots, 0, 1, 0, \dots, 0)$, where the 1 is in the ith spot). Since R is integrally closed in its total fraction ring, it contains each of the e_i.

26.3.I. EXERCISE. Conclude the proof of Lemma 26.3.11, by describing $R = \prod_{i=1}^n Re_i$, and showing that each Re_i is an integrally closed integral domain. Possible hint: Remark 3.6.3. □

We now return to our proof of Serre's criterion. By Lemma 26.3.11, we now know that A is a product of integrally closed integral domains. But $\operatorname{Spec} A$ is connected (A is a local ring!), so A is an integral domain. □

CHAPTER 27

Twenty-seven lines

27.0.13. Why this is Chapter 27. This topic is placed in Chapter 27 for purely numerological reasons. As a result, it invokes one fact that we will only meet in a later chapter: Castelnuovo's Criterion (Theorem 29.7.1) is used in the proof of Proposition 27.4.1. This is a feature rather than a bug: you needn't read the details of the later chapters in order to enjoy this one. In fact, you should probably read this many chapters earlier, taking key facts as "black boxes", to appreciate the value of seemingly technical definitions and theorems before you learn about them in detail. This topic is a fitting conclusion to an energetic introduction to algebraic geometry, even if you do not earlier cover everything topic needed to justify all details.

27.1 Introduction

[picture to be made or found later]

FIGURE 27.1. Twenty-seven lines on a cubic surface

Wake an algebraic geometer in the dead of night, whispering: "27". Chances are, he will respond: "lines on a cubic surface".
— R. Donagi and R. Smith, [**DS**] *(on page 27, of course)*

Since the middle of the nineteenth century, geometers have been entranced by the fact that there are 27 lines on every smooth cubic surface, and by the remarkable structure of the conifiguration of the lines. Their discovery by Cayley and Salmon in 1849 has been called the beginning of modern algebraic geometry, [**Do**, p. 55].

The reason so many people are bewitched by this fact is because it requires some magic, and this magic connects to many other things, including fundamental ideas we have discussed, other beautiful classical constructions (such as Pascal's Mystical Hexagon Theorem 19.4.4, the fact that most smooth quartic plane curves have 28 bitangents, exceptional Lie groups, ...), and many themes in modern algebraic geometry (deformation theory, intersection theory, enumerative geometry, arithmetic and diophantine questions, ...). It will be particularly pleasant for us, as it takes advantage of many of the things we have learned.

You are now ready to be initiated into the secret fellowship of the twenty-seven lines.

27.1.1. Theorem. — *Every smooth cubic surface in $\mathbb{P}^3_{\overline{k}}$ has exactly 27 lines.*

Theorem 27.1.1 is closely related to the following.

27.1.2. Theorem. — *Every smooth cubic surface over $\overline{k}$ is isomorphic to $\mathbb{P}^2$ blown up at 6 points.*

There are many reasons why people consider these facts magical. First, there is the fact that there are *always* 27 lines. Unlike most questions in enumerative geometry, there are no weasel words such as "a general cubic surface" or "most cubic surfaces" or "counted correctly" — as in, "every monic degree d polynomial has d roots — counted correctly". And somehow (and we will see how) it is precisely the smoothness of the surface that makes it work.

Second, there is the magic that you *always* get the blow-up of the plane at six points (§27.4).

Third, there is the magical incidence structure of the 27 lines, which relates to E_6 in Lie theory. The Weyl group of E_6 is the symmetry group of the incidence structure (see Remark 27.3.5). In a natural way, the 27 lines form a basis of the 27-dimensional fundamental representation of E_6.

27.1.3. Structure of this chapter. Throughout this chapter, X will be a smooth cubic surface over an algebraically closed field $\overline{k}$. In §27.2, we establish some preliminary facts. In §27.3, we prove Theorem 27.1.1. In §27.4, we prove Theorem 27.1.2. We remark here that the only input that §27.4 needs from §27.3 is Exercise 27.3.J. This can be done directly by hand (see for example [**Rei**, §7], [**Be**, Thm. IV.13], or [**Sh**, p. 246-7]), and Theorem 27.1.2 readily implies Theorem 27.1.1, using Exercise 27.4.E. We would thus have another, shorter, proof of Theorem 27.1.1. The reason for giving the argument of §27.3 is that it is natural given what we have done so far, it gives you some glimpse of some ideas used more broadly in the subject (the key idea is that a map from one moduli space to another is finite and flat), and it may help you further appreciate and digest the tools we have developed.

27.2 Preliminary facts

By Theorem 14.1.C, there is a 20-dimensional vector space of cubic forms in four variables, so the cubic surfaces in $\mathbb{P}^3$ are parametrized by $\mathbb{P}^{19}$.

27.2.A. EXERCISE. Show that there is an irreducible hypersurface $\Delta \subset \mathbb{P}^{19}$ whose closed points correspond precisely to the singular cubic surfaces over $\overline{k}$. Hint: construct an incidence correspondence $Y \subset \mathbb{P}^{19} \times \mathbb{P}^3$ corresponding to a cubic surface X, along with a singular point of X. Show that Z is a $\mathbb{P}^{15}$-bundle over $\mathbb{P}^3$, and thus irreducible of dimension 18. To show that its image in $\mathbb{P}^{19}$ is "full dimensional" (dimension 18), use Exercise 11.4.A or Proposition 11.4.1, and find a cubic surface singular at precisely one point.

27.2.B. EXERCISE. Show that the any smooth cubic surface X is "anticanonically embedded" — it is embedded by the anticanonical linear series $\mathscr{K}_X^\vee$. Hint: the adjunction formula, Exercise 21.5.B.

27.2.C. EXERCISE. Suppose $X \subset \mathbb{P}^3_k$ is a smooth cubic surface. Suppose C is a curve on X. Show that C is a line if and only if C is a (-1)-curve (Definition 20.2.6). Hint: the adjunction formula again, perhaps in the guise of Exercise 20.2.B(a); and also Exercise 18.6.L.

It will be useful to find a *single* cubic surface with exactly 27 lines:

27.2.D. EXERCISE. Show that the *Fermat cubic surface*

$$x_0^3 + x_1^3 + x_2^3 + x_3^3 = 0 \tag{27.2.0.1}$$

in $\mathbb{P}^3_{\mathbb{C}}$ has precisely 27 lines, each of the form

$$x_0 + \omega x_i = x_j + \omega' x_k = 0,$$

where $\{1, 2, 3\} = \{i, j, k\}$, $j < k$, and ω and ω' are cube roots of -1 (possibly the same). Hint: up to a permutation of coordinates, show that every line in $\mathbb{P}^3$ can be written $x_0 = ax_2 + bx_3$, $x_1 = cx_2 + dx_3$. Show that this line is on (27.2.0.1) if and only if

$$a^3 + c^3 + 1 = b^3 + d^3 + 1 = a^2b + c^2d = ab^2 + cd^2 = 0 \tag{27.2.0.2}$$

Show that if a, b, c, and d are all nonzero, then (27.2.0.2) has no solutions.

27.2.1. $\star$ **The singular cubic surfaces give an irreducible divisor in the parameter space $\mathbb{P}^{19}$ of all cubic surfaces.** This discussion is starred not because it is hard, but because it is not needed in the rest of the chapter. Nonetheless, it is a pretty application of what we have learned, and it foreshadows key parts of the argument in §27.3.

Fix a field k. Bertini's Theorem 12.4.2 shows that the locus of smooth cubics in the parameter space $\mathbb{P}^{19}$ of all cubics is dense and open. The proof suggests more: that the complement Δ of this locus is of pure codimension 1. We now show that this is the case, and even that this "discriminant hypersurface" Δ is also irreducible.

27.2.E. EXERCISE. (Hint for both: recall the solution to Exercise 11.2.J.)
(a) Define the incidence correspondence $Y \subset \mathbb{P}^{19} \times \mathbb{P}^3$ corresponding to the data of a cubic surface X, along with a *singular* point $p \in X$. (This is part of the problem! We need Z as a scheme, not just as a set.) Let μ be the projection $Y \to \mathbb{P}^{19}$.

$$\begin{tikzcd} & Y \arrow[dl, "\mu"'] \arrow[dr, "\mathbb{P}^{15}\text{-bundle}"] & \\ \mathbb{P}^{19} & & \mathbb{P}^3 \end{tikzcd}$$

(b) Show that Y is an irreducible smooth variety of dimension 18, by describing it as a $\mathbb{P}^{15}$-bundle over $\mathbb{P}^3$.

27.2.F. EXERCISE. Show that there exists a cubic surface with a single singular point. Feel free to assume k is your favorite field; the main point is to familiarize yourself with geometric ideas, not the peculiarities of positive characteristic fields. (Hint if char $k \neq 2$: Exercise 12.3.C.)

27.2.G. EXERCISE. Show that $\mu(Y) = \Delta$ is a closed irreducible subset of codimension 1. Hint for the codimension 1 statement: Exercise 11.4.A or Proposition 11.4.1, in combination with Exercise 27.2.F.

Your argument will generalize with essentially no change to deal with degree d hypersurfaces in $\mathbb{P}^n$.

27.3 Every smooth cubic surface (over $\overline{k}$) has 27 lines

We are now ready to prove Theorem 27.1.1. Until Exercise 27.3.K, to avoid distraction, we assume $\operatorname{char} \overline{k} = 0$. (However, the following argument carries through without change if $\operatorname{char} k \neq 3$. The one required check — and the reason for the restriction on the characteristic — is that Exercise 27.2.D works with $\mathbb{C}$ replaced by any such $\overline{k}$.)

27.3.A. EXERCISE. (Hint for both: recall the solution to Exercise 11.2.J.)
(a) Define the incidence correspondence $Z \subset \mathbb{P}^{19} \times \mathbb{G}(1,3)$ corresponding to the data of a line ℓ in $\mathbb{P}^3$ contained in a cubic surface X. (As in Exercise 27.2.E, this is part of the problem.) Let π be the projection $Z \to \mathbb{P}^{19}$.

$$\begin{tikzcd} & Z \arrow[dl, "\pi"'] \arrow[dr, "\mathbb{P}^{15}\text{-bundle}"] & \\ \mathbb{P}^{19} & & \mathbb{G}(1,3) \end{tikzcd}$$

(b) Show that Z is an irreducible smooth variety of dimension 19, by describing it as a $\mathbb{P}^{15}$-bundle over $\mathbb{G}(1,3)$.

27.3.B. EXERCISE. Use the fact that there exists a cubic surface with a finite number of lines (Exercise 27.2.D), and the behavior of dimensions of fibers of morphisms (Exercise 11.4.A or Proposition 11.4.1) to show the following.

(a) Every cubic surface contains a line, i.e., π is surjective. (Hint: show that π is projective.)
(b) "Most cubic surfaces have a finite number of lines": there is a dense open subset $U \subset \mathbb{P}^{19}$ such that the cubic surfaces parametrized by closed points of U have a positive finite number of lines. (Hint: upper semicontinuity of fiber dimension, Theorem 11.4.2.)

The following fact is the key result in the proof of Theorem 27.1.1, and one of the main miracles of the 27 lines, that ensures that the lines stay distinct on a smooth surface. It states, informally, that two lines can't come together without damaging the surface. This is really a result in deformation theory: we are explicitly showing that a line in a smooth cubic surface has no first-order deformations.

27.3.1. Theorem. — *If ℓ is a line in a regular cubic surface X, then $\{\ell \subset X\}$ is a reduced point of the fiber of π.*

Before proving Theorem 27.3.1, we use it to prove Theorem 27.1.1.

27.3.2. *Proof of Theorem 27.1.1.* Now π is a projective morphism, and over $\mathbb{P}^{19} \setminus \Delta$, π has dimension 0, and hence has finite fibers. Hence by Theorem 18.1.9, π is finite over $\mathbb{P}^{19} \setminus \Delta$.

Furthermore, as Z is regular (hence Cohen-Macaulay, §26.2.5) and $\mathbb{P}^{19}$ is regular, the Miracle Flatness Theorem 26.2.11 implies that π is flat over $\mathbb{P}^{19} \setminus \Delta$.

Thus, over $\mathbb{P}^{19} \setminus \Delta$, π is a finite flat morphism, and so the fibers of π (again, away from Δ) always have the same number of points, "counted correctly" (Exercise 24.4.G). But by Theorem 27.3.1, above each closed point of $\mathbb{P}^{19} \setminus \Delta$, each point of the fiber of π counts with multiplicity one. Finally, by Exercise 27.2.D, the Fermat cubic gives an example of one regular cubic surface with precisely 27 lines, so (as $\mathbb{P}^{19} \setminus \Delta$ is connected) we are done. $\square$

We have actually shown that away from Δ, $Z \to \mathbb{P}^{19}$ is a finite étale morphism of degree 27.

27.3.3. $\star$ *Proof of Theorem 27.3.1.* Choose projective coordinates so that the line ℓ is given, in a distinguished affine subset (with coordinates named x, y, z), by the z-axis. (We use affine coordinates to help visualize what we are doing, although this argument is better done in projective coordinates. On a second reading, you should translate this to a fully projective argument.)

27.3.C. EXERCISE. Consider the lines of the form $(x, y, z) = (a, b, 0) + t(a', b', 1)$ (where $(a, b, a', b') \in \mathbb{A}^4$ is fixed, and t varies in $\mathbb{A}^1$). Show that a, b, a', b' can be interpreted as the "usual" coordinates on one of the standard open subsets of the Grassmannian (see §6.7), with $[\ell]$ as the origin.

[picture to be made later: plane $z = 0$, with (a, b) marked; and plane $z = 1$, with $(a + a', b + b')$ marked]

FIGURE 27.2. Parameters for the space of "lines near ℓ", in terms of where they meet the $z = 0$ plane and the $z = 1$ plane

Having set up local coordinates on the moduli space, we can get down to business. Suppose $f(x, y, z)$ is the (affine version) of the equation for the cubic surface X. Because X contains the z-axis ℓ, $f(x, y, z) \in (x, y)$. More generally, the line

$$(x, y, z) = (a, b, 0) + t(a', b', 1) \qquad (27.3.3.1)$$

lies in X precisely when $f(a + ta', b + tb', t)$ is 0 as a cubic polynomial in t. This is equivalent to four equations in a, a', b, and b', corresponding to the coefficients of t^3, t^2, t, and 1. This is better than just a set-theoretic statement:

27.3.D. EXERCISE. Verify that these four equations are local equations for the scheme-theoretic fiber $\pi^{-1}([X])$.

Now we come to the crux of the argument, where we use the regularity of X (along ℓ). We have a specific question in algebra. We have a cubic surface X given by $f = 0$, containing ℓ, and we know that X is regular (including "at ∞", i.e., in $\mathbb{P}^3$). To show that $[\ell] = V(a, a', b, b')$ is a reduced point in the fiber, we work in the ring $\overline{k}[a, a', b, b']/(a, a', b, b')^2$, i.e., we impose the equations

$$a^2 = aa' = \cdots = (b')^2 = 0, \tag{27.3.3.2}$$

and try to show that $a = a' = b = b' = 0$. (It is essential that you understand why we are setting $(a, a', b, b')^2 = 0$. You can also interpret this argument in terms of the derivatives of the functions involved — which after all can be interpreted as forgetting higher-order information and remembering only linear terms in the relevant variables, cf. Exercise 12.1.G. See [**Mu6**, §8D] for a description of this calculation in terms of derivatives.)

Suppose $f(x, y, z) = c_{x^3}x^3 + c_{x^2y}x^2y + \cdots + c_1 1 = 0$, where $c_{x^3}, c_{x^2y}, \cdots \in \overline{k}$. Because $\ell \in X$, i.e., $f \in (x, y)$, we have $c_1 = c_z = c_{z^2} = c_{z^3} = 0$. We now substitute (27.3.3.1) into f, and then apply (27.3.3.2). Only the coefficients of f of monomials involving precisely one x or y survive:

$$\begin{aligned} & c_x(a + a't) + c_{xz}(a + a't)t + c_{xz^2}(a + a't)t^2 \\ & +c_y(b + b't) + c_{yz}(b + b't)t + c_{yz^2}(b + b't)t^2 \\ = \quad & (a + a't)(c_x + c_{xz}t + c_{xz^2}t^2) + (b + b't)(c_y + c_{yz}t + c_{yz^2}t^2) \end{aligned}$$

is required to be 0 as a polynomial in t. (Recall that $c_x, \dots, c_{yz^2}$ are fixed elements of $\overline{k}$.) Let $C_x(t) = c_x + c_{xz}t + c_{xz^2}t^2$ and $C_y(t) = c_y + c_{yz}t + c_{yz^2}t^2$ for convenience.

Now X is regular at $(0, 0, 0)$ precisely when c_x and c_y are not both 0 (as $c_z = 0$). More generally, X is regular at $(0, 0, t_0)$ precisely if $c_x + c_{xz}t_0 + c_{xz^2}t_0^2 = C_x(t_0)$ and $c_y + c_{yz}t_0 + c_{yz^2}t_0^2 = C_y(t_0)$ are not both zero. You should be able to quickly check that X is regular at the point of ℓ "at ∞" precisely if c_{xz^2} and c_{yz^2} are not both zero. We summarize this as follows: X is regular at every point of ℓ precisely if the two quadratics $C_x(t)$ and $C_y(t)$ have no common roots, including "at ∞".

We now use this to force $a = a' = b = b' = 0$ using $(a + a't)C_x(t) + (b + b't)C_y(t) \equiv 0$.

We deal first with the special case where C_x and C_y have two distinct roots, both finite (i.e., c_{xz^2} and c_{yz^2} are nonzero). If t_0 and t_1 are the roots of $C_x(t)$, then substituting t_0 and t_1 into $(a + a't)C_x(t) + (b + b't)C_y(t)$, we obtain $b + b't_0 = 0$, and $b + b't_1 = 0$, from which $b = b' = 0$. Similarly, $a = a' = 0$.

27.3.E. EXERCISE. Deal with the remaining cases to conclude the proof of Theorem 27.3.1. (It is possible to do this quite cleverly. For example, you may be able to re-choose coordinates to ensure that C_x and C_y have finite roots.)

□

27.3.4. The configuration of lines.

By the "configuration of lines" on a cubic surface, we mean the data of which pairs of the 27 lines intersect. We can readily work this out in the special case of the Fermat cubic surface (Exercise 27.2.D). (It can be more enlightening to use the description of X as a blow-up of $\mathbb{P}^2$, see Exercise 27.4.E.) We now show that the configuration is the "same" (interpreted appropriately) for all smooth cubic surfaces.

27.3.F. EXERCISE. Construct a degree 27! finite étale map $Y \to \mathbb{P}^{19} \setminus \Delta$, that parametrizes a cubic surface along with an *ordered list* of 27 distinct lines. Hint: let Y' be the 27th fibered power of Z over $\mathbb{P}^{19} \setminus \Delta$, interpreted as parametrizing a cubic surface with an ordered list of 27 lines, not necessarily distinct. Let Y be the subset corresponding to where the lines are distinct, and show that Y is open and closed in Y', and thus a union of connected components of Y'.

We now make sense of the statement of the fact that configuration of lines on the Fermat surface (call it X_0) is the "same" as the configuration on some other smooth cubic surface (call it X_1). Lift the point $[X_0]$ to a point $y_0 \in Y$. Let Y'' be the connected component of Y containing y_0.

27.3.G. EXERCISE. Show that $Y'' \to \mathbb{P}^{19} \setminus \Delta$ is finite étale.

Choose a point $y_1 \in Y''$ mapping to $[X_1]$. Because Y parametrizes a "labeling" or ordering of the 27 lines on a surface, we now have chosen an identification of the lines on X_0 with those of X_1. Let the lines be $\ell_1, \dots, \ell_{27}$ on X_0, and let the corresponding lines on X_1 be $m_1, \dots, m_{27}$.

27.3.H. EXERCISE (USING STARRED EXERCISE 24.7.D). Show that $\ell_i \cdot \ell_j = m_i \cdot m_j$ for all i and j.

27.3.I. EXERCISE. Show that for each smooth cubic surface $X \subset \mathbb{P}^3_{\overline{k}}$, each line on X meets exactly 10 other lines $\ell_1, \ell_1', \dots, \ell_5, \ell_5'$ on X, where ℓ_i and ℓ_i' meet for each i, and no other pair of the lines meet.

27.3.J. EXERCISE. Show that every smooth cubic surface contains two disjoint lines ℓ and ℓ', such that there are precisely five other lines $\ell_1, \dots, \ell_5$ meeting both ℓ and ℓ'.

27.3.5. *Remark: The Weyl group* $W(E_6)$. The symmetry group of the configuration of lines — i.e., the subgroup of the permutations of the 27 lines preserving the intersection data — magically turns out to be the Weyl group of E_6, a group of order 51840. (You know enough to at least verify that the size of the group is 51840, using the Fermat surface of Exercise 27.2.D, but this takes some work.) It is no coincidence that the degree of Y'' over $\mathbb{P}^{19} \setminus \Delta$ is 51840, and the Galois group of the Galois closure of $K(Z)/K(\mathbb{P}^{19} \setminus \Delta)$ is isomorphic to $W(E_6)$ (see [**H1**, III.3]).

27.3.K. ⋆⋆ EXERCISE. Prove Theorem 27.3.1 in arbitrary characteristic. Begin by figuring out the right statement of Exercise 27.3.A over $\mathbb{Z}$, and proving it. Then follow the argument given in this section, making changes when necessary.

27.3.6. ⋆ **Fano varieties of lines, and Hilbert schemes.** In Exercises 27.3.A and 27.3.K, you constructed a moduli space of lines contained in a X, as a scheme. Your argument can be generalized to any $X \subset \mathbb{P}^N$. This construction is called the *Fano variety of lines* of X, and is an example of a Hilbert scheme.

27.4 Every smooth cubic surface (over $\overline{k}$) is a blown up plane

We now prove Theorem 27.1.2. As stated in §27.1.3, this section is remarkably independent from the previous one; all we will need is Exercise 27.3.J, and it is possible to prove this in other ways.

Suppose X is a smooth cubic surface (over $\overline{k}$). Suppose ℓ is a line on X, and choose coordinates on the ambient $\mathbb{P}^3$ so that ℓ is cut out by x_0 and x_1. Projection from ℓ gives a rational map $\mathbb{P}^3 \dashrightarrow \mathbb{P}^1$ (given by $[x_0, x_1, x_2, x_3] \mapsto [x_0, x_1]$), which extends to a morphism on X. The reason is that this rational map is resolved by blowing up the closed subscheme $V(x_0, x_1)$ (Exercise 22.4.L). But (x_0, x_1) cuts out the Cartier divisor ℓ on X, and blowing up a Cartier divisor does not change X (Observation 22.2.1).

Now choose two disjoint lines ℓ and ℓ' as in Exercise 27.3.J, and consider the morphism $\rho : X \to \mathbb{P}^1 \times \mathbb{P}^1$, where the map to the first $\mathbb{P}^1$ is projection from ℓ, and the map to the second $\mathbb{P}^1$ is the projection from ℓ'. The first $\mathbb{P}^1$ can then be identified with ℓ', and the second with ℓ.

27.4.A. EXERCISE. Show that the morphism ρ is birational. Hint: given a general point of $(p, q) \in \ell' \times \ell$, we obtain a point of X as follows: the line pq in $\mathbb{P}^3$ meets the cubic X at three points by Bézout's theorem 8.2.E: p, q, and some third point $x \in X$; send (p, q) to x. (This idea appeared earlier in the development of the group law on the cubic curve, see Proposition 19.10.3.) Given a general point $x \in X$, we obtain a point $(p, q) \in \ell' \times \ell$ by projecting from ℓ' and ℓ.

In particular, we have shown for the first time that every smooth cubic surface over $\overline{k}$ is rational.

27.4.B. EXERCISE: ρ CONTRACTS PRECISELY $\ell_1, \ldots, \ell_5$. Show that ρ is an isomorphism away from the ℓ_i mentioned in Exercise 27.3.J, and that each $\rho(\ell_i)$ is a point $p_i \in \ell' \times \ell$.

27.4.1. Proposition. — *The morphism $\rho : X \to \mathbb{P}^1 \times \mathbb{P}^1$ is the blow-up at $\mathbb{P}^1 \times \mathbb{P}^1$ at the five p_i.*

Proof. By Castelnuovo's Criterion (Theorem 29.7.1), as the lines ℓ_i are (-1)-curves (Exercise 27.2.C), they can be contracted. More precisely, there is a morphism $\beta : X \to X'$ that is the blow-up of X' at five closed points $p'_1, \ldots, p'_5$, such that ℓ_i is the exceptional divisor at p'_i. We wish to show that X' is $\mathbb{P}^1 \times \mathbb{P}^1$.

The morphism $\rho : X \to \mathbb{P}^1 \times \mathbb{P}^1$ yields a morphism $\rho' : X' \setminus \{p'_1, \ldots p'_5\} \to \mathbb{P}^1 \times \mathbb{P}^1$. We now show that ρ' extends over p'_i for each i, sending p'_i to p_i. Choose a neighborhood of $p_i \in \mathbb{P}^1 \times \mathbb{P}^1$ isomorphic to $\mathbb{A}^2$, with coordinates x and y. Then both x and y pull back to functions on a punctured neighborhoods of p'_i (i.e., there is some open neighborhood U of p'_i such that x and y are functions on $U \setminus \{p'_i\}$). By Algebraic Hartogs's Lemma 11.3.10, they extend over p'_i, and this extension is unique as $\mathbb{P}^1 \times \mathbb{P}^1$ is separated — use the Reduced-to-Separated Theorem 10.2.2 if you really need to. Thus ρ' extends over p'_i. (Do you see why $\rho'(p'_i) = p_i$?)

27.4.C. EXERCISE. Show that the birational morphism $\rho' : X' \to \mathbb{P}^1 \times \mathbb{P}^1$ is invertible. Hint: you can use Zariski's Main Theorem (in the guise of Exercise 29.6.D), but you needn't use something so powerful. Instead, note that the birational map ρ'^{-1} is a morphism away from $p_1, \ldots, p_5$. Use essentially the same argument as in the last paragraph to extend ρ'^{-1} over each p_i.

□

As a consequence we see that X is the blow-up of $\mathbb{P}^1 \times \mathbb{P}^1$ at 5 points. Because the blow-up of $\mathbb{P}^1 \times \mathbb{P}^1$ at one point is isomorphic to the blow-up of $\mathbb{P}^2$ at two points (Exercise 22.4.K), Theorem 27.1.2 then follows. □

27.4.2. Reversing the process.

(This is a more precise version of 22.4.10(iii).) The process can be reversed: we can blow-up $\mathbb{P}^2$ at six points, and embed it in $\mathbb{P}^3$. We first explain why we can't blow up $\mathbb{P}^2$ at just any six points and hope to embed the result in $\mathbb{P}^3$. Because the cubic surface is embedded anticanonically (Exercise 27.2.B), we see that any curve C in X must satisfy $(\mathscr{K}_X \cdot C) < 0$.

27.4.D. EXERCISE. Suppose $\mathbb{P}^2$ is sequentially blown up at $p_1, \dots, p_6$, resulting in smooth surface X.
(a) If p_i lies on the exceptional divisor of the blow-up at p_j $(i > j)$, then show that there is a curve $C \subset X$ isomorphic to $\mathbb{P}^1$, with $(\mathscr{K}_X \cdot C) \geq 0$.
(b) If the p_i are distinct points on $\mathbb{P}^2$, and three of them are collinear, show that there is a curve $C \subset X$ isomorphic to $\mathbb{P}^1$, with $(\mathscr{K}_X \cdot C) \geq 0$.
(c) If the six p_i are distinct points on a smooth conic, show that there is a curve $C \subset X$ isomorphic to $\mathbb{P}^1$, with $(\mathscr{K}_X \cdot C) \geq 0$.

Thus the only chance we have of obtaining a smooth cubic surface by blowing up six points on $\mathbb{P}^2$ is by blowing up six distinct points, no three on a line and not all on a conic.

27.4.3. Proposition. — *The anticanonical map of $\mathbb{P}^2$ blown up at six distinct points, no three on a line and not all on a conic, gives a closed embedding into $\mathbb{P}^3$, as a cubic surface.*

Because we won't use this, we only describe the main steps of the proof: first count sections of the *anticanonical bundle* $\mathscr{K}_{\mathbb{P}^2}^{\vee} \cong \mathscr{O}_{\mathbb{P}^2}(3)$ (there is a 4-dimensional vector space of cubics on $\mathbb{P}^2$ vanishing at the six points, and these correspond to sections of the anticanonical bundle of the blow-up via Exercise 22.4.S(a)). Then show that these sections separate points and tangent vectors of X, thus showing that the anticanonical linear series gives a closed embedding, Theorem 19.1.1. Judicious use of the Cremona transformation (Exercise 6.5.I) can reduce the amount of tedious case-checking in this step.

27.4.E. EXERCISE. Suppose X is the blow-up of $\mathbb{P}^2_k$ at six distinct points $p_1, \dots, p_6$, no three on a line and not all on a conic. Verify that the only (-1)-curves on X are the six exceptional divisors, the proper transforms of the 10 lines p_ip_j, and the proper transforms of the six conics through five of the six points, for a total of 27.

27.4.F. EXERCISE. Solve Exercises 27.3.I and 27.3.J again, this time using the description of X as a blow-up of $\mathbb{P}^2$.

27.4.4. *Remark.* If you blow-up $4 \leq n \leq 8$ points on $\mathbb{P}^2$, with no three on a line and no six on a conic, then the symmetry group of the configuration of lines is a Weyl group, as shown in the following table.

n	4	5	6	7	8
	$W(A_4)$	$W(D_5)$	$W(E_6)$	$W(E_7)$	$W(E_8)$

(If you know about Dynkin diagrams, you may see the pattern, and may be able to interpret what happens for $n = 3$ and $n = 9$.) This generalizes part of Remark 27.3.5, and the rest of it can similarly be generalized.

CHAPTER 28

Cohomology and base change theorems

28.1 Statements and applications

Higher pushforwards are easy to define, but it is hard to get a geometric sense of what they are, or how they behave. For example, given a reasonable morphism $\pi : X \to Y$, and a quasicoherent sheaf on $\mathscr{F}$, you might reasonably hope that the fibers of $R^i\pi_*\mathscr{F}$ are the cohomologies of $\mathscr{F}$ along the fibers. More precisely, given $\psi : q \to Y$ corresponding to the inclusion of a point (better: $\psi : \operatorname{Spec} \kappa(q) \to Y$), yielding the fibered diagram

(28.1.0.1)
$$\begin{tikzcd} X_q \arrow[r, "\psi'"] \arrow[d, "\pi'"'] & X \arrow[d, "\pi"] \\ q \arrow[r, "\psi"] & Y, \end{tikzcd}$$

one might hope that the morphism

$$\boxed{\phi_q^p : \psi^*(R^p\pi_*\mathscr{F}) \to H^p(X_q, (\psi')^*\mathscr{F})}$$

(given in Exercise 18.8.B(a)) is an isomorphism. (Note: $\mathscr{F}|_{X_q}$ and $(\psi')^*\mathscr{F}$ are symbols for the same thing. The first is often preferred, but we sometimes use the second because we will consider more general ψ and ψ'.) We could then picture $R^i\pi_*\mathscr{F}$ as somehow fitting together the cohomology groups of fibers into a coherent sheaf. (Warning: this is too much to hope for, see Exercise 28.1.A.)

It would also be nice if $H^p(X_q, (\psi')^*\mathscr{F})$ was constant, and ϕ_q^p put them together into a nice locally free sheaf (vector bundle) $\psi^*(R^p\pi_*\mathscr{F})$.

There is no reason to imagine that the particular choice of base change $\psi : q \mapsto Y$ should be special. As long as we are dreaming, we may as well hope that in good circumstances, given a fiber diagram (18.8.4.1)

(28.1.0.2)
$$\begin{tikzcd} W \arrow[r, "\psi'"] \arrow[d, "\pi'"'] & X \arrow[d, "\pi"] \\ Z \arrow[r, "\psi"] & Y, \end{tikzcd}$$

the natural morphism

(28.1.0.3)
$$\boxed{\phi_Z^p : \psi^*(R^p\pi_*\mathscr{F}) \to R^p\pi'_*(\psi')^*\mathscr{F}}$$

of sheaves on Z (Exercise 18.8.B(a)) is an isomorphism. (In some cases, we can already address this question. For example, cohomology commutes with flat base change, Theorem 24.2.8, so the result holds if ψ is flat. Also related: if $\mathscr{F}$ is flat over Y, then the Euler characteristic of $\mathscr{F}$ on fibers is locally constant, Theorem 24.7.1.)

There is no point in dreaming if we are not going to try to make our dreams come true. So let's formalize them. Suppose $\mathscr{F}$ is a coherent sheaf on X, $\pi : X \to Y$ is projective, Y (hence X) is Noetherian, and $\mathscr{F}$ is flat over Y. We formalize our dreams into three nice properties that we might wish in this situation. We will see that they are closely related.

(a) Given a fibered square (28.1.0.1), is $\phi^p_q : R^p\pi_*\mathscr{F} \otimes \kappa(q) \to H^p(X_q, \mathscr{F}|_{X_q})$ an isomorphism?
(b) Given a fibered square (28.1.0.2), is $\phi^p_Z : \psi^*(R^p\pi_*\mathscr{F}) \to R^p\pi'_*(\psi')^*\mathscr{F}$ an isomorphism?
(c) Is $R^p\pi_*\mathscr{F}$ locally free?

We turn first to property (a). The dimension of the left side $R^p\pi_*\mathscr{F} \otimes \kappa(q)$ is an upper semicontinuous function of $q \in Y$ by upper semicontinuity of rank of finite type quasicoherent sheaves (Exercise 13.7.J). The Semicontinuity Theorem states that the dimension of the right is also upper semicontinuous. More formally:

28.1.1. Semicontinuity Theorem. — *Suppose $\pi : X \to Y$ is a proper morphism of Noetherian schemes, and $\mathscr{F}$ is a coherent sheaf on X flat over Y. Then for each $p \geq 0$, the function $Y \to \mathbb{Z}$ given by $q \mapsto \dim_{\kappa(q)} H^p(X_q, \mathscr{F}|_{X_q})$ is an upper semicontinuous function of $q \in Y$.*

Translation: cohomology groups are upper semicontinuous in proper flat families. (A proof will be given in §28.2.4.)

28.1.2. *Example.* You may already have seen an example of cohomology groups jumping, in §24.4.13. Here is a simpler example, albeit not of the structure sheaf. Let (E, p_0) be an elliptic curve over a field k, and consider the projection $\pi : E \times E \to E$. Let $\mathscr{L}$ be the invertible sheaf (line bundle) corresponding to the divisor that is the diagonal, minus the section $p_0 \in E$. Then $\mathscr{L}_{p_0}$ is trivial, but $\mathscr{L}_p$ is non-trivial for any $p \neq p_0$ (as we showed in our study of genus 1 curves, in §19.9). Thus $h^0(E, \mathscr{L}_p)$ is 0 in general, but jumps to 1 for $p = p_0$.

28.1.A. EXERCISE. Show that $\pi_*\mathscr{L} = 0$. Thus we cannot picture $\pi_*\mathscr{L}$ as "gluing together" h^0 of the fibers; in this example, cohomology does not commute with "base change" or "taking fibers".

28.1.3. *Side Remark.* In characteristic 0, the cohomology of $\mathscr{O}$ doesn't jump in smooth families. Over $\mathbb{C}$, this is because Betti numbers are constant in connected families, and (21.5.11.1) (from Hodge theory) expresses the Betti constants h^k_{Betti} as sums (over $i + j = k$) of upper semicontinuous (and hence constant) functions $h^j(\Omega^i)$, so the Hodge numbers $h^j(\Omega^i)$ must be constant. The general characteristic 0 case can be reduced to $\mathbb{C}$ by an application of the Lefschetz principle (which also arose in §21.5.9). But cohomology groups of $\mathscr{O}$ (for flat families of varieties) *can* jump in positive characteristic. Also, the example of §24.4.13 shows that the "smoothness" hypothesis cannot be removed.

28.1.4. Grauert's Theorem. If $R^p\pi_*\mathscr{F}$ is locally free (property (c)) and ϕ_q^p is an isomorphism (property (a)), then $h^p(X_q, \mathscr{F}|_{X_q})$ is clearly locally constant. The following is a partial converse.

28.1.5. Grauert's Theorem. — *If $\pi : X \to Y$ is proper, Y is reduced and locally Noetherian, $\mathscr{F}$ is a coherent sheaf on X flat over Y, and $h^p(X_q, \mathscr{F}|_{X_q})$ is a locally constant function of $q \in Y$, then $R^p\pi_*\mathscr{F}$ is locally free, and ϕ_Z^p is an isomorphism for all $\psi : Z \to Y$.*

In other words, if cohomology groups of fibers have locally constant dimension (over a reduced base), then they can be fit together to form a vector bundle, and the fiber of the pushforward is identified with the cohomology of the fiber. (No Noetherian hypotheses are needed.)

By Exercise 5.1.E (on quasicompact schemes, nonempty closed subsets contain closed points) and the Semicontinuity Theorem 28.1.1, if Y is quasicompact, then to check that $h^p(X_q, \mathscr{F}|_{X_q})$ is constant requires only checking at closed points.

Finally, we note that if Y is integral, π is proper, and $\mathscr{F}$ is a coherent sheaf on X flat over Y, then by the Semicontinuity Theorem 28.1.1 there is a dense open subset of Y on which $R^p\pi_*\mathscr{F}$ is locally free (and on which the fiber of the pth pushforward is the pth cohomology of the fiber).

The following statement is even more magical than Grauert's Theorem 28.1.5.

28.1.6. Cohomology and Base Change Theorem. — *Suppose π is proper, Y is locally Noetherian, $\mathscr{F}$ is coherent and flat over Y, and ϕ_q^p is surjective. Then the following hold.*

(i) *There is an open neighborhood U of q such that for any $\psi : Z \to U$, ϕ_Z^p is an isomorphism. In particular, ϕ_q^p is an isomorphism.*

(ii) *Furthermore, ϕ_q^{p-1} is surjective (hence an isomorphism by (i)) if and only if $R^p\pi_*\mathscr{F}$ is locally free in some neighborhood of q (or equivalently, $(R^p\pi_*\mathscr{F})_q$ is a free $\mathscr{O}_{Y,q}$-module, Exercise 13.7.F). This in turn implies that h^p is constant in a neighborhood of q.*

(Proofs of Theorems 28.1.5 and 28.1.6 will be given in §28.2.)

This is amazing: the hypothesis that ϕ_q^p is surjective involves what happens only over *reduced* points, and it has implications over the (possibly nonreduced) scheme as a whole! This might remind you of the local criterion for flatness (Theorem 24.6.2), and indeed that is the key technical ingredient of the proof.

Here are some consequences.

28.1.B. EXERCISE. Use Theorem 28.1.6 to give a second solution to Exercise 24.4.E. (This is a big weapon to bring to bear on this problem, but it is still enlightening; the original solution to Exercise 24.4.E foreshadowed the proof of the Cohomology and Base Change Theorem 28.1.6.)

28.1.C. EXERCISE. Suppose $\pi : X \to Y$ is proper, Y is locally Noetherian, and $\mathscr{F}$ is a coherent sheaf on X, flat over Y. Suppose further that $H^p(X_q, \mathscr{F}|_{X_q}) = 0$ for some $q \in Y$. Show that there is an open neighborhood U of q such that $(R^p\pi_*\mathscr{F})|_U = 0$.

28.1.D. EXERCISE. Suppose $\pi : X \to Y$ is proper, Y is locally Noetherian, and $\mathscr{F}$ is a coherent sheaf on X, flat over Y. Suppose further that $H^p(X_q, \mathscr{F}|_{X_q}) = 0$ for all $q \in Y$. Show that the $(p-1)$st cohomology commutes with arbitrary base change: ϕ_Z^{p-1} is an isomorphism for all $\psi : Z \to Y$.

28.1.E. EXERCISE. Suppose π is proper, Y is locally Noetherian, and $\mathscr{F}$ is coherent and flat over Y. Suppose further that $R^p\pi_*\mathscr{F} = 0$ for $p \geq p_0$. Show that $H^p(X_q, \mathscr{F}|_{X_q}) = 0$ for all $q \in Y$, $p \geq p_0$.

28.1.F. EXERCISE. Suppose π is proper, Y is locally Noetherian, and $\mathscr{F}$ is coherent and flat over Y. Suppose further that Y is reduced. Show that there exists a dense open subset U of Y such that ϕ^p_Z is an isomorphism for all $\psi : Z \to U$ and all p. (Hint: find suitable neighborhoods of the generic points of Y. See Exercise 24.2.M and the paragraph following it.)

28.1.7. An important class of morphisms: Proper, $\mathscr{O}$-connected morphisms $\pi : X \to Y$ of locally Noetherian schemes.

If a morphism $\pi : X \to Y$ satisfies the property that the natural map $\mathscr{O}_Y \to \pi_*\mathscr{O}_X$ is an isomorphism, we say that π is **$\mathscr{O}$-connected**.

28.1.G. EASY EXERCISE. Show that proper $\mathscr{O}$-connected morphisms of locally Noetherian schemes are surjective.

28.1.8. We will soon meet the Zariski Connectedness Lemma 29.5.1, which shows that proper, $\mathscr{O}$-connected morphisms of locally Noetherian schemes have connected fibers. In some sense, this class of morphisms is really the right class of morphisms capturing what we might want by "connected fibers"; this is the motivation for the terminology. The following result gives some evidence for this point of view, in the flat context.

28.1.H. IMPORTANT EXERCISE. Suppose $\pi : X \to Y$ is a proper *flat* morphism of locally Noetherian schemes, whose fibers satisfy $h^0(X_q, \mathscr{O}_{X_q}) = 1$. (Important remark: this is satisfied if π has geometrically connected and geometrically reduced fibers, by §10.3.7.) Show that π is $\mathscr{O}$-connected. Hint: consider

$$\mathscr{O}_Y \otimes \kappa(q) \longrightarrow (\pi_*\mathscr{O}_X) \otimes \kappa(q) \xrightarrow{\phi^0} H^0(X_q, \mathscr{O}_{X_q}) \cong \kappa(q) .$$

The composition is surjective, hence ϕ^0 is surjective, hence it is an isomorphism by the Cohomology and Base Change Theorem 28.1.6(a). Then by the Cohomology and Base Change Theorem 28.1.6(b), $\pi_*\mathscr{O}_X$ is locally free, thus of rank 1. Use Nakayama's Lemma to show that a map of invertible sheaves $\mathscr{O}_Y \to \pi_*\mathscr{O}_X$ that is an isomorphism on closed points is necessarily an isomorphism of sheaves.

28.1.9. $\star$ *Unimportant remark.* This class of proper, $\mathscr{O}$-connected morphisms is not preserved by arbitrary base change, and thus is not "reasonable" in the sense of §7.1.1. But you can show that they are preserved by *flat* base change, using the fact that cohomology commutes with flat base change, Theorem 24.2.8. Furthermore, the conditions of Exercise 28.1.H behave well under base change, and Noetherian conditions can be removed from the Cohomology and Base Change Theorem 28.1.6 (at the expense of finitely presented hypotheses, see §28.2.9), so the class of morphisms $\pi : X \to Y$ that are proper, finitely presented, and flat, with geometrically connected and geometrically reduced fibers, *is* "reasonable" (and useful).

28.1.10. We next address the following question. Suppose $\pi : X \to Y$ is a morphism of schemes. Given an invertible sheaf $\mathscr{L}$ on X, we ask when it is the pullback of an invertible sheaf $\mathscr{M}$ on Y. For this to be true, we certainly need that $\mathscr{L}$ is trivial on the fibers. We will see that if π is a proper $\mathscr{O}$-connected morphism of locally Noetherian schemes, then this often suffices. Given $\mathscr{L}$, we recover $\mathscr{M}$ as $\pi_*\mathscr{L}$; the fibers of $\mathscr{M}$ are one-dimensional, and glue together to form a line bundle. We now begin to make this precise.

28.1.I. EXERCISE. Suppose $\pi : X \to Y$ is a proper, $\mathscr{O}$-connected morphism of locally Noetherian schemes. Show that if $\mathscr{M}$ is any invertible sheaf on Y, then the natural morphism $\mathscr{M} \to \pi_*\pi^*\mathscr{M}$ is an isomorphism. In particular, we can recover $\mathscr{M}$ from $\pi^*\mathscr{M}$ by applying the pushforward π_*.

28.1.11. Proposition. — *Suppose $\pi : X \to Y$ is a* **flat***, proper, $\mathscr{O}$-connected morphism of locally Noetherian schemes. Suppose also that Y is reduced, and $\mathscr{L}$ is an invertible sheaf on X that is trivial on the fibers of π (i.e., $\mathscr{L}_q$ is a trivial invertible sheaf on X_q for all $q \in Y$). Then $\pi_*\mathscr{L}$ is an invertible sheaf on Y (call it $\mathscr{M}$), and the natural map $\pi^*\mathscr{M} \to \mathscr{L}$ is an isomorphism.*

Proof. By Grauert's Theorem 28.1.5, $\pi_*\mathscr{L}$ is locally free of rank 1 (again, call it $\mathscr{M}$), and $\mathscr{M} \otimes_{\mathscr{O}_Y} \kappa(q) \to H^0(X_q, \mathscr{L}_q)$ is an isomorphism. We have a natural map of invertible sheaves $\pi^*\mathscr{M} = \pi^*\pi_*\mathscr{L} \to \mathscr{L}$. To show that it is an isomorphism, we need only show that it is surjective. (Do you see why? If A is a ring, and $\phi : A \to A$ is a surjection of A-modules, why is ϕ an isomorphism?) For this, it suffices to show that it is surjective on the fibers of π. (Do you see why? Hint: if the kernel of the map is not 0, then it is not 0 above some point of Y.) But this follows from the hypotheses. $\square$

Proposition 28.1.11 has some pleasant consequences. For example, if you have two invertible sheaves $\mathscr{A}$ and $\mathscr{B}$ on X that are isomorphic on every fiber of π, then they differ by a pullback of an invertible sheaf on Y: just apply Proposition 28.1.11 to $\mathscr{A} \otimes \mathscr{B}^\vee$. But Proposition 28.1.11 has the unpleasant hypothesis that the target must be reduced. We can get rid of this hypothesis by replacing the use of Grauert's Theorem with the Cohomology and Base Change Theorem 28.1.6. We do this now.

28.1.12. Projective bundles.

28.1.J. EXERCISE. Let X be a locally Noetherian scheme, and let $\mathrm{pr}_1 : X \times \mathbb{P}^n \to X$ be the projection onto the first factor. Suppose $\mathscr{L}$ is a line bundle on $X \times \mathbb{P}^n$, whose degree on every fiber of pr_1 is zero. Use the Cohomology and Base Change Theorem 28.1.6 to show that $(\mathrm{pr}_1)_*\mathscr{L}$ is an invertible sheaf on X. Use Nakayama's Lemma (in some guise) to show that the natural map $\mathrm{pr}_1^*((\mathrm{pr}_1)_*\mathscr{L}) \to \mathscr{L}$ is an isomorphism.

Your argument will apply just as well to the situation where $\mathrm{pr}_1 : X \times \mathbb{P}^n \to X$ is replaced by a $\mathbb{P}^n$-bundle over X, $\mathrm{pr}_1 : Z \to X$; or by $\mathrm{pr}_1 : Z \to X$ which is a smooth morphism whose geometric fibers are integral curves of genus 0.

Furthermore, the locally Noetherian hypotheses can be removed, see §28.2.9.

28.1.K. EXERCISE. Suppose X is a connected Noetherian scheme. Show that $\mathrm{Pic}(X \times \mathbb{P}^n) \cong \mathrm{Pic}\, X \times \mathbb{Z}$. Hint: the map $\mathrm{Pic}\, X \times \mathrm{Pic}\, \mathbb{P}^n \to \mathrm{Pic}(X \times \mathbb{P}^n)$ is given by $(\mathscr{L}, \mathscr{O}(m)) \mapsto \mathrm{pr}_1^*\mathscr{L} \otimes \mathrm{pr}_2^*\mathscr{O}(m)$, where $\mathrm{pr}_1 : X \times \mathbb{P}^n \to X$ and $\mathrm{pr}_2 : X \times \mathbb{P}^n \to \mathbb{P}^n$

are the projections from $X \times \mathbb{P}^n$ to its factors. (The notation $\boxtimes$ is often used for this construction, see Exercise 9.6.A.)

The same argument will show that if Z is a $\mathbb{P}^n$-bundle over X, then $\operatorname{Pic} Z \cong \operatorname{Pic} X \times \mathbb{Z}$. You will undoubtedly also be able to figure out the right statement if X is not connected.

28.1.13. *Remark.* As mentioned in §19.10.1, the Picard group of a scheme often "wants to be a scheme". You may be able to make this precise in the case of $\operatorname{Pic} \mathbb{P}^n_{\mathbb{Z}}$. In this case, the *scheme* $\operatorname{Pic} \mathbb{P}^n_{\mathbb{Z}}$ is "$\mathbb{Z}$ copies of $\operatorname{Spec} \mathbb{Z}$", with the "obvious" group scheme structure. Can you figure out what functor it represents? Can you show that it represents this functor? This will require extending Exercise 28.1.K out of the Noetherian setting, using §28.2.9.

28.1.L. EXERCISE. Suppose $\pi : X \to Y$ is a projective flat morphism over a Noetherian integral scheme, all of whose geometric fibers are isomorphic to $\mathbb{P}^n$ (over the appropriate field). Show that π is a projective bundle if and only if there is an invertible sheaf on X that restricts to $\mathscr{O}(1)$ on all the fibers. (One direction is clear: if it is a projective bundle, then it has a $\mathscr{O}(1)$ which comes from the projectivization, see Exercise 17.2.D. In the other direction, the candidate vector bundle is $\pi_* \mathscr{O}(1)$. Show that it is indeed a locally free sheaf of the desired rank. Show that its projectivization is indeed $\pi : X \to Y$.)

Caution: the map $\pi : \operatorname{Proj} \mathbb{R}[x, y, z]/(x^2 + y^2 + z^2) \to \operatorname{Spec} \mathbb{R}$ shows that not every projective flat morphism over a Noetherian integral scheme, all of whose geometric fibers are isomorphic to $\mathbb{P}^n$, is necessarily a $\mathbb{P}^n$-bundle. However, *Tsen's Theorem* implies that if the target is a *smooth curve over an algebraically closed field*, then the morphism *is* a $\mathbb{P}^n$-bundle (see [**GS**, Thm. 6.2.8]). Example 18.4.5 shows that "curve" cannot be replaced by "5-fold" in this statement — the "universal smooth plane conic" is not a $\mathbb{P}^1$-bundle over the parameter space $U \subset \mathbb{P}^5$ of smooth plane conics. If you wish, you can extend Example 18.4.5 to show that "curve" cannot even be replaced by "surface". (Just replace the $\mathbb{P}^5$ of all conics with a generally chosen $\mathbb{P}^2$ of conics — but then figure out what goes wrong if you try to replace it with a generally chosen $\mathbb{P}^1$ of conics.)

28.1.M. EXERCISE. Suppose $\pi : X \to Y$ is the projectivization of a vector bundle $\mathscr{F}$ over a locally Noetherian scheme (i.e., $X \cong \mathcal{P}roj \operatorname{Sym}^\bullet \mathscr{F}$). Recall from §17.2.3 that for any invertible sheaf $\mathscr{L}$ on Y, $X \cong \mathcal{P}roj \operatorname{Sym}^\bullet(\mathscr{F} \otimes \mathscr{L})$). Show that these are the only ways in which it is the projectivization of a vector bundle. (Hint: recover $\mathscr{F}$ by pushing forward $\mathscr{O}(1)$.)

28.1.14. The Hodge bundle.

28.1.N. EXERCISE (THE HODGE BUNDLE). Suppose $\pi : X \to Y$ is a flat proper morphism of locally Noetherian schemes, and the fibers of π are regular irreducible curves of genus g. Show that $\pi_* \Omega_{X/Y}$ is a locally free sheaf on Y of rank g, and that the construction of π commutes with base change: given a fibered square

(28.1.14.1)
$$\begin{tikzcd} X' \arrow[r, "\psi'"] \arrow[d, "\pi'"'] & X \arrow[d, "\pi"] \\ Y' \arrow[r, "\psi"] & Y, \end{tikzcd}$$

there is an isomorphism $\Omega_{X'/Y'} \cong (\psi')^*\Omega_{X/Y}$. (The locally free sheaf $\pi_*\Omega_{X/Y}$ is called the **Hodge bundle**.) Hint: use the Cohomology and Base Change Theorem 28.1.6 twice, once with $p = 2$, and once with $p = 1$.

28.2 ⋆⋆ Proofs of cohomology and base change theorems

The key to proving the Semicontinuity Theorem 28.1.1, Grauert's Theorem 28.1.5, and the Cohomology and Base Change Theorem 28.1.6 is the following wonderful idea of Mumford (see [**Mu3**, p. 47 Lem. 1]). It turns questions of pushforwards (and how they behave under arbitrary base change) into something computable with vector bundles (hence questions of linear algebra). After stating it, we will interpret it.

28.2.1. Key Theorem. — *Suppose $\pi : X \to \operatorname{Spec} B$ is a proper morphism, and $\mathscr{F}$ is a coherent sheaf on X, flat over $\operatorname{Spec} B$. Then there is a complex*

(28.2.1.1)
$$\cdots \longrightarrow K^{-1} \longrightarrow K^0 \longrightarrow K^1 \longrightarrow \cdots \longrightarrow K^n \longrightarrow 0$$

of finitely generated free B-modules and an isomorphism of functors

(28.2.1.2)
$$H^p(X \times_B A, \mathscr{F} \otimes_B A) \cong H^p(K^\bullet \otimes_B A)$$

for all p, for all ring maps $B \to A$.

Because (28.2.1.1) is an exact sequence of *free* B-modules, all of the information is contained in the maps, which are matrices with entries in B. This will turn questions about cohomology (and base change) into questions about linear algebra. For example, semicontinuity will turn into the fact that ranks of matrices (with functions as entries) drop on closed subsets (§11.4.4(ii)).

Although the complex (28.2.1.1) is infinite, by (28.2.1.2) it has no cohomology in negative degree, even after any ring extension $B \to A$ (as the left side of (28.2.1.2) is 0 for $p < 0$).

The idea behind the proof is as follows: take the Čech complex, produce a complex of finite rank free modules mapping to it "with the same cohomology" (a *quasiisomorphic complex*, §18.2.3). We first construct the complex so that (28.2.1.2) holds for $B = A$, and then show the same complex works for general A. We begin with a lemma.

28.2.2. Lemma. — *Let B be a Noetherian ring. Suppose $C^\bullet$ is a complex of B-modules such that $H^i(C^\bullet)$ are finitely generated B-modules, and such that $C^p = 0$ for $p > n$. Then there exists a complex $K^\bullet$ of finite rank free B-modules such that $K^p = 0$ for $p > n$, and a homomorphism of complexes $\alpha : K^\bullet \to C^\bullet$ such that α induces isomorphisms $H^i(K^\bullet) \to H^i(C^\bullet)$ for all i.*

Proof. We build this complex inductively. (This may remind you of Hint 23.3.3.) Assume we have defined $(K^p, \alpha^p, \delta^p)$ for $p \geq m+1$ (as in (28.2.2.1)) such that the squares commute, and the top row is a complex, and α^q defines an isomorphism of cohomology $H^q(K^\bullet) \to H^q(C^\bullet)$ for $q \geq m+2$ and a surjection $\ker \delta^{m+1} \to H^{m+1}(C^\bullet)$, and the K^p are finite rank free B-modules. (Our base case is $m = p$: take $K^n = 0$ for $n > p$.)

(28.2.2.1)
$$\begin{tikzcd} & & & K^{m+1} \arrow[r, "\delta^{m+1}"] \arrow[d, "\alpha^{m+1}"] & K^{m+2} \arrow[r] \arrow[d, "\alpha^{m+2}"] & \cdots \\ \cdots \arrow[r] & C^{m-1} \arrow[r] & C^m \arrow[r, "\delta^m"'] & C^{m+1} \arrow[r, "\delta^{m+1}"'] & C^{m+2} \arrow[r] & \cdots . \end{tikzcd}$$

(We sloppily use δ^q for the horizontal morphisms in both rows.)

We construct $(K^m, \delta^m, \alpha^m)$. Choose generators of $H^m(C^\bullet)$, say $c_1, \dots, c_M$. Let

$$D^{m+1} := \ker\left(\ker(\delta^{m+1}) \xrightarrow{\alpha^{m+1}} H^{m+1}(C^\bullet)\right)$$

(where the δ^{m+1} is the differential of the top complex $K^\bullet$). Choose generators of D^{m+1}, say $d_1, \dots, d_N$. Let $K^m = B^{\oplus(M+N)}$. Define $\delta^m : K^m \to K^{m+1}$ by sending the last N generators to $d_1, \dots d_N$, and the first M generators to 0. Define α^m by sending the first M generators of $B^{\oplus(M+N)}$ to (lifts of) $c_1, \dots, c_M$, and sending the last N generators to arbitrarily chosen lifts of the $\alpha^{m+1}(d_i)$ (as the $\alpha^{m+1}(d_i)$ are 0 in $H^{m+1}(C^\bullet)$, and thus lie in the image of δ^m), so the square (with upper left corner K^m) commutes. Then by construction, we have completed our inductive step:

$$\begin{tikzcd} & & K^{m} \arrow[r, "\delta^{m}"] \arrow[d, "\alpha^{m}"] & K^{m+1} \arrow[r, "\delta^{m+1}"] \arrow[d, "\alpha^{m+1}"] & K^{m+2} \arrow[r] \arrow[d, "\alpha^{m+2}"] & \cdots \\ \cdots \arrow[r] & C^{m-1} \arrow[r] & C^m \arrow[r, "\delta^m"'] & C^{m+1} \arrow[r, "\delta^{m+1}"'] & C^{m+2} \arrow[r] & \cdots . \end{tikzcd}$$

□

28.2.3. Lemma. — *Suppose $\alpha : K^\bullet \to C^\bullet$ is a morphism of complexes of* **flat** *B-modules inducing isomorphisms of cohomology (a quasiisomorphism, §18.2.3). Then "this quasiisomorphism commutes with arbitrary change of base ring": for every B-algebra A, the maps $H^p(C^\bullet \otimes_B A) \to H^p(K^\bullet \otimes_B A)$ are isomorphisms.*

Proof. The mapping cone $M^\bullet$ of $\alpha : K^\bullet \to C^\bullet$ is exact by Exercise 1.7.E. Then $M^\bullet \otimes_B A$ is still exact, by Exercise 24.3.F. But $M^\bullet \otimes_B A$ is the mapping cone of $\alpha \otimes_B A : K^\bullet \otimes_B A \to C^\bullet \otimes_B A$, so by Exercise 1.7.E, $\alpha \otimes_B A$ induces an isomorphism of cohomology (i.e., is a quasiisomorphism) too. □

Proof of Key Theorem 28.2.1. Choose a finite affine covering of X. Take the Čech complex $C^\bullet$ for $\mathscr{F}$ with respect to this cover. Recall that Grothendieck's Coherence Theorem 18.9.1 (which had Noetherian hypotheses) showed that the cohomology of $\mathscr{F}$ is coherent. (That Theorem required serious work. If you need Theorem 28.2.1 only in the projective case, the analogous statement with projective hypotheses, Theorem 18.8.1(d), was much easier.) Apply Lemma 28.2.2 to get the nicer variant $K^\bullet$ of the same complex $C^\bullet$. By Lemma 28.2.3, if we tensor with A and take cohomology, we get the same answer whether we use $K^\bullet$ or $C^\bullet$. $\square$

We now use Theorem 28.2.1 to prove some of the fundamental results stated earlier: the Semicontinuity Theorem 28.1.1, Grauert's Theorem 28.1.5, and the Cohomology and Base Change Theorem 28.1.6. In the course of proving Semicontinuity, we will give a new proof of Theorem 24.7.1, that Euler characteristics are locally constant in flat families (that applies more generally in proper situations).

28.2.4. Proof of the Semicontinuity Theorem 28.1.1. The result is local on Y, so we may assume Y is affine. Let $K^\bullet$ be a complex as in Key Theorem 28.2.1.

Then for $q \in Y$,

$$\begin{aligned} \dim_{\kappa(q)} H^p(X_q, \mathscr{F}|_{X_q}) &= \dim_{\kappa(q)} \ker(\delta^p \otimes_B \kappa(q)) - \dim_{\kappa(q)} \operatorname{im}(\delta^{p-1} \otimes_B \kappa(q)) \\ &= \dim_{\kappa(q)} (K^p \otimes_B \kappa(q)) - \dim_{\kappa(q)} \operatorname{im}(\delta^p \otimes_B \kappa(q)) \\ &\qquad - \dim_{\kappa(q)} \operatorname{im}(\delta^{p-1} \otimes_B \kappa(q)) \end{aligned} \tag{28.2.4.1}$$

Now $\dim_{\kappa(q)} \operatorname{im}(\delta^p \otimes_B \kappa(q))$ is a lower semicontinuous function on Y. (Reason: the locus where the dimension is less than some number N is obtained by setting all $N \times N$ minors of the matrix $K^p \to K^{p+1}$ to 0; cf. §11.4.4(ii)).) The same is true for $\dim_{\kappa(q)} \operatorname{im}(\delta^{p-1} \otimes_B \kappa(q))$. The result follows. $\square$

28.2.5. A new proof (and extension to the proper case) of Theorem 24.7.1 that Euler characteristics of flat sheaves are locally constant.

If $K^\bullet$ were finite "on the left" as well — if $K^p = 0$ for $p \ll 0$ — then we would have a short proof of Theorem 24.7.1. By taking alternating sums (over p) of (28.2.4.1), we would have that

$$\chi(X_q, \mathscr{F}|_{X_q}) = \sum (-1)^p h^p(X_q, \mathscr{F}|_{X_q}) = \sum (-1)^p \operatorname{rank} K^p,$$

which is locally constant. The only problem is that the sums are infinite. We patch this problem as follows. Define $J^\bullet$ by $J^p = K^p$ for $p \geq 0$, $J^p = 0$ for $p < -1$, and $J^{-1} := \ker(K^0 \to K^1)$. Combine the $J^\bullet$ into a complex, by defining $J^p \to J^{p+1}$ as the obvious map induced by $K^\bullet$. We have a map of complexes $J^\bullet \to K^\bullet$. Clearly this induces an isomorphism on cohomology (as $J^\bullet$ patently has the same cohomology as $K^\bullet$ at step $p \geq 0$, and both have 0 cohomology for $p < 0$). Thus the composition $\beta : J^\bullet \to K^\bullet \to C^\bullet$ induces an isomorphism on cohomology as well.

Now J^{-1} is coherent (as it is the kernel of a map of coherent modules). Consider the mapping cone $M^\bullet$ of $\beta : J^\bullet \to C^\bullet$:

$$0 \to J^{-1} \to C^{-1} \oplus J^0 \to C^0 \oplus J^1 \to \cdots \to C^{n-1} \oplus J^n \to C^n \to 0.$$

From Exercise 1.7.E, as $J^\bullet \to C^\bullet$ induces an isomorphism on cohomology, the mapping cone has no cohomology — it is exact. All terms in it are flat except possibly J^{-1} (the C^p are flat by assumption, and J^i is free for $i \neq -1$). Hence J^{-1} is flat too, by Exercise 24.3.G. But flat coherent sheaves are locally free (Theorem 24.4.7). Then Theorem 24.7.1 follows from

$$\chi(X_q, \mathscr{F}|_{X_q}) = \sum (-1)^p h^p(X_q, \mathscr{F}|_{X_q}) = \sum (-1)^p \operatorname{rank} J^p.$$

$\square$

28.2.6. Proof of Grauert's Theorem 28.1.5 and the Cohomology and Base Change Theorem 28.1.6.

Thanks to Theorem 28.2.1.2, Theorems 28.1.5 and 28.1.6 are now statements about complexes of free modules over a Noetherian ring. We begin with some general comments on dealing with the cohomology of a complex

$$\cdots \longrightarrow K^p \xrightarrow{\delta^p} K^{p+1} \longrightarrow \cdots .$$

We define some notation for functions on a complex (most of which already appeared in §23.3.7).

- Let Z^p be the kernel of the pth differential of a complex, so for example $Z^pK^\bullet = \ker \delta^p$.
- Let B^{p+1} be the image of the pth differential, so for example $B^{p+1}K^\bullet = \operatorname{im} \delta^p$.
- Let W^{p+1} be the cokernel of the pth differential, so for example $W^{p+1}K^\bullet = \operatorname{coker} \delta^p$.
- As usual, let H^p be the homology at the pth step.

We have exact sequences (cf. (1.6.5.3) and (1.6.5.4))

$$0 \longrightarrow Z^p \longrightarrow K^p \longrightarrow K^{p+1} \longrightarrow W^{p+1} \longrightarrow 0 \tag{28.2.6.1}$$

$$0 \longrightarrow Z^p \longrightarrow K^p \longrightarrow B^{p+1} \longrightarrow 0 \tag{28.2.6.2}$$

$$0 \longrightarrow B^p \longrightarrow Z^p \longrightarrow H^p \longrightarrow 0 \tag{28.2.6.3}$$

$$0 \longrightarrow B^p \longrightarrow K^p \longrightarrow W^p \longrightarrow 0 \tag{28.2.6.4}$$

$$0 \longrightarrow H^p \longrightarrow W^p \longrightarrow B^{p+1} \longrightarrow 0. \tag{28.2.6.5}$$

We proceed by a series of exercises, some of which were involved in the proof of the FHHF Theorem (Exercise 1.6.H). Suppose $C^\bullet$ is any complex in an abelian category $\mathscr{A}$ with enough projectives, and suppose F is any right-exact functor from $\mathscr{A}$.

28.2.A. EXERCISE (COKERNELS COMMUTE WITH RIGHT-EXACT FUNCTORS). Describe an *isomorphism* $\gamma^p : FW^pC^\bullet \xrightarrow{\sim} W^pFC^\bullet$. (Hint: consider $C^{p-1} \to C^p \to W^pC^\bullet \to 0$.)

28.2.B. EXERCISE.
(a) Describe a map $\beta^p : FB^pC^\bullet \to B^pFC^\bullet$. Hint: (28.2.6.4) induces

$$\begin{tikzcd}
R^1FW^pC^\bullet \arrow[r] & FB^pC^\bullet \arrow[r] \arrow[d, dotted, "\beta^p"] & FC^p \arrow[r] \arrow[d, "="] & FW^pC^\bullet \arrow[r] \arrow[d, "\gamma^p"', "\sim"] & 0 \\
0 \arrow[r] & B^pFC^\bullet \arrow[r] & FC^p \arrow[r] & W^pFC^\bullet \arrow[r] & 0.
\end{tikzcd}$$

(b) Show that β^p is surjective. Possible hint: use Exercise 1.7.B, a weaker version of the snake lemma, to get an exact sequence

$$\boxed{\begin{aligned} &R^1FC^p \longrightarrow R^1FW^pC^\bullet \longrightarrow \ker\beta^p \longrightarrow 0 \longrightarrow \ker\gamma^p \\ &\qquad\qquad\longrightarrow \operatorname{coker}\beta^p \longrightarrow 0 \longrightarrow \operatorname{coker}\gamma^p \longrightarrow 0. \end{aligned}}$$

28.2.C. EXERCISE.
(a) Describe a map $\alpha^p : FZ^pC^\bullet \to Z^pFC^\bullet$. Hint: use (28.2.6.2) to induce

$$\begin{tikzcd} R^1FB^{p+1}C^\bullet \arrow[r] & FZ^pC^\bullet \arrow[r] \arrow[d, dotted, "\alpha^p"] & FC^p \arrow[r] \arrow[d, "="] & FB^{p+1}C^\bullet \arrow[r] \arrow[d, "\beta^{p+1}"] & 0 \\ 0 \arrow[r] & Z^pFC^\bullet \arrow[r] & FC^p \arrow[r] & B^{p+1}FC^\bullet \arrow[r] & 0 \end{tikzcd}$$

(b) Use Exercise 1.7.B to get an exact sequence

$$\boxed{\begin{aligned} &R^1FC^\bullet \longrightarrow R^1FB^{p+1}C^\bullet \longrightarrow \ker\alpha^p \longrightarrow 0 \longrightarrow \ker\beta^{p+1} \\ &\qquad\qquad\longrightarrow \operatorname{coker}\alpha^p \longrightarrow 0 \longrightarrow \operatorname{coker}\beta^{p+1} \longrightarrow 0. \end{aligned}}$$

28.2.D. EXERCISE.
(a) Describe a map $\epsilon^p : FHK^p \to HFK^p$. (This is the FHHF Theorem for right-exact functors, Exercise 1.6.H(a).) Hint: (28.2.6.3) induces

$$\begin{tikzcd} R^1FH^pC^\bullet \arrow[r] & FB^pC^\bullet \arrow[r] \arrow[d, "\beta^p"] & FZ^pC^\bullet \arrow[r] \arrow[d, "\alpha^p"] & FH^pC^\bullet \arrow[r] \arrow[d, dotted, "\epsilon^p"] & 0 \\ 0 \arrow[r] & B^pFC^\bullet \arrow[r] & Z^pFC^\bullet \arrow[r] & H^pFC^\bullet \arrow[r] & 0 \end{tikzcd}$$

(b) Use Exercise 1.7.B to get an exact sequence:

$$\boxed{\begin{aligned} &R^1FZ^pC^\bullet \longrightarrow R^1FH^pC^\bullet \longrightarrow \ker\beta^p \longrightarrow \ker\alpha^p \longrightarrow \ker\epsilon^p \\ &\qquad\qquad\longrightarrow \operatorname{coker}\beta^p \longrightarrow \operatorname{coker}\alpha^p \longrightarrow \operatorname{coker}\epsilon^p \longrightarrow 0. \end{aligned}}$$

28.2.7. *Back to the theorems we want to prove.* Recall the properties we discussed at the start of §28.1.

(a) Given a fibered square (28.1.0.1), is $\phi^p_q : R^p\pi_*\mathscr{F} \otimes \kappa(q) \to H^p(X_q, \mathscr{F}|_{X_q})$ an isomorphism?
(b) Given a fibered square (28.1.0.2), is $\phi^p_Z : \psi^*(R^p\pi_*\mathscr{F}) \to R^p\pi'_*(\psi')^*\mathscr{F}$ an isomorphism?
(c) Is $R^p\pi_*\mathscr{F}$ locally free?

We reduce to the case Y and Z are both affine, say $Y = \operatorname{Spec} B$. We apply our general results of §28.2.6 to the complex (28.2.1.1) of Theorem 28.2.1.

28.2.E. EXERCISE. Suppose $W^p K^\bullet$ and $W^{p+1} K^\bullet$ are flat. Show that the answer to (b), and hence (a), is yes. Show that the answer to (c) is yes if Y is reduced or locally Noetherian. Hint: (You will take F to be the functor $(\cdot) \otimes_B A$, where A is some B-algebra.) Use (28.2.6.4) (shifted) to show that $B^{p+1} K^\bullet$ is flat, and then (28.2.6.5) to show that $H^p K^\bullet$ is flat. By Exercise 28.2.A, the construction of the cokernel $W^\bullet$ behaves well under base change. The flatness of B^{p+1} and H^p imply that their constructions behave well under base change as well — apply F to (28.2.6.4) and (28.2.6.5) respectively. (If you care, you can check that $Z^p K^\bullet$ is also locally free, and behaves well under base change.)

28.2.F. EXERCISE. Prove Grauert's Theorem 28.1.5. (Reminder: you won't need Noetherian hypotheses.) Hint: By (28.2.4.1), $W^p K^\bullet$ and $W^{p+1} K^\bullet$ have constant rank. But finite type quasicoherent sheaves having constant rank on a reduced scheme are locally free (Exercise 13.7.K), so we can invoke Exercise 28.2.E. Conclude that $H^p K^\bullet$ is flat of constant rank, and hence locally free.

28.2.8. *Proof of the Cohomology and Base Change Theorem 28.1.6.* Keep in mind that we now have locally Noetherian hypotheses. We have reduced to the case Y and Z are both affine, say $Y = \operatorname{Spec} B$. Let $F = \cdot \otimes_B \kappa(q)$. The key input is the local criterion for flatness (Theorem 24.6.2): $R^1 F W^p K^\bullet = 0$ if and only if $F W^p K^\bullet$ is flat at y (and similarly with W replaced by other letters). In particular, $R^1 F K^p = 0$ for all p. Also keep in mind that if a coherent sheaf on a locally Noetherian scheme (such as $\operatorname{Spec} B$) is flat at a point q, then it is flat in a neighborhood of that point, by Corollary 24.4.7 (flat = locally free for such sheaves).

28.2.G. EXERCISE. Look at the boxed snakes in §28.2.6 (with $C^\bullet = K^\bullet$), and show the following in order, starting from the assumption that $\operatorname{coker} \epsilon^p = 0$:

- $\operatorname{coker} \alpha^p = 0$, $\ker \beta^{p+1} = 0$, $R^1 F W^{p+1} K^\bullet = 0$;
- $W^{p+1} K^\bullet$ is flat, $B^{p+1} K^\bullet$ is flat (use (28.2.6.4) with the indexing shifted by one), $Z^p K^\bullet$ is flat (use (28.2.6.3));
- $R^1 F B^{p+1} K^\bullet = 0$;
- $\ker \alpha^p = 0$, $\ker \epsilon^p = 0$.

It might be useful for later to note that

$$R^1 F W^p K^\bullet \cong \ker \beta^p \cong R^1 F H^p K^\bullet$$

At this point, we have shown that ϕ^p_q is an isomorphism — part of of part (i) of the theorem.

28.2.H. EXERCISE. Prove part (i) of the Cohomology and Base Change Theorem 28.1.6.

Also, ϕ^{p-1}_q surjective implies $W^p K^\bullet$ is flat (in the same way that you showed ϕ^p_q surjective implies $W^{p+1} K^\bullet$ is flat), so we get H^p is free by Exercise 28.2.E, yielding half of (ii).

28.2.I. EXERCISE. For the other direction of (ii), shift the grading of the last two boxed snakes down by one, to obtain further isomorphisms

$$\ker \beta^p \cong \operatorname{coker} \alpha^{p-1} \cong \operatorname{coker} \epsilon^{p-1}.$$

For the other direction of (a), note that if the stalks $W^p K^\bullet$ and $W^{p+1} K^\bullet$ at y are flat, then they are locally free (as they are coherent, by Theorem 24.4.5), and hence $W^p K^\bullet$ and $W^{p+1} K^\bullet$ are locally free in a *neighborhood* of q by Exercise 13.7.F. Thus the stalks of $W^p K^\bullet$ and $W^{p+1} K^\bullet$ are flat in a neighborhood of q, and the same argument applies for any point in this neighborhood to show that $W^{p+1} K^\bullet$, $B^{p+1} K^\bullet$, and $Z^p K^\bullet$ are all flat.

28.2.J. EXERCISE. Use this to show the following, possibly in order:

- $R^1 F C^{p+1} = R^1 F B^{p+1} = R^1 Z^p = 0$.
- $\ker \beta^{p+1} = 0$, $\operatorname{coker} \alpha^p = 0$, $\operatorname{coker} \epsilon^p = 0$.

28.2.K. EXERCISE. Put all the pieces together and finish the proof of part (ii) of the Cohomology and Base Change Theorem 28.1.6. □

28.2.9. ⋆ **Removing Noetherian conditions.**

It can be helpful to have versions of the theorems of §28.1 without Noetherian conditions; important examples come from moduli theory, and will be discussed in the next section. Noetherian conditions can often be exchanged for finite presentation conditions. We begin with an extension of Exercise 9.3.H.

28.2.L. EXERCISE. Suppose $\pi : X \to \operatorname{Spec} B$ is a finitely presented morphism, and $\mathscr{F}$ is a finitely presented quasicoherent sheaf on X. Show that there exists a base change diagram of the form

(28.2.9.1)
$$\begin{tikzcd} \mathscr{F} \arrow[d, dash] & \mathscr{F}' \arrow[d, dash] \\ X \arrow[r, "\sigma"] \arrow[d, "\pi"'] & X' \arrow[d, "\pi'"] \\ \operatorname{Spec} B \arrow[r, "\rho"] & \operatorname{Spec} \mathbb{Z}[x_1, \dots, x_N]/I \end{tikzcd}$$

where N is some integer, $I \subset \mathbb{Z}[x_1, \dots, x_N]$, and π' is finitely presented (= finite type as the target is Noetherian, see §7.3.17), and a finitely presented (= coherent) quasicoherent sheaf $\mathscr{F}'$ on X' with $\mathscr{F} \cong \sigma^* \mathscr{F}$.

28.2.10. *Properties of π'.* (The ideal I appears in the statement of Exercise 28.2.L not because it is needed there, but to make the statement of this remark correct.) If π is proper, then diagram (28.2.9.1) can be constructed so that π' is also proper (using [**Gr-EGA**, IV$_3$.8.10.5]). Furthermore, if $\mathscr{F}$ is flat over $\operatorname{Spec} B$, then (28.2.9.1) can be constructed so that $\mathscr{F}'$ is flat over $\operatorname{Spec} \mathbb{Z}[x_1, \dots, x_N]/I$ (using [**Gr-EGA**, IV$_3$.11.2.6]). This requires significantly more work.

28.2.M. EXERCISE. *Assuming* the results stated in §28.2.10, prove the following results, with the "locally Noetherian" hypotheses removed, and "finite presentation" hypotheses added:

(a) the constancy of Euler characteristic in flat families (Theorem 24.7.1, extended to the proper case as in §28.2.5);
(b) the Semicontinuity Theorem 28.1.1;
(c) Grauert's Theorem 28.1.5 (you will have to show that $\mathbb{Z}[x_1, \dots, x_N]/I$ in (28.2.9.1) can be taken to be reduced); and
(d) the Cohomology and Base Change Theorem 28.1.6.

28.2.11. *Necessity of finite presentation conditions.* The finite presentation conditions are necessary. There is a projective flat morphism to a connected target where the fiber dimension jumps. There is a finite flat morphism where the degree of the fiber is not locally constant. There is a projective flat morphism to a connected target where the fibers are curves, and the arithmetic genus is not constant. See [**Stacks**, tag 05LB] for the first example; the other two use the same idea.

28.3 Applying cohomology and base change to moduli problems

The theory of moduli relies on ideas of cohomology and base change. We explore this by examining two special cases of one of the primordial moduli spaces, the Hilbert scheme: the Grassmannian, and the fact that degree d hypersurfaces in projective space are "parametrized" by another projective space (corresponding to degree d polynomials, see Remark 4.5.3).

As suggested in §24.1, the Hilbert functor $\operatorname{Hilb}_B \mathbb{P}^n$ of $\mathbb{P}^n_B$ parametrizes finitely presented closed subschemes of $\mathbb{P}^n_B$, where B is an arbitrary scheme. More precisely, it is a contravariant functor sending the scheme B to finitely presented closed subschemes of $X \times_{\mathbb{Z}} B$ flat over B (and sending morphisms $B_1 \to B_2$ to pullbacks of flat families). An early achievement of Grothendieck was the construction of the Hilbert scheme, which can then be cleverly used to construct many other moduli spaces.

28.3.1. Theorem (Grothendieck). — $\operatorname{Hilb}_{\mathbb{Z}} \mathbb{P}^n$ *is representable by a scheme locally of finite type.*

(Grothendieck's original argument is in [**Gr5**]. A readable construction is given in [**Mu2**], and in [**FGIKNV**, Ch. 5].)

28.3.A. EASY EXERCISE. Assuming Theorem 28.3.1, show that $\operatorname{Hilb}_B \mathbb{P}^n$ representable, by showing that it is represented by $\operatorname{Hilb}_{\mathbb{Z}} \mathbb{P}^n \times_{\mathbb{Z}} B$. Thus the general case follows from the "universal" case of $B = \mathbb{Z}$.

28.3.B. EXERCISE. Assuming Theorem 28.3.1, show that $\operatorname{Hilb}_{\mathbb{Z}} \mathbb{P}^n$ is the disjoint union of schemes $\operatorname{Hilb}^{p(m)}_{\mathbb{Z}} \mathbb{P}^n$, each one corresponding to finitely presented closed subschemes of $\mathbb{P}^n_{\mathbb{Z}}$ whose fibers have fixed Hilbert polynomial $p(m)$. Hint: Corollary 24.7.3.

28.3.2. Theorem (Grothendieck). — *Each* $\operatorname{Hilb}^{p(m)}_{\mathbb{Z}} \mathbb{P}^n$ *is projective over* $\mathbb{Z}$.

In order to get some feel for the Hilbert scheme, we discuss two important examples, without relying on Theorem 28.3.1.

28.3.3. The Grassmannian.

We have defined the Grassmannian $G(k,n)$ twice before, in §6.7 and §16.7. The second time involved showing the representability of a (contravariant) functor (from sheaves to sets), of rank k locally free quotient sheaves of a rank n free sheaf.

We now consider a parameter space for a more geometric problem. The space will again be $G(k,n)$, but because we won't immediately know this, we invent some temporary notation. Let $G'(k,n)$ be the contravariant functor (from schemes to sets) which assigns to a scheme B the set of *finitely presented* closed subschemes of $\mathbb{P}^{n-1}_B$, flat over B, whose fiber over any point $b \in B$ is a (linearly embedded) $\mathbb{P}^{k-1}_{\kappa(b)}$ in $\mathbb{P}^{n-1}_{\kappa(b)}$:

(28.3.3.1)
$$\begin{tikzcd} X \arrow[r, hook, "\text{cl. subsch.}"] \arrow[d, "\text{flat, f. pr.}"'] & \mathbb{P}^{n-1}_B \arrow[dl, "\pi"] \\ B & \end{tikzcd}$$

(This describes the map to sets; you should think through how pullback makes this into a contravariant functor.)

28.3.4. Theorem. — *The functor* $G'(k,n)$ *is represented by* $G(k,n)$.

Translation: there is a natural bijection between diagrams of the form (28.3.3.1) (where the fibers are $\mathbb{P}^{k-1}$'s) and diagrams of the form (16.7.0.1) (the diagrams that $G(k,n)$ parametrizes, or represents).

One direction is easier. Suppose we are given a diagram of the form (16.7.0.1) over a scheme B,

(28.3.4.1)
$$\mathscr{O}_B^{\oplus n} \twoheadrightarrow \mathscr{Q},$$

where $\mathscr{Q}$ is locally free of rank k. Applying $\mathcal{P}roj_B$ to the $\operatorname{Sym}^\bullet$ construction on both $\mathscr{O}_B^{\oplus n}$ and $\mathscr{Q}$, we obtain a closed embedding

(28.3.4.2)
$$\begin{tikzcd} \mathcal{P}roj_B(\operatorname{Sym}^\bullet \mathscr{Q}) \arrow[rr, hook] \arrow[dr] & & \mathcal{P}roj_B\left(\operatorname{Sym}^\bullet \mathscr{O}_B^{\oplus n}\right) \arrow[dl] & = \mathbb{P}^{n-1} \times B \\ & B & & \end{tikzcd}$$

(as, for example, in Exercise 17.2.H).

The fibers are linearly embedded $\mathbb{P}^{k-1}$'s (as base change, in this case to a point of B, commutes with the $\mathcal{P}roj$ construction, Exercise 17.2.E). Note that $\mathcal{P}roj(\operatorname{Sym}^\bullet \mathscr{Q})$ is flat and finitely presented over B, as it is a projective bundle. We have constructed a diagram of the form (28.3.3.1).

We now need to reverse this. The trick is to produce (28.3.4.1) from our geometric situation (28.3.3.1), and this is where cohomology and base change will be used.

Given a diagram of the form (28.3.3.1) (where the fibers are $\mathbb{P}^{k-1}$'s), consider the closed subscheme exact sequence for X:

$$0 \to \mathscr{I}_X \to \mathscr{O}_{\mathbb{P}^{n-1}_B} \to \mathscr{O}_X \to 0.$$

Tensor this with $\mathscr{O}_{\mathbb{P}_B^{n-1}}(1)$:

$$0 \to \mathscr{I}_X(1) \to \mathscr{O}_{\mathbb{P}_B^{n-1}}(1) \to \mathscr{O}_X(1) \to 0. \tag{28.3.4.3}$$

Note that $\mathscr{O}_X(1)$ restricted to each fiber of π is $\mathscr{O}(1)$ on $\mathbb{P}^{k-1}$ (over the residue field), for which all higher cohomology vanishes (§18.3).

28.3.C. EXERCISE. Show that $R^i\pi_*\mathscr{O}_X(1) = 0$ for $i > 0$, and $\pi_*\mathscr{O}_X(1)$ is locally free of rank k. Hint: use the Cohomology and Base Change Theorem 28.1.6. Either use the non-Noetherian discussion of §28.2.9 (which we haven't proved), or else just assume B is locally Noetherian.

28.3.D. EXERCISE. Show that the long exact sequence obtained by applying π_* to (28.3.4.3) is just a short exact sequence of locally free sheaves

$$0 \to \pi_*\mathscr{I}_X(1) \to \pi_*\mathscr{O}_{\mathbb{P}_B^{n-1}}(1) \to \pi_*\mathscr{O}_X(1) \to 0.$$

of ranks $n-k$, n, and k respectively, where the middle term is canonically identified with $\mathscr{O}^{\oplus n}$.

The surjection $\mathscr{O}^{\oplus n} \to \pi_*\mathscr{O}_X(1)$ is precisely a diagram of the sort we wished to construct, (16.7.0.1).

28.3.E. EXERCISE. Close the loop, by using these two "inverse" constructions to show that $G(k,n)$ represents the functor $G'(k,n)$.

28.3.5. Hypersurfaces.
Ages ago (in Remark 4.5.3), we informally said that hypersurfaces of degree d in $\mathbb{P}^n$ are parametrized by a $\mathbb{P}^{\binom{n+d}{d}}$. We now make this precise. We work over a base $\mathbb{Z}$ for suitable generality. You are welcome to replace $\mathbb{Z}$ by a field of your choice, but by the same argument as in Easy Exercise 28.3.A, all other cases are obtained from this one by base change.

Define the contravariant functor $H_{d,n}$ from schemes to sets as follows. To a scheme B, we associated a closed subscheme $X \hookrightarrow \mathbb{P}_B^n$, flat and finitely presented over B, all of whose fibers are degree d hypersurfaces in $\mathbb{P}^n$ (over the appropriate residue field). To a morphism $B_1 \to B_2$, we obtain a map $H_{d,n}(B_2) \to H_{d,n}(B_1)$ by pullback.

28.3.6. Proposition. — *The functor $H_{d,n}$ is represented by $\mathbb{P}^{\binom{n+d}{d}-1}$.*

As with the case of the Grassmannian, one direction is easy, and the other requires cohomology and base change.

28.3.F. EASY EXERCISE. Over $\mathbb{P}^{\binom{n+d}{d}-1}$, described a closed subscheme $\mathscr{X} \hookrightarrow \mathbb{P}^n \times \mathbb{P}^{\binom{n+d}{d}-1}$ that will be the universal hypersurface. Show that $\mathscr{X}$ is flat and finitely presented over $\mathbb{P}^{\binom{n+d}{d}-1}$. (For flatness, you can use the local criterion of flatness on the source, Exercise 24.6.G, but it is possible to deal with it easily by working by hand.)

Thus given any morphism $B \to \mathbb{P}^{\binom{n+d}{d}-1}$, by pullback, we have a degree d hypersurface X over B (an element of $H_{d,n}(B)$).

Our goal is to reverse this process: from a degree d hypersurface $\pi : X \to B$ over B (an element of $H_{d,n}(B)$), we want to describe a morphism $B \to \mathbb{P}^{\binom{n+d}{d}-1}$.

Consider the closed subscheme exact sequence for $X \hookrightarrow \mathbb{P}^n_B$, twisted by $\mathscr{O}_{\mathbb{P}^n_B}(d)$:

$$0 \to \mathscr{I}_X(d) \to \mathscr{O}_{\mathbb{P}^n_B}(d) \to \mathscr{O}_X(d) \to 0. \tag{28.3.6.1}$$

28.3.G. EXERCISE (CF. EXERCISE 28.3.C). Show that the higher pushforwards (by π) of each term of (28.3.6.1) is 0, and that the long exact sequence of pushforwards of (28.3.6.1) is

$$0 \to \pi_* \mathscr{I}_X(d) \to \pi_* \mathscr{O}_{\mathbb{P}^n_B}(d) \to \pi_* \mathscr{O}_X(d) \to 0.$$

where the middle term is free of rank $\binom{n+d}{d}$ (whose summands can be identified with degree d monomials in the projective variables $x_1, \dots, x_n$ (see Exercise 8.2.K), and the left term $\pi_* \mathscr{I}_X(d)$ is locally free of rank 1 (basically, a line bundle).

(It is helpful to interpret the middle term $\mathscr{O}_B^{\oplus \binom{n+d}{d}}$ as parametrizing homogeneous degree d polynomials in $n+1$ variables, and the rank 1 subsheaf of $\pi_* \mathscr{I}_X(d)$ as "the equation of X". This will motivate what comes next.)

Taking the dual of the injection $\pi_* \mathscr{I}_X(d) \hookrightarrow \mathscr{O}_B^{\oplus \binom{n+d}{d}}$, we have a surjection

$$\mathscr{O}_B^{\oplus \binom{n+d}{d}} \twoheadrightarrow \mathscr{L}$$

from a free sheaf onto an invertible sheaf $\mathscr{L} = (\pi_* \mathscr{I}_X(d))^\vee$, which (by the universal property of projective space) yields a morphism $B \to \mathbb{P}^{\binom{n+d}{d}-1}$.

28.3.H. EXERCISE. Close the loop: show that these two constructions are inverse, thereby proving Proposition 28.3.6.

28.3.7. *Remark.* The proof of the representability of the Hilbert scheme shares a number of features of our arguments about the Grassmannian and the parameter space of hypersurfaces.

CHAPTER 29

Power series and the Theorem on Formal Functions

29.1 Introduction

Power series are a central tool in analytic geometry. Their analog in algebraic geometry, completion, is similarly useful. We will only touch on some aspects of the subject.

In §29.2, we deal with some algebraic preliminaries. In §29.3, we use completions to (finally) give a definition of various singularities, such as nodes. (We won't use these definitions in what follows.) In §29.4, we state the main technical result of the chapter, the Theorem on Formal Functions 29.4.2. The subsequent three chapters give applications. In §29.5, we prove Zariski's Connectedness Lemma 29.5.1 and the Stein Factorization Theorem 29.5.3. In §29.6, we prove a commonly used version of (Grothendieck's version of) Zariski's Main Theorem 29.6.1; we rely on §29.5. In §29.7, we prove Castelnuovo's criterion for contracting (-1)-curves, which we used in Chapter 27. The proof of Castelnuovo's criterion also uses §29.5. Finally, in §29.8, we prove the Theorem on Formal Functions 29.4.2.

There are deliberately many small sections in this chapter, so you can see that they tend not to be as hard at they look, with the exception of the proof of the Theorem on Formal Functions itself, and possibly Theorem 29.2.6 relating completion to exactness and flatness.

29.2 Algebraic preliminaries

Suppose A is a ring, and $J_1 \supset J_2 \supset \cdots$ is a decreasing sequence of ideals. Then we may take the limit $\varprojlim A/J_n$. This ring is often denoted $\hat{A}$; the sequence of ideals is left implicit. The most important case of completion is if $J_n = I^n$, where I is an ideal of A. The limit $\varprojlim A/I^n$ is called the **I-adic completion of A**, or **the completion of A along I** or **at I**.

29.2.1. *Example.* We define $k[[x_1, \dots, x_n]]$ as the completion of $k[x_1, \dots, x_n]$ at the maximal ideal $(x_1, \dots, x_n)$. This is the ring of **formal power series** in n variables over k.

The p-adic numbers (Exercise 1.4.3) are another example.

29.2.A. EXERCISE. Suppose that $J'_1 \supset J'_2 \supset \cdots$ is a decreasing series of ideals that is *cofinal* with J_n. In other words, for every J_n, there is some J'_N with $J_n \supset J'_N$, and for every J'_n, there is some J_N with $J'_n \supset J_N$. Show that there is a canonical isomorphism $\varprojlim A/J'_n \cong \varprojlim A/J_n$. Thus what matters is less the specific sequence of ideals than the "cofinal" equivalence class.

29.2.2. *Preliminary remarks.* We have an obvious morphism $A \to \hat{A}$.

If we put the discrete topology on A/J_n, then $\hat{A}$ naturally has the structure of a **topological ring** (a ring that is a topological space, where all the ring operations are continuous). We then can have the notion of a **topological module** M over a topological ring A' — a module over the underlying ring A', with a topology, such that the action of A' on M is continuous.

In the case of completion at an ideal I, if $A \to \hat{A}$ is an injection, we say that A is I**-adically separated** or **complete with respect to** I, although we won't use this phrase. (For example, the Krull Intersection Theorem implies that if I is a proper ideal of a Noetherian integral domain or a Noetherian local ring, then A is I-adically separated, see Exercise 12.9.A(b).)

29.2.B. EXERCISE. Suppose $\mathfrak{m}$ is a maximal ideal of a ring A. Show that the completion of A at $\mathfrak{m}$ is canonically isomorphic to the completion of $A_{\mathfrak{m}}$ at $\mathfrak{m}$.

If $(A, \mathfrak{m})$ is a local ring, then the natural map $A \to \hat{A}$ is an injection: anything in the kernel must lie in $\cap \mathfrak{m}^i$, which is 0 by the Krull Intersection Theorem (Exercise 12.9.A(b)). Thus "no information is lost by completing", just as analytic functions are (locally) determined by their power series expansion.

29.2.C. EXERCISE. Suppose that $(A, \mathfrak{m})$ is Noetherian local ring containing its residue field k (i.e., it is a k-algebra), of dimension d. Let $x_1, \dots, x_n$ be elements of A whose images are a basis for $\mathfrak{m}/\mathfrak{m}^2$. Show that the map of k-algebras

$$k[[t_1, \dots, t_n]] \to \hat{A} \tag{29.2.2.1}$$

defined by $t_i \mapsto x_i$ is a surjection. (First explain why there *is* such a map!) As usual, for local rings, the completion is assumed to be at the maximal ideal.

Exercise 29.2.C is a special case of the *Cohen Structure Theorem*. (See [**E**, §7.4] for more on this topic.)

29.2.D. EXERCISE. Let X be a locally Noetherian scheme over k, let $p \in X$ be a rational (k-valued) point. Suppose $p \in X$ is regular of codimension d (i.e., $\dim \mathscr{O}_{X,p} = d$). Describe an isomorphism $\hat{\mathscr{O}}_{X,p} \cong k[[x_1, \dots, x_d]]$ as topological rings. (Hint: As in Exercise 29.2.C, choose d elements of $\mathfrak{m}$ that restrict to a basis of $\mathfrak{m}/\mathfrak{m}^2$; these will be your $x_1, \dots, x_d$. Show that the map (29.2.2.1) has no kernel. It may help to identify $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ with $\operatorname{Sym}^n(\mathfrak{m}/\mathfrak{m}^2)$ using Theorem 22.3.8.)

The converse also holds: if $(A, \mathfrak{m}, k)$ is a Noetherian local ring that is a k-algebra, and $\hat{A} \cong k[[x_1, \dots, x_d]]$, then A is a regular local ring of dimension d; see [**AtM**, Prop. 11.24] for the key step.

29.2.3. *Remark.* Suppose p is a smooth rational (k-valued) point of a k-variety of dimension d, and $f \in \mathscr{O}_{X,p}$ (f is a "local function"). By way of the isomorphism of Exercise 29.2.D, we can interpret f as an element of $k[[x_1, \dots, x_d]]$. This should be interpreted as the "power series expansion of f at p", in the "local coordinates x_1, $\dots$, x_d".

29.2.4. Aside: Some more geometric motivation.

Exercise 29.2.D may give some motivation for completion: "in the completion, regular schemes look like affine space". This is often stated in the suggestive language of "formally locally, regular schemes are isomorphic to affine space"; this will be made somewhat more precise in §29.4.1.

We now give a little more geometric motivation for completion, that we will not use later on. Recall from §12.6.1 that étale morphisms are designed to look like "local isomorphisms" in differential geometry. But Exercise 12.6.G showed that this metaphor badly fails in one important way. More precisely, let $Y = \operatorname{Spec} k[t]$, and $X = \operatorname{Spec} k[u, 1/u]$, and let $p \in X$ be given by $u = 1$, and $q \in Y$ be given by $t = 1$. The morphism $\pi : X \to Y$ induced by $t \mapsto u^2$ is étale, and $\pi(p) = q$. But there is no open neighborhood of p that π maps isomorphically onto a neighborhood of q. However, the following Exercise shows that π induces an isomorphism of completions (an isomorphism of "formal neighborhoods").

29.2.E. EXERCISE. Continuing the notation of the previous paragraph, show that π induces an isomorphism of completions $\widehat{\mathscr{O}}_{Y,q} \to \widehat{\mathscr{O}}_{X,p}$.

29.2.F. EXERCISE. Suppose $Y = \operatorname{Spec} A$, and $X = \operatorname{Spec} A[t]/(f(t))$, and $\pi : X \to Y$ is the morphism induced by $A \to A[t]/(f(t))$. Suppose $p \in X$ is a closed point, $q \in Y$ is a closed point, $\pi(p) = q$, and π is étale at p and induces an isomorphism of residue fields. Show that π induces an isomorphism of completions $\widehat{\mathscr{O}}_{Y,q} \to \widehat{\mathscr{O}}_{X,p}$. (If you are not familiar with Hensel's Lemma, you will rediscover its central idea in the course of solving this exercise.)

With a little more care, you can show more generally that if $\pi : X \to Y$ is an étale morphism, $\pi(p) = q$, and π induces an isomorphism of residue fields at p, then π induces an isomorphism of completions $\widehat{\mathscr{O}}_{Y,q} \to \widehat{\mathscr{O}}_{X,p}$. (You may even wish to think about how to remove the hypothesis of isomorphism of residue fields.)

These ideas are close to the definition of "formal étaleness" discussed in §25.2.6.

You can interpret these results as a statement that "the implicit function theorem works formally locally", even though it doesn't work Zariski-locally. (The étale topology is somewhere between these two; it is partially designed so that the implicit function theorem in this sense always holds, essentially by fiat. But this is not the place to discuss the étale topology.)

29.2.5. Completion and exactness.

We conclude this section with an interesting and useful statement.

29.2.6. Theorem. — *Suppose A is a Noetherian ring, and $I \subset A$ is an ideal. For any A-module M, let $\hat{M} = \varprojlim M/I^j M$ be the completion of M with respect to I.*

(a) The completion $\hat{A}$ (of A with respect to I) is flat over A.

(b) If M is finitely generated, then the natural map $\hat{A} \otimes_A M \to \hat{M}$ is an isomorphism.

(c) If $0 \to M \to N \to P \to 0$ is a short exact sequence of finitely generated A-modules, then $0 \to \hat{M} \to \hat{N} \to \hat{P} \to 0$ is exact. (Thus completion preserves exact sequences of finitely generated modules, by Exercise 1.6.E.)

We will use (a) in §29.3, and (b) in §29.8.

29.2.7. *Remark.* Before proving Theorem 29.2.6, we make some remarks. Parts (a) and (b) together clearly imply part (c), but we will use (c) to prove (a) and (b). Also, note a delicate distinction (which helps me remember the statement): if $0 \to M \to N \to P \to 0$ is an exact sequence of A-modules, *not necessarily finitely generated*, then

$$0 \to \hat{A} \otimes_A M \to \hat{A} \otimes_A N \to \hat{A} \otimes_A P \to 0 \tag{29.2.7.1}$$

is *always* exact, but

$$0 \to \hat{M} \to \hat{N} \to \hat{P} \to 0 \tag{29.2.7.2}$$

need *not* be exact — and when it *is* exact, it is often because the modules are finitely generated, and thus (29.2.7.2) is really (29.2.7.1).

Caution: completion is not always exact. Consider the exact sequence of $k[t]$-modules

$$0 \longrightarrow \oplus_{n=1}^{\infty} k[t] \xrightarrow{\times(t,t^2,t^3,\dots)} \oplus_{n=1}^{\infty} k[t] \longrightarrow \oplus_{n=1}^{\infty} k[t]/(t^n) \longrightarrow 0.$$

After completion, the sequence is no longer exact in the middle: $(t^2, t^3, t^4, \dots)$ maps to 0, but is not in the image of the completion of the previous term.

⋆⋆ *Proof.* The key step is to prove (c), which we do through a series of exercises. Suppose that $0 \to M \to N \to P \to 0$ is a short exact sequence of finitely generated A-modules.

29.2.G. EXERCISE. Show that $\hat{N} \to \hat{P}$ is surjective. Hint: consider an element of $\hat{P}$ as a sequence $(p_j \in P/I^j P)_{j \geq 0}$, where the image of p_{j+1} is p_j, cf. Exercise 1.4.A. Build a preimage $(n_j \in N/I^j N)_{j \geq 0}$ by induction on j.

We now wish to identify $\ker(\hat{N} \to \hat{P})$ with $\hat{M}$.

29.2.H. EXERCISE. Show that for each $j \geq 0$,

$$0 \to M/(M \cap I^j N) \to N/I^j N \to P/I^j P \to 0 \tag{29.2.7.3}$$

is exact. (Possible hint: show that $0 \to M \cap I^j N \to M \to N/I^j N \to P/I^j P \to 0$ is exact.)

The short exact sequences (29.2.7.3) form an inverse system as j varies. Its limit is left-exact (because limits always are), but it is also right-exact by Exercise 1.6.M, as the "transition maps on the left" $M/(M \cap I^{j+1} N) \to M/(M \cap I^j N)$ are clearly surjective. Thus

$$0 \to \varprojlim M/(M \cap I^j N) \to \hat{N} \to \hat{P} \to 0 \tag{29.2.7.4}$$

is exact. To complete the proof of (c), it suffices (by Exercise 29.2.A) to show that the sequence of submodules $I^j M$ is cofinal with the sequence $M \cap I^j N$, so that the term $\varprojlim M/(M \cap I^j N)$ on the left of (29.2.7.4) is naturally identified with $\hat{M}$.

29.2.I. EXERCISE. Prove this. Hint: clearly $I^j M \subset M \cap I^j N$. By Corollary 12.9.4 to the Artin-Rees Lemma 12.9.3, for some integer s, $M \cap I^{j+s} N = I^j (M \cap I^s N)$ for all $j \geq 0$, and clearly $I^j (M \cap I^s N) \subset I^j M$.

This completes the proof of part (c) of Theorem 29.2.6.

For part (b), present M as

$$A^{\oplus m} \xrightarrow{\alpha} A^{\oplus n} \longrightarrow M \longrightarrow 0 \tag{29.2.7.5}$$

where α is an $m \times n$ matrix with coefficients in A. Completion is exact by part (c), and commutes with direct sums, so

$$\hat{A}^{\oplus m} \longrightarrow \hat{A}^{\oplus n} \longrightarrow \hat{M} \longrightarrow 0$$

is exact. Tensor product is right-exact, and commutes with direct sums, so

$$\hat{A}^{\oplus m} \longrightarrow \hat{A}^{\oplus n} \longrightarrow \hat{A} \otimes_A M \longrightarrow 0$$

is exact as well. Notice that the maps from $\hat{A}^{\oplus m}$ to $\hat{A}^{\oplus n}$ in both right-exact sequences are the same; they are both α. Thus their cokernels are identified, and (b) follows.

Finally, to prove (a), we need to extend the ideal-theoretic criterion for flatness (Theorem 24.4.1) slightly. Recall (§24.4.2) that it is equivalent to the fact that an A-module M is flat if and only if for all ideals I, the natural map $I \otimes_A M \to M$ is an injection.

29.2.J. EXERCISE (STRONGER FORM OF THE IDEAL-THEORETIC CRITERION FOR FLATNESS). Show that an A-module M is flat if and only if for all *finitely generated ideals* I, the natural map $I \otimes_A M \to M$ is an injection. (Hint: if there is a counterexample for an ideal J that is not finitely generated, use it to find another counterexample for an ideal I that *is* finitely generated.)

By this criterion, to prove (a) it suffices to prove that the multiplication map $I \otimes_A \hat{A} \to \hat{A}$ is an injection for all finitely generated ideals I. But by part (b), this is the same showing that $\hat{I} \to \hat{A}$ is an injection; and this follows from part (c). $\square$

29.3 Defining types of singularities

Singularities are best defined in terms of completions. As an important first example, we finally define "node".

29.3.1. *Definition.* Suppose X is a dimension 1 variety over $\overline{k}$, and $p \in X$ is a closed point. We say that X has a **node** at p if the completion of $\mathscr{O}_{X,p}$ at $\mathfrak{m}_{X,p}$ is isomorphic (as topological rings) to $\overline{k}[[x,y]]/(xy)$.

29.3.A. EXERCISE. Suppose $k = \overline{k}$ and $\operatorname{char} k \neq 2$. Show that the curve in $\mathbb{A}^2_k$ cut out by $(y^2 - x^2 - x^3)$, (which we have been studying repeatedly since Figure 7.4) has a node at the origin.

29.3.B. EXERCISE. Suppose $k = \overline{k}$ and $\operatorname{char} k \neq 2$, and we have $f(x,y) \in k[x,y]$. Show that $\operatorname{Spec} k[x,y]/(f(x,y))$ has a node at the origin if and only if f has no terms of degree 0 or 1, and the degree 2 terms are not a perfect square. (This generalizes Exercise 29.3.A.)

The definition of node outside the case of varieties over algebraically closed fields is more problematic, and we give some possible ways forward. For varieties over a non-algebraically closed field k, one can always base change to the closure $\overline{k}$. As an alternative approach, if p is a k-valued point of a variety over k (not necessarily algebraically closed), then we could take the same Definition 29.3.1; this might reasonably be called a *split node*, because the branches (or more precisely, the tangent directions) are distinguished. Those singularities that are not split nodes, but which become nodes after base change to $\overline{k}$ (such as the origin in $\operatorname{Spec} \mathbb{R}[x,y]/(x^2+y^2)$) might reasonably be called *non-split nodes*.

29.3.2. Other singularities. We may define other singularities similarly. To avoid complications, we do so over an algebraically closed field.

29.3.3. *Definition.* Suppose X is a variety over $\overline{k}$, and p is a closed point of X, where $\operatorname{char} k \neq 2,3$. We say that X has a **cusp** (resp. **tacnode, triple point**) at p if $\widehat{\mathscr{O}}_{X,p}$ is isomorphic to the completion of the curve $\operatorname{Spec} \overline{k}[x,y]/(y^2-x^3)$ (resp. $\operatorname{Spec} \overline{k}[x,y]/(y^2-x^4)$, $\operatorname{Spec} \overline{k}[x,y]/(y^3-x^3)$). We say that X has an **ordinary multiple point of multiplicity** m, or **ordinary** m**-fold point**, if $\widehat{\mathscr{O}}_{X,p}$ is isomorphic to the completion of the curve $\operatorname{Spec} \overline{k}[x,y]/(f(x,y))$ where f is a homogeneous polynomial of degree m with no repeated roots. (You can see quickly that an ordinary 2-fold point — or *ordinary double point* — is precisely a node, and an ordinary 3-fold point is a triple point.)

29.3.C. TRIVIAL EXERCISE. (For this exercise, work over an algebraically closed field for simplicity.) Define A_n curve singularity (see §22.4.4). Define A_n surface singularity (see Exercise 22.4.F). Define D_n, E_6, E_7, and E_8 surface and curve singularities (see Remark 22.4.5).

29.3.4. Using this definition. We now give an example of how this definition can be used.

29.3.D. EXERCISE. Suppose X is a $\overline{k}$-variety with a node at a closed point p. Show that the blow up of X at p yields a morphism $\beta : \tilde{X} \to X$, where the exceptional divisor $\beta^{-1}p$ consists of two reduced smooth points. Hint: use the fact that completion is flat (Theorem 29.2.6(a)), and that blowing up commutes with flat base change (Exercise 24.2.P(a)), to turn this into a calculation on the "formal model" of the node, $\operatorname{Spec} k[[x,y]]/(xy)$.

29.3.E. EXERCISE. Continuing the terminology of the previous exercise, describe an exact sequence

$$0 \to \mathscr{O}_X \to \beta_* \mathscr{O}_{\tilde{X}} \to \mathscr{O}_p \to 0. \tag{29.3.4.1}$$

Hint: cohomology commutes with flat base change, Theorem 24.2.8 (or more simply, the affine case version Exercise 18.8.B(b)).

29.3.F. EXERCISE. We continue the terminology of the previous two exercises. If X is a pure-dimensional reduced projective curve, show that $p_a(X) = p_a(\tilde{X}) + 1$.

Thus "resolving a node" of a curve reduces the arithmetic genus of the curve by 1. If you wish, you can readily show that "resolving a cusp" reduces the genus by 1, and "resolving a tacnode" reduces the genus by 2. In general, for each type of curve singularity, the contribution it makes to the genus — the difference between the genus of the curve and that of its normalization, when the curve is projective — is called the δ-**invariant**. Thus δ for a node or cusp is 1, and δ for a tacnode is 2.

29.3.G. EXERCISE. Show that for *any* singularity type (other than a smooth point), $\delta > 0$.

29.3.H. EXERCISE. Show that a reduced irreducible degree d plane curve can have at most $\binom{d-1}{2}$ singularities.

29.3.I. EXERCISE. Here is a trick for speedily working out the genus of a nodal (projective reduced) curve when it comes up in examples. (The description will be informal.) Assuming it is possible, draw a sketch on a piece of paper (i.e., the real plane) where nodes "look like nodes", and a curve of genus g is drawn with g holes, as in Figure 29.1. (In Figure 29.1, we see three irreducible components, of geometric genus 0, 2, and 3 respectively. The first two components are smooth, while the third meets itself twice. There are also three additional nodes, joining each pair of the irreducible components.) Then if the curve is connected, count the number of holes visible in the picture. Prove that this recipe works. (Feel free to figure out what to do if the curve is not connected.)

[picture to be made later]

FIGURE 29.1. This curve has genus 8

This can be quite useful. For example, the sketches in Figure 29.2 can remind you that a degree d plane curve has arithmetic genus $\binom{d-1}{2}$, the curve on $\mathbb{P}^1 \times \mathbb{P}^1$ coming from a section of $\mathscr{O}(a, b)$ (Exercise 9.6.A) has genus $(a-1)(b-1)$, and a curve in class $2C + 2F$ on the Hirzebruch surface $\mathbb{F}_2$ (using the language of Exercise 20.2.M) has arithmetic genus 3. In each case, we are taking a suitable curve in the class, taking advantage of the fact that arithmetic genus is constant in flat families (Crucial Exercise 24.7.B).

[picture to be made later]

FIGURE 29.2. These curves have arithmetic genus 6, 6, and 3 respectively

29.3.5. Other definitions. If $\overline{k} = \mathbb{C}$, then definitions of this sort agree with the analytic definitions (see [**Ar1**, §1]). For example, a complex algebraic curve singularity is a node if and only if it is *analytically* isomorphic to a neighborhood of $xy = 0$ in $\mathbb{C}^2$. There is also a notion of isomorphism "étale-locally", which we do not define here. Once again, this leads to the same definition of these types of singularities (see [**Ar2**, §2]).

29.4 The Theorem on Formal Functions

Suppose $\pi : X \to Y$ is proper morphism of locally Noetherian schemes, and $\mathscr{F}$ is a coherent sheaf on X, so $R^i\pi_*\mathscr{F}$ is a coherent sheaf (Theorem 18.9.1). Fix a point

$q \in Y$. We already have a sense that there is an imperfect relationship between the fiber of $R^i\pi_*\mathscr{F}$ at q and the cohomology of $\mathscr{F}$ "on the fiber X_q", and much of Chapter 28 was devoted to making this precise.

The Theorem of Formal Functions deals with this issue in a different way. Rather than comparing the fiber of $R^i\pi_*\mathscr{F}$ with the cohomology of $\mathscr{F}$ restricted to the fiber, it gives a precise isomorphism between information on "infinitesimal thickenings" of the situation.

29.4.1. We now make this more precise. If Z is a closed subscheme of W cut out by ideal sheaf $\mathscr{I}$, we say that the closed subscheme cut out by $\mathscr{I}^{n+1}$ is the **nth order formal** (or **infinitesimal**) **neighborhood** (or **thickening**) of Z (in W). The phrase "formal neighborhood" without mention of an "order" refers to the information contained in all of these neighborhoods at once, often in the form of a limit of the sort we will soon describe. (This also leads us to the notion of *formal schemes*, an important notion we will nonetheless not need.)

We turn now to our situation of interest. Rather than the fiber of $R^i\pi_*\mathscr{F}$ at q, we consider its completion: we take the stalk at q, and complete it at $\mathfrak{m}_{Y,q}$. Rather than the cohomology of $\mathscr{F}$ along the fiber, we consider the cohomology of $\mathscr{F}$ when restricted to the nth formal neighborhood X_n of the fiber, and take the limit. The Theorem of Formal Functions says that these are canonically identified.

Before we state the Theorem of Formal Functions, we need to be precise about what these limits are, and how they are defined. To concentrate on the essential, we do this in the special case where q is a *closed* point of Y, and leave the (mild) extension to the general case to you (Exercise 29.4.A).

We deal first with $R^i\pi_*\mathscr{F}$. It is helpful to note that by Exercise 29.2.B, we can compute the restriction to the nth formal neighborhood either by restricting/tensoring to the stalk $\mathscr{O}_{Y,q}$, or in any affine neighborhood of $q \in Y$. For convenience, we pick an affine neighborhood $\operatorname{Spec} B \subset Y$ of q. (Again, by Exercise 29.2.B, it won't matter which affine open subset we take.) Define $X_B := \pi^{-1}(\operatorname{Spec} B)$ for convenience. Let $\mathfrak{m} \subset B$ be the maximal ideal corresponding to q. Then $R^i\pi_*(\mathscr{F})\hat{}_q$ is canonically the completion of the coherent B-module $H^i(X_B, \mathscr{F}|_{X_B})$ at $\mathfrak{m}$:

$$R^i\pi_*(\mathscr{F})\hat{}_q = \varprojlim H^i(X_B, \mathscr{F}|_{X_B})/\mathfrak{m}^n H^i(X_B, \mathscr{F}|_{X_B}).$$

We turn next to the cohomology of $\mathscr{F}$ on thickenings of the fiber. We have closed embeddings $X_n \hookrightarrow X_{n+1}$, and thus maps of cohomology groups $H^i(X_{n+1}, \mathscr{F}|_{X_{n+1}}) \to H^i(X_n, \mathscr{F}|_{X_n})$. We have base change maps

$$H^i(X_B, \mathscr{F}|_{X_B})/\mathfrak{m}^n H^i(X_B, \mathscr{F}|_{X_B}) \to H^i(X_n, \mathscr{F}|_{X_n}) \tag{29.4.1.1}$$

(see (28.1.0.3)) such that the square

$$\begin{tikzcd}
H^i(X_B, \mathscr{F}|_{X_B})/\mathfrak{m}^{n+1} H^i(X_B, \mathscr{F}|_{X_B}) \arrow[r] \arrow[d] & H^i(X_{n+1}, \mathscr{F}|_{X_{n+1}}) \arrow[d] \\
H^i(X_B, \mathscr{F}|_{X_B})/\mathfrak{m}^n H^i(X_B, \mathscr{F}|_{X_B}) \arrow[r] & H^i(X_n, \mathscr{F}|_{X_n})
\end{tikzcd}$$

commutes. (Do you see why? Basically, this is again because they are base change maps.) Thus we have an induced map of limits:

$$R^i\pi_*(\mathscr{F})\hat{}_q \to \varprojlim H^i(X_n, \mathscr{F}_n). \tag{29.4.1.2}$$

29.4.A. EXERCISE. Extend the previous discussion to the case where q is not closed.

The Theorem of Formal Functions states that this is an isomorphism.

29.4.2. Theorem on Formal Functions. — *Suppose $\pi : X \to Y$ is a proper morphism of locally Noetherian schemes, $\mathscr{F}$ is a coherent sheaf on X, and $q \in Y$. Then (29.4.1.2) is an isomorphism for all $i \geq 0$.*

Warning: the Theorem on Formal Functions does *not* imply anything about the maps "at finite level", i.e., (29.4.1.1).

The proof of Theorem 29.4.2 is quite subtle, and is postponed to the double-starred §29.8. We first give some important applications.

29.5 Zariski's Connectedness Lemma and Stein Factorization

We now state and prove Zariski's Connectedness Lemma, which was mentioned in §28.1.8.

29.5.1. Zariski's Connectedness Lemma. — *If a proper morphism $\pi : X \to Y$ of locally Noetherian schemes is $\mathscr{O}$-connected (see §28.1.7), then $\pi^{-1}(q)$ is connected for every $q \in Y$.*

The proof requires the following lemma.

29.5.A. EASY EXERCISE (THE COMPLETION OF A LOCAL RING IS A LOCAL RING). Suppose $(A, \mathfrak{m})$ is a Noetherian local ring. Show that the completion of A along $\mathfrak{m}$, $\hat{A} := \varprojlim A/\mathfrak{m}^n$, is a local ring, with maximal ideal $\mathfrak{m}\hat{A}$. Hint: show that any element of $\hat{A}$ not in the maximal ideal is invertible.

Proof. Assume otherwise that there is some $q \in Y$ such that $\pi^{-1}(q)$ is not connected, say $\pi^{-1}(q) = X_1 \coprod X_2$ (where X_1 and X_2 are nonempty open subsets). Then the nth order formal neighborhood of $\pi^{-1}(q)$, having the same topological space, is also disconnected. We use the useful trick of idempotents (Remark 3.6.3). Let e_1 be the function 1 on X_1 and 0 on X_2, and let e_2 be the function 1 on X_2 and 0 on X_1. These functions makes sense for any order formal neighborhood, and they have natural images in the inverse limit. Thus we get elements $e_1, e_2 \in \hat{\mathscr{O}}_y$ with $e_1 + e_2 = 1$, $e_1 e_2 = 0$. Now, $\hat{\mathscr{O}}_y$ is a local ring (Exercise 29.5.A). But $e_1 e_2 = 0$ implies that neither is invertible, so both are in the maximal ideal, and hence $e_1 + e_2$ can't be 1. We thus have a contradiction. □

29.5.2. Stein factorization. We next show the existence of a Stein factorization. We could have given this construction long before, but now Zariski's Connectedness Lemma 29.5.1 will give us some impressive consequences.

29.5.3. Stein Factorization Theorem. — *Any proper morphism $\pi : X \to Y$ of locally Noetherian schemes can be factored into $\beta \circ \alpha$, where $\alpha : X \to Y'$ satisfies $\mathscr{O}_{Y'} = \alpha_* \mathscr{O}_X$ (hence has connected fibers, by Zariski's Connectedness Lemma 29.5.1), and $\beta : Y' \to Y$ is* *a finite morphism.*

(29.5.3.1)
$$\begin{tikzcd} X \arrow[rr, "\mathscr{O}\text{-conn.}", "\alpha"'] \arrow[dr, "\pi"'] & & Y' \arrow[dl, "\beta", "\text{finite}"'] \\ & Y & \end{tikzcd}$$

We note that by the Cancellation Property 10.1.19 for projective morphisms, if π is projective, then so is α (as δ_β, being a closed embedding, is projective).

Although it is not in the statement of the theorem, the proof produces a *specific* factorization, which is called the **Stein factorization** of π. The picture to have in mind is that the Stein factorization (roughly) contracts the continuous parts of the fibers, in some canonical way.

As usual, the Noetherian hypotheses can be removed (see [**Stacks**, tag 03H2]).

Proof of the Stein Factorization Theorem 29.5.3. By Grothendieck's Coherence Theorem 18.9.1, $\pi_* \mathscr{O}_X$ is a coherent sheaf (of algebras) on Y. Define $Y' := \mathscr{S}pec\, \pi_* \mathscr{O}_X$, so (as $\pi_* \mathscr{O}_X$ is finite type) the structure morphism $\beta : Y' \to Y$ is finite. We have a factorization

$$\begin{tikzcd} X \arrow[rr, "\alpha"'] \arrow[dr, "\pi"'] & & Y' \arrow[dl, "\beta", "\text{finite}"'] \\ & Y & \end{tikzcd}$$

Finally, we show that α is $\mathscr{O}$-connected. By the quasicoherence of both $\mathscr{O}_{Y'}$ and $\alpha_* \mathscr{O}_X$, it suffices to check on an affine cover of Y'. We choose as our affine cover of Y' a pullback of an affine cover of Y. If U is an affine open subset of Y', then on $\beta^{-1}(U)$, on the level of rings, $\mathscr{O}_{Y'} \to \alpha_* \mathscr{O}_X$ is precisely the isomorphism $\Gamma(U, \pi_* \mathscr{O}_X) \to \Gamma(U, \pi_* \mathscr{O}_X)$. ∎

29.5.B. EXERCISE. Suppose $\pi : X \to Y$ is a proper morphism of locally Noetherian schemes, that is an isomorphism over a dense open subset of Y. Suppose further that Y is normal. Show that π is $\mathscr{O}$-connected. (This applies, for example, to blow-ups of smooth varieties under smooth centers.) Hence, by Zariski's Connectedness Lemma 29.5.1, for every $q \in Y$, $\pi^{-1}(q)$ is connected. This is called **Zariski's Main Theorem** (the version for birational morphisms). Hint: Exercise 9.7.P.

29.5.C. EXERCISE. Suppose $\pi : X \to Y$ and $\pi' : X \to Y'$ are two proper morphisms of locally Noetherian schemes, each contracting the same *connected* set of X (and an isomorphism elsewhere), and suppose Y and Y' are normal. Show that π and π' are the same (or more precisely, that there is an isomorphism $i : Y \to Y'$ such that $\pi' = i \circ \pi$). Informal translation: If $X \to Y$ has connected fibers, then Y is determined by just knowing the locus contracted on X. Hint: Identify Y and Y' as topological spaces, and then identify π and π' as maps of topological spaces. Use Stein factorization to recover the structure sheaf as $\pi_* \mathscr{O}_X$.

29.5.D. EXERCISE. Show that the construction of the Stein factorization of a morphism $\pi : X \to Y$ of locally Noetherian schemes commutes with any flat base change $Z \to Y$ of locally Noetherian schemes.

29.5.4. Resolution of singularities of curves.

If C is a reduced projective curve over a field k, then we can resolve its singularities by normalization: $\nu : \tilde{C} \to C$. (See Remark 22.4.6 for some discussion of resolution of singularities, and Theorem 12.5.8 for the reason why normal curves are nonsingular.) But as we have seen from examples, it requires luck and insight to figure out how to take an integral closure, and I often have neither. We now take the (inspired) guesswork out of desingularization by explaining how to desingularize by blowing up. The algorithm is simple: we find a singular point, then blow it up, then look for more singular points to repeat the process. Clearly if there are no singular points to be found, then we are done; the problem is to show that this is guaranteed to terminate. We do this by making use of an integer invariant: the arithmetic genus. We must show (i) that if $p_a(C) = p_a(\tilde{C})$, then ν is an isomorphism — C is already nonsingular; (ii) if $C' \to C$ is a finite morphism, then $p_a(C') \leq p_a(C)$ (so $p_a(\tilde{C})$ will be a lower bound for the arithmetic genus throughout the process); and (iii) that blowing up a singular point decreases the arithmetic genus.

To set up these result, we consider the following situation. Suppose $\pi : C' \to C$ is a finite morphism (where C is as described in the previous paragraph, and C' is a reduced curve), that is an isomorphism away from a finite subset of C. Then the pullback map $\mathscr{O}_C \to \pi_* \mathscr{O}_{C'}$ has no kernel, as $\mathscr{O}_C$ is torsion-free, and the map is an isomorphism on the generic points of the components of C. The cokernel $\mathscr{G}$ is supported on a finite set of points, and thus $H^i(C, \mathscr{G}) = 0$ for $i > 0$, and if $H^i(C, \mathscr{G}) = 0$, then $\mathscr{G} = 0$. From the exact sequence

$$0 \to \mathscr{O}_C \to \pi_* \mathscr{O}_{C'} \to \mathscr{G} \to 0,$$

and the fact that π is affine, $\chi(C', \mathscr{O}_{C'}) = \chi(C, \mathscr{O}_C) + \chi(C, \mathscr{G})$, from which $p_a(C) \leq p_a(C')$, with equality if and only if π is $\mathscr{O}$-connected (i.e., $\mathscr{G} = 0$).

29.5.E. EXERCISE. Show that if π is $\mathscr{O}$-connected, then π is an isomorphism. Hint: use Zariski's Connectedness Lemma 29.5.1 to show that the fibers are connected, and hence that π is bijective. Identify C and C' as topological spaces. Then use the isomorphism $\mathscr{O}_C \to \pi_* \mathscr{O}_C$ to identify C and C' as ringed spaces.

To complete our strategy, it remains to show that if $\beta : C' \to C$ is the blow-up of C at a point p, and β is an isomorphism, then C is already nonsingular at that point. But if β is an isomorphism, then p is the exceptional divisor, and hence effective Cartier, and hence cut out locally by a singular equation. But then $\mathfrak{m}_{C,p} \subset \mathscr{O}_{C,p}$ is principal, so $\mathscr{O}_{C,p}$ is a regular local ring of dimension 1.

(You should verify that we have completed our strategy. Can you extend this argument to the case where C is not projective?)

Resolution of singularities for surfaces will be briefly discussed in §29.7.4.

29.6 Zariski's Main Theorem

Zariski's Main Theorem is misnamed in all possible ways (although not horribly so). It is not a single *Theorem*, because there are many results that go by this name, and they often seem quite unrelated. What they have in common is that they are intellectual descendants of a particular result of Zariski. This result is not the *Main* Theorem of Zariski's remarkable career; the name comes because it was the Main Theorem of a particular paper, [**Z**]. And finally, it is not *Zariski's* in any stronger sense than this; the modern versions are due to Grothendieck.

We have already seen one version of Zariski's Main Theorem: the "birational" form of Exercise 29.5.B. The goal of this section is to prove a strong statement that is close to the optimal version of the one most often used.

Before we start, we note that the fiber of a proper morphism over a point $q \in Y$ is a proper (hence finite type) scheme over the residue field $\kappa(q)$, and we say that a point is *isolated in its fiber* if it forms a component of the fiber of dimension 0.

29.6.1. Zariski's Main Theorem (Grothendieck's form). — *Suppose $\pi : X \to Y$ is a proper morphism of locally Noetherian schemes.*
(a) The set of points of X that are isolated in their fiber forms an open subset $X_0 \subset X$.
(b) The morphism $\pi|_{X_0} : X_0 \to Y$ factors into an open embedding followed by a finite morphism.

(29.6.1.1)
$$\begin{tikzcd} X_0 \arrow[r, hook, "\text{open}"] \arrow[dr, "\pi"'] & Y' \arrow[d, "\text{finite}"] \\ & Y \end{tikzcd}$$

As with Stein factorization, the proof of the theorem yields a *specific* factorization.

If π is a morphism of *varieties*, we already know part (a), by upper semicontinuity of fiber dimension Theorem 11.4.2(a). For this reason, the proof in the case of varieties is easier, and (even if you are interested in the more general case) you are advised to read the proof in this simpler case, to concentrate on the main ideas. But even in the case of varieties, we will need some of the ideas arising in the proof of (a), so you should read it until advised to skip ahead to (b).

Proof. (a) Take the Stein factorization (29.5.3.1) of π:

$$\begin{tikzcd} X \arrow[r, "\mathscr{O}\text{-conn.}", "\alpha"'] \arrow[dr, "\pi"'] & Y' \arrow[d, "\beta", "\text{finite}"'] \\ & Y. \end{tikzcd}$$

By the finiteness of β, a point of X is an isolated point in its fiber of π if and only if it is an isolated point in its fiber of α. Thus we may replace by π by α. (Basically, we are reducing to the case where π is $\mathscr{O}$-connected, but we prefer to carefully call the morphism under consideration α so we can return to considering π in the proof of (b).)

Because fibers of proper $\mathscr{O}$-connected morphisms of locally Noetherian schemes are connected (Zariski's Connectedness Lemma 29.5.1), a point is isolated in its fiber of α if and only if it *is* its fiber.

If you are considering only the variety case, at this point you should jump to the proof of part (b).

To show that X_0 is open, we will identify it with the locus where $\Omega_\alpha = 0$ (the unramified locus of α — not of π!), which is open by Exercise 21.6.H.

29.6.A. EASY EXERCISE. If $\Omega_\alpha|_p = 0$, show that p is isolated in its fiber of α (and hence of π).

We now show the other direction. Suppose p is isolated in its fiber, and let $r \in Y'$ be $\alpha(p)$. We will show that α induces an isomorphism $\alpha^\sharp : \mathscr{O}_{Y',r} \to \mathscr{O}_{X,p}$, so $\Omega_{\mathscr{O}_{X,p}/\mathscr{O}_{Y',r}} = 0$. As the construction of Ω behaves well with respect to localization on the source and target (Exercise 21.2.L), this implies that $(\Omega_\alpha)_p = 0$, so Ω_α is 0 at p.

The stalk $\mathscr{O}_{Y',r}$ is obtained by taking the limit (of sections of $\mathscr{O}$) over all open subsets of Y' containing r. For any such open subset $V \subset Y'$, the condition $\alpha_* \mathscr{O}_X = \mathscr{O}_{Y'}$ gives an isomorphism $\Gamma(\mathscr{O}_{Y'}, V) \cong \Gamma(\mathscr{O}_X, \alpha^{-1}(V))$ (compatible with inclusions). We will show that the system of open subsets $\alpha^{-1}(V)$ (as V varies through neighborhoods of r), which each necessarily contain p, is cofinal with the system of neighborhoods of p in X. To do this, we must show that for any neighborhood U of p in X, there is a neighborhood V of q in Y such that $\alpha^{-1}(V) \subset U$. To do this, note that $X \setminus U$ is closed in X, so $\alpha(X \setminus U)$ is closed in Y (as α is proper), so its complement $V := Y \setminus \alpha(X \setminus U)$ is open. Simply note that $\alpha^{-1}(V) \subset U$ (implicitly using that α, being proper and $\mathscr{O}$-connected, is surjective, see Exercise 28.1.G), and we are done.

(b) In our proof of (a), we have established that α gives a bijection between X_0 and an open subset of Y'. Notice that this is furthermore a homeomorphism: any open subset of Y' pulls back to an open subset of X_0 by the continuity of α; and if U is an open subset of X_0, then $\alpha(X_0 \setminus U)$ is a closed subset of Y', so its complement (which is $\alpha(U)$, using that proper $\mathscr{O}$-connected morphisms are surjective, Exercise 28.1.G) is open.

Using this isomorphism of topological spaces, the condition of $\mathscr{O}$-connectedness of α (restricted to $\alpha(X_0)$) shows that α gives an isomorphism of ringed spaces (i.e., of schemes) $X_0 \to \alpha(X_0)$. We have found our desired factorization. □

29.6.B. EXERCISE. Prove Zariski's Main Theorem with Y affine, say $\operatorname{Spec} A$, and X a quasiprojective A-scheme. (Hint: by the Definition 4.5.9 of quasiprojective A-scheme, we can find an open embedding $X \hookrightarrow X'$ into a projective A-scheme. Apply Theorem 29.6.1 to $X' \to \operatorname{Spec} A$.)

For the next two applications (Exercises 29.6.D and 29.6.E), we will need an annoying hypothesis, which is foreshadowed by the previous exercise. We say that a morphism $\pi : X \to Y$ of locally Noetherian schemes satisfies (†) if for all affines U_i in some open cover of Y, the restriction $\pi|_{\pi^{-1}(U_i)} : \pi^{-1}(U_i) \to U_i$ factors as an open embedding into a scheme proper over U_i. A morphism satisfying (†) is necessarily separated and finite type. It turns out that this is sufficient: Nagata's compactification theorem states that *every* separated finite type morphism of Noetherian schemes $\pi : X \to Y$ can be factored into an open embedding into a scheme proper over Y. (See [**Lü**] for a proof. The Noetherian hypotheses can be replaced by the condition that Y is quasicompact and quasiseparated, see [**Co**].)

29.6.C. EXERCISE. Show that if $\pi : X \to Y$ is a morphism of varieties over k, and X is an open subset of a proper variety Z, then π satisfies (†). (Hint: X is an open subscheme of its closure in $Z \times_k Y$.)

29.6.D. EXERCISE. Suppose $\pi : X \to Y$ is a quasifinite birational morphism of integral locally Noetherian schemes, and that Y is normal. Suppose further π satisfies $(\dagger)$. Show that π is an open embedding. (Hence if furthermore π is a bijection, then π must be an isomorphism.)

29.6.E. EXERCISE. Suppose $\pi : X \to Y$ is a quasifinite morphism of Noetherian schemes satisfying $(\dagger)$. Show that π factors as an open embedding into a finite morphism.

$$\begin{tikzcd} X \arrow[rr, hook, "\text{open}"] \arrow[dr, "\pi"'] & & Y' \arrow[dl, "\text{finite}"] \\ & Y & \end{tikzcd}$$

This shows that it is reasonable to think of quasifinite morphisms as "open subsets of finite morphisms".

As a final application of Zariski's Main Theorem, we can finally prove a characterization of finite morphisms that we have mentioned a number of times.

29.6.2. Theorem. — *Suppose $\pi : X \to Y$ is a morphism of locally Noetherian schemes. The following are equivalent.*

(a) *π is finite.*
(b) *π is affine and proper.*
(c) *π is proper and quasifinite.*

As usual, Noetherian conditions can be removed: [**Gr-EGA**, IV_3.8.11.1] shows that proper, quasifinite, locally finitely presented morphisms are finite.

Proof. Finite morphisms are affine by definition, and proper by Proposition 10.3.3, so (a) implies (b).

To show that (b) implies (c), we need only show that affine proper morphisms have finite fibers (as the "finite type" part of the definition of quasifiniteness is taken care of by properness). This was shown in Exercise 18.9.A.

Finally, we assume (c), that π is proper and quasifinite, and show (a), that π is finite. By Zariski's Main Theorem 29.6.1, we have a factorization (29.6.1.1).

$$\begin{tikzcd} X_0 \arrow[rr, hook, "\text{open}", "\gamma"'] \arrow[dr, "\pi"'] & & Y' \arrow[dl, "\text{finite}"] \\ & Y & \end{tikzcd}$$

Now $X = X_0$. (Do you see why? Hint: Exercises 7.4.D followed by Exercise 7.3.H.) By the Cancellation Theorem 10.1.19 for proper morphisms (using that $Y' \to Y$ is finite hence separated), γ is proper.

29.6.F. EXERCISE. Show that proper quasicompact open embeddings are closed embeddings. Hint: use Corollary 8.3.5 to show that the image is open and closed in the target.

Applying this to our situation, we see that $X \to Y$ is the composition of two finite morphisms $\gamma : X \hookrightarrow Y'$ and $Y' \to Y$, and is thus finite itself. $\square$

29.6.3. $\star\star$ **Other versions of Zariski's Main Theorem.**

The Noetherian conditions in Theorem 29.6.1 can be relaxed (see [**Gr-EGA**, IV_3.8.12.6] or [**GW**, Thm. 12.73]): if $\pi : X \to Y$ is a separated morphism of finite type, and Y is quasicompact and quasiseparated, then the set of points X_0 isolated in their fiber is open in X, and for every quasicompact open subset U of X_0, there is a factorization

$$\begin{tikzcd} X \arrow[rr, "\gamma"] \arrow[dr, "\pi|_U"'] & & Y' \arrow[dl, "\beta", "\text{finite}"'] \\ & Y & \end{tikzcd}$$

such that β is finite, $\gamma|_U$ is a quasicompact open embedding, and $\gamma^{-1}(\gamma(U)) = U$.

There are also topological and power series forms of Zariski's Main Theorem, see [**Mu7**, §III.9].

29.7 Castelnuovo's criterion for contracting (-1)-curves

We now prove a result used in Chapter 27. We showed in Exercise 22.4.O that if we blow up a regular point of a surface at a (reduced) point, the exceptional divisor is a (-1)-curve (see Definition 20.2.6: it is isomorphic to $\mathbb{P}^1$, and has normal bundle $\mathscr{O}_C(-1)$). Castelnuovo's criterion is the converse: if we have a quasiprojective surface containing a (-1)-curve, that surface is obtained by blowing up another surface at a reduced regular point. We say that we can *blow down* the (-1)-curve.

29.7.1. Theorem (Castelnuovo's Criterion). — *Let $C \subset X$ be a (-1)-curve on a smooth projective surface over k. Then there exists a birational morphism $\pi : X \to Y$ such that Y is a smooth projective surface, $\pi(C)$ is a k-valued point, π is the blow-up of Y at that point, and C is the exceptional divisor of the blow-up.*

By Exercise 29.5.C, there is only one way to "blow down C": this contraction is unique.

Proof. The proof is in three steps. *Step 1.* We construct $\pi : X \to Y$ that contracts C to a point q (and is otherwise an isomorphism). At this point we will know that Y is a projective variety. *Step 2.* Then we show that Y is smooth; this is the hard step, and requires the Theorem of Formal Functions 29.4.2. *Step 3.* Finally, we recognize π as the blow-up $\mathrm{Bl}_q Y \to Y$.

Step 1. We do this by finding an invertible sheaf $\mathscr{L}$ whose complete linear series will do the job. We start with a very ample invertible sheaf $\mathscr{H}$ on X such that $H^1(X, \mathscr{H}) = 0$. We can find such an invertible sheaf $\mathscr{H}$ by choosing any very ample invertible sheaf, and then taking a suitably large multiple, invoking Exercise 16.6.C (very ample $\otimes$ very ample = very ample) and Serre vanishing (Theorem 18.1.4(ii)). Let $k = \mathscr{H} \cdot C$; $k > 0$ by the very ampleness of $\mathscr{H}$ (Exercise 20.1.K).

29.7.A. EXERCISE. Show that $H^1(X, \mathscr{H}(iC)) = 0$ for $0 \le i \le k$. Hint: use induction on i, using

$$0 \to \mathscr{H}((i-1)C) \to \mathscr{H}(iC) \to \mathscr{H}(iC)|_C \to 0. \tag{29.7.1.1}$$

Note that $\mathscr{H}(iC)$ as a line bundle on $C \cong \mathbb{P}^1$; which $\mathscr{O}(n)$ is it?

Define $\mathscr{L} := \mathscr{H}(kC)$. We will use the sections of $\mathscr{H}$ to obtain sections of $\mathscr{L}$, via the map $\phi : H^0(X, \mathscr{H}) \to H^0(X, \mathscr{H}(kC))$.

29.7.B. EXERCISE. Show that $\mathscr{L}$ is base-point-free. Hint: To show that $\mathscr{L}$ has no base points away from C, consider the image of ϕ. To show that $\mathscr{L}$ has no base points *on* C, use that from (29.7.1.1) for $i = k$, we have that $H^0(X, \mathscr{L}) \to H^0(C, \mathscr{L}|_C)$ is surjective, and $\mathscr{L}|_C \cong \mathscr{O}_C$.

Thus the complete linear series $|\mathscr{L}|$ yields a morphism π' from X to some projective space $\mathbb{P}^N$.

29.7.C. EXERCISE. Show that π' precisely contracts C to a point q'. More explicitly, show that π' sends C to a point q', and $\pi|_{X \setminus C} : X \setminus C \to \mathbb{P}^N \setminus \{q\}$ is a closed embedding. Hint: To show that π' gives a closed embedding away from C: consider the image of ϕ. To show that C is contracted by π', use the fact that $\deg \mathscr{L}|_C = 0$.

Let Y' be the image of π', so we have a morphism $X \to Y'$ (which we also call π'). Let $\nu : Y \to Y'$ be the normalization of Y', so π' lifts to $\pi : X \to Y$ by the universal property of normalization (§9.7). (We are normalizing in order to use Exercise 29.5.B in Step 2.)

29.7.D. EASY EXERCISE. Show that $\nu^{-1}(q') \subset Y$ is a single point, which we call q; and that $\pi(C) = q$. (Hint: the image of π is closed, and $\pi(C)$ is connected.)

29.7.E. EXERCISE. Let X be a smooth projective surface, containing $C \subset X$ is congruent to $\mathbb{P}^1_k$. If $C \cdot C < 0$, show that there is a morphism $\pi : X \to X'$ contracting C to a point, and leaving the rest of X unchanged. (This will not be used, and is included to give practice with the argument of Step 1.)

Step 2. We now show that Y is a smooth surface. As Y is normal, we have that $\pi_* \mathscr{O}_X \cong \mathscr{O}_Y$, by Exercise 29.5.B. We will show that Y is smooth at q by showing that $\widehat{\mathscr{O}}_{Y,p} \cong k[[x,y]]$ as topological rings, as then $\dim_k(\mathfrak{m}/\mathfrak{m}^2) = 2$. By the Theorem on Formal Functions 29.4.2, $\widehat{\mathscr{O}}_{Y,p} \cong \varprojlim (C'_n, \mathscr{O}_{C'_n})$, where C'_n is the closed subscheme of X that is the scheme-theoretic pullback of the nth order neighborhood of q.

We are not precisely sure what the nth order neighborhood is, but that is fine; this inverse system is cofinal with $\mathscr{O}(-nC) = \mathscr{I}^n_{C/X}$ (do you see why?), so we can take this inverse limit instead (by Exercise 29.2.A). It suffices to show that for all $n \geq 0$, $H^0(C_n, \mathscr{O}_{C_n}) \cong k[[x,y]]/(x,y)^n$ where C_n is defined as $0 \to \mathscr{O}(-nC) \to \mathscr{O}_X \to \mathscr{O}_{C_n} \to 0$.

We do this by induction on n. As a k-vector space, this is easy, using

$$0 \to \mathscr{I}^n/\mathscr{I}^{n+1} \to \mathscr{O}_{C_{n+1}} \to \mathscr{O}_{C_n} \to 0$$

(the closed subscheme exact sequence (13.5.6.1) for C_n in C_{n+1}), using the canonical isomorphism $H^0(\mathbb{P}^1, \mathscr{O}_{\mathbb{P}^1}(n)) = \operatorname{Sym}^n H^0(\mathbb{P}^1, \mathscr{O}_{\mathbb{P}^1}(1))$. The tricky thing is that we want this isomorphism as *rings*.

29.7.F. EXERCISE. Prove this. (This may remind you of how we found the ring of functions on the total space of an invertible sheaf in Exercise 19.11.E. This is no coincidence.)

Hence we have smoothness, completing Step 2.

Step 3. We now must recognize π as the blow-up of Y at q. As C is an effective Cartier divisor on X, by the universal property of blowing up means that there is a unique morphism α making the following diagram commute

$$\begin{tikzcd} X \arrow[rr, "\alpha"] \arrow[dr, "\pi"'] & & \mathrm{Bl}_q Y \arrow[dl, "\beta"] \\ & Y & \end{tikzcd}$$

By the Cancellation Theorem 10.1.19 for projective morphisms, α is projective.

29.7.G. EXERCISE. Show that α has finite fibers.

As projective morphisms with finite fibers are finite (Theorem 18.1.9), α is a finite morphism. But then α is a birational finite morphism to a normal scheme, and thus an isomorphism by Exercise 9.7.P. $\square$

So far we have only used the Theorem on Formal Functions 29.4.2 for H^0. The next exercise will give you a chance to see it in action for other cohomology groups.

29.7.H. EXERCISE.
(a) Suppose $\beta : X = \mathrm{Bl}_q Y \to Y$ is the blow-up of a smooth surface over k at a smooth k-valued point. Show that $R^1\beta_*\mathscr{O}_X = R^2\beta_*\mathscr{O}_X = 0$.
(b) Suppose further that Y is projective. Show that the natural maps $\beta^* : H^1(Y, \mathscr{O}_Y) \to H^1(X, \mathscr{O}_X)$ and $\beta^* : H^2(Y, \mathscr{O}_Y) \to H^2(X, \mathscr{O}_X)$ are isomorphisms.

29.7.2. Elementary transformations.
Recall the definition of the Hirzebruch surfaces $\mathbb{F}_n$ (over a field k) from Example 17.2.4. If $n > 0$, then on $\mathbb{F}_n$, we have a unique curve E of maximally negative self-intersection (Exercise 20.2.R). It has self-intersection $E \cdot E = -n$. It is a section of the projection $\pi : \mathbb{F}_n \to \mathbb{P}^1$. (If $n = 0$, so $\mathbb{F}_n \cong \mathbb{P}^1 \times \mathbb{P}^1$, E is usually defined as any "constant" section of $\pi : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^1$ — it is not unique.)
Suppose q is a k-valued point of $\mathbb{F}_n$. Let F be the fiber of π containing q.

29.7.I. EXERCISE. Show that $F \cdot F = 0$.

Let $\beta : \mathrm{Bl}_q \mathbb{F}_n \to \mathbb{F}_n$ be the blow-up at q, and let Z be the exceptional divisor (as the name "E" is taken). Let F' be the proper transform of F, and let E' be the proper transform of E.

29.7.J. EXERCISE. Show that $F' \cdot F' = -1$.

Thus by Castelnuovo's criterion, we can blow down F', to obtain a new surface Y.

29.7.K. EXERCISE. Suppose $n > 0$. If $q \notin E$, show that $Y \cong \mathbb{F}_{n-1}$. If $q \in E$, show that $Y \cong \mathbb{F}_{n+1}$. Possible hint: If you knew that Y was a Hirzebruch surface $\mathbb{F}_m$, you could recognize it by the self-intersection of the unique curve of negative self-intersection if $m > 0$.

29.7.L. EXERCISE. Suppose $n = 0$. Show that $Y \cong \mathbb{F}_1$. Hint: Exercise 22.4.K.

This discussion can be generalized to $\mathbb{P}^1$-bundles over more general curves. The following is just a first step in the story.

29.7.M. EXERCISE. Suppose C is a smooth curve over k, $\pi : X \to C$ is a $\mathbb{P}^1$-bundle over C. If q is a k-valued point of C, then we can blow up q and blow down the proper transform of the fiber through q. Show that the result is another $\mathbb{P}^1$-bundle X' over C. (This is called an *elementary transformation* of the ruled surface.)

29.7.N. EXERCISE. Suppose $X = \mathcal{P}roj_C(\mathscr{L} \oplus \mathscr{M})$, and q lies on the section corresponding to

$$0 \to \mathscr{L} \to \mathscr{L} \oplus \mathscr{M} \to \mathscr{M} \to 0.$$

Show that $X' \cong \mathcal{P}roj_C(\mathscr{L}(\pi(q)) \oplus \mathscr{M})$.

Can you speculate (or show) what happens if you are just told that $X \cong \mathcal{P}roj_C \mathscr{V}$, where $\mathscr{V}$ is a locally free sheaf of rank 2, and q lies on a section corresponding to (20.2.7.1)?

29.7.3. Minimal models of surfaces.

A surface over an algebraically closed field is called **minimal** if it has no (-1)-curves. If a surface X is not minimal, then we can choose a (-1)-curve, and blow it down. If the resulting surface is not minimal, we can again blow down a (-1)-curve, and so on. By the finiteness of the Picard number ρ (which we have admittedly not proved, see §18.4.10), as blowing down reduces ρ by 1 (by Exercise 22.4.O), this process must terminate. By this means we construct a *minimal model* of X — a minimal surface birational to X. If minimal models were unique, this would provide a means of classifying surfaces up to birational isomorphism: if X and X' were smooth projective surfaces, to see if they were birational, we would find their minimal models, and see if they were isomorphic.

Sadly, minimal models are not unique. The example of $\mathbb{P}^1 \times \mathbb{P}^1$ and $\mathbb{P}^2$ show that there can be more than one minimal model in a birational equivalence class. (Why are these two surfaces minimal?) Furthermore, the isomorphism between the blow-up of $\mathbb{P}^2$ at two points with the blow-up of $\mathbb{P}^1 \times \mathbb{P}^1$ at one point (Exercise 22.4.K) shows that a smooth surface can have more than one minimal model.

Nonetheless, the failure of uniqueness is well-understood: it was shown by Zariski that the only surfaces with more than one minimal model are those birational to $\mathbb{C} \times \mathbb{P}^1$ for some curve C. This is part of the Enriques-Kodaira classification of surfaces. For more on this, see [**Be**], [**Ba**], or [**BPV**, §6].

29.7.4. Here are two more facts worth mentioning but not proving here. First, singular surfaces can be desingularized by normalizing and then repeatedly blowing up points; this explicit desingularization is analogous to the version for curves in §29.5.4, and is generalized by Hironaka's theorem on resolution of singularities (Remark 22.4.6). Second, birational maps $X \dashrightarrow Y$ between smooth surfaces can be factored into a finite number of blow-ups followed by a finite number of blow-downs. As a consequence, by Exercise 29.7.H, birational projective surfaces have the same $H^i(\mathscr{O})$ for $i = 0, 1, 2$. (Using facts stated in §21.5.10, you can show that over $\mathbb{C}$, this implies that the entire Hodge diamond except for $h^{1,1}$ is a birational invariant of surfaces.)

29.8 ⋆⋆ Proof of the Theorem on Formal Functions 29.4.2

In this section, we will prove the following.

29.8.1. Theorem. — *Suppose $X \to \operatorname{Spec} B$ is a proper morphism, with B Noetherian, and $\mathscr{F}$ is a coherent sheaf on X. Suppose I is an ideal of B, and $\hat{B}$ is the I-adic completion $\varprojlim B/I^n$ of B. Let $X_n := X \times_B (B/I^{n+1})$ (the nth order formal neighborhood of the fiber X_0) and let $\mathscr{F}_n$ be the pullback of $\mathscr{F}$ to X_n. Then for all i, the natural map*

$$H^i(X, \mathscr{F}) \otimes_B \hat{B} \to \varprojlim H^i(X_n, \mathscr{F}_n) \tag{29.8.1.1}$$

is an isomorphism.

29.8.A. EXERCISE. Prove that Theorem 29.8.1 implies Theorem 29.4.2. (This is essentially immediate if q is a closed point, and requires just a little thought if q is not closed; cf. Exercise 29.4.A.)

On the right side of (29.8.1.1), we have morphisms $X_1 \hookrightarrow X_2 \hookrightarrow \cdots \hookrightarrow X$, from which we get restriction maps $H^i(X_1, \mathscr{F}_1) \leftarrow H^i(X_2, \mathscr{F}_2) \leftarrow \cdots \leftarrow H^i(X, \mathscr{F})$. But notice that the X_i are all supported on the same underlying set, and it is helpful to keep in mind that the $\mathscr{F}_i$'s as all sheaves on the same topological space.

We will most often apply this in the case where I is a maximal ideal. But we may as well prove the result in this generality. The result can be relativized (with B and I replaced by a scheme and a closed subscheme), but we won't bother giving a precise statement.

Our proof will only use properness through the fact that the pushforward of coherent sheaves under proper morphisms are coherent (Grothendieck's Coherence Theorem 18.9.1), which required hard work. If you haven't read that proof, you can retreat to the projective case with little loss.

We will show that both sides of (29.8.1.1) are $\hat{B}$-modules, and that the map of (29.8.1.1) is an isomorphism of *topological* $\hat{B}$-modules.

Write $I^n\mathscr{F}$ for the coherent sheaf on X defined as you might expect: if $\mathscr{I}$ is the pullback of $\widetilde{I}$ from $\operatorname{Spec} B$ to X, $I^n\mathscr{F}$ is the image of $\mathscr{I}^n \otimes \mathscr{F} \to \mathscr{F}$. Then $H^i(X_n, \mathscr{F}_n)$ can be interpreted as $H^i(X, \mathscr{F}/I^n\mathscr{F})$ (by §18.1 property **(v)**), so the terms in the limit on the right side of (29.8.1.1) all live on the single space X. From now on we work on X.

The map in the statement of the theorem arises from the restriction $H^i(X, \mathscr{F}) \to H^i(X, \mathscr{F}/I^n\mathscr{F})$. This appears in the long exact sequence associated to

$$0 \to I^n\mathscr{F} \to \mathscr{F} \to \mathscr{F}/I^n\mathscr{F} \to 0,$$

and in particular in the portion

$$H^i(X, I^n\mathscr{F}) \xrightarrow{u_n} H^i(X, \mathscr{F}) \longrightarrow H^i(X, \mathscr{F}/I^n\mathscr{F}) \longrightarrow H^{i+1}(X, I^n\mathscr{F}) \xrightarrow{v_n} H^{i+1}(X, \mathscr{F}).$$

We shrink this further to:
(29.8.1.2)

$$0 \longrightarrow H^i(X, \mathscr{F})/\operatorname{im}(u_n) \longrightarrow H^i(X, \mathscr{F}/I^n\mathscr{F}) \longrightarrow \ker(v_n) \longrightarrow 0$$

and then apply $\varprojlim$ to both sides. Limits are left-exact (§1.6.12), so the result is left-exact. But the situation is even better than that. The transition maps of the left term for our exact sequences,

$$H^i(X, \mathscr{F})/\operatorname{im}(u_{n+1}) \to H^i(X, \mathscr{F})/\operatorname{im}(u_n),$$

are clearly surjective, so by Exercise 1.6.M we have an *exact* sequence

$$0 \to \varprojlim H^i(X, \mathscr{F})/\operatorname{im}(u_n) \to \varprojlim H^i(X, \mathscr{F}/I^n\mathscr{F}) \to \varprojlim \ker(v_n) \to 0. \tag{29.8.1.3}$$

Two key facts will imply our result. *(i)* We will show that $\{\operatorname{im}(u_n)\}$ is cofinal with the I-adic topology on $H^i(X, \mathscr{F})$. Thus by Exercise 29.2.A, we have a natural isomorphism of the left side of (29.8.1.3) with

$$\varprojlim H^i(X, \mathscr{F})/I^n H^i(X, \mathscr{F}) \cong H^i(X, \mathscr{F})\hat{}.$$

Because $H^i(X, \mathscr{F})$ is a finitely generated B-module (by Grothendieck's Coherence Theorem 18.9.1, or more simply Exercise 18.1.B in the projective case), this is isomorphic to $H^i(X, \mathscr{F}) \otimes_B \hat{B}$ by Theorem 29.2.6(b), the left side of (29.8.1.1). *(ii)* We will show that there is some d such that the transition map $\ker(v_{n+d}) \to \ker(v_n)$ is the zero map for $n \gg 0$. This implies that the right term $\varprojlim \ker(v_n)$ in (29.8.1.3) is 0. (Do you see why?) As a result, (29.8.1.3) becomes our desired automorphism. We now begin these two tasks.

(i) We will show that there is some n_0 such that for $n \geq n_0$,

$$I^n H^i(X, \mathscr{F}) \subset \operatorname{im}(u_n) \subset I^{n-n_0} H^i(X, \mathscr{F}).$$

The first inclusion is straightforward: note that $H^i(X, \mathscr{F}) \to H^i(X, \mathscr{F}/I^n\mathscr{F})$ is B-linear with kernel $\operatorname{im}(u_n)$, and target annihilated by I^n.

We will make use of the graded ring $R_\bullet := B \oplus I \oplus I^2 \oplus \cdots$ (the Rees algebra, defined in §12.9.1). It is graded ring generated in degree 1. Because I is a finitely generated ideal of B, $R_\bullet$ is a finitely generated B-algebra, and thus Noetherian.

For each $a \in I^m$, we have a commutative diagram of quasicoherent sheaves

$$\begin{tikzcd} I^n\mathscr{F} \arrow[r, "\times a"] \arrow[d, hook] & I^{n+m}\mathscr{F} \arrow[d, hook] \\ \mathscr{F} \arrow[r, "\times a"] & \mathscr{F} \end{tikzcd}$$

which induces a commutative diagram of B-modules

$$\begin{tikzcd} H^i(X, I^n\mathscr{F}) \arrow[r, "\times a"] \arrow[d, hook] & H^i(X, I^{n+m}\mathscr{F}) \arrow[d, hook] \\ H^i(X, \mathscr{F}) \arrow[r, "\times a"] & H^i(X, \mathscr{F}). \end{tikzcd}$$

In this way, $\oplus_{\geq 0} H^i(X, I^n\mathscr{F})$ and $\oplus_{\geq 0} H^i(X, I^n\mathscr{F})$ are graded $R_\bullet$-modules (where a is considered as an element of R_m). Then $\oplus_{n \geq 0} \operatorname{im}(u_n)$ is the image of the map of $R_\bullet$-modules

$$\oplus\, H^i(X, I^n\mathscr{F}) \to \oplus H^i(X, \mathscr{F}), \tag{29.8.1.4}$$

and is thus itself an $R_\bullet$-module.

We will show that $\oplus_{n \geq 0} H^i(X, I^n\mathscr{F})$ is a *finitely generated* $R_\bullet$-module. This will imply that its image $\oplus_{n \geq 0} \operatorname{im}(u_n)$ is also finitely generated, and thus by Proposition 12.9.2, for some n_0 $\operatorname{im}(u_{n+1}) = I \operatorname{im}(u_n)$ for $n \geq n_0$, which in turn implies (by induction) that for $n \geq n_0$, $\operatorname{im}(u_n) \subset I^{n-n_0} H^i(X, \mathscr{F})$, completing task *(i)*.

Now cohomology of quasicoherent sheaves commutes with direct sums (§18.1 **(vii)**), so $\oplus_{n \geq 0} H^i(X, I^n\mathscr{F}) = H^i(X, \oplus_{n \geq 0} I^n\mathscr{F})$.

29.8.B. EXERCISE. Consider the fibered diagram

$$\begin{tikzcd} X \times_B R_\bullet \arrow[r, "\sigma"] \arrow[d, "\rho"'] & X \arrow[d] \\ \operatorname{Spec} R_\bullet \arrow[r] & \operatorname{Spec} B. \end{tikzcd}$$

Interpret $\oplus_n H^i(X, \oplus I^n \mathscr{F})$ as $\rho_* \sigma^* \mathscr{F}$. (Be sure to check that the $R_\bullet$-action determined by the pushforward agrees with the $R_\bullet$-action on $\oplus_n H^i(X, I^n \mathscr{F})$ described above!)

But $\sigma^* \mathscr{F}$ is coherent (pullback takes finite type sheaves to finite type sheaves, and all schemes here are locally Noetherian), so $\rho_* \sigma^* \mathscr{F}$ is coherent (proper pushforward of coherent sheaves on Noetherian schemes are coherent, Theorem 18.9.1). We have now completed task *(i)*.

(ii) For the same reason that $R_\bullet$-module $\oplus_{n \geq 0} H^i(X, I^n \mathscr{F})$ is a coherent $R_\bullet$-module, $\oplus_{n \geq 0} H^{i+1}(X, I^n \mathscr{F})$ is a coherent $R_\bullet$-module. (Just replace i with $i+1$ throughout the discussion in *(i)*.)

Notice that $\oplus_{n \geq 0} \ker \nu_n$ is a the kernel of the map of graded $R_\bullet$-modules

$$\oplus_{n \geq 0} H^{i+1}(X, I^n \mathscr{F}) \xrightarrow{\oplus_{n \geq 0} \nu_n} \oplus_{n \geq 0} H^{i+1}(X, \mathscr{F}),$$

(which is just (29.8.1.4) with i replaced by $i+1$) and thus itself has the structure of a graded $R_\bullet$-module. It is a submodule of the module $H^{i+1}(X, \mathscr{I}^n \mathscr{F})$, which is finitely generated (for the same reason as $H^i(X, \mathscr{I}^n \mathscr{F})$, see *(i)*), so $\oplus_{n \geq 0} \ker \nu_n$ is finitely generated as well. We invoke Proposition 12.9.2 a second time, to show that there is some d such that $\ker(\nu_{n+1}) = I \ker(\nu_n)$ for $n \geq d$, from which

$$\ker(\nu_{A+B}) = I^A \ker \nu_B \quad \text{for} \quad A \geq 0 \quad \text{and} \quad B \geq d. \tag{29.8.1.5}$$

For any $a \in I^d$, the map

$$\times a : I^n \mathscr{F} \to I^{n+d} \mathscr{F}$$

induces a map $H^{i+1}(X, I^n \mathscr{F}) \to H^{i+1}(X, I^{n+d} \mathscr{F})$, which sends $\ker(\nu_n)$ to $\ker(\nu_{n+d})$. (This can be interpreted as coming from the action of $R_\bullet$ on the module $\oplus \ker(\nu_n)$, taking a as an element of R_d.) As a runs through I^d, the B-linear spans of the images of

$$\ker(\nu_n) \longrightarrow H^{i+1}(X, I^n \mathscr{F}) \xrightarrow{\times a} H^{i+1}(X, I^{n+d} \mathscr{F})$$

for $n \geq d$ is $I^d \ker(\nu_n) = \ker(\nu_{n+d})$ by (29.8.1.5). The composition

$$H^{i+1}(X, I^n \mathscr{F}) \xrightarrow{\times a} H^{i+1}(X, I^{n+d} \mathscr{F}) \longrightarrow H^{i+1}(X, I^n \mathscr{F})$$

(the latter map coming from inclusion of sheaves) is exactly the map

$$\times a : H^{i+1}(X, I^n \mathscr{F}) \to H^{i+1}(X, I^n \mathscr{F}).$$

Thus the image of $\ker(\nu_{n+d}) \to \ker(\nu_n)$ is exactly the B-linear span of the images of the composition maps

$$\ker(\nu_n) \hookrightarrow H^{i+1}(X, I^n \mathscr{F}) \xrightarrow{\times a} H^{i+1}(X, I^n \mathscr{F})$$

as a runs through I^d.

It thus suffices to show that the multiplication by $a \in I^d$ on $H^{i+1}(X, I^n \mathscr{F})$ annihilates $\ker(\nu_n)$ for $n \geq d$. Because $\ker(\nu_n) = I^{n-d} \ker(\nu_d)$ for all $n \geq d$ ((29.8.1.5) again), it suffices to show that any element a of I^d annihilates $\ker(\nu_d)$. But $\ker(\nu_d)$ is the image of $H^i(X, \mathscr{F}/I^d \mathscr{F})$ (from the long exact sequence, see (29.8.1.2)), and thus is annihilated by any element of I^d, as desired. □

CHAPTER 30

$\star$ Proof of Serre duality

Réfléchissant un peu à ton théorème de dualité, je m'aperçois que sa formulation générale est à peu près évidente, et d'ailleurs je viens de vérifier qu'elle se trouve implicitement (dans le cas de l'espace projectif) dans ton théorème donnant les $T^q(M)$ par des Ext. (J'ai bien l'impression, salaud, que tes §3 et 4 du Chap. 3 peuvent se faire aussi sans aucun calcul).

Thinking a bit about your duality theorem, I notice that its general form is almost obvious, and in fact I just checked that (for a projective space) it is implicitly contained in your theorem giving the $T^q(M)$ in terms of Exts. (I have the impression, you bastard, that §3 and 4 in your Chap. 3 could be done without any computation).

— *A. Grothendieck, letter to J.-P. Serre* [**GrS**, p. 19]

30.1 Introduction

We first met Serre duality in §18.4 (Theorem 18.5.1), and we have repeatedly seen how useful it is. We will prove Theorem 18.5.1 (Corollary 30.3.9, combined with Exercise 30.4.I, see Remark 30.4.9), as well as stronger versions, and we will be left with a desire to prove even more. We give three statements (*Serre duality for vector bundles*; *Serre duality for* Hom; and *Serre duality for* Ext), in two versions (*functorial* and *trace*). (These names are idiosyncratic and nonstandard.) We give several variants for a number of reasons. First, the easier statements will be easier to prove, and the hardest statements we won't be able to prove here. Second, they may help give you experience in knowing how to know what to hope for, and what to try to prove.

Throughout this chapter, X will be a projective k-scheme of pure dimension n. We will want a coherent sheaf $\boxed{\omega}$ (or with more precision, $\boxed{\omega_X}$, or even better, $\omega_{X/k}$) on X, the **dualizing sheaf**, which will play a role in the statements of duality. For the best statements, we will want a **trace morphism**

$$\boxed{t : H^n(X, \omega_X) \to k}. \tag{30.1.0.6}$$

30.1.1. Desideratum: the determinant of the cotangent bundle is dualizing for smooth varieties. If X is smooth, we will want $\boxed{\omega_X = \mathscr{K}_X}$ in this case (recall $\mathscr{K}_X = \det \Omega_X$) — the miracle that the canonical bundle is Serre-dualizing (§18.5.2). In particular, ω_X is an invertible sheaf. This will be disturbingly harder to prove than the basic duality statements we show; we will only get to it later (Exercise 30.4.I). But we will prove more, for example that ω_X is an invertible sheaf if X is a regular embedding in a smooth variety (Exercise 30.4.H).

30.1.2. Desideratum: Serre duality for vector bundles. The first version of duality, which (along with Desideratum 30.1.1) gives Theorem 18.5.1, is the following: if $\mathscr{F}$ is locally free, then we have a functorial isomorphism

$$H^i(X, \mathscr{F}) \cong H^{n-i}(X, \mathscr{F}^\vee \otimes \omega_X)^\vee. \tag{30.1.2.1}$$

More precisely, we want to construct a *particular* isomorphism (30.1.2.1), or equivalently, a particular perfect pairing $H^i(X, \mathscr{F}) \times H^{n-i}(X, \mathscr{F}^\vee \otimes \omega_X) \to k$. This isomorphism will be *functorial in* $\mathscr{F}$, i.e., it gives a natural isomorphism of covariant functors

$$H^i(X, \cdot) \xrightarrow{\sim} H^{n-i}(X, \cdot^\vee \otimes \omega_X)^\vee. \tag{30.1.2.2}$$

We call this **functorial Serre duality for vector bundles**.

Better still, there should be a cup product in cohomology, which can be used to construct a map $H^i(X, \mathscr{F}) \times H^{n-i}(X, \mathscr{F}^\vee \otimes \omega_X) \to H^n(X, \omega_X)$. This should be functorial in $\mathscr{F}$ (in the sense that we get a natural transformation of functors $H^i(X, \cdot) \to H^{n-i}(X, \cdot^\vee \otimes \omega_X) \otimes H^n(X, \omega_X)$). Combined with the trace map (30.1.0.6), we get a map $H^i(X, \mathscr{F}) \times H^{n-i}(X, \mathscr{F}^\vee \otimes \omega_X) \to k$, which should yield (30.1.2.2). We call this the **trace version of Serre duality for vector bundles**. (This was hinted at after the statement of Theorem 18.5.1.)

In fact, the trace version of Serre duality for vector bundles (and hence the functorial version) is true when X is Cohen-Macaulay, and in particular when X is smooth. We will prove the functorial version (see Corollary 30.3.9). We will give an indication of how the trace version can be proved in Remark 30.3.14.

30.1.3. Desideratum: duality for more general X and $\mathscr{F}$. A weaker sort of duality will hold with weaker hypotheses. We will show that (without Cohen-Macaulay hypotheses) for any coherent sheaf $\mathscr{F}$ on a pure n-dimensional projective k-scheme X, there is a *functorial isomorphism*

$$\operatorname{Hom}(\mathscr{F}, \omega_X) \xrightarrow{\sim} H^n(X, \mathscr{F})^\vee. \tag{30.1.3.1}$$

We call this **functorial Serre duality for** Hom.

In parallel with Serre duality for vector bundles, we have a natural candidate for the perfect pairing:

$$\operatorname{Hom}(\mathscr{F}, \omega_X) \times H^n(X, \mathscr{F}) \longrightarrow H^n(X, \omega_X) \xrightarrow{t} k, \tag{30.1.3.2}$$

where the trace map t is some linear functional on $H^n(X, \omega_X)$. If this composition is a perfect pairing, we say that X (with the additional data of (ω_X, t)) satisfies the **trace version of Serre duality for** Hom. Unlike the trace version of Serre duality for vector bundles, we already know what the "cup product" map $\operatorname{Hom}(\mathscr{F}, \omega_X) \times H^n(X, \mathscr{F}) \to H^n(X, \omega_X)$ is: an element $[\sigma : \mathscr{F} \to \omega_X]$ of $\operatorname{Hom}(\mathscr{F}, \omega_X)$ induces — by covariance of $H^n(X, \cdot)$, see §18.1 — a map $H^n(X, \mathscr{F}) \to H^n(X, \omega_X)$. The resulting pairing is clearly functorial in $\mathscr{F}$, so the trace version of Serre duality for Hom implies the functorial version. Unlike the case of Serre duality for vector bundles, the functorial version of Serre duality for Hom also implies the trace version:

30.1.A. EXERCISE. Show that the functorial version of Serre duality for Hom implies the trace version. In other words, the trace map is already implicit in the functorial isomorphism $\mathrm{Hom}(\mathscr{F}, \omega_X) \to H^n(X, \mathscr{F})^\vee$. Hint: consider the commuting diagram (coming from functoriality)

$$\begin{tikzcd} \mathrm{Hom}(\mathscr{F}, \omega_X) \arrow[r] \arrow[d] & H^n(X, \mathscr{F})^\vee \arrow[d] \\ \mathrm{Hom}(\omega_X, \omega_X) \arrow[r] & H^n(X, \omega_X)^\vee. \end{tikzcd}$$

30.1.4. *Definition.* Suppose X is a projective k-scheme of pure dimension n. A coherent sheaf $\omega = \omega_X = \omega_{X/k}$ along with a map $t : H^n(X, \omega_X) \to k$ is called **dualizing** if the natural map (cf. (30.1.3.1))

$$\mathrm{Hom}(\mathscr{F}, \omega_X) \times H^n(X, \mathscr{F}) \longrightarrow H^n(X, \omega_X) \xrightarrow{t} k \tag{30.1.4.1}$$

is a perfect pairing. We call ω_X the **dualizing sheaf** and t the **trace map**. (The earlier discussion of ω_X and t was aspirational. This now is a definition.) If X has such (ω_X, t), we say that X **satisfies Serre duality** (for Hom). The following proposition justifies the use of the word "the" (as opposed to "a") in the phrase "the dualizing sheaf".

30.1.5. Proposition. — *If a dualizing sheaf and trace (ω_X, t) exists for X, this data is unique up to unique isomorphism.*

Proof. Suppose we have two such (ω_X, t) and (ω'_X, t'). From the two morphisms

$$\begin{gathered} \mathrm{Hom}(\mathscr{F}, \omega_X) \times H^n(X, \mathscr{F}) \longrightarrow H^n(X, \omega_X) \xrightarrow{t} k \\ \\ \mathrm{Hom}(\mathscr{F}, \omega'_X) \times H^n(X, \mathscr{F}) \longrightarrow H^n(X, \omega'_X) \xrightarrow{t'} k, \end{gathered} \tag{30.1.5.1}$$

we get a natural bijection $\mathrm{Hom}(\mathscr{F}, \omega_X) \cong \mathrm{Hom}(\mathscr{F}, \omega'_X)$, which is functorial in $\mathscr{F}$. By the typical universal property argument (Exercise 1.3.Z), this induces a (unique) isomorphism $\omega_X \cong \omega'_X$. From (30.1.5.1), under this isomorphism, the two trace maps t and t' must be the same too. □

We will prove the functorial and trace versions of Serre duality for Hom in Corollary 30.3.11. The special case of projective space will be a key ingredient; we prove this case now.

30.1.6. *Serre duality (for* Hom*) for projective space.* Define $\omega_{\mathbb{P}^n_k}$ (or just ω for convenience) as $\mathscr{O}_{\mathbb{P}^n_k}(-n-1)$. Let t be any isomorphism $H^n(\mathbb{P}^n_k, \omega_{\mathbb{P}^n_k}) \to k$ (Theorem 18.1.3). As the notation suggests, $(\omega_{\mathbb{P}^n_k}, t)$ will be dualizing for projective space $\mathbb{P}^n_k$.

30.1.B. EXERCISE. Suppose $\mathscr{F} = \mathscr{O}_{\mathbb{P}^n_k}(m)$. Show that the natural map (30.1.4.1) is a perfect perfect pairing. (Hint: do this by hand. See the discussion after Theorem 18.1.3.) Hence show that if $\mathscr{F}$ is a direct sum of line bundles on $\mathbb{P}^n_k$, the natural map (30.1.4.1) is a perfect pairing.

30.1.7. Proposition. — *The functorial version (and hence the trace version by Exercise 30.1.A) of Serre duality for* Hom *holds for* $\mathbb{P}^n_k$.

Proof. We wish to show that the (functorial) natural map

$$\operatorname{Hom}(\cdot, \omega_{\mathbb{P}^n_k}) \xrightarrow{\sim} H^n(\mathbb{P}^n_k, \mathscr{F})^\vee$$

is an isomorphism (cf. (30.1.3.1)). Fix a coherent sheaf $\mathscr{F}$ on $\mathbb{P}^n_k$. By Theorem 15.3.1, we can present $\mathscr{F}$ as

$$0 \longrightarrow \mathscr{G} \longrightarrow \mathscr{E} \longrightarrow \mathscr{F} \longrightarrow 0 \tag{30.1.7.1}$$

where $\mathscr{E}$ is a finite direct sum of line bundles, and $\mathscr{G}$ is coherent. Applying the left-exact functor $\operatorname{Hom}(\cdot, \omega_{\mathbb{P}^n_k})$ to (30.1.7.1), we have the exact sequence

$$0 \longrightarrow \operatorname{Hom}(\mathscr{F}, \omega_{\mathbb{P}^n_k}) \longrightarrow \operatorname{Hom}(\mathscr{E}, \omega_{\mathbb{P}^n_k}) \longrightarrow \operatorname{Hom}(\mathscr{G}, \omega_{\mathbb{P}^n_k}). \tag{30.1.7.2}$$

Taking the long exact sequence in cohomology for (30.1.7.1) and dualizing, we have the exact sequence

$$0 \longrightarrow H^n(\mathbb{P}^n_k, \mathscr{F})^\vee \longrightarrow H^n(\mathbb{P}^n_k, \mathscr{E})^\vee \longrightarrow H^n(\mathbb{P}^n_k, \mathscr{G})^\vee \tag{30.1.7.3}$$

The (functorial) pairing (30.1.3.2) gives to a map from (30.1.7.2) to (30.1.7.3):
(30.1.7.4)

$$\begin{tikzcd}
0 \arrow[r] & 0 \arrow[r] & H^n(\mathbb{P}^n_k, \mathscr{F})^\vee \arrow[r] & H^n(\mathbb{P}^n_k, \mathscr{E})^\vee \arrow[r] & H^n(\mathbb{P}^n_k, \mathscr{G})^\vee \\
0 \arrow[r] \arrow[u, "\alpha"] & 0 \arrow[r] \arrow[u, "\beta"] & \operatorname{Hom}(\mathscr{F}, \omega_{\mathbb{P}^n_k}) \arrow[r] \arrow[u, "\gamma"] & \operatorname{Hom}(\mathscr{E}, \omega_{\mathbb{P}^n_k}) \arrow[r] \arrow[u, "\delta"] & \operatorname{Hom}(\mathscr{G}, \omega_{\mathbb{P}^n_k}) \arrow[u, "\epsilon"]
\end{tikzcd}$$

Maps α and β are obviously isomorphisms, and Exercise 30.1.B shows that δ is an isomorphism. Thus by the subtle version of the five lemma (Exercise 1.7.D, as β and δ are injective and α is surjective), γ is injective. This shows that the natural map $\operatorname{Hom}(\mathscr{F}', \omega_{\mathbb{P}^n_k}) \to H^n(\mathbb{P}^n_k, \mathscr{F}')^\vee$ is injective for *all* coherent sheaves $\mathscr{F}'$, and in particular for $\mathscr{F}' = \mathscr{G}$. Thus ϵ is injective. Then by the dual of the subtle version of the five lemma (as β and δ are surjective, and ϵ is injective), γ is surjective. $\square$

30.1.8. *Mathematical puzzle.* Here is a puzzle to force you to confront a potentially confusing point. We will see that Desideratum 30.1.1 holds for $\mathbb{P}^1$, so $\omega_{\mathbb{P}^1} \cong \Omega_{\mathbb{P}^1}$. What then is the trace map $t : H^1(\mathbb{P}^1, \Omega_{\mathbb{P}^1}) \to k$? The Čech complex for $H^1(\mathbb{P}^1, \Omega_{\mathbb{P}^1})$ (with the usual cover of $\mathbb{P}^1$) is given by

$$0 \longrightarrow \Omega_{\mathbb{P}^1}(U_0) \times \Omega_{\mathbb{P}^1}(U_1) \xrightarrow{\alpha} \Omega_{\mathbb{P}^1}(U_0 \cap U_1) \longrightarrow 0. \tag{30.1.8.1}$$

If $U_0 = \operatorname{Spec} k[x]$, and $U_0 \cap U_1 = \operatorname{Spec} k[x, 1/x]$, then the differentials on $U_0 \cap U_1$ are those of the form $f(x)\ dx$ where $f(x)$ is a Laurent polynomial (for example: $(x^{-3} + x^{-1} + 3 + 17x^4)\ dx$). To compute $H^1(\mathbb{P}^1, \Omega_{\mathbb{P}^1})$, we need to find which such differentials on $U_0 \cap U_1$ are in the image of α in (30.1.8.1). Clearly any term of the form $x^i\ dx$ (for $i \geq 0$) extends to a differential on U_0, and thus is in the image of α. A short calculation shows that any term of the form $x^i\ dx$ ($i < -1$) extends to a differential on U_1. Thus the cokernel of α can be described as the one-dimensional k-vector space generated by $x^{-1}\ dx$. We have an obvious isomorphism to k: take the coefficient of $x^{-1}\ dx$, which can be interpreted as "take the residue at 0". But there is another choice, which is equally good: take the residue at ∞ — certainly there is no reason to privilege 0 over ∞ (or U_0 or U_1)! But these two residues are *not* the same — they add to 0 (as you can quickly calculate — you may also believe it because of the Residue Theorem in the theory of Riemann surfaces). So: which one is the trace?

30.1.9. Desideratum: a stronger version of duality, involving Ext**.** The vector bundle and Hom versions of Serre duality have a common extension. If we have an isomorphism of functors

$$\mathrm{Ext}^i(\cdot, \omega_X) \xrightarrow{\sim} H^{n-i}(X, \cdot)^\vee , \tag{30.1.9.1}$$

we say that X satisfies **functorial Serre duality for** Ext. (The case $i = 0$ is functorial Serre duality for Hom, or by Exercise 30.1.A, the trace version.)

In Exercise 30.2.I, we will find that the functorial version of Serre duality for Ext (resp. the trace version) implies the functorial version of Serre duality or vector bundles (resp. the trace version). Thus to prove Theorem 18.5.1, it suffices to prove functorial Serre duality for Ext, and Desideratum 30.1.1 (see Remark 30.4.9). We will prove the for functorial version of Serre duality for Ext when X is Cohen-Macaulay in Corollary 30.3.13.

30.1.10. *Functorial Serre duality for* Ext *holds for projective space.*

We now prove that functorial Serre duality for Ext holds for projective space. We will use the machinery of universal δ-functors (§23.2.6), so you may wish to either quickly skim that section, or else ignore this discussion.

30.1.C. EXERCISE. Show that $(\mathrm{Ext}^i_{\mathbb{P}^n_k}(\cdot, \omega_{\mathbb{P}^n_k}))$ is a (contravariant) universal δ-functor. Hint: Ext is not a derived functor in its first argument, so you can't use the "projective" version of Corollary 23.2.10. Instead, use Theorem 23.2.8, and the existence of a surjection $\mathscr{O}(m)^{\oplus N} \to \mathscr{F}$ for each $\mathscr{F}$, for some $m < 0$.

30.1.D. EXERCISE. Show that $(H^{n-i}(\mathbb{P}^n_k, \cdot)^\vee)$ is a universal δ-functor. (What are the δ-maps?) Hint: try the same idea as in the previous exercise.

Proposition 30.1.7 gives an isomorphism of functors $\mathrm{Ext}^0_{\mathbb{P}^n_k}(\cdot, \omega_{\mathbb{P}^n_k}) \cong H^n(\mathbb{P}^n_k, \cdot)^\vee$, so by the Definition 23.2.7 of universal δ-functor, we have an isomorphism of δ-functors $(\mathrm{Ext}^i_{\mathbb{P}^n_k}(\cdot, \omega_{\mathbb{P}^n_k})) \cong (H^{n-i}(\mathbb{P}^n_k, \cdot)^\vee)$, thereby proving functorial Serre duality for Ext for $\mathbb{P}^n_k$.

30.1.11. *Trace version of Serre duality for* Ext*.* As with the previous versions, the functoriality of functorial Serre duality for Ext should come from somewhere. We should expect a natural cup product $\mathrm{Ext}^i(\mathscr{F}, \omega_X) \times H^{n-i}(X, , \mathscr{F}) \to H^n(X, \omega_X)$, which coupled with the trace map (30.1.0.6) should yield the isomorphism (30.1.9.1). We call this the **trace version of Serre duality for** Ext. We will not be able to prove the trace version, as we will not define this cup product. But we will give some indication of how it works in §30.2.3.

30.1.12. Necessity of Cohen-Macaulay hypotheses. We remark that the Cohen-Macaulay hypotheses are necessary everywhere they are stated. The following example applies in all cases. Let X be the union of two 2-planes in $\mathbb{P}^4_k$, meeting at a point p. If there were a coherent dualizing sheaf ω_X on X, then for $d \gg 0$, we would have $h^1(X, \mathscr{O}_X(-d)) = h^1(X, \omega_X(d))$ by Serre duality, which must be 0 by Serre vanishing (Theorem 18.1.4(ii)). We show that this is not the case.

30.1.E. EXERCISE. Let H be a hyperplane in $\mathbb{P}^4$ not passing through p; say H is hyperplane $x_0 = 0$. Let Z be the intersection of dH (the divisor $x_0^d = 0$) with X. Show (cheaply) that for $d \geq 0$, $h^0(Z, \mathscr{O}_Z) > 1$. Show that $h^1(X, \mathscr{O}_X) = 0$ (perhaps using an easy Čech cover, or perhaps by comparing the cohomology of X to that of its normalization, as we did in with curves in (29.3.4.1)). Use the exact sequence

$$0 \to \mathscr{O}_X(-Z) \to \mathscr{O}_X \to \mathscr{O}_Z \to 0$$

to show that $h^1(X, \mathscr{O}_X(-d)) > 0$.

30.2 Ext groups and Ext sheaves for $\mathscr{O}$-modules

Recall that for any ringed space X, the category $Mod_{\mathscr{O}_X}$ has enough injectives (Theorem 23.4.1). Thus for any $\mathscr{O}_X$-module $\mathscr{F}$ on X, we may define

$$\operatorname{Ext}^i_X(\mathscr{F}, \cdot) : Mod_{\mathscr{O}_X} \to Mod_{\Gamma(\mathscr{O}_X)}$$

as the ith right derived functor of $\operatorname{Hom}_X(\mathscr{F}, \cdot)$, and we have a corresponding long exact sequence for $\operatorname{Ext}^i_X(\mathscr{F}, \cdot)$. We similarly define a sheaf version of this

$$\mathcal{E}xt^i_X(\mathscr{F}, \cdot) : Mod_{\mathscr{O}_X} \to Mod_{\mathscr{O}_X}$$

as a right derived functor of $\mathcal{H}om_X(\mathscr{F}, \cdot)$. In both cases, the subscript X is often omitted when it is clear from the context, although this can be dangerous when more than one space is relevant to the discussion. (We saw Ext functors for A-modules in §23.2.4.)

Warning: it is not clear (and in fact not true, see §23.4.6) that $Mod_{\mathscr{O}_X}$ has enough projectives, so we cannot define Ext^i as a derived functor in its left argument. Nonetheless, we will see that it behaves as though it *is* a derived functor — it is "computable by acyclics", and has a long exact sequence (Remark 30.2.1).

Another warning: with this definition, it is not clear that if $\mathscr{F}$ and $\mathscr{G}$ are quasicoherent sheaves on a scheme, then the $\mathcal{E}xt^i(\mathscr{F}, \mathscr{G})$ are quasicoherent, and indeed the aside in Exercise 13.7.A(a) points out this need not be true even for $i = 0$. But Exercise 30.2.F will reassure you.

Exercise 23.5.A (an injective $\mathscr{O}_X$-module, when restricted to an open subset $U \subset X$, is injective on U) has a number of useful consequences.

30.2.A. EXERCISE. Suppose $\mathscr{I}$ is an injective $\mathscr{O}_X$-module. Show that $\mathcal{H}om_{\mathscr{O}_X}(\cdot, \mathscr{I})$ is an exact contravariant functor. (A related fact: $\operatorname{Hom}_{\mathscr{O}_X}(\cdot, \mathscr{I})$ is exact, by the definition of injectivity, Exercise 23.2.C(a).)

30.2.B. EXERCISE. Suppose X is a ringed space, $\mathscr{F}$ and $\mathscr{G}$ are $\mathscr{O}_X$-modules, and U is an open subset. Describe a canonical isomorphism $\mathcal{E}xt^i_X(\mathscr{F}, \mathscr{G})|_U \cong \mathcal{E}xt^i_U(\mathscr{F}|_U, \mathscr{G}|_U)$. Hint: take an injective resolution of $\mathscr{G}$ on X, and restrict it to U. Use Exercise 30.2.A to show that the result is an injective resolution of $\mathscr{G}|_U$.

30.2.C. EXERCISE. Suppose X is a ringed space, and $\mathscr{G}$ is an $\mathscr{O}_X$-module.
(a) Show that

$$\mathcal{E}xt^i(\mathscr{O}_X, \mathscr{G}) = \begin{cases} \mathscr{G} & \text{if } i = 0, \text{and} \\ 0 & \text{otherwise.} \end{cases}$$

(b) Describe a canonical isomorphism $\operatorname{Ext}^i(\mathscr{O}_X, \mathscr{G}) \cong H^i(X, \mathscr{G})$.

30.2.D. EXERCISE. Use Exercise 30.2.C(a) to show that if $\mathscr{E}$ is a locally free sheaf on X, then $\mathcal{E}xt^i(\mathscr{E}, \mathscr{G}) = 0$ for $i > 0$.

In the category of modules over rings, we like projectives more than injectives, because free modules are easy to work with. It would be wonderful if locally free sheaves on schemes were always projective, but sadly this is not true. Otherwise, the Euler exact sequence for $\mathbb{P}^1_k$ (Theorem 21.4.6)

$$0 \to \mathscr{O}_{\mathbb{P}^1_k}(-2) \to \mathscr{O}_{\mathbb{P}^1_k}(-1) \oplus \mathscr{O}_{\mathbb{P}^1_k}(-1) \to \mathscr{O}_{\mathbb{P}^1_k} \to 0$$

would split, but Theorem 18.5.6 (or a short direct argument) implies that it doesn't. Nonetheless, we can still compute with locally free sheaves, as shown in the following exercise.

30.2.E. IMPORTANT EXERCISE. Suppose X is a ringed space, and

$$\cdots \longrightarrow \mathscr{E}_1 \longrightarrow \mathscr{E}_0 \longrightarrow \mathscr{F} \longrightarrow 0 \tag{30.2.0.1}$$

is a resolution of an $\mathscr{O}_X$-module $\mathscr{F}$ by locally free sheaves. (Of course we are most interested in the case where X is a scheme, and $\mathscr{F}$ is quasicoherent, or even coherent.) Let $\mathscr{E}_\bullet$ denote the truncation of (30.2.0.1), where $\mathscr{F}$ is removed. Describe an isomorphism $\mathcal{E}xt^i(\mathscr{F}, \mathscr{G}) \cong H^i(\mathcal{H}om(\mathscr{E}_\bullet, \mathscr{G}))$. In other words, $\mathcal{E}xt^\bullet(\mathscr{F}, \mathscr{G})$ can be computed by taking a locally free resolution of $\mathscr{F}$, truncating, applying $\mathcal{H}om(\cdot, \mathscr{G})$, and taking homology. Hint: choose an injective resolution

$$0 \longrightarrow \mathscr{G} \longrightarrow \mathscr{I}_0 \longrightarrow \mathscr{I}_1 \longrightarrow \cdots$$

and consider the spectral sequence whose E_0 term is

$$\begin{tikzcd}
\vdots & \vdots & \\
\mathcal{H}om(\mathscr{E}_0, \mathscr{I}_1) \arrow[r] \arrow[u] & \mathcal{H}om(\mathscr{E}_1, \mathscr{I}_1) \arrow[r] \arrow[u] & \cdots \\
\mathcal{H}om(\mathscr{E}_0, \mathscr{I}_0) \arrow[r] \arrow[u] & \mathcal{H}om(\mathscr{E}_1, \mathscr{I}_0) \arrow[r] \arrow[u] & \cdots .
\end{tikzcd}$$

This result is important: to compute $\mathcal{E}xt$, we can compute it using finite rank locally free resolutions. You can work affine by affine (by Exercise 30.2.B), and on each affine you can use a free resolution of the left argument. As another consequence of Exercise 30.2.E:

30.2.F. EXERCISE. Suppose $\mathscr{F}$ and $\mathscr{G}$ are coherent sheaves on a quasiprojective k-scheme X. Show that $\mathcal{E}xt^i_X(\mathscr{F}, \mathscr{G})$ is a coherent sheaf as well. (Your argument will work on any scheme for which there always exist resolutions by finite rank locally free sheaves.)

30.2.1. *Remark.* The statement "$\operatorname{Ext}^i(\mathscr{F},\mathscr{G})$ *behaves like a derived functor in the first argument*" is true in a number of ways. We can compute it using a resolution of $\mathscr{F}$ by locally frees, which are acyclic for $\operatorname{Ext}^i(\cdot,\mathscr{G})$. And we even have a corresponding long exact sequence, as shown in the next exercise.

30.2.G. EXERCISE. Suppose $0 \to \mathscr{F}'' \to \mathscr{F} \to \mathscr{F}' \to 0$ is an exact sequence of $\mathscr{O}_X$-modules on a ringed space X. For any $\mathscr{O}_X$-module $\mathscr{G}$, describe a long exact sequence

$$\begin{tikzcd} 0 \arrow[r] & \operatorname{Hom}(\mathscr{F}'',\mathscr{G}) \arrow[r] & \operatorname{Hom}(\mathscr{F},\mathscr{G}) \arrow[r] & \operatorname{Hom}(\mathscr{F}',\mathscr{G}) \\ \arrow[r] & \operatorname{Ext}^1(\mathscr{F}'',\mathscr{G}) \arrow[r] & \operatorname{Ext}^1(\mathscr{F},\mathscr{G}) \arrow[r] & \operatorname{Ext}^1(\mathscr{F}',\mathscr{G}) \arrow[r] & \cdots . \end{tikzcd}$$

Hint: take an injective resolution $0 \to \mathscr{G} \to \mathscr{I}^0 \to \cdots$. Use the fact that if $\mathscr{I}$ is injective, then $\operatorname{Hom}(\cdot,\mathscr{I})$ is exact (the definition of injectivity, Exercise 23.2.C(a)). Hence get a short exact sequence of complexes

$$0 \to \operatorname{Hom}(\mathscr{F}'',\mathscr{I}^\bullet) \to \operatorname{Hom}(\mathscr{F},\mathscr{I}^\bullet) \to \operatorname{Hom}(\mathscr{F}',\mathscr{I}^\bullet) \to 0$$

and take the long exact sequence in cohomology.

Here are two useful exercises.

30.2.H. EXERCISE. Suppose X is a ringed space, $\mathscr{F}$ and $\mathscr{G}$ are $\mathscr{O}_X$-modules, and $\mathscr{E}$ is a locally free sheaf on X. Describe isomorphisms

$$\mathcal{E}xt^i(\mathscr{F}\otimes\mathscr{E}^\vee,\mathscr{G}) \cong \mathcal{E}xt^i(\mathscr{F},\mathscr{G}\otimes\mathscr{E}) \cong \mathcal{E}xt^i(\mathscr{F},\mathscr{G})\otimes\mathscr{E}$$
$$\text{and} \qquad \operatorname{Ext}^i(\mathscr{F}\otimes\mathscr{E}^\vee,\mathscr{G}) \cong \operatorname{Ext}^i(\mathscr{F},\mathscr{G}\otimes\mathscr{E}).$$

Hint: show that if $\mathscr{I}$ is injective then $\mathscr{I}\otimes\mathscr{E}$ is injective.

30.2.I. EXERCISE. If $\mathscr{F}$ is a locally free sheaf on X and ω_X is any coherent sheaf on X, describe an isomorphism $\operatorname{Ext}^i(\mathscr{F},\omega_X) \cong H^i(X,\mathscr{F}^\vee\otimes\omega_X)$ (functorial in $\mathscr{F}$). Show that functorial Serre duality for Ext implies functorial Serre duality for vector bundles. (You may wish to ponder the trace versions as well.) As a consequence, by §30.1.10, functorial Serre duality for vector bundles holds for projective space. Hint: Exercises 30.2.H and 30.2.C(b).

30.2.2. The local-to-global spectral sequence for Ext**.**
The "sheaf" $\mathcal{E}xt$ and "global" Ext are related by a spectral sequence. This is a straightforward application of the Grothendieck composition-of-functors spectral sequence, once we show that $\mathcal{H}om(\mathscr{F},\mathscr{I})$ is acyclic for the functor Γ.

30.2.J. EXERCISE. Suppose $\mathscr{I}$ is an injective $\mathscr{O}_X$-module. Show that $\mathcal{H}om(\mathscr{F},\mathscr{I})$ is *flasque* (and thus Γ-acyclic by Exercise 23.4.G). Hint: suppose $j: U \hookrightarrow V$ is an inclusion of open subsets. We wish to show that $\mathcal{H}om(\mathscr{F},\mathscr{I})(V) \to \mathcal{H}om(\mathscr{F},\mathscr{I})(U)$ is surjective. Note that $\mathscr{I}|_V$ is injective on V (Exercise 23.5.A). Apply the exact functor $\operatorname{Hom}_V(\cdot,\mathscr{I}|_V)$ to the inclusion $j_!(\mathscr{F}|_U) \hookrightarrow \mathscr{F}|_V$ of sheaves on V (Exercise 2.6.G).

30.2.K. EXERCISE (THE LOCAL-TO-GLOBAL SPECTRAL SEQUENCE FOR Ext). Suppose X is a ringed space, and $\mathscr{F}$ and $\mathscr{G}$ are $\mathscr{O}_X$-modules. Describe a spectral sequence with E_2-term $H^j(X, \mathcal{E}xt^i(\mathscr{F}, \mathscr{G}))$ abutting to $\operatorname{Ext}^{i+j}(\mathscr{F}, \mathscr{G})$. (Hint: use the Grothendieck composition-of-functors spectral sequence, Theorem 23.3.5. Recall that $\operatorname{Hom}(\mathscr{F}, \cdot) = \Gamma(\mathcal{H}om(\mathscr{F}, \cdot))$, Exercise 2.3.C.)

30.2.3. $\star\star$ **Composing Ext's (and H^i's): the Yoneda cup product.**

It is useful and reassuring to know that Ext's can be composed, in a reasonable sense. We won't need this, and so just outline the ideas, so you can recognize them in the future should you need them. For more detail, see [**Gr3**, §2] or [**Cart**].

If $\mathscr{C}$ is an abelian category, and $A_\bullet$ and $B_\bullet$ are complexes in $\mathscr{C}$, then define $\operatorname{Hom}_\bullet(A_\bullet, B_\bullet)$ as the integer-graded group of *graded homomorphisms*: the elements of $\operatorname{Hom}_n(A_\bullet, B_\bullet)$ are the maps from the complex $A_\bullet$ to $B_\bullet$ shifted "to the right by n". Define $\delta : \operatorname{Hom}_\bullet(A_\bullet, B_\bullet)$ by

$$\delta(u) = du + (-1)^{n+1} ud$$

for each $u \in \operatorname{Hom}_n(A_\bullet, B_\bullet)$ (where d sloppily denotes the differential in both $A_\bullet$ and $B_\bullet$). Then $\delta^2 = 0$, turning $\operatorname{Hom}_\bullet(A_\bullet, B_\bullet)$ into a complex. Let $H^\bullet(A_\bullet, B_\bullet)$ be the cohomology of this complex. If $C_\bullet$ is another complex in $\mathscr{C}$, then composition of maps of complexes yields a map $\operatorname{Hom}_\bullet(A_\bullet, B_\bullet) \times \operatorname{Hom}_\bullet(B_\bullet, C_\bullet) \to \operatorname{Hom}_\bullet(A_\bullet, C_\bullet)$ which induces a map on cohomology:

$$H^\bullet(A_\bullet, B_\bullet) \times H^\bullet(B_\bullet, C_\bullet) \to H^\bullet(A_\bullet, C_\bullet) \tag{30.2.3.1}$$

which can be readily checked to be associative. In particular, $H^\bullet(A_\bullet, A_\bullet)$ has the structure of a graded associative *non-commutative* ring (with unit), and $H^\bullet(A_\bullet, B_\bullet)$ (resp. $H^\bullet(B_\bullet, A_\bullet)$) has a natural graded left-module (resp. right-module) structure over this ring. The cohomology groups $H^\bullet(A_\bullet, B_\bullet)$ are functorial in both $A_\bullet$ and $B_\bullet$. A short exact sequence of complexes $0 \to A'_\bullet \to A_\bullet \to A''_\bullet \to 0$ induces long exact sequences

$$\cdots \longrightarrow H^i(A'_\bullet, B_\bullet) \longrightarrow H^i(A_\bullet, B_\bullet) \longrightarrow H^i(A''_\bullet, B_\bullet) \longrightarrow H^{i+1}(A'_\bullet, B_\bullet) \longrightarrow \cdots$$

and

$$\cdots \longrightarrow H^\bullet(B_\bullet, A''_\bullet) \longrightarrow H^\bullet(B_\bullet, A_\bullet) \longrightarrow H^\bullet(B_\bullet, A'_\bullet) \longrightarrow \cdots .$$

Suppose now that $\mathscr{C}$ has enough injectives. Suppose $A, B \in \mathscr{C}$, and let $I^A_\bullet$ be any injective resolution of A (more precisely: take an injective resolution of A, and remove the "leading" A), and similarly for $I^B_\bullet$. Then it is a reasonable exercise to describe canonical isomorphisms

$$H^\bullet(I^A_\bullet, I^B_\bullet) \cong H^\bullet(A, I^B_\bullet) \cong \operatorname{Ext}^\bullet(A, B)$$

where in the middle term, the "A" is interpreted as a complex that is zero, except the 0th piece is A.

Then the map (30.2.3.1) induces a (graded) map

$$\operatorname{Ext}^\bullet(A, B) \times \operatorname{Ext}^\bullet(B, C) \to \operatorname{Ext}^\bullet(A, C) \tag{30.2.3.2}$$

extending the natural map $\operatorname{Hom}(A, B) \times \operatorname{Hom}(B, C) \to \operatorname{Hom}(A, C)$. (Of course, one must show that the map (30.2.3.2) is independent of choice of injective resolutions of B and C.)

In particular, in the category of $\mathscr{O}$-modules on a ringed space X, we have (using Exercise 30.2.C(b)) a natural map

$$H^i(X, \mathscr{F}) \times \operatorname{Ext}^j(\mathscr{F}, \mathscr{G}) \to H^{i+j}(X, \mathscr{G}).$$

This is the source of the natural map in the trace version of Hom for Ext, discussed at the end of §30.1.

30.3 Serre duality for projective k-schemes

We now prove various versions of Serre duality for projective k-schemes, by leveraging what we know about Serre duality for projective space.

30.3.1. $\pi^!_{sh}$.
The key construction is a *right adjoint* to the pushforward π_*, when π is an affine morphism. This is surprising, as we usually think of π_* as a *right-adjoint* (to the pullback π^*), not a left adjoint. This can be seen as very roughly analogous to the surprising occasional *left-adjoint* to the pullback: extension by 0 (Exercise 2.6.G).

30.3.2. We begin with the ring-theoretic version. Suppose $B \to A$ is a ring morphism, M is an A-module, and N is a B-module. Note that $\operatorname{Hom}_B(A, N)$ naturally has the structure of an A-module. Also, M naturally carries the natural structure of a B-module; when we wish to emphasize its structure as a B-module, we sometimes call it M_B (see Exercise 1.5.E).

Consider the map

$$\operatorname{Hom}_A(M, \operatorname{Hom}_B(A, N)) \to \operatorname{Hom}_B(M_B, N) \tag{30.3.2.1}$$

defined as follows. Given $m \in M$, and an element ϕ of $\operatorname{Hom}_A(M, \operatorname{Hom}_B(A, N))$, send m to $\phi_m(1)$.

30.3.A. EXERCISE.
(a) Show that (30.3.2.1) is a homomorphism of B-modules.
(b) Show that (30.3.2.1) is a bijection. Thus $(M \mapsto M_B, N \mapsto \operatorname{Hom}_B(A, N))$ is an adjoint pair $\mathit{Mod}_A \leftrightarrow \mathit{Mod}_B$.
(c) Show that this bijection (30.3.2.1) behaves well with respect to localization of B with respect to an element of B.

Exercise 30.3.A(c) implies that this naturally "sheafifies" to a construction for an affine morphism $\pi : X \to Y$.

30.3.B. EXERCISE. Suppose $\pi : X \to Y$ is an affine morphism.
(a) Explain how the map

$$\pi^!_{sh} : \mathit{QCoh}_Y \longrightarrow \mathit{QCoh}_X$$

$$\mathscr{G} \longmapsto \pi^{-1}\mathscr{H}om_Y(\pi_*\mathscr{O}_X, \mathscr{G})$$

(where $\mathcal{H}om_Y(\pi_*\mathscr{O}_X, \mathscr{G})$ is interpreted as a quasicoherent sheaf on X via Exercise 17.1.E, and in this guise is denoted $\pi^{-1}\mathcal{H}om_Y(\pi_*\mathscr{O}_X, \mathscr{G})$) globalizes the construction of Exercise 30.3.A, yielding a covariant functor $QCoh_Y \to QCoh_X$. (Caution: the notation $\pi^!_{sh}$ is nonstandard, and is introduced only for the purposes of the arguments we will give.)
(b) Describe a natural isomorphism of quasicoherent sheaves on Y

$$\pi_*\pi^!_{sh}\mathscr{G} \cong \mathcal{H}om_Y(\pi_*\mathscr{O}_X, \mathscr{G}).$$

(c) Show that $(\pi_*, \pi^!_{sh})$ is an adjoint pair between $QCoh_X$ and $QCoh_Y$.

Caution: we have defined $\pi^!_{sh}$ only for categories of quasicoherent sheaves. If π is a finite morphism, and Y (and hence X) is locally Noetherian (the case that will be relevant for us), then $\pi^!_{sh}$ is a covariant functor from the category of *coherent* sheaves on Y to *coherent* sheaves on X. We may show this affine-locally, using the notation of §30.3.2. As A and N are both coherent B-modules, $\mathrm{Hom}_B(A, N)$ is a coherent B-module (cf. Exercise 13.7.A(b)), hence a finitely generated B-module, and hence a finitely generated A-module, hence a coherent A-module.

Thus if π is a finite morphism of locally Noetherian schemes, $(\pi_*, \pi^!_{sh})$ is an adjoint pair between Coh_X and Coh_Y.

30.3.3. If $\mathscr{F} \in QCoh_X$ and $\mathscr{G} \in QCoh_Y$, then there is a natural isomorphism

$$\pi_*\mathcal{H}om_X(\mathscr{F}, \pi^!_{sh}\mathscr{G}) \to \mathcal{H}om_Y(\pi_*\mathscr{F}, \mathscr{G}), \tag{30.3.3.1}$$

which affine-locally is the isomorphism (30.3.2.1) described in Exercise 30.3.A.

If $\mathscr{G}$ is an $\mathscr{O}_X$-module (not necessarily quasicoherent), it is in general *not* clear how to make sense of this construction to define an $\mathscr{O}_Y$-module $\pi^!_{sh}\mathscr{G}$. (Try it an see!) However, in the special case where π is a closed embedding, we *can* make sense of $\pi^!_{sh}\mathscr{G}$, as discussed in the next exercise.

30.3.C. EXERCISE (USED IN §30.4). Suppose $\pi : X \to Y$ is a *closed embedding* of schemes and $\mathscr{G}$ is an $\mathscr{O}_Y$-module.
(a) Explain why $\pi^!_{sh}(\mathscr{G}) := \mathcal{H}om_Y(\pi_*\mathscr{O}_X, \mathscr{G})$ naturally has the structure of an $\mathscr{O}_X$-module. Hint: if $\mathscr{I}$ is the ideal sheaf of X, explain how $\pi^!_{sh}(\mathscr{G})$ (over some open subset $U \subset Y$) is annihilated by "functions vanishing on X" (elements of $\mathscr{I}(U)$). Hence we have defined a map $\pi^!_{sh} : Mod_{\mathscr{O}_Y} \to Mod_{\mathscr{O}_X}$ extending the map of Exercise 30.3.B(a).
(b) Show that $(\pi_*, \pi^!_{sh})$ is an adjoint pair between $Mod_{\mathscr{O}_X}$ and $Mod_{\mathscr{O}_Y}$.
(c) Show that $\pi^!_{sh}$ sends injective $\mathscr{O}_Y$-modules to injective $\mathscr{O}_X$-modules. (Hint: π_* is exact; use Exercise 23.5.B.)

30.3.4. *Remark.* If π is finite and flat, which is the case most of interest to us, $\pi^!_{sh}$ agrees (in the only possible sense of the word) with $\pi^!$, one of Grothendieck's "six operations", and indeed this motivates our notation ("a variant of $\pi^!$ for sheaves"). But $\pi^!$ naturally lives in the world of derived categories, so we will not discuss it here. (For more on $\pi^!$, see [**KS1**, Ch. III].)

We now apply the machinery of $\pi^!_{sh}$ to Serre duality.

30.3.5. Projective n-dimensional schemes are covers of $\mathbb{P}^n$.

30.3.6. Proposition. — *Suppose X is a projective k-scheme of pure dimension n.*

(a) *There exists a finite morphism $\pi : X \to \mathbb{P}^n$.*
(b) *If furthermore X is Cohen-Macaulay, then π is flat.*

Proof. Part (b) follows from part (a) by the Miracle Flatness Theorem 26.2.11, so we will prove part (a).

Choose a closed embedding $j : X \hookrightarrow \mathbb{P}^N$. For simplicity of exposition, first assume that k is an infinite field. By Exercise 11.3.C(d), there is a linear space L of codimension $n+1$ (one less than complementary dimension) disjoint from X. Projection from L yields a morphism $\pi : X \to \mathbb{P}^n$. The morphism π is affine ($\mathbb{P}^N \setminus L \to \mathbb{P}^n$ is affine, and the closed embedding $X \hookrightarrow \mathbb{P}^N \setminus L$ is, like all closed embeddings, affine) and projective (by the Cancellation Theorem 10.1.19 for projective morphisms), so π is finite (projective affine morphisms of locally Noetherian schemes are finite, Corollary 18.1.8).

30.3.D. Exercise. Prove Proposition 30.3.6(a), without the assumption that k is infinite. Hint: using Exercise 11.3.C(c), show that there is some d such that there is an intersection of $N-n-1$ degree d hypersurfaces missing X. Then apply the above argument to the dth Veronese embedding of $\mathbb{P}^N$ (§8.2.6).

$\square$

30.3.7. Serre duality on X via $\pi^!_{sh}$.

Suppose $\pi : X \to Y$ is a finite morphism of projective k-schemes of pure dimension n, and we have a coherent sheaf ω_Y on Y. (We will soon apply this in the case where $Y = \mathbb{P}^n_k$, but we may as well avoid distraction and needless specificity.)

30.3.8. Proposition. — *Suppose π is flat. If functorial Serre duality for* Ext *holds for Y, with dualizing sheaf ω_Y, then functorial Serre duality for vector bundles holds for X, with dualizing sheaf $\pi^!_{sh}(\omega_Y)$.*

Note the mismatch of the hypotheses and the conclusion: we use functorial Serre duality for Ext in the hypothesis, and get functorial Serre duality only for vector bundles in the conclusion.

Proof. For each i, and each finite rank locally free sheaf $\mathscr{F}$ on X, we have isomorphisms (functorial in $\mathscr{F}$):

$$\begin{aligned} H^{n-i}(X, \mathscr{F}^\vee \otimes \pi^!_{sh}\omega_Y) &\cong H^{n-i}(Y, \pi_*(\mathscr{F}^\vee \otimes \pi^!_{sh}\omega_Y)) \quad \text{(affineness of } \pi, \text{ §18.1 } \textbf{(v)}) \\ &\cong H^{n-i}(Y, \pi_*(\mathcal{H}om_X(\mathscr{F}, \pi^!_{sh}\omega_Y))) \quad \text{(Exercise 13.7.B)} \\ &\cong H^{n-i}(Y, \mathcal{H}om_Y(\pi_*\mathscr{F}, \omega_Y)) \quad \text{(equ. (30.3.3.1))} \\ &\cong H^{n-i}(Y, (\pi_*\mathscr{F})^\vee \otimes \omega_Y) \quad \text{(Exercises 24.4.E and 13.7.B)} \\ &\cong H^i(Y, \pi_*\mathscr{F})^\vee \quad \text{(functorial Serre duality for Ext on } Y) \\ &\cong H^i(X, \mathscr{F})^\vee \quad \text{(affineness of } \pi, \text{ §18.1 } \textbf{(v)}) \end{aligned}$$

$\square$

30.3.9. Corollary. — *Functorial Serre duality for vector bundles holds for every Cohen-Macaulay equidimensional projective k-scheme.*

Proof. Combine Proposition 30.3.8 with Proposition 30.3.6(b) and §30.1.10. ☐

30.3.10. Proposition. — *If the trace version of Serre duality for* Hom *holds for* Y *with dualizing sheaf* (ω_Y, t_Y)*, then functorial Serre duality for* Hom *holds for* X *with dualizing sheaf* $\pi^!_{sh}\omega_Y$*.*

Recall from Exercise 30.1.A that for Serre duality for Hom, the trace version is equivalent to the functorial version. We state the Proposition in this awkward way because we use t_Y, and conclude functoriality for Serre duality for Hom for X. But the trace morphism for X will come cheaply as the diagonal arrow in (30.3.10.1).

Note that we have no flatness hypotheses on π (unlike the corresponding propositions for Serre duality for vector bundles and, soon, for Ext).

Proof. The following diagram commutes, and is functorial in $\mathscr{F}$.
(30.3.10.1)

$$\begin{tikzcd}
\mathrm{Hom}_X(\mathscr{F}, \pi^!_{\mathrm{sh}}\omega_Y) \times H^n(X, \mathscr{F}) \arrow[r] \arrow[dd] & H^n(X, \pi^!_{\mathrm{sh}}\omega_Y) \arrow[d] \arrow[ddr, "t_X"] & \\
& H^n(Y, \pi_*\pi^!_{\mathrm{sh}}\omega_Y) \arrow[d] & \\
\mathrm{Hom}_Y(\pi_*\mathscr{F}, \omega_Y) \times H^n(Y, \pi_*\mathscr{F}) \arrow[r] & H^n(Y, \omega_Y) \arrow[r, "t_Y"] & k
\end{tikzcd}$$

(The left vertical arrow is an isomorphism; it is an isomorphism on each factor.) The commutativity of the diagram relies on the adjointness of $(\pi_*, \pi^!_{\mathrm{sh}})$, and the affineness of π. ☐

30.3.11. Corollary. — *The functorial and trace versions of Serre duality for* Hom *hold for all equidimensional projective* k*-schemes.*

Proof. Combine Proposition 30.3.10 with Proposition 30.3.6(a) and Proposition 30.1.7. ☐

30.3.12. Proposition. — *If* X *is a Cohen-Macaulay projective* k*-scheme of pure dimension* n*, and* $\pi : X \to \mathbb{P}^n$ *is a finite flat morphism (as in Proposition 30.3.6(b)), then functorial Serre duality for* Ext *holds for* X *with dualizing sheaf* $\omega_X := \pi^!_{sh}(\omega_{\mathbb{P}^n})$*.*

Notice that unlike Proposition 30.3.8 and 30.3.10, we did not state this for general $\pi : X \to Y$. We will use the fact that the target is $\mathbb{P}^n$.

Proof. The argument will parallel that of §30.1.10: we will show an isomorphism of δ-functors

(30.3.12.1) $$\mathrm{Ext}^i_X(\mathscr{F}, \pi^!_{\mathrm{sh}}(\omega_{\mathbb{P}^n})) \to H^{n-i}(X, \mathscr{F})^\vee.$$

We already have an isomorphism for $i = 0$ (Corollary 30.3.11), so it suffices to show that both sides of (30.3.12.1) are *universal* δ-functors. We do this by verifying the criterion of Theorem 23.2.8. For any coherent sheaf $\mathscr{F}$ on X we can find a surjection $\mathscr{O}(-m)^{\oplus N} \to \mathscr{F}$ for $m \gg 0$ (and some N), so it suffices to show that for $m \gg 0$,

$$\mathrm{Ext}^i_X(\mathscr{O}(-m), \pi^!_{\mathrm{sh}}(\omega_{\mathbb{P}^n})) = 0 \quad \text{and} \quad H^{n-i}(X, \mathscr{O}(-m))^\vee = 0$$

for $i > 0$ and $m \gg 0$. For the first, we have (for $i > 0$ and $m \gg 0$)

$$\begin{aligned}
\operatorname{Ext}^i_X(\mathscr{O}(-m), \pi^!_{sh}(\omega_{\mathbb{P}^n})) &\cong H^i(X, \pi^!_{sh}(\omega_{\mathbb{P}^n})(m)) \quad \text{(Exercise 30.2.H)} \\
&= 0 \qquad \text{(Serre vanishing, Thm. 18.1.4(b))}
\end{aligned}$$

For the second, we have

$$\begin{aligned}
H^{n-i}(X, \mathscr{O}_X(-m)) &= H^{n-i}(X, \mathscr{O}_X \otimes \pi^* \mathscr{O}(-m)) \\
&\cong H^{n-i}(\mathbb{P}^n, (\pi_* \mathscr{O}_X) \otimes \mathscr{O}(-m))
\end{aligned}$$

by the projection formula (Exercise 18.8.E(b)). Now $\pi_* \mathscr{O}_X$ is locally free by Exercise 24.4.E, so by functorial Serre duality for vector bundles on $\mathbb{P}^n$, we have (for $m \gg 0$)

$$\begin{aligned}
H^{n-i}(\mathbb{P}^n, (\pi_* \mathscr{O}_X) \otimes \mathscr{O}(-m)) &\cong H^i(\mathbb{P}^n, (\pi_* \mathscr{O}_X)^\vee \otimes \mathscr{O}(m) \otimes \omega_{\mathbb{P}^n}) \\
&\cong H^i(\mathbb{P}^n, ((\pi_* \mathscr{O}_X)^\vee \otimes \omega_{\mathbb{P}^n}) \otimes \mathscr{O}(m)) \\
&= 0 \qquad \text{(Serre vanishing, Thm. 18.1.4(b))}
\end{aligned}$$

□

30.3.13. Corollary. — *Functorial Serre duality for* Ext *holds for all equidimensional Cohen-Macaulay projective* k*-schemes.*

Proof. Combine Proposition 30.3.12 with Proposition 30.3.6(b) and the projective space case of §30.1.10. □

30.3.14. *Remark: the trace version of Serre duality for vector bundles.* Here is an outline of how the above proof can be extended to prove the trace version of Serre duality for vector bundles on all equidimensional Cohen-Macaulay projective k-schemes. Continuing the notation of Proposition 30.3.12, the cup product of §30.2.3 followed by the trace map gives a pairing

$$\operatorname{Ext}^i_X(\mathscr{F}, \omega_X) \times H^{n-i}(X, \mathscr{F}) \xrightarrow{\cup} H^n(X, \omega_X) \xrightarrow{t} k\ ,$$

which is functorial in $\mathscr{F}$. This induces a functorial map (30.3.12.1). One shows that this particular map is a map of δ-functors. Then by the above argument, this particular map is an isomorphism.

30.4 The adjunction formula for the dualizing sheaf, and $\omega_X = \mathscr{K}_X$

The dualizing sheaf ω behaves well with respect to slicing by effective Cartier divisors, and this will be useful to obtaining Desideratum 30.1.1, i.e., the miracle that the canonical bundle is Serre-dualizing (§18.5.2). In order to show this, we give a description for the dualizing sheaf for a subvariety of a variety satisfying Serre duality for Ext.

30.4.1. *Preliminary Observation.* Suppose $\pi : X \to Y$ is a closed embedding. Then for any $\mathscr{O}_Y$-module, $\mathscr{E}xt^i_Y(\pi_* \mathscr{O}_X, \mathscr{F})$ naturally has the structure of an $\mathscr{O}_X$-module. (Reason: we compute this by taking an injective resolution of $\mathscr{F}$ by $\mathscr{O}_Y$-modules, then truncating, then applying $\mathscr{H}om_Y(\pi_* \mathscr{O}_X, \cdot)$. But for any $\mathscr{O}_Y$-module

$\mathscr{G}$, $\mathscr{H}om_Y(\pi_* \mathscr{O}_Y, \mathscr{G})$ has the structure of an $\mathscr{O}_X$-module.) To emphasize its structure as an $\mathscr{O}_X$-module, we write it as $\pi^{-1} \mathscr{E}xt^i_Y(\pi_* \mathscr{O}_X, \mathscr{F})$.

30.4.A. EXERCISE. Show that if Y (and hence X) is locally Noetherian, and $\mathscr{G}$ is a coherent sheaf on Y, then $\pi^{-1} \mathscr{E}xt^i_Y(\pi_* \mathscr{O}_X, \mathscr{G})$ is a coherent sheaf on X. (Hint: $\mathscr{E}xt^i_Y(\pi_* \mathscr{O}_X, \mathscr{G})$ is a coherent sheaf on Y, Exercise 30.2.F.)

30.4.2. The dualizing sheaf for a subvariety in terms of the dualizing sheaf of the ambient variety.

For the rest of this section, $\pi : X \to Y$ will be a closed embedding of pure-dimensional projective k-schemes of dimension n and N respectively, where Y satisfies functorial Serre duality for Ext. Let $r = N - n$ (the codimension of X in Y). By Corollary 30.3.11, X satisfies functorial Serre duality for Hom. We now identify ω_X in terms of ω_Y.

30.4.3. Theorem. — *We have* $\omega_X = \pi^{-1} \mathscr{E}xt^r_Y(\pi_* \mathscr{O}_X, \omega_Y)$.

Before we prove Theorem 30.4.3, we explain what happened to the "earlier" $\mathscr{E}xt^r_Y(\pi_* \mathscr{O}_X, \omega_Y)$'s.

30.4.4. Proposition. — *Suppose that* $\pi : X \hookrightarrow Y$ *is a closed embedding of pure-dimensional projective* k*-schemes of dimension* n *and* N *respectively, and* Y *satisfies functorial Serre duality for* Ext. *Then for all* $i < r := N - n$, $\mathscr{E}xt^i_Y(\pi_* \mathscr{O}_X, \omega_Y) = 0$.

Proof. As $\mathscr{E}xt^i_Y(\mathscr{O}_X, \omega_Y)$ is coherent (Exercise 30.2.F), it suffices to show that $\mathscr{E}xt^i_Y(\mathscr{O}_X, \omega_Y) \otimes \mathscr{O}(m)$ has no nonzero global sections for $m \gg 0$ (as for any coherent sheaf $\mathscr{G}$ on Y, $\mathscr{G}(m)$ is generated by global sections for $m \gg 0$ by Serre's Theorem A, Theorem 15.3.8). By Exercise 30.2.H,

$$H^j(Y, \mathscr{E}xt^i_Y(\mathscr{O}_X, \omega_Y)(m)) = H^j(Y, \mathscr{E}xt^i_Y(\mathscr{O}_X, \omega_Y(m)))$$

For $m \gg 0$, by Serre vanishing, $H^j(Y, \mathscr{E}xt^i_Y(\mathscr{O}_X, \omega_Y)(m)) = 0$. Thus by the local-to-global sequence for Ext (Exercise 30.2.K),

$$H^0(Y, \mathscr{E}xt^i_Y(\mathscr{O}_X, \omega_Y(m))) = \mathrm{Ext}^i_Y(\mathscr{O}_X, \omega_Y(m))).$$

By Exercise 30.2.H again, then functorial Serre duality for Ext on Y,

$$\mathrm{Ext}^i_Y(\mathscr{O}_X, \omega_Y(m))) = \mathrm{Ext}^i_Y(\mathscr{O}_X(-m), \omega_Y) = H^{N-i}(Y, \mathscr{O}_X(-m))$$

which is 0 if $N - i > n$, as the cohomology of a quasicoherent sheaf on a projective scheme vanishes in degree higher than the dimension of the sheaf's support (dimensional vanishing, Theorem 18.2.6). □

The most difficult step in the proof of Theorem 30.4.3 is the following.

30.4.5. Lemma. — *Suppose that* $\pi : X \to Y$ *is a closed embedding of codimension* r *of equidimensional projective* k*-schemes, and* Y *satisfies functorial Serre duality for* Ext. *We have an isomorphism, functorial in* $\mathscr{F} \in Coh_X$*:*

$$\mathrm{Hom}_X(\mathscr{F}, \pi^{-1} \mathscr{E}xt^r_Y(\mathscr{O}_X, \omega_Y)) \cong \mathrm{Ext}^r_Y(\pi_* \mathscr{F}, \omega_Y). \tag{30.4.5.1}$$

Proof. Choose an injective resolution

$$0 \to \omega_Y \to \mathscr{I}^0 \to \mathscr{I}^1 \to \mathscr{I}^2 \to \cdots$$

of ω_Y. To compute the right side of (30.4.5.1), we drop the ω_Y from this resolution, and apply $\mathrm{Hom}_Y(\pi_* \mathscr{F}, \cdot)$, to obtain

$$0 \to \mathrm{Hom}_Y(\pi_* \mathscr{F}, \mathscr{I}^0) \to \mathrm{Hom}_Y(\pi_* \mathscr{F}, \mathscr{I}^1) \to \mathrm{Hom}_Y(\pi_* \mathscr{F}, \mathscr{I}^2) \to \cdots,$$

or (by adjointness of π_* and $\pi^!_{\mathrm{sh}}$ for $\mathscr{O}$-modules when π is a closed embedding, Exercise 30.3.C(b)):

$$0 \to \mathrm{Hom}_X(\mathscr{F}, \pi^!_{\mathrm{sh}} \mathscr{I}^0) \to \mathrm{Hom}_X(\mathscr{F}, \pi^!_{\mathrm{sh}} \mathscr{I}^1) \to \mathrm{Hom}_X(\mathscr{F}, \pi^!_{\mathrm{sh}} \mathscr{I}^2) \to \cdots.$$

Motivated by this, consider the complex

$$0 \to \pi^!_{\mathrm{sh}} \mathscr{I}^0 \to \pi^!_{\mathrm{sh}} \mathscr{I}^1 \to \pi^!_{\mathrm{sh}} \mathscr{I}^2 \to \cdots. \tag{30.4.5.2}$$

Note first that (30.4.5.2) is indeed a complex, and second that $\pi^!_{\mathrm{sh}} \mathscr{I}^i$ are injectives $\mathscr{O}_X$-modules (Exercise 30.3.C(c)).

30.4.B. EXERCISE. Show that the cohomology of (30.4.5.2) at the ith step is $\mathcal{E}xt^i_Y(\pi_* \mathscr{O}_X, \omega_Y)$.

Thus by Proposition 30.4.4, the complex (30.4.5.2) is exact before the rth step.

30.4.C. EXERCISE. Show that there exists a direct sum decomposition $\pi^!_{\mathrm{sh}} \mathscr{I}_r = \mathscr{J} \oplus \mathscr{K}$, so that

$$0 \to \pi^!_{\mathrm{sh}} \mathscr{I}^0 \to \pi^!_{\mathrm{sh}} \mathscr{I}^1 \to \cdots \to \pi^!_{\mathrm{sh}} \mathscr{I}^{r-1} \to \mathscr{J} \to 0$$

is exact. Hint: work out the case $r = 1$ first. Another hint: Notice that if

$$0 \longrightarrow \mathscr{I}^0 \xrightarrow{\alpha} \mathscr{F}$$

is exact and $\mathscr{I}^0$ is injective, then there is a map $\beta : \mathscr{F} \to \mathscr{I}^0$ "splitting" α, allowing us to write $\mathscr{F}$ as a direct sum $\mathscr{I}^0 \oplus \mathscr{J}^1$. If $\mathscr{F}$ is furthermore injective, then $\mathscr{J}^1$ is injective too. This is the beginning of an induction.

30.4.D. EXERCISE. (We continue the notation of the previous exercise.) Identify $\ker(\mathscr{K} \to \pi^!_{\mathrm{sh}} \mathscr{I}^{r+1})$ with $\pi^{-1} \mathcal{E}xt^r_Y(\pi_* \mathscr{O}_X, \omega_Y)$.

30.4.E. EXERCISE. Put together the pieces above to complete the proof of Lemma 30.4.5. □

We are now ready to prove Theorem 30.4.3.

30.4.6. *Proof of Theorem 30.4.3.* Suppose $\mathscr{F}$ is a coherent sheaf on X. We wish to find an isomorphism $\mathrm{Hom}_X(\mathscr{F}, \pi^{-1} \mathcal{E}xt^r_Y(\pi_* \mathscr{O}_X, \omega_Y)) \cong H^n(X, \mathscr{F})^\vee$, functorial in $\mathscr{F}$. By Lemma 30.4.5, we have a functorial isomorphism $\mathrm{Hom}_X(\mathscr{F}, \pi^{-1} \mathcal{E}xt^r_Y(\pi_* \mathscr{O}_X, \omega_Y)) \cong \mathrm{Ext}^r_Y(\mathscr{F}, \omega_Y)$. By functorial Serre duality for Ext on Y, we have a functorial isomorphism $\mathrm{Ext}^r_Y(\mathscr{F}, \omega_Y) \cong H^{N-r}(Y, \mathscr{F})^\vee$. As $\mathscr{F}$ is a coherent sheaf on X, and $N - r = n$, this is precisely $H^n(X, \mathscr{F})^\vee$. □

30.4.7. Applying Theorem 30.4.3. We first apply Theorem 30.4.3 in the special case where X is an effective Cartier divisor on Y. We can compute the dualizing sheaf $\pi^{-1} \mathcal{E}xt^1_Y(\pi_* \mathscr{O}_X, \omega_Y)$ by computing $\mathcal{E}xt^1_Y(\pi_* \mathscr{O}_X, \omega_Y)$ using any locally free resolution (on Y) of $\pi_* \mathscr{O}_X$ (Exercise 30.2.E). But $\pi_* \mathscr{O}_X$ has a particularly simple resolution, the closed subscheme exact sequence (13.5.6.1) for X:

$$0 \to \mathscr{O}_Y(-X) \to \mathscr{O}_Y \to \pi_* \mathscr{O}_X \to 0. \tag{30.4.7.1}$$

We compute $\mathscr{E}xt^\bullet(\pi_* \mathscr{O}_X, \omega_Y)$ by truncating $\pi_* \mathscr{O}_X$, and applying $\mathscr{H}om(\cdot, \omega_Y)$: $\mathscr{E}xt^\bullet(\pi_* \mathscr{O}_X, \omega_Y)$ is the cohomology of

$$0 \to \mathscr{H}om_Y(\mathscr{O}_Y, \omega_Y) \to \mathscr{H}om(\mathscr{O}_Y(-X), \omega_Y) \to 0,$$

$$\text{i.e.} \qquad 0 \to \omega_Y \to \omega_Y \otimes \mathscr{O}_Y(X) \to 0.$$

We immediately see that $\mathscr{E}xt^i(\pi_* \mathscr{O}_X, \omega_Y) = 0$ if $i \neq 0, 1$. Furthermore, $\mathscr{E}xt^0(\pi_* \mathscr{O}_X, \omega_Y) = 0$ by Proposition 30.4.4. (Unimportant aside: you can use this to show that ω_Y has no embedded points.)

We now consider $\omega_X = \operatorname{coker}(\omega_Y \to \omega_Y \otimes \mathscr{O}_Y(X))$. Tensoring (30.4.7.1) with the invertible sheaf $\mathscr{O}_Y(X)$, and then tensoring with ω_Y, yields

$$\omega_Y \to \omega_Y \otimes \mathscr{O}_Y(X) \to \omega_Y \otimes \mathscr{O}_Y(X)|_X \to 0$$

The right term is often (somewhat informally) written as $\omega_Y(X)|_X$. Thus $\omega_X = \operatorname{coker}(\omega_Y \to \omega_Y \otimes \mathscr{O}_Y(X)) = \omega_Y(X)|_X$, and this identification is *canonical*.

We have shown the following.

30.4.8. Proposition (the adjunction formula). — *Suppose that Y is a Cohen-Macaulay projective scheme of pure dimension n (which satisfies functorial Serre duality for* Ext *by Corollary 30.3.13), and X is an effective Cartier divisor on Y. (Hence X is Cohen-Macaulay by §26.2.4, and thus satisfies functorial Serre duality for* Ext *by Corollary 30.3.13.) Then $\omega_X = \omega_Y(X)|_X$. In particular, if ω_Y is an invertible sheaf on Y, then ω_X is an invertible sheaf on X.*

As an immediate application, we have the following.

30.4.F. Exercise. Suppose X is a complete intersection in $\mathbb{P}^n$, of hypersurfaces of degrees $d_1, \dots, d_r$. (Note that X is Cohen-Macaulay by Proposition 26.2.6, and thus satisfies functorial duality for Ext, by Corollary 30.3.13.) Show that $\omega_X \cong \mathscr{O}_X(-n-1+\sum d_i)$. If furthermore X is smooth, show that $\omega_X \cong \det \Omega_X$. (Hint for the last sentence: use the adjunction formula for $\mathscr{K}$, Exercise 21.5.B).

But we can say more.

30.4.G. Exercise. Suppose Y is a smooth pure-dimensional variety, and $\pi : X \hookrightarrow Y$ is a codimension r regular embedding with normal sheaf $\mathscr{N}_{X/Y}$ (which is locally free, by Proposition 21.2.16(b)). Suppose $\mathscr{L}$ is an invertible sheaf on Y.
(a) Show that $\mathscr{E}xt^i_Y(\pi_* \mathscr{O}_X, \mathscr{L}) = 0$ if $i \neq r$.
(b) Describe a *canonical* isomorphism $\mathscr{E}xt^r(\pi_* \mathscr{O}_X, \mathscr{L}) \cong (\det \mathscr{N}_{X/Y}) \otimes_{\mathscr{O}_X} \mathscr{L}|_X$.

(Note that because Exercise 30.4.G has nothing explicitly to do with duality, we have no projectivity assumptions on Y; it is a completely local question.) From Exercise 30.4.G we deduce the following.

30.4.H. Important Exercise. Suppose X is a codimension r regular embedding in $\mathbb{P}^n_k$. (Then X is Cohen-Macaulay by Proposition 26.2.6, and thus satisfies functorial Serre duality for Ext by Corollary 30.3.13.) Show that

$$\omega_X \cong \pi^{-1} \mathscr{E}xt^r_{\mathbb{P}^n}(\pi_* \mathscr{O}_X, \omega_{\mathbb{P}^n}) \cong (\det \mathscr{N}_{X/Y}) \otimes \omega_{\mathbb{P}^n}|_X.$$

In particular, ω_X is an invertible sheaf.

30.4.I. IMPORTANT EXERCISE. Suppose X is a smooth pure codimension r subvariety of $\mathbb{P}^n_k$ (and hence a regular embedding, by Exercise 12.2.K). Show that $\omega_X \cong \mathscr{K}_X$. Hint: both sides satisfy adjunction (see Exercise 21.5.B for adjunction for Ω): they are isomorphic to $(\det \mathscr{N}_{X/Y}) \otimes \omega_{\mathbb{P}^n}|_X \cong (\det \mathscr{N}_{X/Y}) \otimes (\mathscr{K}_{\mathbb{P}^n})|_X$.

30.4.9. *Remark.* As long promised, the version of Serre duality given in Theorem 18.5.1 now follows by combining Corollary 30.3.9 with Exercise 30.4.I.

Aux yeux de ces amateurs d'inquiétude et de perfection, un ouvrage n'est jamais achevé, *mot qui pour eux n'a aucun sens, mais* abandonné; *et cet abandon, qui le livre aux flammes ou au public (et qu'il soit l'effet de la lassitude ou de l'obligation de livrer), leur est une sorte d'*accident, *comparable à la rupture d'une réflexion . . .*

In the eyes of those lovers of anxiety and perfection, a work is never finished — *a word that for them has no meaning — but* abandoned*; and this abandonment, whether to the flames or to the public (and which is the result of weariness or an obligation to deliver) is a kind of accident to them, like the breaking off of a reflection . . .*

— *P. Valéry*, [**Val**, p. 1497]

Bibliography

[Al] P. Aluffi, *Algebra, Chapter 0*, Grad. Stud. in Math. **104**, Amer. Math. Soc., Providence, RI, 2009.

[ACGH] E. Arbarello, M. Cornalba, P. Griffiths, and J. Harris, *Geometry of Algebraic Curves, Vol. I*, Grund. Math. Wiss. **267**, Springer-Verlag, New York, 1985.

[Ar1] M. Artin, *On the solutions of analytic equations*, Invent. Math. **5** (1968), no. 4, 277–291.

[Ar2] ———, *Algebraic approximation of structures over complete local rings*, Publ. Math. IHES **36** (1969), 23–58.

[Ar3] ———, *Lipman's proof of resolution of singularities for surfaces*, in *Arithmetic geometry (Storrs, Conn., 1984)*, Springer-Verlag, New York, 1986, 267–287.

[Ar4] ———, *Algebra*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1991.

[At1] M. F. Atiyah, *On analytic surfaces with double points*, Proc. Roy. Soc. London Ser. A **247** (1958), 237-244.

[At2] ———, *Mathematics in the 20th century*, Amer. Math. Monthly **108** (2001), no. 7, 654–666.

[AtM] M. F. Atiyah and I. G. Macdonald, *Introduction to Commutative Algebra*, Addison-Wesley Publ. Co., Reading, MA, 1969.

[Ba] L. Bădescu, *Algebraic Surfaces*, V. Maşek trans., Universitext, Springer-Verlag, New York, 2001.

[BCDKT] M. Baker, B. Conrad, S. Dasgupta, K. Kedlaya, and J. Teitelbaum, *p-adic geometry*, D. Savitt and D. Thakur ed., Univ. Lect. Series **45**, Amer. Math. Soc., Providence, RI, 2008.

[BPV] W. Barth, C. Peters, and A. van de Ven, *Compact Complex Surfaces*, Ergeb. Math. Grenz. **4**, Springer-Verlag, Berlin, 1984.

[Be] A. Beauville, *Complex Algebraic Surfaces*, 2nd ed., R. Barlow trans., Cambridge U.P., Cambridge, 1996.

[BL] C. Birkenhake and H. Lange, *Complex Abelian Varieties*, 2nd ed., Grund. Math. Wiss. **302**, Springer-Verlag, Berlin, 2004.

[BLR] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron Models*, Ergeb. Math. Grenz. **21**, Springer-Verlag, Berlin, 1990.

[Bo] N. Bourbaki, *Éléments de mathématique, Fasc. XXX: Algèbre Commutative*, Actualités Scientifiques et Industrielles, no. 1308, Hermann, Paris, 1964.

[BGS] S. Boissière, O. Gabber, and O. Serman, *Sur le produit de variétés localement factorielles ou $\mathbb{Q}$-factorielles*, arxiv/1104.1861v1, preprint 2011.

[BP] R. Brown and T. Porter, *Analogy, concepts and methodology, in mathematics*, UWB Math Preprint, `www.maths.bangor.ac.uk/research/ftp/rpam/06_08.pdf`, May 26, 2006.

[Carr] L. Caroll, *Alice's Adventures in Wonderland*, Macmillan, London, 1865.

[Cart] P. Cartier, *Les groupes* $\mathrm{Ext}^s(A, B)$, in *Séminaire A. Grothendieck: Algèbre homologique*, t. 1, 1957, no. 3.

[Co] B. Conrad, *Deligne's notes on Nagata compactifications*, J. Ramanujan Math. Soc. **22** (2007), no. 3, 205–257.

[Da] E. D. Davis, *Ideals of the principal class, R-sequences and a certain monoidal transformation*, Pacific J. Math. **20** (1967), 197–205.

[DW] R. Dedekind and H. Weber, *Theory of algebraic functions of one variable*, J. Stillwell trans. History of Math. **39**, Amer. Math. Soc., Providence, RI; London Math. Soc. London, 2012.

[De] P. Deligne, *Courbes elliptiques: formulaire d'après Tate*, in *Modular Functions of One Variable IV (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, 53–73, Lect. Notes in Math. **476**, Springer-Verlag, Berlin, 1975.

[DeI] P. Deligne and L. Illusie, *Relèvements modulo* p^2 *et décomposition du complexe de de Rham*, Invent. Math. **89** (1987), no. 2, 247–270.

[Di] J. Dieudonné, *On regular sequences*, Nagoya Math. J. **27** (1966), 355–356.

[Do] I. Dolgachev, *Luigi Cremona and cubic surfaces*, in *Luigi Cremona (1830–1903)*, 55–70, Incontr. Studio, **36**, Instituto Lombardo di Scienze e Lettere, Milan, 2005.

[DS] R. Donagi and R. C. Smith, *The structure of the Prym map*, Acta Math. **146** (1981), no. 1–2, 25–102.

[DZ] H. Duan and X. Zhao, *Schubert calculus and the integral cohomology of exceptional Lie groups*, arXiv:0711.2541v10, preprint 2007.

[DF] D. Dummit and R. Foote, *Abstract Algebra*, 3rd ed., Wiley and Sons, Inc., Hoboken, NJ, 2004.

[E] D. Eisenbud, *Commutative Algebra with a View to Algebraic Geometry*, Grad. Texts in Math. **150**, Springer-Verlag, New York, 1995.

[EGH] D. Eisenbud, M. Green, and J. Harris, *Cayley-Bacharach theorems and conjectures*, Bull. AMS **33** (1996), no. 3, 295–324.

[EH] D. Eisenbud and J. Harris, *The Geometry of Schemes*, Grad. Texts in Math. **197**, Springer-Verlag, New York, 2000.

[Ek] P. Eklof, *Lefschetz's principle and local functors*, Proc. Amer. Math. Soc. **37** (1973), no. 2, 1973.

[EE] E. Enochs and S. Estrada, *Relative homological algebra in the category of quasi-coherent sheaves*, Adv. Math. **194** (2005), 284–295.

[FGIKNV] B. Fantechi, L. Göttsche, L. Illusie, S. Kleiman, N. Nitsure and A. Vistoli, *Fundamental algebraic geometry: Grothendieck's FGA explained*, Mathematical Surveys and Monographs 123, Amer. Math. Soc., Providence, RI, 2005.

[FK] H. M. Farkas and I. Kra, *Riemann Surfaces*, Grad. Texts in Math. **71**, Springer-Verlag, New York, 1980.

[FR] G. Frey and H.-G. Rück, *The strong Lefschetz principle in algebraic geometry*, manuscripta math. **55** (1986), no. 3–4, 385–401.

[F] W. Fulton, *Intersection Theory*, 2nd ed., Ergeb. Math. Grenz. **2**, Springer-Verlag, New York, 1998.

[Ge] S. Germain, *Œuvres Philosophiques de Sophie Germain* (nov. éd.), Librarie de Firmin-Didot et Cie, Paris, 1896, `http://fr.wikisource.org/wiki/OEuvres_philosophiques_de_Sophie_Germain`.

[GS] P. Gille and T. Szamuely, *Central Simple Algebras and Galois Cohomology*, Cambridge U.P., Cambridge, 2006.

[GW] U. Görtz and T. Wedhorn, *Algebraic Geometry I*, Vieweg + Teubner, Wiesbaden, 2010. (Appendices C, D, and E, collecting properties of schemes and morphisms and their relationships, are alone worth the price.)

[GR] H. Grauert and R. Remmert, *Coherent Analytic Sheaves*, Grund. Math. Wiss. **265**, Springer-Verlag, Berlin, 1984.

[GH1] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Wiley-Interscience, New York, 1978.

[GH2] ———, *On the Noether-Lefschetz theorem and some remarks on codimension two cycles*, Math. Ann. **271** (1985), 31–51.

[Gr1] A. Grothendieck, *Sur quelques points d'algèbre homologique*, Tôhoku Math. J. (2) **9** (1957), 119–221.

[Gr2] ———, *Sur la classification des fibrés holomorphes sur la sphère de Riemann*, Amer. J. Math. **79** (1957), 121–138.

[Gr3] ———, *Théorèmes de dualité pour les faisceaux algébriques cohérents*, Séminaire N. Bourbaki **4**, exp. no. 149, Soc. Math. France, Paris, 1995, 169–193.

[Gr4] ———, *The cohomology theory of abstract algebraic varieties*, in *Proc. Internat. Congress Math. (Edinburgh, 1958)*, Cambridge U.P., New York, 1960.

[Gr5] ———, *Fondements de la géométrie algébrique*, in *Collected Bourbaki Talks*, Secrétariat Mathématique, Paris, 1962.

[Gr6] ———, *Récoltes et Semailles*, unpublished.

[Gr-EGA] ——— (w. J. Dieudonné), *Éléments de Géométrie Algébrique*, Publ. Math. IHES, Part I: **4** (1960); Part II: **8** (1961); Part III: **11** (1961), 17 (1963); Part **IV**: 20 (1964), 24 (1965), 28 (1966), 32 (1967).

[Gr-EGA′] ——— (w. J. Dieudonné), *Éléments de Géométrie Algébrique I: Le langage des schémas*, Grund. Math. Wiss. **166** (2nd ed.), Springer-Verlag, Berlin-New York, 1971.

[GrS] A. Grothendieck and J.-P. Serre, *Grothendieck-Serre Correspondence*, P. Colmez and J.-P. Serre ed., C. Maclean trans., Amer. Math. Soc. and Soc. Math. de France, Paris, 2004.

[H1] J. Harris, *Galois groups of enumerative problems*, Duke Math. J. **46** (1979), no. 4, 685724.

[H2] ———, *Algebraic Geometry: A First Course*, Grad. Texts in Math. **133**, Springer-Verlag, New York, 1995.

[Ha1] R. Hartshorne, *Algebraic Geometry*, Grad. Texts in Math. **52**, Springer-Verlag, New York-Heidelberg, 1977.

[Ha2] ———, *Ample Subvarieties of Algebraic Varieties*, Lect. Notes in Math. **156**, Springer-Verlag, Berlin-New York, 1970.

[Hat] A. Hatcher, *Algebraic Topology*, Cambridge U.P., Cambridge, 2002.

[HM] M. Hazewinkel and C. Martin, *A short elementary proof of Grothendieck's theorem on algebraic vector bundles over the projective line*, J. Pure Appl. Algebra **25** (1982), no. 2, 207–211.

[He] R. C. Heitmann, *Examples of noncatenary rings*, Trans. Amer. Math. Soc. **247** (1979), 125–136.

[Hil] D. Hilbert, *Ueber die Theorie der algebraischen Formen*, Math. Ann. **36** (1890), 473–534.

[Hir] H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero. I, II*, Ann. of Math. (2) 79 (1964), 109-203; ibid. (2) 79 (1964), 205-326.

[Ig] J. Igusa, *On some problems in abstract algebraic geometry*, Proc. Nat. Acad. Sci. **41** (1955), 964–967.

[Ii] S. Iitaka, *Algebraic Geometry: An Introduction to Birational Geometry of Algebraic Varieties*, Grad. Texts in Math. **76**, Springer-Verlag, New York-Berlin, 1982.

[Il] L. Illusie, *Frobenius and Hodge degeneration*, in *Introduction to Hodge Theory*, SMF/AMS Texts and Monographs **8**, 99–149, 2002.

[dJ] A. J. de Jong, *Smoothness, semi-stability and alterations*, IHES Publ. Math. **83** (1996), 51–93.

[KN] S. Kaji and M. Nakagawa, *The Chow rings of the algebraic groups of type* E_6, E_7, *and* E_8, arXiv:0709.3702v3, preprint 2007.

[Ka] I. Kaplansky, *Commutative Rings* (rev. ed.), U. of Chicago Press, Chicago, Ill.-London, 1974.

[KS1] M. Kashiwara and P. Schapira, *Sheaves on Manifolds*, Grund. Math. Wiss. **292**, Springer-Verlag, Berlin, 1994.

[KS2] ———, *Categories and sheaves*, Grund. Math. Wiss. **332**, Springer-Verlag, Berlin, 2006.

[KPV] K. Kedlaya, B. Poonen, and R. Vakil, *The William Lowell Putnam Mathematical Competition 1985–2000: Problems, Solutions, and Commentary*, Math. Ass. of Amer., Washington, DC, 2002.

[Kl1] S. Kleiman, *Toward a numerical theory of ampleness*, Ann. Math. (2) **84** (1966), no. 3, 293–344.

[Kl2] ———, *The transversality of a general translate*, Compositio Math. **28** (1974), no. 3, 287–297.

[Kl3] ———, *Cartier divisors versus invertible sheaves*, Comm. Alg. **28** (2000), no. 12, 5677–5678.

[Ko1] J. Kollár, *Rational Curves on Algebraic Varieties*, Ergeb. Math. Grenz. **32**, Springer-Verlag, Berlin, 1996.

[Ko2] ———, *Lectures on Resolutions of Singularities*, Annals of Math. Stud. **166**, Princeton U.P., Princeton, NJ, 2007.

[Lan] S. Lang, *Algebra* (rev. 3rd ed.), Grad. Texts in Math. **211**, Springer-Verlag, New York, 2002.

[Laz] R. Lazarsfeld, *Positivity in Algebraic Geometry I: Classical setting: line bundles and linear series*, Ergeb. Math. Grenz. **48**, Springer-Verlag, Berlin, 2004.

[Lef] S. Lefschetz, *L'analysis situs et la géométrie algébrique*, Gauthier-Villars, Paris, 1924.

[Leo] R. Leowontin, *The Triple Helix: Gene, Organism, and Environment*, Harvard U.P., Cambridge MA, 2000.

[Lit] J. E. Littlewood, *Littlewood's Miscellany*, Cambridge U.P., Cambridge, 1986.

[Liu] Q. Liu, *Algebraic Geometry and Arithmetic Curves*, R. Erné trans., Oxford Grad. Texts in Math. **6**, Oxford U.P., Oxford, 2002.

[Lü] W. Lütkebohmert, *On compactification of schemes*, Manuscripta Math. **80** (1), 95–111.

[Mac] S. Mac Lane, *Categories for the Working Mathematician*, 2nd ed., Grad. Texts in Math. **5**, Springer-Verlag, New York, 1998.

[Mat1] H. Matsumura, *Commutative Algebra*, 2nd ed., Math. Lecture Note Series **56**, Benjamin/Cummings Publ. Co., Inc., Reading, MA, 1980.

[Mat2] ———, *Commutative Ring Theory*, M. Reid trans., 2nd ed, Cambridge Stud. in Adv. Math. **8**, Cambridge U.P., Cambridge, 1989.

[MO7477] Answers by t3suji and G. Elencwajg to `http://mathoverflow.net/questions/7477/`

[MO33489] http://mathoverflow.net/questions/33489/
[MO68421] Answers by G. Elencwajg and M. Kim to http://mathoverflow.net/questions/68421/
[MO90551] Answer by M. Brandenburg to http://mathoverflow.net/questions/90551/
[MO93716] Answer by C. Rezk to http://mathoverflow.net/questions/93716/
[MO129242] Answer by L. Moret-Bailly to http://mathoverflow.net/questions/129242/
[Mc] C. McLarty, *The Rising Sea: Grothendieck on simplicity and generality*, in J. J. Gray and K. H. Parshall eds., *Episodes in the History of Modern Algebra (1800–1950)*, Amer. Math. Soc., Providence, RI, 2007.
[MS] E. Miller and B. Sturmfels, *Combinatorial Commutative Algebra*, Grad. Texts in Math. **227**, Springer-Verlag, New York, 2005.
[Mu1] D. Mumford, *Pathologies of modular algebraic surfaces*, Amer. J. Math. **83** (no. 2), Apr. 1961, 339–342.
[Mu2] ———, *Lectures on Curves on an Algebraic Surface*, Princeton U.P., Princeton NJ, 1966.
[Mu3] ———, *Abelian Varieties*, TIFR Studies in Math. 5, Oxford U.P., London, 1970.
[Mu4] ———, *Curves and Their Jacobians*, Univ. of Michigan Press, Ann Arbor, MI, 1975.
[Mu5] ———, *Hilbert's fourteenth problem — the finite generation of subrings such as rings of invariants*, in *Mathematical Developments Arising from Hilbert Problems, Part 2*, F. Browder ed., Amer. Math. Soc., Providence RI, 1976.
[Mu6] ———, *Algebraic Geometry I: Complex Projective Varieties*, Classics in Math., Springer-Verlag, Berlin, 1995.
[Mu7] ———, *The Red Book of Varieties and Schemes* (2nd expanded ed.), Lect. Notes in Math. **1358**, Springer-Verlag, Berlin, 1999. (Caution: in order to keep readers on their toes, a number of typos were introduced between the first and second editions.)
[Mur] D. Murfet, http://therisingsea.org/notes/ModulesOverAScheme.pdf
[OSS] C. Okonek, M. Schneider, and H. Spindler, *Vector Bundles on Complex Projective Spaces*, corrected reprint of the 1988 edition, Birkhäuser/Springer Basel AG, Basel, 2011.
[Po1] H. Poincaré, *Analysis situs*, J. de l'École Polytechnique (2) **1** (1895), 1–123.
[Po2] H. Poincaré, *Papers on topology: Analysis situs and its five supplements*, J. Stillwell trans., Amer. Math. Soc., Providence, RI, 2010.
[P1] B. Poonen, *Varieties without extra automorphisms, I: Curves*, Math. Res. Lett. **7** (2000), no. 1, 67–76.
[P2] ———, *Rational Points on Varieties*, http://www-math.mit.edu/~poonen/papers/Qpoints.pdf, preprint.
[Ra] M. Raynaud, *Contre-exemple au vanishing theorem en caractéristique* $p > 0$, in *C. P. Ramanujam — A Tribute*, 273–278, TIFR Studies in Math. **8**, Springer, Berlin-New York, 1978.
[RS] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, Inc., New York, 1980.
[Rei] M. Reid, *Undergraduate Algebraic Geometry*, London Math. Soc. Student Texts 12, Cambridge U.P., Cambridge, 1988.
[Rem] R. Remmert, *Local theory of complex spaces*, in *Several complex variables VII: Sheaf-theoretical methods in complex analysis*, Encl. of Math. Sci. **74**, Springer-Verlag, Berlin, 1994.
[Ro] F. Rohrer, goo.gl/OHSDC.
[RV] M. Roth and R. Vakil, *The affine stratification number and the moduli space of curves*, CRM Proceedings and Lecture Notes **38**, Univ. de Montréal, 2004, 213–227.
[Schm] H. L. Schmid, *Über die Automorphismen eines algebraischen Funktionenkörpers von Primzahlcharacteristik*, J. Reine Angew. Math. **179** (1938), 5–14.
[Schw] K. Schwede, *Gluing schemes and a scheme without closed points*, in *Recent Progress in Arithmetic and Algebraic Geometry*, 157–172, Contemp. Math. **386**, Amer. Math. Soc., Providence, RI, 2005.
[Se1] J.-P. Serre, *Faisceaux algébriques cohérents*, Ann. of Math. (2) **61** (1955), 197–278.
[Se2] ———, *Géométrie algébrique et géométrie analytique*, Ann. de l'institut Fourier **6** (1956), 1–42.
[Se3] ———, *Sur la topologie des variétés algébriques en caractéristique* p, in *Symposium Internacional de Topología Algebraica, International Symposium on Algebraic Topology*, 24–53, Univ. Nacional Autónoma de México and UNESCO, Mexico City, 1958.
[Se4] ———, *A Course in Arithmetic*, Grad. Texts in Math. **7**, Springer-Verlag, New York, 1973.
[SGA2] A. Grothendieck, *Séminaire de Géométrie Algébrique du Bois Marie 1962: Cohomologie locale des faisceaux cohérents et Théorèmes de Lefschetz locaux et globaux (SGA2)*, North-Holland Publ. Co., Amsterdam, 1968.

[SGA6] P. Berthelot, A. Grothendieck, and L. Illusie eds., *Séminaire de Géométrie Algébrique du Bois Marie 1966–67: Théorie des intersections et théorème de Riemann Roch (SGA 6)*, Lect. Notes in Math. **225**, Springer-Verlag, Berlin-New York, 1971.

[Sh] I. Shafarevich, *Basic Algebraic Geometry 1: Varieties in Projective Space*, 2nd ed., M. Reid trans., Springer-Verlag, Berlin, 1994.

[Si1] J. Silverman, *The Arithmetic of Elliptic Curves*, 2nd ed., Grad. Texts in Math. **206**, Springer-Verlag, New York-Berlin, 2009.

[Si2] ———, *Advanced Topics in the Arithmetic of Elliptic Curves*, Grad. Texts in Math. **151**, Springer-Verlag, New York-Berlin, 1994.

[Stacks] The Stacks Project Authors, *Stacks Project*, `http://math.columbia.edu/algebraic_geometry/stacks-git`. (To search for tags, go to `http://math.columbia.edu/algebraic_geometry/stacks-git/query.php`.)

[Stic] H. Stichtenoth, *Über die Automorphismengruppe eines algebraischen Funktionenkörpers von Primzahlcharakteristik. I. Eine Abschätzung der Ordnung der Automorphismengruppe*, Arch. Math. (Basel) **24** (1973), 527–544.

[T] A. Tarski, *A decision method for elementary algebra and geometry*, RAND Corp., Santa Monica, CA, 1948.

[Vak] R. Vakil, *Foundations of Algebraic Geometry*, preprint 2013.

[Val] P. Valéry, *Au sujet du Cimetière Marin*, in *Œuvres*, édition Jean Hytier, tome I, Gallimard, *La Pléiade*, Paris, 1957.

[vdW] B. L. van der Waerden, *Algebra* vol. II, J. R. Schulenberger trans., Springer-Verlag, New York, 1991.

[Vo] C. Voisin, *Hodge Theory and Complex Algebraic Geometry I* and *II*, L. Schneps trans., Cambridge Stud. in Adv. Math. **76** and **77**, Cambridge U.P., Cambridge, 2007.

[Wei] C. Weibel, *An Introduction to Homological Algebra*, Cambridge Stud. in Adv. Math. **38**, Cambridge U.P., Cambridge, 1994.

[Wey] H. Weyl, *Philosophy of Mathematics and Natural Science*, Princeton U.P., Princeton, NJ, 2009.

[Wh] H. Whitney, *Geometric Integration Theory*, Princeton U.P., Princeton, NJ, 1957.

[Wł] J. Włodarczyk, *Simple constructive weak factorization*, in *Algebraic Geometry — Seattle 2005*, 957–1004, Proc. Sympos. Pure Math. **80**, Part 2, Amer. Math. Soc., Providence, RI, 2009.

[Z] O. Zariski, *Foundations of a general theory of birational correspondences*, Trans. Amer. Math. Soc. **53** (1943), 490–542.

Index

VERY IMPORTANT: This is not yet a proper index (despite appearances). Once the content of these notes has settled down, I will begin compiling a proper index. I will also say that bold-faced numbers are often where the topic is defined, 11
(-1)-curve, **535**, 721
$(M : x)$, **634**, 637
0-ring, 17
A-scheme, 157
$A[[x]]$, 123
A^1, 100, 101
A^1_Q bold, 103
A^1_R bold, 102
A^2_k, 103
A^n, 104
A^n_X, 437
A_n curve singularity, **712**
A_n singularity, **712**
A_n singularity (curve, surface), **596**
$A_\bullet(I)$, **353**
Coh_X, **376**
D_n singularity, **712**
E_n singularity, **712**
Ext^i and sheaf Ext^i for O-modules, **734**
$GL_n(A)$, **196**
G_m blackboard bold, 194
I-adic completion, **707**
I-adically separated, **708**
I-filtration, **353**
I-filtration of a module, **353**
I-stable filtration, **353**
$I(S)$, 124
Id_n, **202**
$K(X)$, 154
L mathcal with bars, **404**
M-regular sequence, **237**
$M_\bullet(I)$, **353**
N mathfrak, **110**
$N^1(X)$, **472**
$O(a, b)$ oh, 419
PV, **149**
P^1_k, 139
P^n_A, 141
P^n_X, 440
P^n_k, 141
SL_n, **196**
$Supp(s)$, 80
$V(S)$, 112
$Aut(\cdot)$, 26
$\Gamma_\bullet(F)$ cal, 405
Mor, **26**
$\Omega^i_{X/Y}$, **569**
Ω_f, **558**
Pic X, 360
$\mathrm{Pic}^0 X$, **653**
Prin X, 391
Proj underline, 439
$\mathbb{Q}$-Cartier, **394**
$\mathbb{Q}$-line bundle, **472**
Spec Z bold, 181
Spec A, 99
Spec Z bold, 101, 136
Spec underline, 435
Supp F mathcal, 89
$\mathbb{Z}^{\geq 0}$-graded ring, **145**
$\boxtimes$, **419**, 425, 463
Ab_X, Ab_X^{pre}, 77
$Mod_{\mathscr{O}_X}$, $Mod^{pre}_{\mathscr{O}_X}$, 77
$Sets_X$, $Sets_X^{pre}$, 77
δ-functor, 611
δ, Δ, 274
δ-functor, **611**
$\ell(M)$, **470**
$\hookrightarrow$, **200**, **221**, **224**
$\kappa(X)$, **571**
$\mathcal{K}$, **570**
$\mathscr{L}(D)$, 393
$\mathscr{O}_X$, **76**
$\mathscr{O}_{X,p}$, **76**
μ_n, **196**
$\mathscr{O}$-connected, **692**
$\mathscr{O}(D)$, 389
$\mathscr{O}_v$, 340
$\omega_{X/k}$, **731**
$\oplus$, **47**

$\otimes$, **34**
π-ample, **445**
π-very ample, **440**
π^*
for locally free sheaves, **360**
π^{-1}, 87
π^{-1}, inverse image sheaf, 87
π_*, 75
$\mathbb{P}^n$-bundle, **440**
$\rho(X)$, **472**
$\sqrt{I}$, 113
$\times_A$, 241
$\widetilde{M}$, **129**
d-tuple embedding, 336
d-tuple embedding, 418, 419, 478, 479, 481, 495, 504
d-uple embedding, **229**
h^i, 453
k-smooth, **328**, **562**
$k[[x_1, \dots, x_n]]$, **707**
k^p, **259**
k^s, **259**
n-plane, **227**
p_a, 468
p_g, **570**
x-axis, **113**
étale, 655, **657**
étale topology, 709
étale morphism, 343
étale topology, 369, 656
"reasonably behaved" morphisms, 199
"the" vs. "a", 31

abelian category, 25, 48
abelian cone, 437
abelian semigroup, **45**
abelian variety, **291**
above, 250
absolute Frobenius, **213**
abut, **60**
acyclic, **613**
additive functor, **47**
additive category, **47**
adeles, 124
adjoint, **43**, 405
adjoint pair, **43**
adjoint functors, 43
adjugate matrix, 202
adjunction formula, 535, **570**
AF+BG Theorem, **672**
affine cone, 232
affine line, 101
Affine Communication Lemma, 156
affine cone, 143, **231**, 231, 232, 300, 308, 331, 591
affine line, 100
affine line with doubled origin, 139
affine morphism, 206
affine morphisms as Spec underline, 437
affine morphisms separated, 279
affine open, **135**
affine plane, 103
affine scheme, 99, 127
affine space, 104
affine topology/category, 365
affine variety, 157
affine-local, **156**
affine-local on the source, **200**
affine-local on the target, **200**
algebraic group, **290**
Algebraic Hartogs's Lemma, 137
Algebraic Hartogs's Lemma, 141, 159
algebraic space, 422
algebraically equivalent line bundles, **653**
alteration, 597
ample cone, **542**
analytification, **158**, **182**
André-Quillen homology, 552
ann M, **169**
anticanonical bundle, 687
arithmetic genus, 468
arrow, **26**
Artin-Rees Lemma, 339
Artin-Schreier cover, 579
Artinian, **318**, 319, 471
ascending chain condition, **122**
associated graded, 324
associated graded, 540
associated point, 167
associated point of a scheme, **376**
associated point of a coherent sheaf, **376**
associated prime, 169
assumptions on graded rings, 145
asymptotic Riemann-Roch, **532**
Atiyah flop, **600**
automorphism, 26
axiom of choice, 110
axiom of choice, 17, 110, 121, 607, 609, 611
axis, **113**

B'ezout's theorem, 227
B'ezout's Theorem, 230, 268, 334, 419, 479–481, 500, 501, 509, 510, 518, 529, 534, 686
Bézout's Theorem, 480
Bézout's theorem, 334
Bézout's theorem for plane curves, 479
backslash-mid, 411
base, **179**, 250
base locus, **403**
base scheme, **179**, 250, **404**, 404
base change, 250
base change diagram, 250
base locus, **404**, 404
base of a topology, **90**
base point, **403**
base ring, **145**

base-point, **404**, 404
base-point-free, **403**, **404**, 404
base-point-free locus, **403**
base-point-free with respect to π, **444**
Bertini's Theorem, 336
birational, **188**
birational (rational) map, **187**
blow up, **585**
blow-up, 253, **585**
blowing down, 721
blowing up, 586
boundary, **50**
branch divisor, **576**
branch locus, **574**
branch point, 502

c1(L) cup, **540**
Calabi-Yau, 313, **571**
Calabi-Yau varieties, 571
Cancellation Theorem for morphisms, 281
canonical bundle, 473, **570**
canonical curve, 508
canonical embedding, 508
canonical sheaf, **570**
Cartan-Eilenberg resolution, 615, **617**
Cartesian diagram, 250
Cartesian diagram/square, **36**
cat: Sch, 179
category, 25
category of open sets on X, 72
category of ringed spaces, **176**
catenary, **303**, 318
catenary ring, **303**
Cayley-Bacharach Theorem, 500
Cech cohomology fix, 458
Cech complex fix, 458
center of a blow-up, **585**
change of base, 250
Chevalley's theorem, 216
Chevalley's Theorem, 216, 255
Chinese Remainder Theorem, 142
class group, 362, 391
class group in number theory, 362
closed map, **287**
closed morphism, **219**
closed point, *see also* point, closed, 151
closed subscheme, 135, **221**
closed embedding, **221**
closed embedding affine-local, 221
closed immersion, **221**
closed morphism, **209**, 211
closed points, 105
closed points, existence of, 151
closed subscheme exact sequence, **374**
CM, **670**
cocycle condition, **92**
cocycle condition for transition functions, **358**
codim, **297**
codimension, 297
cofibered product, 242
cofinal, 707
Cohen Structure Theorem, 708
Cohen-Macaulay, 473, 480, **670**
coherent sheaf, 375, **376**
cohomological criterion for affineness, 463
cohomological criterion for affineness, 454
cohomology of a double complex, 58
cokernel, 48
colimit, 41
complete (k-scheme), **287**
complete intersection, **240**
complete linear series, **404**, 404
complete with respect to I, **708**
completion, 41, **707**
complex, **50**
composition series, 170
composition series, **471**
concrete category, **27**
cone, 231, 323, 324, 391, 394, 597
cone over smooth quadric surface, 396
cone over smooth quadric surface, 161, 324
cone over smooth quadric surface, 142, 162, 163
cone over smooth quadric surface, 266
cone over smooth quadric surface, 600
cone over smooth quadric surface, 301, 331, 394, 599
cone over the quadric surface, 161
conic, 226
connected, **116**, 151
connected component, **120**
connected component, 122, 260
connecting homomorphism, 454
conormal module, **554**
conormal sheaf, 555
constant (pre)sheaf, **74**, 83
constructible, 255
constructible set, 216
constructible subset of a Noetherian topological space, **215**
convergence of spectral sequence, **60**
coproduct, 37, **42**
coproduct of schemes, 422
coproduct of schemes, 241
corank, **328**
cotangent sheaf, 547, **558**
cotangent complex, 552
cotangent space, 322
cotangent vector, 547
cotangent vector = differential, 322
counit of adjunction, **44**
counterexamples, list of, 130
covariant, **28**
covers, **70**
Cremona transformation, 424
Cremona transformation, **191**

criterion for flatness, infinitesimal, **649**
crumpled plane, 269, 674
crumpled plane, **343**, 674
cubic, 226
cubic surface, 509
cubic surface [later inventory], 306
cubic surface, 509
curve, 296
cusp, 268, 333, 595, **712**
cycle, **50**

dévissage, 530
Dedekind domain, 269, **343**
Dedekind domain, 362
deformation theory, 326, 573, 628, 637
deformation to the normal cone, **602**
degenerate, **419**
degree of line bundle on curve, **467**
degree 1 point, *see also* point, degree 1
degree d map, 419
degree of a point, *see also* point, degree of
degree of a projective scheme, 479
degree of a projective scheme, 479
degree of a rational map, **301**
degree of coherent sheaf on curve, 469
degree of divisor on projective curve, **467**
degree of finite morphism, 381
degree of invertible sheaf on $\mathbb{P}^n_k$, 392
delta functor, 611
depth, **667**
derivation, 556
derived functor, 610
derived functor cohomology, 453
descending chain condition, **121**
descending chain condition, 319
descent, 249
desingularization, **584**, 595
det $\mathcal{F}$, **372**
det cF, **372**
determinant, 219
determinant (line) bundle, **372**
determinant bundle, **372**
determinant map $GL_n \to G_m$, **196**
diagonal, **274**
diagonal morphism δ, **274**
diagonalizing quadrics, 163
different ideal, **575**
differential = cotangent vector, 322
dilation, 585
dimension, 295
dimensional vanishing of cohomology, 454
direct limit, **41**
direct image sheaf, **75**
discrete topology, **74**, 81, 141, 210, 624
discrete topology, 74
discrete valuation, **340**
discrete valuation ring, **341**
discriminant ideal, **575**
discriminant hypersurface, 681
disjoint union (of schemes), 135
disjoint union of schemes, **135**
distinguished affine base, **365**
distinguished open set, 106, 115
divisor of zeros and poles, 388
domain of definition of rational map, **285**
domain of definition of a rational function, **169**
dominant, **187**
dominant rational map, 187
dominating, **187**
double complex, **57**
dual numbers, 110, 637
dual variety, **338**
dual coherent sheaf, **377**
dual numbers, **102**, 629, 631, 637
dual of a locally free sheaf, 360
dual of an $\mathscr{O}_X$-module, 78
dual projective space, **335**
dualizing sheaf, 455, **729**, **731**
DVR, **341**

effaceable, **612**
effective Cartier divisor, 236, 287
effective Cartier divisor, 745
effective Cartier divisor, **236**, 237, 597, 742
effective Cartier divisor, **236**, 253, 307, 385, 393, 394, 396, 397, 555, 584–588, 590–592, 598
effective Cartier divisor, relative, **650**
effective Weil divisor, 388
Eilenberg-Zilber Theorem, **463**
Eisenstein's criterion, 333, 578
elementary transformation, 724
elliptic curve, 361
elliptic curve, 511
embedded point, **167**
embedded points, 166
embedding, 655
enough projectives, **610**, 613, 698, 734
enough injectives, 610
enough projectives, 610
epi morphism, 48
epimorphism, **38**
equational criterion for flatness, **636**
equidimensional, 296
equivalence of categories, **30**
espace étalé, 83
espace étalé, 69, **75**, 88
essentially of finite presentation, **650**
essentially surjective, 31
etale, **344**
Euler test, 333
Euler characteristic, 387, 466
Euler exact sequence, 567
exact, **50**
exceptional divisor, 253, **585**

exceptional divisor, 586
excision exact sequence, 392
exponential exact sequence, 84, 86
Ext functors, **610**
extending the base field, **249**
extension by zero, **89**
extension of an ideal, **247**
exterior algebra, 372
exterior algebra, **372**

factorial, 161, 391, 393
faithful functor, **29**
faithful functor, 39
faithful module, **205**
faithfully flat, **643**
faithfully flat, **643**
Faltings' Theorem (Mordell's Conjecture), 505
Fano, 313, **571**
Fano variety of lines, 685
Fermat curve, 333
Fermat cubic surface, 681
fiber, **137**
fiber diagram, 250
fiber above a point, **251**
fiber of $\mathscr{O}$-module, **137**
fibered diagram/square, **36**
fibered product of schemes, 241
fibration, **344**
fibre, 137, 251
fibred product, 35
field-valued point, *see also* point, field-valued
final object, 31
finite implies projective, 444
finite presentation, 214
finite extension of rings, 210
finite flat morphism, 253, 638, 702, 741
finite module, 208
finite morphism is closed, 211
finite morphism is quasifinite, 211
finite morphisms are affine, 208
finite morphisms are projective, 444
finite morphisms separated, 279
finite presentation, 375
finite type, 212
finite type A-scheme, 157
finite type (quasicoherent) sheaf, **376**
finite union of closed subschemes, **223**
finitely generated, 375
finitely generated domain, **300**
finitely generated field extension, **189**
finitely generated graded ring (over A), **145**
finitely generated modules over discrete valuation rings, 379
finitely generated sheaf, **376**
finitely presented, **213**
finitely presented module, **53**
finitely presented algebra, 213, 549
flabby sheaf, 619
flag variety, **198**, 433
flasque, 624, 736
flasque sheaf, **619**
flasque sheaves are Γ-acyclic, 620
flat, 54, **416**, 450, 606, 627
flat limit, **640**
flat A-module, 629
flat (at p), **631**
flat morphism, 631
flat of relative dimension n, **645**
flat of relative dimension n, **645**
flat quasicoherent sheaf, **631**
flat quasicoherent sheaf over a base, 631
flat ring morphism, 629
flex, **515**
forgetful functor, 28
formal power series, **707**
formal neighborhood, **714**
formal scheme, 714
formally étale, 662, 709
formally smooth, 662
formally unramified, **574**
formally unramified, 662
fraction field $K(\cdot)$, **32**
fractional ideal, **362**
fractional linear transformations, 418
free sheaf, **357**, **358**
Freyd-Mitchell Embedding Theorem, 49
Frobenius morphism, 213, 561
Frobenius morphism, 213, 493, 749
Frobenius reciprocity, 44
full functor, **29**, 39
full rank, **163**
fully faithful functor, **29**
function field $K(\cdot)$, 154
function (on ringed space), **75**
function field, **154**, 169, 300
functions on a scheme, **100**, 136
functor, **28**
functor category, 39
functor of points, **29**
functor of points, **183**, 183
fundamental point, **285**

GAGA, 572, 627
Gauss's Lemma, 161, 162
Gaussian integers $mathbbZ[i]$, 342
Gaussian integers $mathbbZ[i]$, 334
general fiber, **254**
general point, **119**
general type, 313, **571**
general linear group over A, **196**
generated by global sections, **402**
generated in degree 1, **145**
generic fiber, 640
generic flatness, **646**
generic point, 151
generic fiber, **251**

generic fiber, **254**, 318
generic freeness, 646
generic point, *see also* point, generic
generically finite, **254**
generically finite, 575, 597
generically finite implies generally finite, 317
generically separable morphism, 576
generization, **119**, 151
genus, 565
geometric fiber, 259
geometric genus, **570**
geometric genus of a curve, **564**
geometric point, *see also* point, geometric
geometric fiber, **258**
geometric genus, 565
geometric Noether Normalization, 301
geometric point, 259
geometrically connected, 259
geometrically connected/irreducible/integral/reduced fibers, **258**
geometrically integral, 259
geometrically irreducible, 259
geometrically reduced, 259
germ, **72**
germ of function near a point, 135
globally generated, **402**
globally generated with respect to π, **444**
gluability axiom, **73**
gluing along closed subschemes, 422
gluing topological spaces along open subsets, 138
gluing two schemes together along a closed subscheme, 422
Going-Down Theorem for flat morphisms, **644**
Going-Down Theorem for integrally closed domains, **304**
Going-Down Theorem (for flat morphisms), **644**
Going-Up Theorem, **204**
graded ring, **145**
graded ring over A, **145**
graph morphism, gr of a m, **280**
graph of a morphism, **280**
graph of rational map, **285**
Grassmannian, 150, **197**, **198**, 569
Grothendieck topology, **369**, 656
Grothendieck functor, 193
Grothendieck group, **540**
Grothendieck's Theorem, **475**
Grothendieck, calls Serre a bastard, 729
Grothendieck-Riemann-Roch, 532
group schemes, 194
group scheme, **195**
group scheme action, **197**
group variety, **277**, **290**
groupoid, 26
Harris figure, 628
Hartogs's Lemma, 137, 141
Hartogs's Theorem, 361
Hausdorff, 273, 275
height, **297**
Hensel's Lemma, 709
higher direct image sheaf, 485, 486
higher pushforward sheaf, 486
Hilbert polynomial, **477**, 477
Hilbert scheme, 628, 685
Hilbert Basis Theorem, 122
Hilbert function, 477
Hilbert polynomial, 455
Hilbert scheme, **702**, 702
Hilbert Syzygy Theorem, 122
Hirzebruch surface, **440**
Hirzebruch-Riemann-Roch, 532
Hodge bundle, **695**
Hodge theory, 724
Hodge theory, 572
Hom, 47
Hom math symbol, **47**
homogeneous ideal, 145
homogeneous elements, **145**
homogeneous space, **664**
homology, **50**
homomorphism, **47**
homomorphism of local rings, 644
homotopic maps of complexes, 607
Hopf algebra, **197**
horseshoe construction, 617
horseshoe construction, 608
Hurwitz's automorphisms theorem, **580**
hypercohomology, 58
hyperplane, 226, **227**
hyperplane class, 390
hyperplane class in cohomology, 480
hypersurface, 144, 226, 298

I-depth, **667**
ideal denominators, 312
ideal of denominators, 160
ideal sheaf, 222
idempotent, 116
identity axiom, **73**
identity functor id, **28**
image presheaf, **79**
immersion, 655
implicit function theorem, 709
incidence correspondence, 306
incidence variety, 335
incidence correspondence, 335, 680–682
incidence variety, 306
index, **538**
index category, 39, 40
induced reduced subscheme structure, **235**
infinite-dimensional Noetherian ring, 299
infinitesimal neighborhood, **714**

infinitesimal criterion for flatness, **649**
initial object, 31
injective object, 619
injective sheaf, 624
injective limit, **41**
injective object, **610**, 621
injective object in an abelian category, 610
injective sheaves are flasque, 624
integral, 154, **201**
integral closure, 267
integral extension of rings, 201
integral morphism, 211
integral morphism of rings, 201
intersection number, **529**, **534**
intersection of closed subschemes, **223**
intersection of locally closed embeddings, **248**
inverse image, 88
inverse image ideal sheaf, 250
inverse image scheme, 250
inverse image sheaf, **87**
inverse limit, 40
invertible sheaf, **359**
invertible ideal sheaf, 236
invertible sheaf, 362
irreducible, **116**, 151
irreducible (Weil) divisor, 388
irreducible component, 120
irreducible components, 151
irregularity, **573**
irrelevant ideal, 145
isomorphism, 26
isomorphism of schemes, 136

Jacobian, 551
Jacobian criterion, 325
Jacobson radical, 204
joke, 39, 41, 569

K(X), **540**
K3 surface, 570, **571**, 571
K3 surfaces, 571
kernel, 48
Kodaira dimension, **571**
Kodaira vanishing, 455, **572**
Kodaira's criterion for ampleness, 425
Krull, 307
Krull dimension, 295
Krull dimension, 295
Krull Intersection Theorem, **354**
Krull's Height Theorem, **309**
Krull's Principal Ideal Theorem, 307
Kunneth formula for quasicoherent sheaves, **463**

Lüroth's theorem, 579
Lüroth's theorem, **579**
lci morphism, 239
Lefschetz principle, 572, 690
left-adjoint, **43**
left-exact, 52
left-exactness of global section functor, 86
Leibniz rule, 548
length, **470**, 471, 534, 647
length of a module, 647
length of a chain, **303**
length of dim 0 scheme, 534
Leray spectral sequence, 454
limit, 40
line, **227**
line sheaf, 359
line bundle, 358, 359
linear series, 404
linear space, **227**
linear series, **404**, 404
local on the source, **200**
local complete intersection morphism, 239
local constructibility, 255
local criterion for flatness, **647**
local isomorphism, 655
local morphism of local rings, **178**
local on the target, **199**
local on the source, **200**
local on the target, **199**
local ring, **17**
localization, **31**, 106
locally constructible, 255
locally ringed spaces, 178
locally finite type A-scheme, 157
locally (of) finite type, 212
locally closed immersion, **224**
locally closed embedding, **224**
locally free sheaf, **358**
locally free sheaf, **357**, 362
locally integral (temp.), 330
locally Noetherian scheme, 156
locally of finite presentation, 214
locally principal subscheme, **236**
locally principal closed subscheme, 307
locally principal Weil divisor, **391**
locally ringed space, **135**, 178
locus of indeterminacy, **285**
long exact sequence, 52
long exact sequence of higher pushforward sheaves, 486
lotus-eater, 12
lower semicontinuity, **313**
lower semicontinuity
rank of matrix, 317, 695, 697
Lutz-Nagell Theorem, 525
Lying Over Theorem, 301
Lying Over Theorem, 267
Lying Over Theorem, **202**, 203, 204, 211, 350

magic diagram, **37**
map of graded rings, **184**
mapping cone, **64**, 461
mathcal G-mid-sub-X, 411

Max Noether's $AF + BG$ Theorem, **672**
maximal rank, **163**
minimal model of a surface, 724
minimal prime, **125**, 125
minimal surface, **724**
modular curve, 513
module of Kähler differentials, 548
module of relative differentials, 548
moduli space, 505, 513
monic morphism, 48
monoidal transformation, 585
monomorphism, **37**
Mordell's conjecture, 505
Mordell-Weil, **512**
Mordell-Weil Theorem, 524
morphism, **26**
morphism of (pre)sheaves, **76**
morphism of local rings, **178**
morphism of δ-functors, **612**
morphism of (pre)sheaves, 76
morphism of ringed spaces, 176
morphism of ringed spaces, **176**
morphism of schemes, 179
multiplicative group
  see $\mathbb{G}_m$, **194**
multiplicity of a function, **598**

Néron-Severi group, **472**
Néron-Severi Theorem, **472**
Nagata, 299, 395
Nagata's Lemma, 395
Nagell-Lutz Theorem, 525
Nakayama's Lemma, 204
natural transformation, **30**
natural transformation of functors, 39
nef, **472**
nef cone, **473**
nilpotents, **110**, 153
nilradical, **110**, 110, 113
nodal cubic, 638
nodal cubic, 640, 642
node, 268, 333, **712**
Noether-Lefschetz theorem, 306
Noetherian induction, 121
Noetherian module, **123**
Noetherian ring, **122**, 122
Noetherian rings, important facts about, 122, 123
Noetherian scheme, 151, 156
Noetherian topological space, **121**
Noetherian topological space, 123
non-archimedean, **340**
non-archimedean analytic geometry, 369, 382
non-torsion point on elliptic curve, 524, 525
non-torsion $\mathbb{Q}$-point on elliptic curve, 525
non-zerodivisor, **32**
nondegenerate, **419**
nonprojective proper variety, 289, 420, 536
nonprojective proper variety, 654
nonregular, **327**
nonsingular, **327**
normal, 137, 160
normal cone, **592**
normal exact sequence, **563**
normal sheaf, 555
normalization, 267
Nullstellensatz, **104**, 158
number field, **269**
numerical equivalence, **472**
numerically effective, **472**
numerically trivial line bundle, **471**

object, **26**
octic, 226
Oka's Theorem, 382
Oka's Theorem, **376**
open condition
  etaleness, 345
open embedding of ringed spaces, **176**
open immersion of ringed spaces, **176**
open subscheme, **135**
open condition
  ampleness in proper families, 426
  Cohen-Macaulayness, 673
  factorial, 161
  flatness, 646, 647
  reducedness, 236
  reducedness, sometimes, 378
  reducedness, sometimes, 153
  regular, 507
  regular embedding, 240
  smoothness, 345
  unramifiedness, 575
  where $\mathscr{F}$ is globally generated, 403
open cover, **70**
open embedding, **200**
open immersion, **200**
open subscheme, **200**
opposite category, 29
order of zero/pole, **341**
ordinary double point, 712
ordinary m-fold point, **712**
ordinary multiple point, **712**
orientation of spectral sequence, **58**
over, 250

p-adic, **40**, 340
page of spectral sequence, **58**
Pappus's Theorem, **500**
partial flag variety, **198**
partial flag variety, 433
partially ordered set, **27**
partition of unity, 460
Pascal's "Mystical Hexagon" Theorem, **500**
perfect pairing, **373**
pi upper shreik, 739
pi upper shriek sub sh, **739**

$\mathrm{Pic}^{\tau} X$, **472**
Picard group, 360
plane, **227**
plurigenus, **571**
point
  closed, **118**
  degree 1, **158**
  degree of, **158**
  field-valued, **182**, **183**
  generic, **119**
  geometric, **258**
  ring-valued, **182**, **183**
  scheme-valued, **182**, **183**
pole, **341**
Poncelet's Porism, **522**
porism, 522
poset, **27**
presheaf, **71**
presheaf cokernel, **78**
presheaf image, **79**
presheaf kernel, **78**
presheaf quotient, **79**
prevariety (analytic), **158**
primary ideal, 172
prime avoidance, **304**, 308, 310, 632
prime avoidance, 305
prime avoidance (temp. notation), 308
principal divisor, 391
principal Weil divisor, **391**
product, 31, 241
product of quasicoherent ideal sheaves, **397**, 555
products of ideals, **113**
products of ideals, 361
Proj, **146**
projection formula, **533**
projection formula, **416**, **488**
projective cone, **232**
projective coordinates, 143, 144, **148**
projective distinguished open set, **147**
projective space, 141
projective A-scheme, 146
projective X-scheme, 442
projective and quasifinite implies finite, 457
projective bundle, 337, **440**
projective completion, **232**
projective cone, 232
projective coordinates, 144
projective line, 139
projective module, 609
projective modules are summands of free modules, 609
projective modules are flat, 609
projective morphism, 442
projective object in an abelian category, 609
projective space, **148**
projective variety, 158
projectivization of a locally free sheaf, **440**
projectivization of a vector space, **149**
proper, 287
proper transform, 584, **588**
proper transform, **586**
property R_k, **674**
property S_k, **674**
pseudo-morphisms, 187
Puisseux series, 340
pullback diagram, 250
pullback of $\mathscr{O}$-module, **413**
pullback of quasicoherent sheaf, **413**
pure dimension, 296
pure tensor, 432
purity of the branch locus, 576
pushforward sheaf, **75**
pushforward of coherent sheaves, 489
pushforward of quasicoherent sheaves, 409
pushforward sheaf, **75**

QCoh sub X, **369**
qcqs, 205, **206**
Qcqs Lemma, 206
quadratic transformation, 585
quadric, 226
quadric surface, 266
quadric surface, 142, 230, 301, 523
quadric cone, 231
quadric surface, 161–163, 230, 266, 324, 325, 331, 393, 394, 396, 432, 509, 599, 600
quartic, 226
quasicoherent sheaf, **357**
quasicoherent sheaf, 129, 363
quasicoherent sheaves
  morphisms of, **369**
quasicoherent sheaves: product, direct sum, $\wedge$, Sym, cokernel, image, $\otimes$, 373
quasicompact, 151
quasicompact morphism, **205**
quasicompact topological space, 117
quasifinite, 213
quasiinverse, 31
quasiisomorphic, 695
quasiisomorphism, 460
quasiprojective morphism, **602**
quasiprojective scheme, 149
quasiprojective is separated, 280
quasiprojective morphism, **442**
quasiseparated, 278
quasiseparated scheme, **152**
quasiseparated morphism, **205**
quintic, 226
quotient presheaf, **79**
quotient object, 48
quotient sheaf, 83

R sub k, **674**
radical, **113**
radical ideal, **113**
radicial morphism, **264**

ramification locus, **574**
ramification point, 502
ramification divisor, **576**
rank of locally free sheaf, 362
rank of quadratic, **163**
rank of coherent sheaf on curve, 469
rank of finite type quasicoherent sheaf, **380**
rational function, 102
rational map, **186**, **187**
rational function, 75, 100, 101, 169
rational normal curve, 418
rational normal curve, 229
rational normal curve take 1, 117
rational point, 182
rational section, **361**
reduced, 153, 154, 156
reduced ring, **110**
reduced scheme, **153**
reduced subscheme structure, **235**
reducedness is stalk-local, 153
reducible, **116**
reduction, 236
Rees algebra, **353**, 590, 726
reflexive sheaf, **377**, **555**
regular, 321
regular function := function, 169
regular scheme, **327**
regular sequence, **237**
regular embedding, **239**
regular function := function, **75**, **100**
regular implies Cohen-Macaulay, 671
regular local ring, **327**
regular point, 321
regular ring, **327**
regular sequence, 237
regularity, 321
relative i-forms, **569**
relative (co)tangent sheaf, **558**
relative (co)tangent vectors, 547
relative effective Cartier divisor, **650**
relative Proj, **439**
relatively ample with respect to π, **445**
relatively very ample, **424**
relatively ample, **445**
relatively base-point-free, **444**
relatively globally generated, **444**
representable functor, **193**
residue field, **136**
residue field $\kappa(\cdot)$, **135**
Residue Theorem, 733
residue field at a point, **136**
Residue theorem, 451, 467
resolution of singularities, **584**, 597
resolution of singularities, 595
restriction map, **72**
restriction of a quasicoherent sheaf, **411**
restriction of a quasicoherent sheaf, 411
restriction of sheaf to open set, 77
restriction of sheaf to open set $\mathcal{G}$, _U89
resultant, 219
Riemann surface, 450
Riemann-Hurwitz formula, 493
Riemann-Hurwitz formula, 502, 565
Riemann-Hurwitz formula, 503, 504, 506, 512, 515, 566, 570, 575, **578**, 578, 579
Riemann-Roch for coherent sheaves on a curve, 469
Riemann-Roch for surfaces, **535**
right-adjoint, **43**
right-exact, 34, 52
ring scheme, **195**
ring of integers in a number field, **269**
ring scheme, 196
ring-valued point, *see also* point, ringed-valued
ringed space, **75**, **97**
ruled surface, **440**
ruling, 230, 331, 432
ruling on quadric surface, 266
rulings of smooth quadric surface, **230**
rulings on quadric surface, 230
rulings on the quadric surface, 230

S sub k, **674**
saturated module, 407
saturation functor, 405
saturation map, **405**
schematic closure, **235**
schematically dense, **166**
scheme over A, 157
scheme, definition of, 134
scheme-theoretic base locus, **404**
scheme-theoretic inverse image, 250
scheme-theoretic support, **490**
scheme-theoretic closure, **235**
scheme-theoretic pullback, 250
scheme-valued point, *see also* point, scheme-valued
Schubert cell, 197
sections over an open set, 71
Segre embedding, **265**, 419
Segre product, **265**
Segre variety, **265**
semiample, **428**
semicontinuity, **313**, 690
- degree of finite morphism, 317, 381, 495
- dimension of tangent space at closed points of variety, 318, 557
- fiber dimension, 316
- fiber dimension, 317, 458, 495, 682
- rank of cohomology groups of coherent sheaves, 318, 640, 643, 651, 690, 691, 695, 697, 702
- rank of matrix, 317, 695, 697
- rank of sheaf, 317, 380, 381, 562, 663, 690

separable morphism, 576

separably generated, **558**, **563**
separated, 139, 152, **274**
separated over A, **274**
separated presheaf, **73**
separatedness, 135
septic, 226
Serre duality, 455
Serre vanishing, 455, 489
Serre's Theorem A, 426, 743
Serre's Theorem A, **404**
sextic, 226
sheaf, 71
sheaf of relative i-forms, **569**
sheaf Hom (Hom underline), 77
sheaf Hom (Hom underline) of quasicoherent sheaves, 377
sheaf Hom (underline), 77
sheaf determined by sheaf on base, 366
sheaf of ideals, 222
sheaf of relative differentials, 547
sheaf on a base, 90
sheaf on a base, 91
sheaf on a base determines sheaf, 91
sheaf on affine base, 365
sheafification, 79, 81
shriek, 89
signature, **538**
simple module, **471**
singular, 321, **327**
singularity, **327**
site, **369**
skyscraper sheaf, **74**
small resolution, **599**
smooth, 321, 655
smooth morphism, 343, **344**
smooth over a field, **328**
smooth (morphism), **657**
smooth cubic plane curves exist, 333
smooth over a field, **562**
smooth quadric surface, 162, 163
space of sections, **75**
space of sections, 83
specialization, **119**, 151
spectral sequence, 57
spectrum, 99
stack, 92, 369
stalk, **72**
stalk-local, **153**, 156
Stein factorization, **716**
strict transform, **586**
structure morphism, 180
structure sheaf, 97
structure sheaf (of ringed space), **75**
structure sheaf on Spec A, 127
submersion, 25
subobject, 48
subpresheaf, **79**
subscheme cut out by a section of a locally free sheaf, **360**
subsheaf, 83
subvariety, **277**
support, **378**
support of a section, **80**
support of a sheaf, 89
support of a Weil divisor, 388
surface, 296
surjective morphism, **203**, **256**
symbolic power of an ideal, **319**
symmetric algebra, **372**, 372
system of parameters, 314
system of parameters, **310**

tacnode, 268, 333, **712**
tame ramification, **578**
tangent plane, 334
tangent cone, **592**
tangent line, 515
tangent sheaf, **558**
tangent space, 322
tangent vector, 322
tautological bundle, 437
tensor algebra, **372**
tensor algebra $T^{\bullet}_{A}(M)$, 372
tensor product, 33, **34**
tensor product of O-modules, **87**
tensor product of sheaves, 87
Theorem A, **404**, 426, 743
Theorem of the Base, **472**
topological module, **708**
topological ring, **708**
topos, **369**
torsion (sub)module, **374**
torsion quasicoherent sheaf on reduced X, **374**
torsion sheaf, 379
torsion-free, **374**
torsion-free sheaf, 379
torsor, 248
total fraction ring, **169**, 169
total quotient ring, **169**
total space of locally free sheaf, **437**
total transform, **586**, **588**
trace map, **731**
trace morphism, **729**
transcendence basis/degree, **300**
transition functions, **358**
transitive G-action, **664**
trigonal curve, **507**
triple point, **712**
Tsen's Theorem, 694
twisted cubic, 227
twisted cubic, 300
twisted cubic curve, 117
twisting by a line bundle, **401**
two planes meeting at a point, 130, 482, 671, 674, 734

two planes meeting at a point, 474
two planes, 671
two planes meeting at a point, 639
two planes meeting at a point, 639, 650, 669–671, 673, 674

ultrafilter, 118
underline S, 74
underline S^{pre}, 74
uniformizer, 339
uniruled, **313**
unit of adjunction, **44**
universal δ-functor, 611
universal property, 24
universal δ-functor, **612**
universal property of blowing up, 586
universally, **287**
universally closed, **287**
universally injective, **264**, 283
universally open, **260**
unramified, **574**, 655
upper semicontinuity, **313**
upper semicontinuity: dimension of tangent space, 557
upper semicontinuity
  degree of finite type morphism, 317, 381, 495
  dimension of tangent space at closed points of variety, 318, 557
  fiber dimension, 316
  fiber dimension, 317, 458, 495, 682
  rank of cohomology groups of coherent sheaves, 318, 640, 643, 651, 690, 691, 695, 697, 702
  rank of sheaf, 317, 380, 381, 562, 663, 690
upper semicontinuity of rank of finite type sheaf, 380
upper semicontinuous functions in algebraic geometry, 317

valuation, **340**
valuation ring, 340
valuative criterion for flatness, 639
valuative criterion for separatedness, 347
value of a function, 100
value of a quasicoherent sheaf at a point, 380
value of function, **136**
value of function at a point, 135
vanishing set, **112**
vanishing theorems, 466
vanishing scheme, **223**
varieties (classically), 119
variety, 273, **277**
vector bundle, 437
vector bundle associated to locally free sheaf, **437**
vector bundle, 437
Veronese, 418
Veronese embedding, 227, 419
Veronese subring, **185**
Veronese embedding, **229**, 336, 418, 478, 479, 481, 495, 504
Veronese surface, **229**
vertical (co)tangent vectors, 547
very ample, **424**
very ample with respect to π, **440**

Weierstrass normal form, 516
weighted projective space, **231**
Weil divisor, 388
wild ramification, **578**

Yoneda cup product, 737
Yoneda embedding, 39
Yoneda's Lemma, 38
Yoneda's lemma, **245**
Yoneda's lemma, **39**

Zariski topology, **134**
Zariski (co)tangent space, 322
Zariski sheaf, **245**
Zariski site, 245
Zariski tangent space, 321
Zariski topology, 112, **113**
Zariski topology on $\overline{k}^n$, **218**
Zariski's Main Theorem, **716**
Zariski's Main Theorem, 290
zero ring, 17
zero object, 31, 47
zerodivisor, **32**