Introduction to Algebraic Geometry
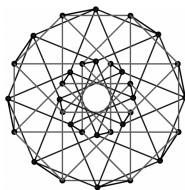
Introduction to Algebraic Geometry

by

Justin R. Smith

Second Edition

This is dedicated to the memory of my wonderful wife, Brigitte.

Also by Justin R. Smith

- *Eye of a Fly* (Kindle edition).
- *The God Virus* (Kindle and paperback)
- *Ohana* (Kindle and paperback)
- *The Accidental Empress* (Kindle and paperback)
- Constance Fairchild Adventures (published by Silver Leaf Books):
    - *The Mills of God*, hardcover.
    - *The Well of Souls*, (Kindle edition and paperback).

Foreword

> "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."
> —David Mumford in [**122**].

This book is intended for self-study or as a textbook for graduate students or advanced undergraduates. It presupposes some basic knowledge of point-set topology and a solid foundation in linear algebra. Otherwise, it develops all of the commutative algebra, sheaf-theory and cohomology needed to understand the material. This is the kind of background students might have at a school that emphasizes applied mathematics, or one where enrollment is not sufficient to run separate courses in commutative algebra.

The first chapter is an introduction to the algebraic approach to solving a classic geometric problem. It develops concepts that are useful and interesting on their own, like the Sylvester matrix and resultants of polynomials. It concludes with a discussion of how problems in robots and computer vision can be framed in algebraic terms.

Chapter 2 on page 35 develops classical affine algebraic geometry, providing a foundation for scheme theory and projective geometry. it also develops the theory of Gröbner bases and applications of them to the robotics problems from the first chapter.

Chapter 3 on page 117 studies the local properties of affine varieties — material that is relevant for projective varieties as well.

Chapter 4 on page 161 is an introduction to the language of schemes and general varieties. It attempts motivate these concepts by showing that certain natural operations on varieties can lead to objects that are schemes but not varieties.

Chapter 5 on page 219 covers projective varieties, using material from chapter 3 on open affines. In the section on Grassmanians, it has a complete treatment of interior products.

In the section on intersection theory, it revisits the classical problem introduced in chapter 1 and provides a modern treatment.

In chapter 6, the book culminates with two proofs of the Riemann-Roch theorem. The first is classical (Brill-Noether) and reasonably straightforward — introducing some elegant geometric concepts and results. The second proof is the modern one using the heavy machinery of sheaf cohomology and Serre Duality. Both are included because they give an instructor flexibility in approaching this subject. In particular, the sheaf cohomology of the second proof gives students a good idea of how the subject is done today.

Appendix A on page 341 develops almost all of the commutative algebra needed to understand the rest of the book (specialized material is provided as needed): students are only required to have an understanding of linear algebra and the concept of a group. Students with some commutative algebra can skip it and refer back to it as needed (page-references are used throughout the book to facilitate this). It ends with a brief treatment of category theory.

Appendix B on page 495 is an introduction to sheaves, in preparation for structure sheaves of schemes and general varieties. It also develops the theory of vector-bundles over an affine variety.

Appendix C on page 507 develops the topological concept of vector bundles.

Appendix D on page 519 develops basic concepts of homological algebra and applies this to sheaves and rings. It culminates with a proof of the Serre Duality theorem.

Sections marked with a "dangerous bend" symbol are more advanced and may be skipped on a first reading.

Answers to roughly half of the exercises are found at the end of the book.

Chapters 1 and 2 (with a sidelong glance at appendix A) may be suitable for a semester of an undergraduate course. Appendix A has been used as the text for the second semester of an abstract algebra course.

Chapters 3 and 4 (or even chapter 5, skipping chapter 4) could make up the text for a second semester.

I am grateful to Patrick Clarke and Thomas Yu for many helpful and interesting discussions. Their insights and comments have improved this book considerably. I am also grateful to people at mathoverflow.net for their comments. The list includes (but is not limited to): Matthew Emerton, Will Swain, Nick Ramsey, and Angelo Vistoli.

I am indebted to Noel Robinson for pointing out a gap in the proof of corollary 2.8.30 on page 111. Correcting it entailed adding the material on *catenary rings* in section 2.8.2 on page 100. His questions also resulted in simpler and (hopefully) clearer proof of theorem 2.5.27 on page 86.

I am grateful to Darij Grinberg for his extremely careful reading of the manuscript. He identified several significant errors.

I am also grateful to Matthias Ettrich and the many other developers of the software, LyX — a free front end to LaTeX that has the ease of use of a word processor, with spell-checking, an excellent equation editor, and a thesaurus. I have used this software for years and the current version is more polished and bug-free than most commercial software.

LyX is available from http://www.lyx.org.

Contents

List of Figures

Introduction to Algebraic Geometry

CHAPTER 1

A classical result

Awake! for Morning in the Bowl of Night
Has flung the Stone that puts the Stars to Flight:
And Lo! the Hunter of the East has caught
The Sultan's Turret in a Noose of Light.

—*The Rubaiyat* of Omar Khayyam, verse I

Algebraic geometry is a branch of mathematics that combines techniques of abstract algebra with the language and the problems of geometry.

It has a long history, going back more than a thousand years. One early (circa 1000 A.D.) notable achievement was Omar Khayyam's[1] proof that the roots of a cubic equation could be found via the intersection of a parabola and a circle (see [**89**]).

It occupies a central place in modern mathematics and has multiple connections with fields like complex analysis, topology and number theory. Today, algebraic geometry is applied to a diverse array of fields including theoretical physics, control theory, cryptography (see section 6.2.2 on page 322), and algebraic coding theory — see [**32**]. Section 1.6 on page 28 describes an application to robotics.

1.1. Bézout's Theorem

We begin with a classical result that illustrates how algebraic geometry approaches geometric questions. It was stated (without proof) by Isaac Newton in his proof of lemma 28 of volume 1 of his *Principia Mathematica* and was discussed by MacLaurin (see [**107**]) and Euler (see [**45**]) before Bézout's published proof in [**17**].

> Étienne Bézout (1730–1783) was a French algebraist and geometer credited with the invention of the determinant (in [**15**]).

Let's examine graphs of polynomials and the points where they intersect. Two linear curves — straight lines — intersect in a single point. A line and a quadratic curve intersect in two points and figure 1.1.1 on page 2 shows the intersections between

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

and

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

at 4 points.

---

[1]*The* Omar Khayyam who wrote the famous poem, *The Rubaiyat* — see [**48**].
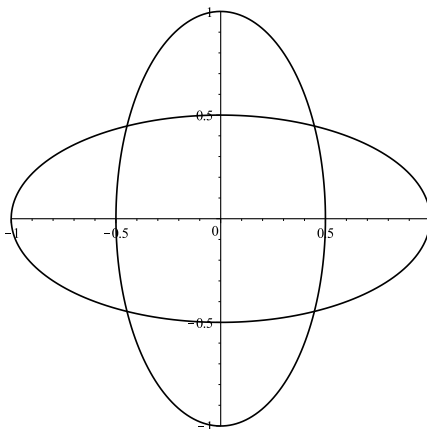
FIGURE 1.1.1.   A quadratic intersecting a quadratic

Figure 1.1.2 shows the intersection of the quadratic curve

$$\frac{3}{5}\,(x-2)^2 + 25\,\left(y - \frac{1}{10}\right)^2 = 1$$

with the cubic curve

$$y = x^3 - 6\,x^2 + 11\,x - 6$$

at 6 points.
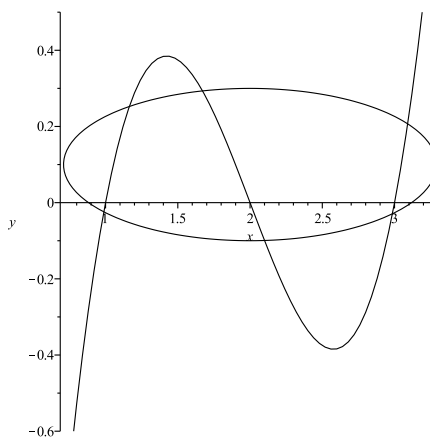
FIGURE 1.1.2.   Intersection of a quadratic and cubic

Leonhard Euler (1707 – 1783) was, perhaps, the greatest mathematician all time. Although he was born in Switzerland, he spent most of his life in St. Petersburg, Russia and Berlin, Germany. He originated the notation $f(x)$ for a function and made contributions to mechanics, fluid dynamics, optics, astronomy, and music theory. His final work, "Treatise on the Construction and Steering of Ships," is a classic whose ideas on shipbuilding are still used to this day.

To do justice to Euler's life would require a book considerably longer than the current one — see the article [**53**]. His collected works fill more than 70 volumes and, after his death, he left enough manuscripts behind to provide publications to the Journal of the Imperial Academy of Sciences (of Russia) for 47 years.

We arrive at the conjecture (that Newton, MacLaurin and Euler regarded as self-evident):

CONJECTURE 1.1.1. *If $f(x, y)$ is a polynomial of degree n and $g(x, y)$ is a polynomial of degree m, the curves defined by $f = 0$ and $g = 0$ intersect in $m \cdot n$ points.*

As soon as we state this conjecture, some problems immediately become clear:

(1) The set of points that satisfy $x^2 - y^2 = 0$ consists of the union of the line $y = x$ and the line $y = -x$. So, if $f(x, y) = x^2 - y^2$ and $g(x, y) = x - y$ then the intersection of the curves defined by them has an infinite number of points, namely the *entire line $y = x$.*

(2) the parabola $y = x^2 + 1$ does not intersect the line $y = 0$ at all.

(3) two parallel lines $y = 2x + 1$ and $y = 2x + 3$ do not intersect.

The first problem can be solved by requiring f and g to have *no common factor.*

In light of the second and third problem, one might be tempted to limit our conjecture to giving an *upper bound* to the number of intersections. On the other hand, the second problem can be easily solved by passing to the complex numbers — where we get two intersections $(\pm i, 0)$.

We can deal with the third problem too, but we need to address another important geometric development.

1.2. The projective plane

Projective geometry extends euclidean geometry by adding "points at infinity" where parallel lines intersect. It arose in an effort to understand the mathematical aspects of perspective — in effect, it is the geometry of the *visual* world. It was also one of the first non-Euclidean geometries to be studied.

The early development of projective geometry in the 1600's was along the lines of Euclidean geometry, with axioms and theorems that analyzed intersections of lines and triangles — see [**71**].

The modern approach is quantitative, and projective geometry is used heavily in computer graphics today — see [**76**]. A high-end graphics card that allows one to play computer games is actually performing millions of three-dimensional projective transformations per second.

When one looks at a scene (without binocular vision!) one sees along *lines of sight.* All points along a line from the eye to the horizon are equivalent in that they contribute one point to the two-dimensional image one sees.

DEFINITION 1.2.1. The *real projective plane* is the space of equivalence classes

$$\mathbb{R}P^2 = \{\mathbb{R}^3 \setminus (0,0,0)\} / \sim$$

where $\sim$ is the equivalence relation

(1.2.1) $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$

if there exists a $t \neq 0$ such that $x_2 = t \cdot x_1$, $y_2 = t \cdot y_1$, $z_2 = t \cdot z_1$. We denote points of $\mathbb{R}P^2$ by so-called *homogeneous coordinates*: $[x : y : z]$, where the colons indicate that the *ratios* between the coordinates are the only things that are significant.

REMARK. Homogeneous coordinates first appeared in the 1827 monograph, *Der barycentrische Calcül*, by Möbius (see [**154**]).

August Ferdinand Möbius (1790–1868) was a German mathematician and astronomer popularly known for his discovery of the Möbius strip (although he made many other contributions to mathematics, including Möbius transformations, the Möbius function in combinatorics and the Möbius inversion formula).

In more generality we have

DEFINITION 1.2.2. The projective spaces

$$\mathbb{R}P^n = \{\mathbb{R}^{n+1} \setminus \text{Origin}\} / \sim$$

and

$$\mathbb{C}P^n = \{\mathbb{C}^{n+1} \setminus \text{Origin}\} / \sim$$

are sets of equivalence classes via an $n + 1$-dimensional version of equation 1.2.1 where t is in $\mathbb{R}$ and $\mathbb{C}$, respectively.

It is not hard to see that these projective spaces can be broken down into unions of ordinary Euclidean space and other projective spaces:

PROPOSITION 1.2.3. *We have inclusions*

$$(x_1, \ldots, x_n) \quad \mapsto \quad [x_1 : \cdots : x_n : 1]$$
$$\mathbb{R}^n \quad \hookrightarrow \quad \mathbb{R}P^n$$
$$\mathbb{C}^n \quad \hookrightarrow \quad \mathbb{C}P^n$$

so a point $[x_1 : \cdots : x_{n+1}] \in \mathbb{C}P^n$ with $x_{n+1} \neq 0$ corresponds to the point

$$\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}} \right) \in \mathbb{C}^n$$

We also have inclusions

$$[x_1 : \cdots : x_n] \quad \mapsto \quad [x_1 : \cdots : x_n : 0]$$
$$\mathbb{R}P^{n-1} \quad \hookrightarrow \quad \mathbb{R}P^n$$
$$\mathbb{C}P^{n-1} \quad \hookrightarrow \quad \mathbb{C}P^n$$

and

$$\mathbb{R}P^n \quad = \quad \mathbb{R}^n \cup \mathbb{R}P^{n-1}$$
(1.2.2) $$\mathbb{C}P^n \quad = \quad \mathbb{C}^n \cup \mathbb{C}P^{n-1}$$

where the embedded copies of $\mathbb{R}P^{n-1}$ and $\mathbb{C}P^{n-1}$ are called the "spaces at infinity".

Every point, $[x_1 : \cdots : x_{n+1}]$, of $\mathbb{C}P^n$ is either in the image of $\mathbb{C}^n$ (if $x_{n+1} \neq 0$) or in the image of $\mathbb{C}P^{n-1}$ (if $x_{n+1} = 0$).

REMARK. The reason for the term "space at infinity" is that the point

$$[x_1 : \cdots : x_{n+1}] \in \mathbb{C}P^n$$

with $x_{n+1} \neq 0$ corresponds to the point

$$\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}} \right) \in \mathbb{C}^n$$

and this point moves out to infinity as we let $x_{n+1} \to 0$.

PROOF. If $[x_1 : \cdots : x_{n+1}] \in \mathbb{R}P^n$ is a point, then there are two *mutually exclusive* possibilities: $x_{n+1} = 0$ or $x_{n+1} \neq 0$.

Among points for which $x_{n+1} = 0$, we have $[x_1 : \cdots : x_n : 0] \sim [x_1' : \cdots : x_n' : 0] \in \mathbb{R}P^n$ if and only if $[x_1 : \cdots : x_n] \sim [x_1' : \cdots : x_n'] \in \mathbb{R}P^{n-1}$. It follows that such points are in the image of $\mathbb{R}P^{n-1} \hookrightarrow \mathbb{R}P^n$, as defined above.

Among points for which $x_{n+1} \neq 0$, we have

$$[x_1 : \cdots : x_n : x_{n+1}] \sim \left[\frac{x_1}{x_{n+1}} : \cdots : \frac{x_n}{x_{n+1}} : 1 \right] \in \mathbb{R}P^n$$

so that such points are in the image of the map $\mathbb{R}^n \hookrightarrow \mathbb{R}P^n$, as defined above.

Since *every* point of $\mathbb{R}P^n$ is in the image of $\mathbb{R}P^{n-1} \hookrightarrow \mathbb{R}P^n$ or $\mathbb{R}^n \hookrightarrow \mathbb{R}P^n$, it follows that $\mathbb{R}P^n = \mathbb{R}^n \cup \mathbb{R}P^{n-1}$. The proof for $\mathbb{C}P^n$ is similar. $\square$

Returning to Bézout's theorem, we need to understand algebraic *curves* in projective spaces.

DEFINITION 1.2.4. A polynomial $F(x_1, \ldots, x_n)$ is *homogeneous of degree k* if

$$F(tx_1, \ldots, tx_n) = t^k F(x_1, \ldots, x_n)$$

for all $t \in \mathbb{C}$.

REMARK 1.2.5. Note that the set of solutions of an equation

$$F(x_1, \ldots, x_{n+1}) = 0$$

where F is homogeneous of any degree is naturally well-defined over $\mathbb{C}P^n$ since any multiple of a solution is also a solution.

PROPOSITION 1.2.6. *There is a 1-1 correspondence between degree k polynomials on $\mathbb{C}^n$ and homogeneous polynomials of degree k on $\mathbb{C}^{n+1}$ that are not divisible by x_{n+1}. It sends the polynomial $p(x_1, \ldots, x_n)$ over $\mathbb{C}^n$ to the homogeneous polynomial*

$$\bar{p}(x_1, \ldots, x_{n+1}) = x_{n+1}^k p\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}} \right)$$

over $\mathbb{C}^{n+1}$, and sends the homogeneous polynomial $\bar{p}(x_1, \ldots, x_{n+1})$ to $\bar{p}(x_1, \ldots, x_n, 1)$.

PROOF. It is clear that $\bar{p}(x_1, \ldots, x_{n+1})$ is homogeneous of degree k:

$$
\begin{aligned}
\bar{p}(tx_1, \ldots, tx_{n+1}) &= t^k x_{n+1}^k p\left(\frac{tx_1}{tx_{n+1}}, \ldots, \frac{tx_n}{tx_{n+1}}\right) \\
&= t^k x_{n+1}^k p\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}\right) \\
&= t^k \bar{p}(x_1, \ldots, x_{n+1})
\end{aligned}
$$

Furthermore, since $p(x_1, \ldots, x_n)$ is of degree k, there is at least one monomial of total degree k

$$
x_1^{\alpha_1} \cdots x_n^{\alpha_n}
$$

with $\sum_{i=1}^{n} \alpha_i = k$, which will give rise to

$$
x_{n+1}^k \cdot \left(\left(\frac{x_1}{x_{n+1}}\right)^{\alpha_1} \cdots \left(\frac{x_n}{x_{n+1}}\right)^{\alpha_n}\right) = x_1^{\alpha_1} \cdots x_n^{\alpha_n}
$$

so $\bar{p}(x_1, \ldots, x_{n+1})$ will have at least one monomial that does not contain a factor of x_{n+1} and $\bar{p}(x_1, \ldots, x_{n+1})$ will not be a multiple of x_{n+1}.

It is easy to see that if $p(x_1, \ldots, x_n)$ is a polynomial of degree k, converting it to $\bar{p}(x_1, \ldots, x_{n+1})$ and then setting the final variable to 1 will regenerate it:

$$
\bar{p}(x_1, \ldots, x_n, 1) = p(x_1, \ldots, x_n)
$$

Conversely, if we start with a homogeneous polynomial, $g(x_1, \ldots, x_{n+1})$ of degree k that is not divisible by x_{n+1}, then the defining property of a homogeneous polynomial implies that

$$
g(x_1, \ldots, x_{n+1}) = x_{n+1}^k g\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}, 1\right)
$$

We claim that $g(x_1, \ldots, x_n, 1)$ is a polynomial of degree k. This follows from the fact that the original g was not divisible by X_{n+1} so there exists a monomial in g of degree k that does not contain X_{n+1}. $\qquad\square$

An equation for a line

$$
x_2 = ax_1 + b
$$

in $\mathbb{C}^2$ gives rise to an equation involving homogeneous coordinates, $[x_1 : x_2 : x_3]$, in $\mathbb{C}P^2$

$$
\frac{x_2}{x_3} = a\frac{x_2}{x_3} + b
$$

or

(1.2.3) $$x_2 = ax_1 + bx_3$$

It follows that lines in $\mathbb{C}^2$ extend uniquely to lines in $\mathbb{C}P^2$.

It is encouraging that:

PROPOSITION 1.2.7. *Two distinct lines*

$$
\begin{aligned}
x_2 &= a_1 x_1 + b_1 x_3 \\
x_2 &= a_2 x_1 + b_2 x_3
\end{aligned}
$$

in $\mathbb{C}P^2$ intersect in one point. If they are parallel with $a_1 = a_2 = a$ then they intersect at the point $[x_1 : ax_1 : 0]$ at infinity.

REMARK. This matches what we experience in looking at a landscape: parallel lines always meet at the horizon, and their common slope determines *where* they meet.

PROOF. If the lines are not parallel (i.e., $a_1 \neq a_2$) then they intersect in $\mathbb{C}^2 \subset \mathbb{C}P^2$ in the usual way. If $a_1 = a_2 = a$, we have $0 = (b_1 - b_2)x_3$ so $x_3 = 0$. We get a point of intersection $[x_1 : ax_1 : 0] \in \mathbb{C}P^1 \subset \mathbb{C}P^2$ — "at infinity". $\square$

To find the solutions of
$$F(x_1, x_2, x_3) = 0$$
in $\mathbb{C}P^2$, proposition 1.2.3 on page 4 implies that we must consider two cases:
   (1) $[x_1 : x_2 : x_3] \in \mathbb{C}^2 \subset \mathbb{C}P^2$. We set $x_3 = 1$ and solve $F(x_1, x_2, 1) = 0$.
   (2) $[x_1 : x_2 : x_3] \in \mathbb{C}P^1 \subset \mathbb{C}P^2$. In this case, we set $x_3 = 0$ and solve $F(x_1, x_2, 0) = 0$. This is a homogeneous polynomial in x_1 and x_2 and solving it involves two cases again:
      (a) $F(x_1, 1, 0) = 0$
      (b) $F(x_1, 0, 0) = 0$. In this case, we only care whether a *nonzero* value of x works because $[x : 0 : 0] = [1 : 0 : 0]$ defines a *single point* of $\mathbb{C}P^2$.

Here's an example:

If we start with a "circle," C, in $\mathbb{C}^2$ defined by
$$x_1^2 + y_2^2 - 1 = 0$$
we can extend it to a curve $\bar{C}$ in $\mathbb{C}P^2$ by writing
$$x_3^2 \left(\left(\frac{x_1}{x_3} \right)^2 + \left(\frac{x_2}{x_3} \right)^2 - 1 \right) = 0$$
or
$$x_1^2 + x_2^2 = x_3^2$$
The first case gives us back the original equation $C \subset \mathbb{C}^2$. The second case ($x_3 = 0$) gives us

(1.2.4)
$$x_1^2 + x_2^2 = 0$$

The first sub-case, 2a, gives us
$$x_1^2 + 1 = 0$$
so we get $x_1 = \pm i$ and $x_2 = 1$. The second sub-case is
$$x_1^2 + 0 = 0$$
which is unacceptable since at least *one* of the three variables x_1, x_2, x_3 must be nonzero.

So the curve $\bar{C} \subset \mathbb{C}P^2$ includes C and the two points at infinity $[\pm i : 1 : 0]$.

DEFINITION 1.2.8. A *circle in* $\mathbb{C}P^2$ is defined to be the extension to $\mathbb{C}P^2$ of a curve in $\mathbb{C}^2$ defined by a equation
$$(x_1 - a)^2 + (x_2 - b)^2 = r^2$$
for $a, b, r \in \mathbb{C}$.

It is interesting that:
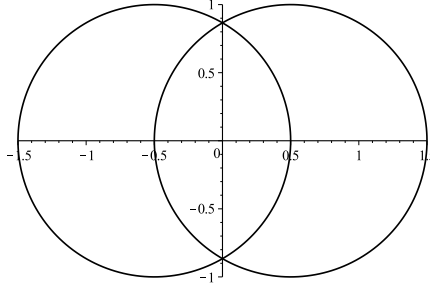
FIGURE 1.2.1.   Two circles

PROPOSITION 1.2.9. *All circles in $\mathbb{C}P^2$ include the two points $[\pm i\colon 1\colon 0]$.*

PROOF. Let

$$(x_1 - a)^2 + (x_2 - b)^2 = r^2$$

be a circle in $\mathbb{C}^2$. This becomes

$$x_3^2 \left(\left(\frac{x_1}{x_3} - a \right)^2 + \left(\frac{x_2}{x_3} - b \right)^2 - r^2 \right) = 0$$

or

$$(x_1 - ax_3)^2 + (x_2 - bx_3)^2 = r^2 x_3^2$$

If we set $x_3 = 0$, we get equation 1.2.4 on the preceding page.   □

This is encouraging because, a first glance, it would appear that the circles in Figure 1.2.1 only intersect at two points. Proposition 1.2.9 shows that they actually intersect at two other points as well, giving a total of *four* intersections as predicted by Conjecture 1.1.1 on page 3.

LEMMA 1.2.10. *Let $F(x_1, x_2)$ be homogeneous of degree k. Then F factors into k linear factors*

$$F(x_1, x_2) = \prod_{i=1}^{k} (\alpha_i x_1 - \beta_i x_2)$$

with $\alpha_i, \beta_i \in \mathbb{C}$.

PROOF. The idea is that F is naturally defined over $\mathbb{C}P^1 = \mathbb{C} \cup [1\colon 0]$. Let k_0 be the highest power of x_2 such that

$$F = x_2^{k_0} G(x_1, x_2)$$

Then G will be homogeneous of degree $k - k_0$ and

$$G(x_1, x_2) = x_2^{k - k_0} G \left(\frac{x_1}{x_2}, 1 \right)$$

and we solve

$$G \left(\frac{x_1}{x_2}, 1 \right) = g(z) = 0$$

in the usual way to get a factorization

$$g(z) = \prod_{i=1}^{k-k_0} (\alpha_i z - \beta_i)$$

Now replace z by x_1/x_2 and multiply by $x_2^{k-k_0}$ to get the factorization

$$F(x_1, x_2) = x_2^{k_0} \prod_{i=1}^{k-k_0} (\alpha_i x_1 - \beta_i x_2)$$

$\square$

EXERCISES.

1. Convert the equation

$$x^2 + 3xy + 25 = 0$$

into an equation in $\mathbb{C}P^2$ and describe the point-set it defines.

2. Find points of intersection of the parabolas

$$\begin{aligned} y &= x^2 + 1 \\ y &= x^2 + 2 \end{aligned}$$

in $\mathbb{C}P^2$.

3. Factor the homogeneous polynomial

$$x^3 + 6x^2 y + 11xy^2 + 6y^3$$

4. Convert $x^2 + y^2 + 9 = 0$ to an equation in $\mathbb{C}P^2$ and describe the point-set it defines.

5. If A is an $(n+1) \times (n+1)$ invertible matrix, show that

$$A : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$$

on homogeneous coordinates, defines a continuous map

$$\bar{A} : \mathbb{C}P^n \to \mathbb{C}P^n$$

6. Why did the matrix in the previous problem have to be *invertible*?

7. Which $(n+1) \times (n+1)$ matrices

$$A : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$$

have the property that the induced map

$$\bar{A} : \mathbb{C}P^n \to \mathbb{C}P^n$$

preserves $\mathbb{C}^n \subset \mathbb{C}P^n$ (as in proposition 1.2.3 on page 4).

8. Suppose A is the $(n+1) \times (n+1)$ matrix

$$A = \begin{bmatrix} 1 & \cdots & 0 & z_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & z_n \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

The induced map

$$\bar{A} \colon \mathbb{C}P^n \to \mathbb{C}P^n$$

preserves $\mathbb{C}^n \subset \mathbb{C}P^n$. What effect does $\bar{A}$ have on $\mathbb{C}^n$?

1.3. The Sylvester Matrix

In order to pursue these ideas further, we need some more algebraic machinery.

We begin by trying to answer the question:

Given polynomials

(1.3.1)	$f(x) = a_n x^n + \cdots + a_0$
(1.3.2)	$g(x) = b_m x^m + \cdots + b_0$

when do they have a common root?

An initial (but not very helpful) answer is provided by:

LEMMA 1.3.1. *If $f(x)$ is a nonzero degree n polynomial and $g(x)$ is a nonzero degree m polynomial, they have a common root if and only if there exist nonzero polynomials $r(x)$ of degree $\leq m - 1$ and $s(x)$ of degree $\leq n - 1$ such that*

$$(1.3.3) \qquad r(x)f(x) + s(x)g(x) = 0$$

REMARK. Note that the conditions on the degrees of $r(x)$ and $s(x)$ are important. Without them, we could just write

$$\begin{aligned} r(x) &= g(x) \\ s(x) &= -f(x) \end{aligned}$$

and *always* satisfy equation 1.3.3.

PROOF. Suppose $f(x), g(x)$ have a common root, α. Then we can set

$$\begin{aligned} r(x) &= g(x)/(x - \alpha) \\ s(x) &= -f(x)/(x - \alpha) \end{aligned}$$

and satisfy equation 1.3.3.

On the other hand, if equation 1.3.3 is satisfied it follows that $r(x)f(x)$ and $s(x)g(x)$ are degree $t \leq n + m - 1$ polynomials that have the *same t* factors

$$x - \alpha_1, \ldots, x - \alpha_t$$

since they cancel each other out. This set (of factors) of size t includes the n factors of $f(x)$ and the m factors of $g(x)$. The pigeonhole principal implies that at least 1 of these factors must be common to $f(x)$ and $g(x)$. And this common factor implies the existence of a common root. $\square$

Suppose

$$\begin{aligned} r(x) &= u_{m-1}x^{m-1} + \cdots + u_0 \\ s(x) &= v_{n-1}x^{n-1} + \cdots + v_0 \end{aligned}$$

Then

(1.3.4)
$$r(x) \cdot f(x) = \sum_{i=0}^{n+m-1} x^i c_i$$

$$s(x) \cdot g(x) = \sum_{i=0}^{n+m-1} x^i d_i$$

where $c_i = \sum_{j+k=i} u_j a_k$. We can compute the coefficients $\{c_i\}$ by matrix products

$$[u_{m-1}, \ldots, u_0] \begin{bmatrix} a_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} = c_{n+m-1}$$

and

$$[u_{m-1}, \ldots, u_0] \begin{bmatrix} a_{n-1} \\ a_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} = c_{n+m-2}$$

or, combining the two,

$$[u_{m-1}, \ldots, u_0] \begin{bmatrix} a_n & a_{n-1} \\ 0 & a_n \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} = [c_{n+m-1}, c_{n+m-2}]$$

where the subscripts of a_k increase from top to bottom and those of the u_j increase from left to right.

On the other end of the scale

$$[u_{m-1}, \ldots, u_0] \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_0 \end{bmatrix} = c_0$$

and

$$[u_{m-1}, \ldots, u_0] \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_0 \\ a_1 \end{bmatrix} = c_1$$

so we get

$$[u_{m-1}, \ldots, u_0] \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ a_0 & 0 \\ a_1 & a_0 \end{bmatrix} = [c_1, c_0]$$

This suggests creating a matrix

$$M_1 = \begin{bmatrix} a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & a_n & \cdots & a_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n & \cdots & a_1 & a_0 \end{bmatrix}$$

of m rows and $n + m$ columns. The top row contains the coefficients of $f(x)$ followed by $m - 1$ zeros and each successive row is the one above shifted to the right. We stop when a_0 reaches the rightmost column. Then

$$\begin{bmatrix} u_{m-1} & \cdots & u_0 \end{bmatrix} M_1 = \begin{bmatrix} c_{n+m-1} & \cdots & c_0 \end{bmatrix} = [\mathbf{c}]$$

so we get the coefficients of $r(x)f(x)$. In like fashion, we can define a matrix with n rows and $n + m$ columns

$$M_2 = \begin{bmatrix} b_m & b_{m-1} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & b_m & \cdots & b_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_m & \cdots & b_1 & b_0 \end{bmatrix}$$

whose top row is the coefficients of $g(x)$ followed by $n - 1$ zeros and each successive row is shifted one position to the right, with b_0 on the right in the bottom row. Then

$$\begin{bmatrix} v_{n-1} & \cdots & v_0 \end{bmatrix} M_2 = [d_{n+m-1}, \ldots, d_0] = [\mathbf{d}]$$

— a vector of the coefficients of $s(x)g(x)$. If we combine the two together, we get an $(n + m) \times (n + m)$-matrix

$$S = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$$

with the property that

(1.3.5) $\begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix} S = [\mathbf{c} + \mathbf{d}]$

— an $n + m$ dimensional vector of the coefficients of $r(x)f(x) + s(x)g(x)$, where

(1.3.6) $\begin{aligned} \mathbf{u} &= \begin{bmatrix} u_{m-1} & \cdots & u_0 \end{bmatrix} \\ \mathbf{v} &= \begin{bmatrix} v_{n-1} & \cdots & v_0 \end{bmatrix} \end{aligned}$

It follows that S reduces the question of the existence of a common root of $f(x)$ and $g(x)$ to *linear algebra:* The equation

(1.3.7) $\begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix} S = [0]$

has a nontrivial solution if and only if $\det(S) = 0$.

DEFINITION 1.3.2. If

$$f(x) = a_n x^n + \cdots + a_0$$
$$g(x) = b_m x^m + \cdots + b_0$$

are two polynomials, their *Sylvester Matrix* is the $(n+m) \times (n+m)$-matrix

$$S(f,g,x) = \begin{bmatrix} a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & a_n & \cdots & a_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n & \cdots & a_1 & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & b_m & \cdots & b_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_m & \cdots & b_1 & b_0 \end{bmatrix}$$

and its determinant $\det(S(f,g,x)) = \mathrm{Res}(f,g,x)$ is called the *resultant* of f and g.

> James Joseph Sylvester (1814–1897) was an English mathematician who made important contributions to matrix theory, invariant theory, number theory and other fields.

The reasoning above shows that:

PROPOSITION 1.3.3. *The polynomials $f(x)$ and $g(x)$ have a common root if and only if $\mathrm{Res}(f,g,x) = 0$.*

PROOF. Equations 1.3.2 on page 10 and 1.3.5 on the preceding page imply that the hypothesis of lemma 1.3.1 on page 10 is satisfied if and only if $\det(S(f,g,x)) = 0$. □

EXAMPLE. For instance, suppose

$$\begin{aligned} f(x) &= x^2 - 2x + 5 \\ g(x) &= x^3 + x - 3 \end{aligned}$$

Then the Sylvester matrix is

$$M = \begin{bmatrix} 1 & -2 & 5 & 0 & 0 \\ 0 & 1 & -2 & 5 & 0 \\ 0 & 0 & 1 & -2 & 5 \\ 1 & 0 & 1 & -3 & 0 \\ 0 & 1 & 0 & 1 & -3 \end{bmatrix}$$

and the resultant is 169, so these two polynomials have no common roots.

There are many interesting applications of the resultant. Suppose we are given parametric equations for a curve

$$\begin{aligned} x &= \frac{f_1(t)}{g_1(t)} \\ y &= \frac{f_2(t)}{g_2(t)} \end{aligned}$$

where f_i and g_i are polynomials, and want an implicit equation for that curve, i.e. one of the form

$$F(x,y) = 0$$

This is equivalent to finding x, y such that the polynomials

$$\begin{aligned} f_1(t) - xg_1(t) &= 0 \\ f_2(t) - yg_2(t) &= 0 \end{aligned}$$

have a common root (in t). So the condition is

$$\text{Res}(f_1(t) - xg_1(t), f_2(t) - yg_2(t), t) = 0$$

This resultant will be a polynomial in x and y. We have *eliminated* the variable t — in a direct generalization of Gaussian elimination — and the study of such algebraic techniques is the basis of Elimination Theory — see section 2.3 on page 45. This develops the theory of *Gröbner Bases*, which allow one to perform arbitrarily *many* eliminations in a single step.

For example, let

$$\begin{aligned} x &= t^2 \\ y &= t^2(t+1) \end{aligned}$$

Then the Sylvester matrix is

$$\begin{bmatrix} 1 & 0 & -x & 0 & 0 \\ 0 & 1 & 0 & -x & 0 \\ 0 & 0 & 1 & 0 & -x \\ 1 & 1 & 0 & -y & 0 \\ 0 & 1 & 1 & 0 & -y \end{bmatrix}$$

and the resultant is

$$\text{Res}(t^2 - x, t^2(t+1) - y, t) = -x^3 + y^2 - 2\,yx + x^2$$

and it is not hard to verify that

$$-x^3 + y^2 - 2\,yx + x^2 = 0$$

after plugging in the parametric equations for x and y.

EXERCISES.

1. Compute an implicit equation for the curve defined parametrically by

$$\begin{aligned} x &= t/(1+t^2) \\ y &= t^2/(1-t) \end{aligned}$$

2. Compute an implicit equation for the curve

$$\begin{aligned} x &= t/(1-t^2) \\ y &= t/(1+t^2) \end{aligned}$$

3. Compute an implicit equation for the curve

$$\begin{aligned} x &= (1-t)/(1+t) \\ y &= t^2/(1+t^2) \end{aligned}$$

4. Solve the equations

$$\begin{aligned} x^2 + y^2 &= 1 \\ x + 2y - y^2 &= 1 \end{aligned}$$

by computing a suitable resultant to eliminate y.

5. Find implicit equations for x, y, and z if

$$\begin{aligned} x &= s+t \\ y &= s^2 - t^2 \\ z &= 2s - 3t^2 \end{aligned}$$

Hint: Compute resultants to eliminate s from every pair of equations and then eliminate t from the resultants.

1.4. Application to Bézout's Theorem

In this section, we will denote homogeneous coordinates of $\mathbb{C}P^2$ by $[x:y:z]$.

We will apply the resultant to computing intersections in $\mathbb{C}P^2$. Suppose we have two homogeneous polynomials

$$F(x,y,z)$$
$$G(x,y,z)$$

of degrees n and m, respectively, and we want to compute common zeros.

We regard these as polynomials in one variable with coefficients that are polynomials in the others:

$$(1.4.1) \qquad F(x,y,z) = a_n(x,y)z^n + \cdots + a_0(x,y)$$
$$(1.4.2) \qquad G(x,y,z) = b_m(x,y)z^m + \cdots + b_0(x,y)$$

If we set the *resultant* of *these* polynomials and to zero, we get conditions x and y must satisfy for there to exist a value of z that makes the *original* polynomials equal to zero.

The key step in the proof of Bézout's Theorem is

THEOREM 1.4.1. *The resultant of the polynomial F, in 1.4.1 and G, in 1.4.2 is a homogeneous polynomial in x and y of degree nm (see definition 1.2.4 on page 5)*

PROOF. If $R(x,y)$ is this resultant, we will show that $R(tx,ty) = t^{nm}R(x,y)$ for all t, which will prove the conclusion.

Since F is homogeneous of degree n, the degree of $a_i(x, y)$ will be $n - i$. Similar reasoning shows that the degree of $b_i(x, y)$ is $m - i$. It follows that

$$R(tx, ty) = \det \begin{bmatrix} a_n & ta_{n-1} & \cdots & t^n a_0 & 0 & \cdots & 0 \\ 0 & a_n & \cdots & t^{n-1} a_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n & \cdots & t^{n-1} a_1 & t^n a_0 \\ b_m & tb_{m-1} & \cdots & t^m b_0 & 0 & \cdots & 0 \\ 0 & b_m & & t^{m-1} b_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_m & \cdots & t^{m-1} b_1 & t^m b_0 \end{bmatrix}$$

Now do the following operations to the top and bottom halves of this matrix:

Multiply the second row by t, the third by t^2 and the mth row (in the top half) by t^{m-1} and the nth row (in the bottom half) by t^{n-1}.

We get

$$t^{N+M} R(tx, ty) =$$

$$\det \begin{bmatrix} a_n & ta_{n-1} & \cdots & t^n a_0 & 0 & \cdots & 0 \\ 0 & ta_n & \cdots & t^n a_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n & \cdots & t^{n+m-2} a_1 & t^{n+m-1} a_0 \\ b_m & tb_{m-1} & \cdots & t^m b_0 & 0 & \cdots & 0 \\ 0 & tb_m & \ddots & t^m b_1 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & t^{n-1} b_m & \cdots & t^{m+n-2} b_1 & t^{m+n-1} b_0 \end{bmatrix}$$

where

$$N = 1 + \cdots + n - 1 = n(n-1)/2$$
$$M = 1 + \cdots + m - 1 = m(m-1)/2$$

This new matrix is the same as the original Sylvester matrix with the second *column* multiplied by t, the third by t^2 and so on. It follows that

$$t^{N+M} R(tx, ty) = t^Z R(x, y)$$

where

$$Z = 1 + \cdots + n + m - 1 = (m+n)(m+n-1)/2$$

We conclude that

$$R(tX, tY) = t^{\frac{(m+n)(m+n-1)}{2} - \frac{n(n-1)}{2} - \frac{m(m-1)}{2}} R(x, y) = t^{nm} R(x, y)$$

$$\square$$

COROLLARY 1.4.2. *The resultant of the polynomial F in 1.4.1 on page 15 and G in 1.4.2 on page 15 factors into nm linear factors*

$$(1.4.3) \qquad \mathrm{Res}(F, G, z) = \prod_{i=1}^{nm} (\alpha_i x - \beta_i y)$$

PROOF. This follows immediately from theorem 1.4.1 on page 15 and lemma 1.2.10 on page 8. □

Note that each of these factors defines the equation of a line through the origin like $y = \alpha_i x / \beta_i$. These are lines from the origin to the intersections between the two curves defined by equation 1.4.6 and 1.4.7.

If there is a *nonzero value* of z associated with a given factor $(\alpha_i x - \beta_i y)$ then the intersection point is

$$\left(x : \frac{\alpha_i x}{\beta_i} : z \right) = \left(\frac{x}{z} : \frac{\alpha_i x}{\beta_i z} : 1 \right)$$

although it is usually easier to simply find where the line $y = \alpha_i x / \beta_i$ intersects both curves than to solve for z.

Here is an example:

EXAMPLE 1.4.3. We compute the intersections between the parabola

$$(1.4.4) \qquad y = x^2 + 1$$

and the circle

$$(1.4.5) \qquad x^2 + y^2 = 4$$

We first translate these to formulas in $\mathbb{C}P^2$:

$$z^2 \left(\frac{y}{z} - \left(\frac{x}{z} \right)^2 - 1 \right) = 0$$

$$(1.4.6) \qquad yz - x^2 - z^2 = 0$$

and

$$z^2 \left(\left(\frac{x}{z} \right)^2 + \left(\frac{y}{z} \right)^2 - 2 \right) = 0$$

$$(1.4.7) \qquad x^2 + y^2 - 4z^2 = 0$$

Now we regard equations 1.4.6 and 1.4.7 as polynomials in z whose coefficients are polynomials in x and y. The Sylvester matrix is

$$\begin{bmatrix} -1 & y & -x^2 & 0 \\ 0 & -1 & y & -x^2 \\ -4 & 0 & x^2 + y^2 & 0 \\ 0 & -4 & 0 & x^2 + y^2 \end{bmatrix}$$

and the resultant is

$$6\,y^2 x^2 - 3\,y^4 + 25\,x^4$$
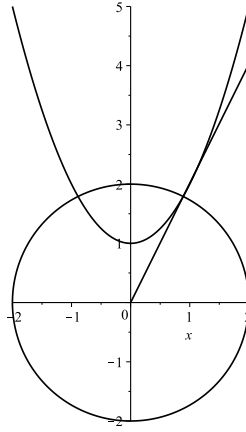
FIGURE 1.4.1.   Intersection in example 1.4.3 on the previous page.

and we get the factorization (using a computer algebra system like Maple, Maxima or Sage):

$$25 \left(x - \frac{1}{5} iy \sqrt{3 + 2\sqrt{21}} \right) \left(x + \frac{1}{5} iy \sqrt{3 + 2\sqrt{21}} \right)$$
$$\left(x - \frac{1}{5} y \sqrt{-3 + 2\sqrt{21}} \right) \left(x + \frac{1}{5} y \sqrt{-3 + 2\sqrt{21}} \right)$$

Setting factors to 0 gives us lines through intersections between the parabola and circle. For instance, setting the third factor to 0 gives:

$$x = \frac{y}{5} \sqrt{-3 + 2\sqrt{21}}$$

and figure 1.4.1 shows the intersection of this line with the two curves.

To compute the *point* of intersection, we plug it into equation 1.4.4 on the preceding page and get

$$y = \frac{3}{2} + \frac{1}{6} \sqrt{21}$$
$$y = -\frac{1}{2} + \frac{1}{2} \sqrt{21}$$

If we plug it into equation 1.4.5 on the previous page, we get

$$y = -\frac{1}{2} + \frac{1}{2} \sqrt{21}$$
$$y = \frac{1}{2} - \frac{1}{2} \sqrt{21}$$

so the *common value* is $y = -\frac{1}{2} + \frac{1}{2}\sqrt{21}$ and the corresponding x value is

$$\left(-\frac{1}{2} + \frac{1}{2}\sqrt{21}\right)\left(\frac{1}{5}\sqrt{-3 + 2\sqrt{21}}\right)$$

$$= -\frac{1}{10}\sqrt{-3 + 2\sqrt{21}} + \frac{1}{10}\sqrt{42 - 3\sqrt{21}}$$

We have found *one intersection point* between the two curves:

$$\left(-\frac{1}{10}\sqrt{-3 + 2\sqrt{21}} + \frac{1}{10}\sqrt{42 - 3\sqrt{21}}, -\frac{1}{2} + \frac{1}{2}\sqrt{21}\right)$$

and there are clearly three others.

Here's another example:

EXAMPLE 1.4.4. The curves are

$$\begin{aligned} x^2 + y^2 &= 4 \\ xy &= 1 \end{aligned}$$

Extended to $\mathbb{C}P^2$, these become

$$(1.4.8) \qquad \begin{aligned} x^2 + y^2 - 4z^2 &= 0 \\ xy - z^2 &= 0 \end{aligned}$$

and the resultant is

$$\left(4xy - x^2 - y^2\right)^2$$

with a factorization

$$(1.4.9) \qquad (x - (2 + \sqrt{3})y)^2(x - (2 - \sqrt{3})y)^2$$

In this case, we get two factors of multiplicity 2. This represents the symmetry in the graph: a line given by

$$x = (2 + \sqrt{3})y$$

and representing a factor of the resultant, intersects the graph in *two* points

$$\pm\left(\sqrt{2 + \sqrt{3}}, \frac{1}{\sqrt{2 + \sqrt{3}}}\right)$$

— see figure 1.4.2 on the following page

Unfortunately, figure 1.4.2 raises a possible problem:

> A linear factor of equation 1.4.3 might represent more than one intersection. After all, the vanishing of the resultant means that F and G in equations 1.4.1 and 1.4.2 have at least one common root. They are polynomials of degree n and m, respectively, and what is to prevent them from having many common roots? Maybe the curves we are studying have *more* than mn intersections.
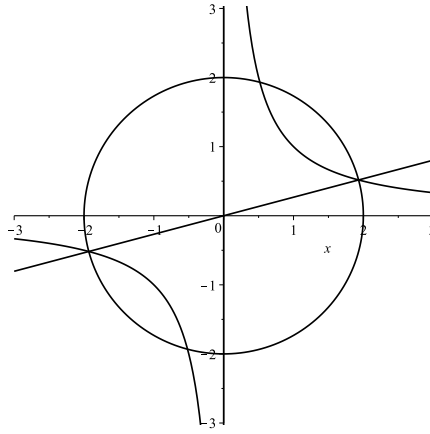
FIGURE 1.4.2.   A factor of multiplicity 2 in example 1.4.4 on
the previous page.
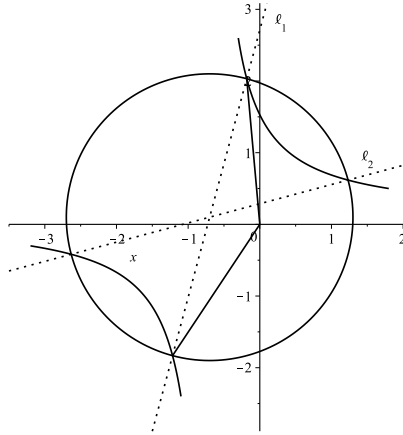
FIGURE 1.4.3.   Perturbation of figure 1.4.2

The key is that the factors in equation 1.4.3 represent *lines* from the origin to the
intersections. Suppose the intersections between the curves are

$$\{(x_i : y_i : z_i)\}$$

and draw lines $\{\ell_j\}$ through each *pair* of points. Now displace both curves by
a small amount (u, v)

$$
\begin{aligned}
F(x - u, y - v, z) &= 0 \\
G(x - u, y - v, z) &= 0
\end{aligned}
$$

so that none of the $\{\ell_j\}$ passes through the origin. The number of intersections
will not change and every linear factor of equation 1.4.3 will represent a *unique*
intersection.
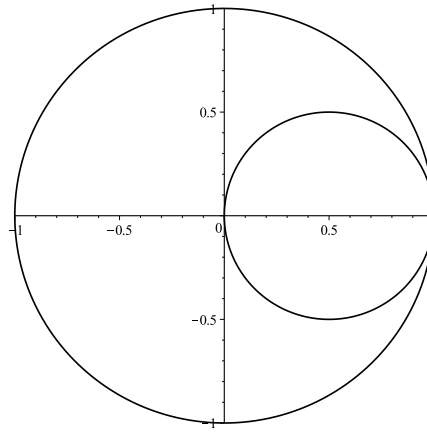
FIGURE 1.4.4.   Intersection of multiplicity 2

For instance, applying a displacement of $(.2, .2)$ to the equations in 1.4.8 translates the intersections by this. The *single* line through two intersections in figure 1.4.2 on the preceding page splits into *two* lines, as in figure 1.4.3 on the facing page. The squared term in expression 1.4.9 on page 19 representing this line splits into two distinct linear factors.

Since corollary 1.4.2 on page 16 implies that the resultant has $n \cdot m$ linear factors, we conclude that:

PROPOSITION 1.4.5. *The curves defined by equations 1.4.1 on page 15 and 1.4.2 have at most mn intersections in* $\mathbb{C}P^2$.

Now we consider another potential problem in counting intersections. Consider the equations

(1.4.10)
$$
\begin{aligned}
x^2 + y^2 &= 4 \\
(x-1)^2 + y^2 &= 1
\end{aligned}
$$

which give

$$
\begin{aligned}
x^2 + y^2 - 4z^2 &= 0 \\
(x-z)^2 + y^2 - z^2 &= 0
\end{aligned}
$$

in $\mathbb{C}P^2$ and have a resultant

$$-4\,x^2 y^2 - 4\,y^4$$

that factors as

$$-4y^2(x^2 + y^2)$$

In this case, the factor $x^2 + y^2 = 0$ represents the two points at infinity that all circles have — see proposition 1.2.9 on page 8. The only other intersection point is $(2, 0)$, which corresponds to the factor y^2 of multiplicity two: setting $y = 0$ in equations 1.4.10 forces x to be 2. Figure 1.4.4 shows that the circles are *tangent* to each other at the point of intersection.

We will perturb the equations in 1.4.10 on the previous page slightly:

$$x^2 + y^2 - 4z^2 = 0$$
$$(x - az)^2 + y^2 - z^2 = 0$$

The resultant becomes

$$9\,x^4 + 18\,x^2y^2 - 10\,x^4a^2 - 4\,x^2y^2a^2$$
$$+ 9\,y^4 + 6\,y^4a^2 + x^4a^4 + 2\,x^2a^4y^2 + y^4a^4$$

which we factor as

$$- (x^2 + y^2)\left(x\sqrt{10\,a^2 - a^4 - 9} - \left(a^2 + 3\right)y\right)$$
$$\left(x\sqrt{10\,a^2 - a^4 - 9} + \left(a^2 + 3\right)y\right)$$

Which gives us two lines of intersection

(1.4.11)
$$y = \pm\frac{\sqrt{10\,a^2 - a^4 - 9}}{a^2 + 3} \cdot x$$

that *coalesce* as $a \to 1$. It follows that the one point of intersection in figure 1.4.4 on the preceding page is the result of two intersections *merging*. We call the result an intersection of *multiplicity two.*

Consider a line joining the two intersections in equation 1.4.11. As the intersections approach each other, this line becomes the tangent line to both curves and the curves must be tangent to each other at this intersection of multiplicity two.

Consequently, we call an intersection that splits into n intersections when we perturb the equations slightly, an *intersection of multiplicity n*. This a *direct generalization* of the concepts of roots of polynomials having a multiplicity larger than 1: A root of the polynomial

$$p(x)$$

of multiplicity n defines an intersection of multiplicity n between the curve

$$y = p(x)$$

and the x-axis.

This definition is not particularly useful and we attempt to improve on it by considering examples.

If we write the circles in the form

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

and expand them in Taylor series, we get, respectively,

$$x = 2 - \frac{1}{4}y^2 - \frac{1}{64}y^4 + O\left(y^6\right)$$
$$x = 2 - \frac{1}{2}y^2 - \frac{1}{8}y^4 + O\left(y^6\right)$$

The 2 as the constant term merely means the curves intersect at (2,0). The first *difference* between the curves is in the y^2-term.
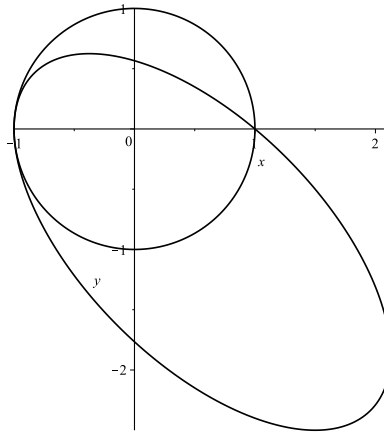
FIGURE 1.4.5.   Intersection of multiplicity 3

Intersections can have higher multiplicities, as this example shows

$$5\,x^2 + 6\,xy + 5\,y^2 + 6\,y - 5 \;=\; 0$$
$$x^2 + y^2 - 1 \;=\; 0$$

If we map into $\mathbb{C}P^2$ and compute the resultant, we get $-36y^4$ which implies that all intersections occur at $y = 0$.

In the graph of these curves in figure 1.4.5 we only see two intersections at $y = 0$. Since the curves are tangent at the intersection point on the left, it must have multiplicity > 1.

We compute Taylor series expansions for x in terms of y:

$$x \;=\; -\sqrt{1 - y^2}$$
$$x \;=\; -\frac{3}{5}\,y - \frac{1}{5}\,\sqrt{-16\,y^2 + 25 - 30\,y}$$

and get

$$x \;=\; -1 + \frac{1}{2}y^2 + \frac{1}{8}y^4 + O\left(x^6\right)$$
$$x \;=\; -1 + \frac{1}{2}y^2 + \frac{3}{10}y^3 + \frac{61}{200}y^4 + \frac{333}{1000}y^5 + O\left(y^6\right)$$

and the difference between them is a multiple of y^3. We will call this an intersection of multiplicity 3.

A rigorous definition of intersection multiplicity requires more algebraic machinery than we have now, but we can *informally* define[2]

---

[2]Using the results of section 3.3.3 on page 134, it is possible to make this definition rigorous.
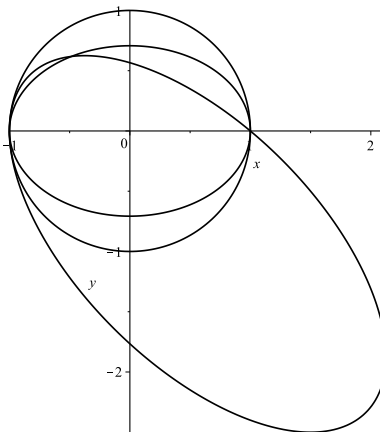
FIGURE 1.4.6.   Intersection of a union of curves

DEFINITION 1.4.6. Let

$$f(x,y) = 0$$
$$g(x,y) = 0$$

be two curves (with f and g polynomials) that intersect at a point p. After a
linear transformation, giving

$$\bar{f}(x,y) = 0$$
$$\bar{g}(x,y) = 0$$

assume that

(1) p is at the origin and
(2) neither the curve defined by $\bar{f}(x,y) = 0$ nor that defined by $\bar{g}(x,y) = 0$
     is tangent to the y-axis.

We define the *multiplicity of the intersection* to be the degree of the lowest term
in the Taylor series for y as a function of x for $\bar{f}$ that differs from a term in y as
a function of x for $\bar{g}$.

REMARK. There are many other ways we could informally define multi-
plicity. For instance, we can compute implicit derivatives

$$\frac{d^n y}{dx^n}$$

at p for both curves and define the intersection multiplicity as $1+$ the largest
value of n for which these derivatives agree.

This definition has the shortcoming that it does not take *unions* of curves
into account. For instance the curve defined by

$$\left(X^2 + Y^2 - 1\right)\left(X^2 + 2Y^2 - 1\right) = 0$$

is a union of a circle and an ellipse. We would expect intersection multiplicities
to *add* in this situation — see figure 1.4.6. We could augment definition 1.4.6 on

page 23 by specifying that the curves

$$f(X,Y) = 0$$
$$g(X,Y) = 0$$

cannot be proper *unions* of other curves, but what exactly would that mean? And how would we test that condition?

We need the algebraic machinery in appendix A on page 341 to answer this and other questions. With that machinery under our belt, we will be able to give a rigorous (and more abstract) definition of intersection multiplicity in section 3.3.4 on page 139. We will finally give a rigorous proof of a generalization of Bézout's Theorem (or a modern version of it) in section 5.8 on page 274.

Despite our flawed notion of intersection multiplicity, we can finally give a correct statement of Bézout's Theorem:

THEOREM 1.4.7. *Let $f(x,y)$ and $g(x,y)$ be polynomials of degree n and m, respectively with no common factor. Then the curves*

$$f(x,y) = 0$$
$$g(x,y) = 0$$

intersect in nm points in $\mathbb{C}P^2$, where intersections are counted with multiplicities.

REMARK. The first published proof (in [**16**], available in English translation as [**17**]) is due to Bézout, but his version did not take multiplicities into account so it was incorrect. In fact, since others had already stated the result without proof, critics called Bézout's Theorem "neither original nor correct".

1.5. The Mystic Hexagram

In this section, we give a simple application of Bézout's theorem.

Blaise Pascal (1623–1662) was a famous French mathematician, physicist, inventor and philosopher who wrote his first mathematical treatise *Essai pour les coniques* at the age of 16. At the age of 31, he abandoned science and studied philosophy and theology.

Pascal's *Essai pour les coniques* proved what Pascal called the *Hexagrammum Mysticum Theorem* ("the Mystic Hexagram"). It is now simply called Pascal's Theorem.

It states:

> If you inscribe a hexagon in any conic section, C, and extend the three pairs of opposite sides until they intersect, the intersections will lie on a line. See figure 1.5.1 on the next page.

Although Pascal's original proof was lost, we can supply a proof using Bézout's theorem.

Suppose C is defined by the quadratic equation

$$C(X,Y) = 0$$

and we respectively label the six lines that form the hexagon, A_1, B_1, A_2, B_2, A_3, B_3 (see figure 1.5.2 on the following page, where the A-lines are solid and
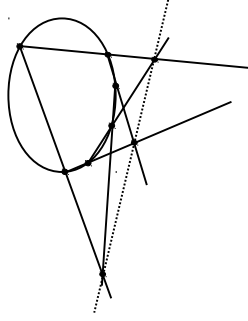
FIGURE 1.5.1.   Hexagrammum Mysticum
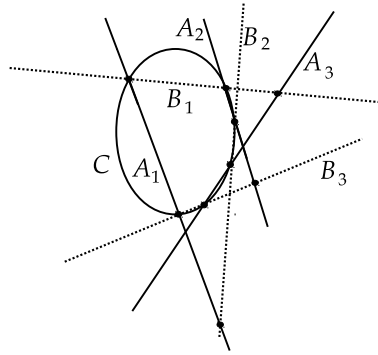
FIGURE 1.5.2.   Grouping the lines into two sets

the B-lines are dotted). By an abuse of notation, we label the linear equations defining the lines the same way

$$A_i(X, Y) = 0 \text{ and } B_i(X, Y) = 0$$

for $i = 1, 2, 3$. Then $A_1 A_2 A_3 = 0$ is a single cubic equation whose solutions are the three A-lines given above and $B_1 B_2 B_3 = 0$ defines the B-lines. Bézout's theorem implies that these two *sets* of lines intersect in 9 points[3], 6 of which lie on the conic section defined by C.

Construct the cubic equation

$$f_\lambda(X, Y) = A_1 A_2 A_3 + \lambda B_1 B_2 B_3$$

where λ is some parameter. No matter *what* we set λ to, the set of points defined by

$$f_\lambda(X, Y) = 0$$

will contain all nine of the intersections of the A-collection of lines with the B-collection — and six of these will be in C.

---

[3]Intersections of A_i with B_j for $i = 1, 2, 3$ and $j = 1, 2, 3$.
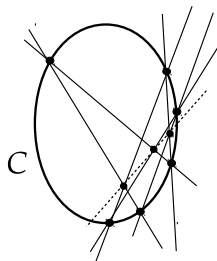
FIGURE 1.5.3.   Another Pascal Line

Now select a value λ_0 of λ that makes f_λ vanish on a *seventh* point of C, distinct from the other six points. We have a cubic function, f_{λ_0}, with a zero-set that intersects the zero-set of a quadratic, C, in 7 points — which seems to violate Bézout's theorem. How is this possible?

The function $f_{\lambda_0}(X, Y)$ must violate the *hypothesis* of Bézout's theorem — i.e., it must have a *common factor* with $C(X, Y)$! In other words,

$$f_{\lambda_0}(X, Y) = C(X, Y) \cdot \ell(X, Y)$$

where $\ell(X, Y)$ is a linear function — since $C(X, Y)$ is quadratic and $f_{\lambda_0}(X, Y)$ is cubic. The zero-set of $f_{\lambda_0}(X, Y)$ still contains *all* of the intersections of the two sets of lines extending the sides of the hexagon. Since six of the intersection lie in the conic section defined by C, the other three must lie in the set

$$\ell(X, Y) = 0$$

In particular, these three intersections lie on a *line* — called a Pascal Line.

Given 6 points on a conic section, there are 60 ways to connect them into "hexagons" — i.e. connect them in such a way that each of the 6 points lies in precisely two lines. Each of these "twisted hexagons" has its own Pascal line. Figure 1.5.3 shows a different way of connecting the *same* six points as in figures 1.5.1 on the facing page and 1.5.2 on the preceding page, and its (dotted) Pascal line.

EXERCISES.

1. How does one get 60 possible hexagons in an ellipse?

2. Suppose the hexagon is *regular* — i.e. all sides are the same length and opposite sides are parallel to each other. Where is the Pascal line in this case?
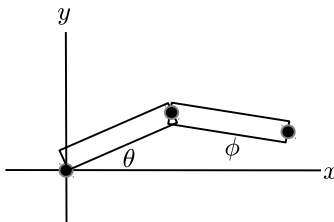
FIGURE 1.6.1.   A simple robot arm

1.6. Robotics

In this section we develop a framework for applying projective spaces to robotics. At this point, we do not have the necessary algebraic machinery to do much beyond setting up the equations. As we develop more of the theory, we will revisit this subject. The excellent book, [**103**], is a good general reference.

The geometry of projective spaces is interesting and useful — even if we are only concerned with $\mathbb{R}^n$: Recall that if we want to represent rotation in $\mathbb{R}^2$ via an angle of θ in the counterclockwise direction, we can use a matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$$

One interesting feature of projective spaces is that linear transformations can represent *displacement* as well as rotation (compare with the solution to exercise 8 on page 10). Regard $\mathbb{R}^2$ as a subspace of $\mathbb{R}P^2$ as in proposition 1.2.3 on page 4, i.e. $(x, y) \mapsto (x : y : 1) \in \mathbb{R}P^2$. The linear transformation

(1.6.1)
$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) & a \\ \sin(\theta) & \cos(\theta) & b \\ 0 & 0 & 1 \end{bmatrix} : \mathbb{R}P^2 \to \mathbb{R}P^2$$

sends

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \in \mathbb{R}^2 \subset \mathbb{R}P^2$$

to

$$\begin{bmatrix} x\cos(\theta) - y\sin(\theta) + a \\ x\sin(\theta) + y\cos(\theta) + b \\ 1 \end{bmatrix} \in \mathbb{R}^2 \subset \mathbb{R}P^2$$

and represents

(1) *rotation* by θ (in a counterclockwise direction), followed by
(2) *displacement* by (a, b).

This feature of projective spaces is heavily used in computer graphics: creating a scene in $\mathbb{R}^3$ is done by creating objects at the *origin* of $\mathbb{R}^3 \subset \mathbb{R}P^3$ and moving them into position (and rotating them) via linear transformations in $\mathbb{R}P^3$.

Suppose we have a simple robot-arm with two links, as in figure 1.6.1.

If we assume that both links are of length ℓ, suppose the second link were attached to the origin rather than at the end of the second link.

Then its endpoint would be at (see equation 1.6.1 on the facing page)

$$
\begin{bmatrix} \ell\cos(\phi) \\ \ell\sin(\phi) \\ 1 \end{bmatrix}
= \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 1 & 0 & \ell \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
$$

$$
= \begin{bmatrix} \cos(\phi) & -\sin(\phi) & \ell\cos(\phi) \\ \sin(\phi) & \cos(\phi) & \ell\sin(\phi) \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
$$

In other words, the effect of moving from the origin to the end of the second link (attached to the origin) is

(1) *displacement* by ℓ — so that $(0,0)$ is moved to $(\ell,0) = (\ell\!:\!0\!:\!1) \in \mathbb{R}P^2$.
(2) *rotation* by ϕ

This is the effect of the *second* link on all of $\mathbb{R}^2$. If we want to compute the effect of *both* links, *insert* the first link into the system — i.e. rigidly attach the second link to the first, displace by ℓ, and rotate by θ. The effect is equivalent to multiplying by

$$
M_2 = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & \ell\cos(\theta) \\ \sin(\theta) & \cos(\theta) & \ell\sin(\theta) \\ 0 & 0 & 1 \end{bmatrix}
$$

It is clear that we can compute the endpoint of any number of links in this manner — always inserting new links at the *origin* and moving the rest of the chain accordingly.

At this point, the reader might wonder

Where does *algebra* enter into all of this?

The point is that we do not have to deal with trigonometric functions until the very last step. If $a, b \in \mathbb{R}$ are numbers with the property that

(1.6.2) $$a^2 + b^2 = 1$$

there is a *unique* angle θ with $a = \cos(\theta)$ and $b = \sin(\theta)$. This enables us to replace the trigonometric functions by real numbers that satisfy equation 1.6.2 and derive purely algebraic equations for

(1) the set of points in $\mathbb{R}^2$ reachable by a robot-arm
(2) strategies for reaching those points (solving for explicit angles).

In the simple example above, let $a_1 = \cos(\theta)$, $b_1 = \sin(\theta)$, $a_2 = \cos(\phi)$, $b_2 = \sin(\phi)$ so that our equations for the endpoint of the second link become

$$
\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
= \begin{bmatrix} a_1 & -b_1 & \ell a_1 \\ b_1 & a_1 & \ell b_1 \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} \ell a_2 \\ \ell b_2 \\ 1 \end{bmatrix}
$$

$$
= \begin{bmatrix} \ell a_1 a_2 - \ell b_2 b_1 + \ell a_1 \\ \ell b_1 a_2 + \ell a_1 b_2 + \ell b_1 \\ 1 \end{bmatrix}
$$
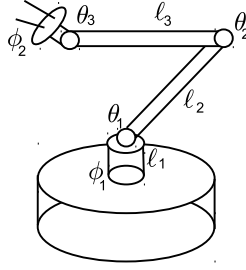
FIGURE 1.6.2.   A more complicated robot arm

It follows that the points (x, y) reachable by this link are those for which the system of equations

$$\ell a_1 a_2 - \ell b_2 b_1 + \ell a_1 - x = 0$$
$$\ell b_1 a_2 + \ell a_1 b_2 + \ell b_1 - y = 0$$
$$a_1^2 + b_1^2 - 1 = 0$$
(1.6.3)
$$a_2^2 + b_2^2 - 1 = 0$$

has *real* solutions (for a_i and b_i). Given values for x and y, we can solve for the set of configurations of the robot arm that will *reach* (x, y). Section 2.3 on page 45 develops the ring-theory needed and example 2.3.18 on page 56 applies this to the robot-arm.

We conclude this chapter with a more complicated robot-arm in figure 1.6.2— somewhat like a Unimation Puma 560[4].

It has:
   (1) A base of height ℓ_1 and motor that rotates the whole assembly by ϕ_1 — with 0 being the positive x-axis.
   (2) An arm of length ℓ_2 that can be moved forward or backward by an angle of θ_1 — with 0 being straight forward (in the positive x-direction).
   (3) A second arm of length ℓ_3 linked to the first by a link of angle θ_2, with 0 being when the second arm is in the same direction as the first.
   (4) A little "hand" of length ℓ_4 that can be inclined from the second arm by an angle of θ_3 and rotated perpendicular to that direction by an angle ϕ_2.

We do our computations in $\mathbb{R}P^3$, start with the "hand" and work our way back to the base. The default position of the hand is on the origin and pointing in the positive x-direction. It displaces the origin in the x-direction by ℓ_4, represented by the matrix

$$D_0 = \begin{bmatrix} 1 & 0 & 0 & \ell_4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

---

[4]In 1985, this type of robot-arm was used to do brain-surgery! See [**101**].

The angle ϕ_2 rotates the hand in the yz-plane, and is therefore represented by

$$
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(\phi_2) & -\sin(\phi_2) & 0 \\
0 & \sin(\phi_2) & \cos(\phi_2) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

or

$$
Z_1 =
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & a_1 & -b_1 & 0 \\
0 & b_1 & a_1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

with $a_1 = \cos(\phi_2)$ and $b_1 = \sin(\phi_2)$. The "wrist" inclines the hand in the xz-plane by an angle of θ_3, given by the matrix

$$
Z_2 =
\begin{bmatrix}
a_2 & 0 & -b_2 & 0 \\
0 & 1 & 0 & 0 \\
b_2 & 0 & a_2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

with $a_2 = \cos(\theta_3)$ and $b_2 = \sin(\theta_3)$ and the composite is

$$
Z_2 Z_1 D_0 =
\begin{bmatrix}
a_2 & -b_2 b_1 & -b_2 a_1 & a_2 \ell_4 \\
0 & a_1 & -b_1 & 0 \\
b_2 & a_2 b_1 & a_2 a_1 & b_2 \ell_4 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

The second arm displaces everything by ℓ_3 in the x-direction, giving

$$
D_1 =
\begin{bmatrix}
1 & 0 & 0 & \ell_3 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

so

$$
D_1 Z_2 Z_1 D_0 =
\begin{bmatrix}
a_2 & -b_2 b_1 & -b_2 a_1 & a_2 \ell_4 + \ell_3 \\
0 & a_1 & -b_1 & 0 \\
b_2 & a_2 b_1 & a_2 a_1 & b_2 \ell_4 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

so and then inclines it by θ_2 in the xz-plane, represented by

$$
Z_3 =
\begin{bmatrix}
a_3 & 0 & -b_3 & 0 \\
0 & 1 & 0 & 0 \\
b_3 & 0 & a_3 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

so that $Z_3 D_1 Z_2 Z_1 D_0$ is

$$\begin{bmatrix} a_3 a_2 - b_3 b_2 & (-a_3 b_2 - b_3 a_2)\, b_1 & (-a_3 b_2 - b_3 a_2)\, a_1 & (a_3 a_2 - b_3 b_2)\, \ell_4 + a_3 \ell_3 \\ 0 & a_1 & -b_1 & 0 \\ b_3 a_2 + a_3 b_2 & (a_3 a_2 - b_3 b_2)\, b_1 & (a_3 a_2 - b_3 b_2)\, a_1 & (b_3 a_2 + a_3 b_2)\, \ell_4 + b_3 \ell_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Continuing in this fashion, we get a huge matrix, Z. To find the endpoint of the robot-arm, multiply

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

(representing the origin of $\mathbb{R}^3 \subset \mathbb{R}P^3$) by Z to get

(1.6.4)
$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = $$

$$\begin{bmatrix} ((a_5 a_3 + b_5 b_4 b_3)\, a_2 + (-a_5 b_3 + b_5 b_4 a_3)\, b_2)\, \ell_4 + (a_5 a_3 + b_5 b_4 b_3)\, \ell_3 + a_5 \ell_2 \\ ((b_5 a_3 - a_5 b_4 b_3)\, a_2 + (-b_5 b_3 - a_5 b_4 a_3)\, b_2)\, \ell_4 + (b_5 a_3 - a_5 b_4 b_3)\, \ell_3 + b_5 \ell_2 \\ (a_4 b_3 a_2 + a_4 a_3 b_2)\, \ell_4 + a_4 b_3 \ell_3 + \ell\ 1 \\ 1 \end{bmatrix}$$

where $a_3 = \cos(\theta_2)$, $b_3 = \sin(\theta_2)$, $a_4 = \cos(\theta_1)$, $b_4 = \sin(\theta_1)$ and $a_5 = \cos(\phi_1)$, $b_5 = \sin(\phi_1)$. Note that $a_i^2 + b_i^2 = 1$ for $i = 1, \ldots, 5$. We are also interested in the *angle* that the hand makes (for instance, if we want to pick something up). To find this, compute

(1.6.5) $\quad Z \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - Z \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = Z \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} =$

$$\begin{bmatrix} (a_5 a_3 + b_5 b_4 b_3)\, a_2 + (-a_5 b_3 + b_5 b_4 a_3)\, b_2 \\ (b_5 a_3 - a_5 b_4 b_3)\, a_2 + (-b_5 b_3 - a_5 b_4 a_3)\, b_2 \\ a_4 b_3 a_2 + a_4 a_3 b_2 \\ 0 \end{bmatrix}$$

The numbers in the top three rows of this matrix are the *direction-cosines* of the hand's direction. We can ask what points the arm can reach with its hand aimed in a particular direction. This question is answered in example 2.3.19 on page 58.

EXERCISES.

1. Find a linear transformation in $\mathbb{R}P^3$ that:

   a. rotates $\mathbb{R}^3 \subset \mathbb{R}P^3$ by an angle of $\pi/4$ in the xy-plane,

   b. displaces $\mathbb{R}^3$ by

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

   c. then rotates by an angle of $\pi/3$ in the xz-plane.

2. In computer graphics, after a scene in $\mathbb{R}P^3$ has been *constructed*, it is *viewed* — i.e., there is a "camera" that photographs the scene with proper perspective. Suppose this camera lies at the origin and is pointed in the positive z direction. Describe the mapping that shows how the scene looks.

How do we handle the situation where the camera is *not* at the origin and pointed in the z-direction?

3. Consider the spiral given by

$$\begin{aligned} x &= \cos(3t) + t \\ y &= \sin(3t) - t \\ z &= t + 3 \end{aligned}$$

Compute the perspective image of this as viewed in a positive z-direction.

CHAPTER 2

Affine varieties

"Number, the boundary of things-become, was represented, not as before, pictorially by a figure, but symbolically by an equation. 'Geometry' altered its meaning; the coordinate system as a picturing disappeared and the point became an entirely abstract number-group."
—Oswald Spengler, chapter 2 (The Meaning of Number), from *The Decline of the West* ([**156**]).

2.1. Introduction

Algebraic geometry concerns itself with objects called *algebraic varieties.* These are essentially solution-sets of systems of algebraic equations, like the curves studied in chapter 1.

Although restricting our attention to algebraic varieties might seem limiting, it has long been known that more general objects like compact smooth manifolds are diffeomorphic to real varieties — see [**125**][1] and [**166**]. The paper [**2**] even shows that many piecewise-linear manifolds, including ones with *no* smooth structure are homeomorphic to real varieties.

We begin with *algebraic sets,* whose geometric properties are completely characterized by a basic algebraic invariant called the *coordinate ring.* The main objects of study — algebraic varieties — are the result of gluing together multiple affine sets.

Throughout this discussion, *k* will denote a fixed *algebraically closed* field (see definition A.2.26 on page 399). In classical algebraic geometry $k = \mathbb{C}$.

In general, the reader should be familiar with the concepts of *rings* and *ideals* — see section A.1 on page 341.

DEFINITION 2.1.1. An *n-dimensional affine space*, $\mathbb{A}^n = k^n$, regarded as a space in which geometric objects can be defined. An *algebraic set* $\mathcal{V}(S)$ in k^n is the set of common zeros of some set S of polynomials in $k[X_1, \ldots, X_m]$:

$$\mathcal{V}(S) = \{(a_1, \ldots, a_n) \in \mathbb{A}^n | f(a_1, \ldots, a_n) = 0 \text{ for all } f(X_1, \ldots, X_n) \in S\}$$

REMARK. It is not hard to see that if the set of polynomials is larger, the set of common zeros will be smaller, i.e.,

$$S \subset S' \implies \mathcal{V}(S) \supset \mathcal{V}(S')$$

If $\mathfrak{a}$ is the *ideal* generated by the polynomials in S, we have $\mathcal{V}(\mathfrak{a}) = \mathcal{V}(S)$ so algebraic sets are described as $\mathcal{V}(\mathfrak{a})$ for some ideal $\mathfrak{a} \subseteq k[X_1, \ldots, X_m]$ (see definition A.1.19 on page 349).

---

[1]Written by John Nash, the character of the film "A beautiful mind."

Recall that all ideals in $k[X_1, \ldots, X_n]$ are *finitely generated* by theorem A.1.50 (the Hilbert Basis Theorem).
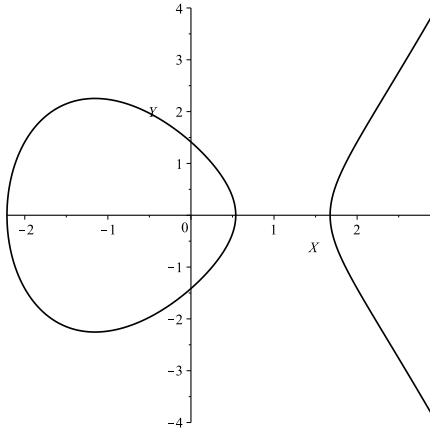
FIGURE 2.1.1.   An elliptic curve

EXAMPLE. For instance, we have
(1) If S is a system of homogeneous linear equations, then $\mathcal{V}(S)$ is a subspace of $\mathbb{A}^n$.
(2) If S consists of the single equation

$$Y^2 = X^3 + aX + b \text{ where } 4a^3 + 27b^2 \neq 0$$

then $\mathcal{V}(S)$ is an *elliptic curve* — studied in detail in section 6.2 on page 313. The quantity, $4a^3 + 27b^2$ is the discriminant (see definition A.1.55 on page 366) of the cubic polynomial $Y^2 = X^3 + aX + b$. Its non-vanishing guarantees that the polynomial has no repeated roots — see corollary A.1.56 on page 366. Figure 2.1.1 shows the elliptic curve $Y^2 = X^3 - 2X + 1$. Elliptic curves over finite fields form the basis of an important cryptographic system — see section 6.2.2 on page 322.
(3) For the zero-ideal, $\mathcal{V}((0)) = \mathbb{A}^n$.
(4) $\mathcal{V}((1)) = \emptyset$,
(5) The algebraic subsets of $k = \mathbb{A}^1$ itself are finite sets of points since they are roots of polynomials.
(6) The *special linear group*, $SL(n,k) \subset \mathbb{A}^{n^2}$ — the group of $n \times n$ matrices with determinant 1. This is an algebraic set because the determinant is a polynomial of the matrix-elements — so that $SL(n,k)$ is the set of zeros of the polynomial, $\det(A) - 1$ for $A \in \mathbb{A}^{n^2}$. This is an example of an algebraic group, an algebraic set that is also a group under a multiplication-map that can be expressed as polynomial functions of the coordinates.

(7) If A is an $n \times m$ matrix whose entries are in $k[X_1, \ldots, X_t]$ and $r \geq 0$ is an integer, then define $\mathcal{R}(A, r)$, the *rank-variety* (also called a *determinantal variety*),

$$\mathcal{R}(A, r) = \begin{cases} \mathbb{A}^t & \text{if } r \geq \min(n, m) \\ p \in \mathbb{A}^t & \text{such that } \operatorname{rank}(A(p)) \leq r \end{cases}$$

This is an algebraic set because the statement that the rank of A is $\leq r$ is the same as saying the determinants of all $(r+1) \times (r+1)$ submatrices are 0.

Here are some basic properties of algebraic sets and the ideals that generate them:

PROPOSITION 2.1.2. *Let* $\mathfrak{a}, \mathfrak{b} \subset k[X_1, \ldots, X_n]$ *be ideals. Then*

(1) $\mathfrak{a} \subset \mathfrak{b} \implies \mathcal{V}(\mathfrak{a}) \supset \mathcal{V}(\mathfrak{b})$
(2) $\mathcal{V}(\mathfrak{a}\mathfrak{b}) = \mathcal{V}(\mathfrak{a} \cap \mathfrak{b}) = \mathcal{V}(\mathfrak{a}) \cup \mathcal{V}(\mathfrak{b})$
(3) $\mathcal{V}(\sum \mathfrak{a}_i) = \bigcap \mathcal{V}(\mathfrak{a}_i)$

PROOF. For statement 2 note that

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

For the reverse inclusions, let $x \notin \mathcal{V}(\mathfrak{a}) \cup \mathcal{V}(\mathfrak{b})$. Then there exist $f \in \mathfrak{a}$ and $g \in \mathfrak{b}$ such that $f(x) \neq 0$ and $g(x) \neq 0$. Then $fg(x) \neq 0$ so $x \notin \mathcal{V}(\mathfrak{a}\mathfrak{b})$. □

It follows that the algebraic sets in $\mathbb{A}^n$ satisfy the axioms of the *closed sets* in a *topology*.

DEFINITION 2.1.3. The *Zariski topology* on $\mathbb{A}^n$ has closed sets that are algebraic sets. Complements of algebraic sets will be called *distinguished open sets*.

REMARK. Oscar Zariski originally introduced this concept in [**177**].
This topology has some distinctive properties:

- every algebraic set is compact in this topology.
- algebraic maps (called *regular* maps) are continuous. The converse is not necessarily true, though. See exercise 3 on page 45.

The Zariski topology is also *extremely coarse* i.e, has very "large" open sets. To see this, recall that the *closure,* $\bar{S}$ of a subset $S \subset X$ of a space is the smallest closed set that contains it — i.e., the intersection of all closed sets that contain S.

Now suppose $k = \mathbb{C}$ and $S \subset \mathbb{A}^1 = \mathbb{C}$ is an arbitrarily line segment, as in figure 2.1.2 on the next page. Then we claim that $\bar{S} = \mathbb{C}$ in the Zariski topology.
Let $\mathfrak{I} \subset \mathbb{C}[X]$ be the ideal of all polynomials that vanish on S. Then the closure of S is the set of points where the polynomials in $\mathfrak{I}$ all vanish — i.e., $\mathcal{V}(\mathfrak{I})$. But nonzero polynomials vanish on finite sets of points and S is infinite. It follows that $\mathfrak{I} = (0)$ i.e., the only polynomials that vanish on S are identically zero. Since $\mathcal{V}((0)) = \mathbb{C}$, we get that the closure of S is *all* of $\mathbb{C}$, as is the closure of *any* infinite set of points.
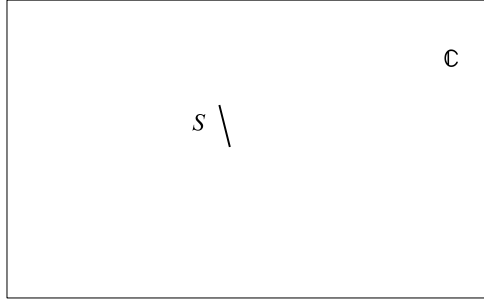
FIGURE 2.1.2.   Closure in the Zariski topology

DEFINITION 2.1.4. For a subset $W \subseteq \mathbb{A}^n$, define
$$\mathcal{I}(W) = \{f \in k[X_1, \ldots, X_n] | f(P) = 0 \text{ for all } P \in W\}$$
It is not hard to see that:

PROPOSITION 2.1.5. *The set $\mathcal{I}(W)$ is an ideal in $k[X_1, \ldots, X_n]$ with the properties:*

(1) $V \subset W \implies \mathcal{I}(V) \supset \mathcal{I}(W)$
(2) $\mathcal{I}(\varnothing) = k[X_1, \ldots, X_n]; \mathcal{I}(k^n) = 0$
(3) $\mathcal{I}(\bigcup W_i) = \bigcap \mathcal{I}(W_i)$
(4) *The Zariski closure of a set $X \subset \mathbb{A}^n$ is exactly $\mathcal{V}(\mathcal{I}(X))$.*

EXERCISES.

1. Show that the Zariski topology on $\mathbb{A}^2$ does not coincide with the product-topology of $\mathbb{A}^1 \times \mathbb{A}^1$ (the Cartesian product).

2. If $V \subset \mathbb{A}^n$ is an algebraic set and $p \notin V$ is a point of $\mathbb{A}^n$, show that any line, ℓ, through p intersects V in a finite number of points (if it intersects it at all).

3. If
$$0 \to M_1 \to M_2 \to M_3 \to 0$$
is a short exact sequence of modules over $k[X_1, \ldots, X_n]$, show that
$$\mathcal{V}(\mathrm{Ann}(M_2)) = \mathcal{V}(\mathrm{Ann}(M_1)) \cup \mathcal{V}(\mathrm{Ann}(M_3))$$

(see definition A.1.73 on page 374 for $\mathrm{Ann}(*)$). This example has applications to the Hilbert polynomial in section 5.7.2 on page 267.

4. If $V = \mathcal{V}\left((X_1^2 + X_2^2 - 1, X_1 - 1)\right)$, what is $\mathcal{I}(V)$?

5. If $V = \mathcal{V}\left((X_1^2 + X_2^2 + X_3^2)\right)$, determine $\mathcal{I}(V)$ when the characteristic of k is 2.

6. Find the ideal $\mathfrak{a} \subset k[X, Y]$ such that $\mathcal{V}(\mathfrak{a})$ is the union of the coordinate-axes.

7. Find the ideal $\mathfrak{a} \subset k[X, Y, Z]$ such that $\mathcal{V}(\mathfrak{a})$ is the union of the three coordinate-axes.

8. If $V \subset \mathbb{A}^2$ is defined by $Y^2 = X^3$, show that every element of $k[V]$ can be uniquely written in the form $f(X) + g(X)Y$.

9. If V is as in exercise 8, show that the map

$$T \mapsto (T^2, T^3)$$

is a 1-1 and onto regular map from $\mathbb{A}^1$ to V that is *not* an isomorphism.

2.2. Hilbert's Nullstellensatz

2.2.1. The weak form. Hilbert's *Nullstellensatz* (in English, "zero-locus theorem") was a milestone in the development of algebraic geometry, making precise the connection between algebra and geometry.

> David Hilbert (1862–1943) was one of the most influential mathematicians in the 19th and early 20th centuries, having contributed to algebraic and differential geometry, physics, and many other fields.

The Nullstellensatz completely characterizes the correspondence between algebra and geometry of affine varieties. It is usually split into two theorems, called the *weak* and *strong* forms of the Nullstellensatz. Consider the question:

> *When do the equations*

$$g(X_1, \ldots, X_n) = 0, g \in \mathfrak{a}$$

> *have a common zero (or are consistent)?*

This is clearly impossible if there exist $f_i \in k[X_1, \ldots, X_n]$ such that $\sum f_i g_i = 1$ — or $1 \in \mathfrak{a}$, so $\mathfrak{a} = k[X_1, \ldots, X_n]$. The weak form of Hilbert's Nullstellensatz essentially says that this is the *only* way it is impossible. Our presentation uses properties of *integral extensions* of rings (see section A.4.1 on page 419).

LEMMA 2.2.1. *Let F be an infinite field and suppose $f \in F[X_1, \ldots, X_n]$, $n \geq 2$ is a polynomial of degree $d > 0$. Then there exist $\lambda_1, \ldots, \lambda_{n-1} \in F$ such that the coefficient of X_n^d in*

$$f(X_1 + \lambda_1 X_n, \ldots, X_{n-1} + \lambda_{n-1} X_n, X_n)$$

is nonzero.

PROOF. If f_d is the homogeneous component of f of degree d (i.e., the sum of all monomials of degree d), then the coefficient of X_n^d in $f(X_1 + \lambda_1 X_n, \ldots, X_{n-1} + \lambda_{n-1} X_n, X_n)$ is $f_d(\lambda_1, \ldots, \lambda_{n-1}, 1)$. Since F is infinite, there is a point $(\lambda_1, \ldots, \lambda_{n-1}) \in F^{n-1}$ for which $f_d(\lambda_1, \ldots, \lambda_{n-1}, 1) \neq 0$ (a fact that is easily established by induction on the number of variables). $\square$

The following result is called the *Noether Normalization Theorem* or Lemma. It was first stated by Emmy Noether in [**129**] and further developed in [**130**].

Besides helping us to prove the Nullstellensatz, it will be used in important geometric results like theorem 2.5.12 on page 77.

THEOREM 2.2.2 (Noether Normalization). *Let F be an infinite field and suppose $A = F[r_1, \ldots, r_m]$ is a finitely generated F-algebra that is an integral domain with generators $r_1 \ldots, r_m$. Then for some $q \leq m$, there are algebraically independent elements $y_1, \ldots, y_q \in A$ such that the ring A is integral (see definition A.4.3 on page 420) over the polynomial ring $F[y_1, \ldots, y_q]$.*

REMARK. Recall that an *F*-algebra is a vector space over *F* that is also a ring. The r_i generate it as a *ring* (so the vector space's dimension over *F* might be $> m$).

PROOF. We prove this by induction on *m*. If the r_i are algebraically independent, simply set $y_i = r_i$ and we are done. If not, there is a nontrivial polynomial $f \in F[x_1, \ldots, x_m]$, say of degree *d* such that

$$f(r_1, \ldots, r_m) = 0$$

and lemma 2.2.1 on the preceding page that there a polynomial of the form

$$r_m^d + g(r_1, \ldots, r_m) = 0$$

If we regard this as a polynomial of r_m with coefficients in $F[r_1, \ldots, r_{m-1}]$ we get

$$r_m^d + \sum_{i=1}^{d-1} g_i(r_1, \ldots, r_{m-1}) r_m^i = 0$$

which implies that r_m is integral over $F[r_1, \ldots, r_{m-1}]$. By the inductive hypothesis, $F[r_1, \ldots, r_{m-1}]$ is integral over $F[y_1, \ldots, y_q]$, so statement 2 of proposition A.4.5 on page 421 implies that r_m is integral over $F[y_1, \ldots, y_q]$ as well. □

We are now ready to prove:

THEOREM 2.2.3 (Hilbert's Nullstellensatz (weak form)). *The maximal ideals of $k[X_1, \ldots, X_n]$ are precisely the ideals*

$$\Im(a_1, \ldots, a_n) = (X_1 - a_1, X_2 - a_2, \ldots, X_n - a_n)$$

for all points

$$(a_1, \ldots, a_n) \in \mathbb{A}^n$$

Consequently every proper ideal $\mathfrak{a} \subset k[X_1, \ldots, X_n]$ has a 0 in $\mathbb{A}^n$.

REMARK. See proposition A.1.20 on page 350 and lemma A.1.30 on page 355 for a discussion of the properties of maximal ideals.

PROOF. Clearly

$$k[X_1, \ldots, X_n] / \Im(a_1, \ldots, a_n) = k$$

The projection

$$k[X_1, \ldots, X_n] \to k[X_1, \ldots, X_n] / \Im(a_1, \ldots, a_n) = k$$

is a homomorphism that evaluates polynomial functions at the point $(a_1, \ldots, a_n) \in \mathbb{A}^n$. Since the quotient is a field, the ideal $\Im(a_1, \ldots, a_n)$ is maximal (see lemma A.1.30 on page 355).

We must show that *all* maximal ideals are of this form, or equivalently, if

$$\mathfrak{m} \subset k[X_1, \ldots, X_n]$$

is any maximal ideal, the quotient field is *k*.

Suppose $\mathfrak{m}$ is a maximal ideal and

$$K = k[X_1, \ldots X_n]/\mathfrak{m}$$

is a field. If the transcendence degree of K over k is d, the Noether Normalization Theorem 2.2.2 on the preceding page implies that K is integral over

$$k[y_1, \ldots, y_d]$$

where $y_1, \ldots, y_d$ are a transcendence basis. Proposition A.4.9 on page 422 implies that $k[y_1, \ldots, y_d]$ must also be a field. The only way for this to happen is for $d = 0$. So K must be an algebraic extension of k, which implies that it must equal k because k is algebraically closed.

The final statement follows from the fact that every proper ideal is contained in a maximal one, say $\mathfrak{I}(a_1, \ldots, a_n)$ so its zero-set contains at least the point $(a_1, \ldots, a_n)$. $\square$

2.2.2. The strong form. The strong form of the Nullstellensatz gives the precise correspondence between ideals and algebraic sets. It implies the weak form of the Nullstellensatz, but the two are usually considered separately.

Hilbert's strong Nullstellensatz describes which ideals in $k[X_1, \ldots, X_n]$ occur as $\mathcal{I}(P)$ when P is an algebraic set.

PROPOSITION 2.2.4. *For any subset $W \subset \mathbb{A}^n$, $\mathcal{V}(\mathcal{I}W)$ is the smallest algebraic subset of $\mathbb{A}^n$ containing W. In particular, $\mathcal{V}(\mathcal{I}W) = W$ if W is algebraic.*

REMARK. In fact, $\mathcal{V}(\mathcal{I}W)$ is the *Zariski closure* of W.

PROOF. Let $V = \mathcal{V}(\mathfrak{a})$ be an algebraic set containing W. Then $\mathfrak{a} \subset \mathcal{I}(W)$ and $\mathcal{V}(\mathfrak{a}) \supset \mathcal{V}(\mathcal{I}W)$. $\square$

THEOREM 2.2.5 (Hilbert's Nullstellensatz). *For any ideal $\mathfrak{a} \in k[X_1, \ldots, X_n]$, $\mathcal{IV}(\mathfrak{a}) = \sqrt{\mathfrak{a}}$ (see definition A.1.43 on page 359). In particular, $\mathcal{IV}(\mathfrak{a}) = \mathfrak{a}$ if $\mathfrak{a}$ is radical.*

PROOF. If f^n vanishes on $\mathcal{V}(\mathfrak{a})$, then f vanishes on it too so that $\mathcal{IV}(\mathfrak{a}) \supset \sqrt{\mathfrak{a}}$. For the reverse inclusion, we have to show that if h vanishes on $\mathcal{V}(\mathfrak{a})$, then $h^r \in \mathfrak{a}$ for some exponent r.

Suppose $\mathfrak{a} = (g_1, \ldots, g_m)$ and consider the system of $m + 1$ equations in $n + 1$ variables, $X_1, \ldots, X_m, Y$:

$$\begin{aligned} g_i(X_1, \ldots, X_n) &= 0 \\ 1 - Yh(X_1, \ldots, X_n) &= 0 \end{aligned}$$

If $(a_1, \ldots, a_n, b)$ satisfies the first m equations, then $(a_1, \ldots, a_m) \in V(\mathfrak{a})$. Consequently $h(a_1, \ldots, a_n) = 0$ and the equations are *inconsistent*.

According to the weak Nullstellensatz (see theorem 2.2.3 on the facing page), the ideal generated by the left sides of these equations generate the whole ring $k[X_1, \ldots, X_n, Y]$ and there exist $f_i \in k[X_1, \ldots, X_n, Y]$ such that

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

Now regard this equation as an identity in $k(X_1, \ldots, X_n)[Y]$ — polynomials in Y with coefficients in the field of fractions of $k[X_1, \ldots, X_n]$. After substituting h^{-1} for Y, we get

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

Clearly

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

2.4.9 for some N_i.

Let N be the largest of the N_i. On multiplying our equation by h^N, we get

$$h^N = \sum(\text{polynomial in } X_1, \ldots, X_n) \cdot g_i$$

so $h^N \in \mathfrak{a}$. $\square$

Hilbert's Nullstellensatz precisely describes the correspondence between algebra and geometry:

COROLLARY 2.2.6. *The map $\mathfrak{a} \mapsto \mathcal{V}(\mathfrak{a})$ defines a 1-1 correspondence between the set of radical ideals in $k[X_1, \ldots, X_n]$ and the set of algebraic subsets of $\mathbb{A}^n$.*

PROOF. We know that $\mathcal{IV}(\mathfrak{a}) = \mathfrak{a}$ if $\mathfrak{a}$ is a radical ideal and that $\mathcal{V}(\mathcal{I}W) = W$ if W is an algebraic set. It follows that $\mathcal{V}(*)$ and $\mathcal{I}(*)$ are inverse maps. $\square$

COROLLARY 2.2.7. *The radical of an ideal in $k[X_1, \ldots, X_n]$ is equal to the intersection of the maximal ideals containing it.*

REMARK. In general rings, the radical is the intersections of all *prime* ideals that contain it (corollary 2.2.7). The statement given here is true for algebras over algebraically closed fields.

PROOF. Let $\mathfrak{a} \subset k[X_1, \ldots X_n]$ be an ideal. Because maximal ideals are radical, every maximal ideal containing $\mathfrak{a}$ also contains $\sqrt{\mathfrak{a}}$, so

$$\sqrt{\mathfrak{a}} \subset \bigcap_{\mathfrak{m} \supset \mathfrak{a}} \mathfrak{m}$$

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

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

so

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

If $f \in \mathfrak{m}_P$ for all $P \in \mathcal{V}(\mathfrak{a})$, then f vanishes on $V(\mathfrak{a})$ so $f \in \mathcal{IV}(\mathfrak{a}) = \sqrt{\mathfrak{a}}$. It follows that

$$\sqrt{\mathfrak{a}} \supset \bigcap_{P \in \mathcal{V}(\mathfrak{a})} \mathfrak{m}_P$$

$\square$

REMARK. This result allows us to directly translate between geometry and algebra:

(1) "Since $\mathcal{V}(\mathfrak{a})$ is the *union* of the points *contained in* it, $\sqrt{\mathfrak{a}}$ is the *intersection* of the maximal ideals *containing* it."
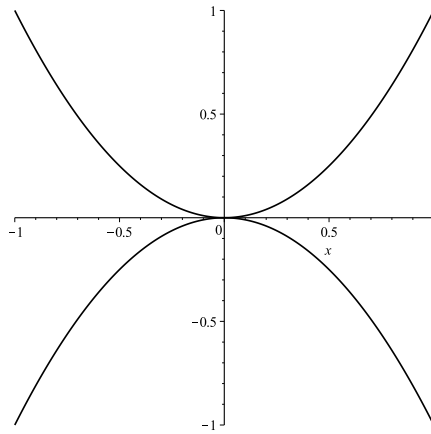
FIGURE 2.2.1.   An intersection of multiplicity 2

(2) Because $\mathcal{V}\left((0)\right) = k^n$

$$\mathcal{I}(k^n) = \mathcal{I}\mathcal{V}\left((0)\right) = \sqrt{0} = 0$$

— only the zero polynomial vanishes on all of k^n.
(3) The 1-1 correspondence is order-inverting so the maximal proper radical ideals correspond to the minimal nonempty algebraic sets.
(4) But the maximal proper radical ideals are the maximal ideals and the minimal nonempty algebraic sets are one-point sets.
(5) Let W and W' be algebraic sets. Then $W \cap W'$ is the largest algebraic subset contained in W and W' — so $\mathcal{I}(W \cap W')$ must be the smallest radical ideal containing both $\mathcal{I}(W)$ and $\mathcal{I}(W')$. It follows that

$$\mathcal{I}(W \cap W') = \sqrt{\mathcal{I}(W) + \mathcal{I}(W')}$$

EXAMPLE 2.2.8. Let $W = \mathcal{V}\left(X^2 - Y\right)$ and $W' = \mathcal{V}\left(X^2 + Y\right)$.

Then $I(W \cap W') = \sqrt{(X^2, Y)} = (X, Y)$ (assuming the characteristic of k is $\neq 2$).

So $W \cap W' = \{(0,0)\}$.

When considered at the intersection of $Y = X^2$ and $Y = -X^2$ it has multiplicity 2 (see definition 1.4.6 on page 23) — see figure 2.2.1 and definition 1.4.6.

LEMMA 2.2.9. *If V is an algebraic subset of $\mathbb{A}^n$, then*

(1) *The points of V are closed in the Zariski topology (thus V is a T_1-space).*
(2) *Every* ascending *chain of open subsets $U_1 \subset U_2 \subset \cdots$ of V eventually becomes constant — hence every* descending *chain of closed sets eventually becomes constant.*
(3) *Every open covering has a finite subcovering.*

REMARK. Topological spaces satisfying Condition 2 above are called *noetherian*. This is equivalent to:

"Every nonempty set of closed subsets of V has a minimal element."

Spaces satisfying condition 3 are called compact (although the Bourbaki group requires compact spaces to be Hausdorff, so they call such spaces *quasicompact*).

PROOF. Let $\{(a_1, \ldots, a_n)\}$ be the algebraic set defined by the ideal $(X_1 - a_1, \ldots, X_n - a_n)$.

A sequence $V_1 \supset V_2 \supset \ldots$ gives rise to a sequence of radical ideals $\mathcal{I}(V_1) \subset \mathcal{I}(V_2) \subset \cdots$ which must eventually become constant by theorem A.1.50 on page 362.

Let $V = \bigcup_{i \in I} U_i$. If $V \neq U_1$, there exists an $i_1 \in I$ such that $U_1 \subsetneq U_1 \cup U_{i_1}$. If $V \neq U_1 \cup U_{i_1}$, continue this process. By statement 2, this must stop in a finite number of steps. $\qquad\square$

DEFINITION 2.2.10. A function $f \colon \mathbb{A}^n \to \mathbb{A}^m$ is a *regular mapping* if it is of the form

$$f(X_1, \ldots X_n) = \begin{bmatrix} F_1(X_1, \ldots, X_n) \\ \vdots \\ F_m(X_1, \ldots, X_n) \end{bmatrix}$$

for $F_1, \ldots, F_m \in k[X_1, \ldots, X_n]$.

If $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ are algebraic sets and $f \colon \mathbb{A}^n \to \mathbb{A}^m$ is a regular mapping such that

$$f(V) \subset W$$

then we call $\bar{f} = f|V \colon V \to W$ a *regular mapping from V to W*.

Although the Zariski topology is very coarse — implying that it is difficult for a map to be continuous in this topology — there is an important class of continuous maps:

PROPOSITION 2.2.11. *If $f \colon V \subset \mathbb{A}^n \to W \subset \mathbb{A}^m$ is a regular map of algebraic sets, then f is continuous in the Zariski topology.*

PROOF. The map, f, is continuous if $f^{-1}(K) \subset \mathbb{A}^n$ is a closed set whenever $K \subset \mathbb{A}^m$ is closed. Let

$$f(X_1, \ldots X_n) = \begin{bmatrix} F_1(X_1, \ldots, X_n) \\ \vdots \\ F_m(X_1, \ldots, X_n) \end{bmatrix}$$

A closed set $K \subset \mathbb{A}^m$, in the Zariski topology, is defined by a finite set of equations

$$g_1(X_1, \ldots, X_m) = 0$$

$$\vdots$$

$$g_t(X_1, \ldots, X_m) = 0$$

where the g_i are polynomials. $f^{-1}(K)$ is defined by
$$g_1(F_1, \ldots, F_m)(X_1, \ldots, X_n) = 0$$
$$\vdots$$
$$g_t(F_1, \ldots, F_m)(X_1, \ldots, X_n) = 0$$
which is a closed set in $\mathbb{A}^n$. $\square$

EXERCISES.

1. Show that prime ideals are radical.

2. Show that the strong form of the Nullstellensatz implies the weak form.

3. Give an example of a map $f \colon \mathbb{A}^n \to \mathbb{A}^m$ that is continuous in the Zariski topology but not regular.

4. Suppose $f = \begin{bmatrix} f_1(X_1, \ldots, X_n) \\ \vdots \\ f_n(X_1, \ldots, X_n) \end{bmatrix} \colon \mathbb{A}^n \to \mathbb{A}^n$ is a regular map and
$$A_{i,j} = \frac{\partial f_i}{\partial X_j}$$
suppose that $z = \det A_{i,j}$ is never 0. Show that it must be a nonzero constant.

The Inverse function theorem in calculus implies that f has a smooth inverse in a neighborhood of every point.

Jacobi's Conjecture states that such an f has an *global inverse* that is a *regular map*.

The only cases that have been proved are when $k = \mathbb{C}$ and $n = 2$. It has been shown that proving it for $n = 3$ would prove it for all n when $k = \mathbb{C}$ — see [**41**] as a general reference.

5. Find the irreducible components of the algebraic set $X^2 - YZ = XZ - Z = 0$ in $\mathbb{A}^3$.

2.3. Computations in polynomial rings: Gröbner bases

2.3.1. Introduction. As the previous section makes clear, it is important to be able to do calculations in polynomial rings. There is an extensive theory of such calculations that involve computing a special basis for an ideal, called a *Gröbner basis.* Gröbner bases were discovered by Bruno Buchberger (in his thesis, [**21**]) and named after his teacher, Wolfgang Gröbner. He refined this construction in subsequent papers — see [**22, 23**].

One key idea in the theory of Gröbner bases involves imposing an *ordering* on the monomials in a polynomial ring:

DEFINITION 2.3.1. Define an ordering on the elements of $\mathbb{N}^n$ and an induced ordering on the monomials of $k[X_1, \ldots, X_n]$ defined by $\alpha = (a_1, \ldots, a_n) \succ \beta = (b_1, \ldots, b_n)$ implies that

$$\prod X_i^{a_i} \succ \prod X_i^{b_i}$$

The ordering of $\mathbb{N}^n$ must satisfy the conditions:

(1) if $\alpha \succ \beta$ and $\gamma \in \mathbb{N}^n$, then $\alpha + \gamma \succ \beta + \gamma$
(2) $\succ$ is a *well-ordering:* every set of elements of $\mathbb{N}^n$ has a *minimal* element.

For any polynomial $f \in k[X_1, \ldots, X_n]$, let $\text{LT}(f)$ denote its *leading term* in this ordering — the polynomial's highest-ordered monomial with its coefficient.

REMARK. Condition 1 implies that the corresponding ordering of monomials is preserved by multiplication by a monomial. Condition 2 implies that there are no infinite descending sequences of monomials.

DEFINITION 2.3.2. Suppose an ordering has been chosen for the monomials of $k[X_1, \ldots, X_n]$. If $\mathfrak{a} \in k[X_1, \ldots, X_n]$ is an ideal, let $\text{LT}(\mathfrak{a})$ denote the ideal generated by the leading terms of the polynomials in $\mathfrak{a}$.

(1) If $\mathfrak{a} = (f_1, \ldots, f_t)$, then $\{f_1, \ldots, f_t\}$ is a *Gröbner basis* for $\mathfrak{a}$ if

$$\text{LT}(\mathfrak{a}) = (\text{LT}(f_1), \ldots, \text{LT}(f_t))$$

(2) A Gröbner basis $\{f_1, \ldots, f_t\}$ is *minimal* if the leading coefficient of each f_i is 1 and for each i

$$\text{LT}(f_i) \notin (\text{LT}(f_1), \ldots, \text{LT}(f_{i-1}), \text{LT}(f_{i+1}), \ldots \text{LT}(f_t))$$

(3) A Gröbner basis $\{f_1, \ldots, f_t\}$ is *reduced* if the leading coefficient of each f_i is 1 and for each i and *no monomial* of f_i is contained in

$$(\text{LT}(f_1), \ldots, \text{LT}(f_{i-1}), \text{LT}(f_{i+1}), \ldots \text{LT}(f_t))$$

REMARK. There are many different types of orderings that can be used and a Gröbner basis with respect to one ordering will generally not be one with respect to another.

DEFINITION 2.3.3. The two most common orderings used are:

(1) *Lexicographic ordering.* Let $\alpha = (a_1, \ldots, a_n)$, $\beta = (b_1, \ldots, b_n) \in \mathbb{N}^n$. Then $\alpha > \beta \in \mathbb{N}^n$ if, in the vector difference $\alpha - \beta \in \mathbb{Z}^n$, the leftmost nonzero entry is positive — and we define

$$\prod X_i^{a_i} \succ \prod X_i^{b_i}$$

so

$$XY^2 \succ Y^3 Z^4$$

(2) *Graded reverse lexicographic order.* Here, monomials are first ordered by *total degree* — i.e., the sum of the exponents. Ties are resolved lexicographically (in reverse — higher lexicographic order represents a lower monomial).

REMARK. In Graded Reverse Lexicographic order, we get

$$X^4 Y^4 Z^7 \succ X^5 Y^5 Z^4$$

since the total degree is greater. As remarked above, Gröbner bases depend on the ordering, $\succ$: different orderings give different bases and even different *numbers* of basis elements.

Gröbner bases give an algorithmic procedure (detailed later) for deciding whether a polynomial is contained in an ideal and whether two ideals are equal.

To describe Buchberger's algorithm for finding a Gröbner (or standard) basis, we need something called the *division algorithm*. This is a generalization of the usual division algorithm for polynomials of a single variable:

ALGORITHM 2.3.4 (Division Algorithm). *Let $\succ$ be an ordering on the monomials of $k[X_1, \ldots, X_n]$ and let $F = \{f_1, \ldots, f_k\}$ be a set of polynomials. If $f \in k[X_1, \ldots, X_n]$ is some polynomial, the division algorithm computes polynomials $a_1, \ldots, a_s$ such that*

(2.3.1) $$f = a_1 f_1 + \cdots + a_k f_k + R$$

where $R = 0$ or no monomial in R is divisible by $\mathrm{LT}(f_i)$ for any i.

In general, we will be more interested in the remainder, *R, than the "quotients" a_i. We will use the notation*

$$f \to_F R$$

to express the fact that the remainder has a certain value ("f reduces to R"). The algorithm is:

> **function** DIVISION($f, f_1, \ldots, f_k$)
> $a_i \leftarrow 0$
> $r \leftarrow 0$
> $g \leftarrow f$
>   **while** $g \neq 0$ **do**
>   Matched $\leftarrow$ **False**
>   **for** $i = 1, \ldots, k$ **do**
>   **if** $\mathrm{LT}(f_i)|\mathrm{LT}(g)$ **then**
> $h \leftarrow \frac{\mathrm{LT}(g)}{\mathrm{LT}(f_i)}$
> $a_i \leftarrow a_i + h$
> $g \leftarrow g - f_j \cdot h$
>   Matched $\leftarrow$ **True**   ▷ $\mathrm{LT}(g)$ *was* divisible by one of the $\mathrm{LT}(f_i)$
>   **Break**   ▷ Leave the **for**-loop and continue the **While**-loop
>   **end if**
>   **end for**
>   **if not** Matched **then**   ▷ $\mathrm{LT}(g)$ was not divisible by *any* of the $\mathrm{LT}(f_i)$
> $r \leftarrow r + \mathrm{LT}(g)$ ▷ so put it into the remainder
> $g \leftarrow g - \mathrm{LT}(g)$ ▷ Subtract it from f
>   **end if**
>   **end while**
>   **return** $f = a_1 f_1 + \cdots + a_k f_k + r$
>   ▷ where the monomials of r are not divisible by the leading terms of *any* of the f_i
> **end function**

REMARK. As is usual in describing algorithms, $a \leftarrow b$ represents *assignment*, i.e. "take the value in b and *plug* it into a" (the symbol '=' merely states

that two quantities are equal). The symbol ▷ denotes a *comment* — on how the computation is proceeding.

It should be noted that:

PROPOSITION 2.3.5. *The division algorithm terminates in a finite number of steps and, in equation 2.3.1 on the previous page,*

$$(2.3.2) \qquad\qquad LT(f) \succeq LT(a_i f_i)$$

for $i = 1, \ldots, k$.

PROOF. The algorithm requires a finite number of steps since f contains a finite number of monomials and g — initially equal to f — loses one in each iteration of the **While**-loop. The final statement follows from how the a_i are constructed — which implies that

$$LT(a_i f_i) \neq LT(a_j f_j)$$

for $i \neq j$, so no cancellation can occur among the leading terms in equation 2.3.1 on the preceding page. □

EXAMPLE 2.3.6. Let $f = X^2 Y + XY^2 + Y^2$, and let $f_1 = XY - 1$ and $f_2 = Y^2 - 1$.

Assume lexicographic ordering with $X \succ Y$. Then $LT(f_1) \,\big|\, LT(f)$ and we get

$$
\begin{aligned}
h &\leftarrow X\\
a_1 &\leftarrow X\\
g &\leftarrow g - X \cdot (XY - 1)\\
  &= XY^2 + Y^2 + X
\end{aligned}
$$

In the second iteration of the **While**-loop, $LT(f_1) \,\big|\, LT(g)$ and

$$
\begin{aligned}
h &\leftarrow Y\\
a_1 &\leftarrow a_1 + Y\\
  &= X + Y\\
g &\leftarrow g - Y \cdot (XY - 1)\\
  &= Y^2 + X + Y
\end{aligned}
$$

In the third iteration of the **While**-loop, we have $LT(f_1) \nmid LT(g)$ *and* $LT(f_2) \nmid LT(g)$ so

$$
\begin{aligned}
r &\leftarrow X\\
g &\leftarrow g - X\\
  &= Y^2 + Y
\end{aligned}
$$

In the fourth iteration of the **While**-loop, we have $LT(f_1) \nmid LT(g)$ but $LT(f_2) \,\big|\, LT(g)$ so

$$
\begin{aligned}
h &\leftarrow 1\\
a_2 &\leftarrow 1\\
g &\leftarrow g - 1 \cdot (Y^2 - 1)\\
  &= Y + 1
\end{aligned}
$$

Since neither Y nor 1 are divisible by the leading terms of the f_i they are thrown into the remainder and we get

$$f = (X + Y) \cdot f_1 + 1 \cdot f_2 + X + Y + 1$$

Note that our remainder depends on the *order* of the polynomials. If we set $f_1 = Y^2 - 1$ and $f_2 = XY - 1$ we get

$$f = (X + 1) \cdot f_1 + X \cdot f_2 + 2X + 1$$

It turns out that the remainder can vanish with one ordering and not another!

With the Division Algorithm in hand, we can discuss some of the more important properties of Gröbner bases:

PROPOSITION 2.3.7 (Division Property). *Let $\succ$ be an ordering of monomials in $k[X_1, \ldots, X_n]$ and let $\mathfrak{a} = (g_1, \ldots, g_k) \subset k[X_1, \ldots, X_n]$ be an ideal with $G = \{g_1, \ldots, g_k\}$ a Gröbner basis. If $f \in k[X_1, \ldots, X_n]$, then $f \in \mathfrak{a}$ if and only if*

$$f \to_G 0$$

PROOF. If $f \to_G 0$, then $f \in \mathfrak{a}$. Conversely, suppose $f \in \mathfrak{a}$ and $f \to_G R$. If $R \neq 0$ then

$$R = f - \sum_{i=1}^{t} a_i g_i$$

so that $R \in \mathfrak{a}$ and $\mathrm{LT}(R) \in \mathrm{LT}(\mathfrak{a})$ (since G is a Gröbner basis). This contradicts the fact that the leading term of R is not divisible by the leading terms of the g_i. $\qquad \square$

This immediately implies that

COROLLARY 2.3.8. *If $\mathfrak{a} \subset k[X_1, \ldots, X_n]$ is an ideal and B is a minimal Gröbner basis then $\mathfrak{a} = (1)$ if and only if $B = \{1\}$.*

PROOF. If $1 \in \mathfrak{a}$, then

$$1 \to_B 0$$

which can only happen if $1 \in B$. Since B is minimal, $B = \{1\}$. $\qquad \square$

2.3.2. Buchberger's Algorithm. We begin by proving a property that Gröbner bases have.

DEFINITION 2.3.9. *If $\succeq$ is some ordering on the monomials of $k[X_1, \ldots, X_n]$ and $\mathfrak{a} = (f_1, \cdots, f_t)$ is an ideal, let*

(2.3.3)
$$s_{i,j} = \frac{\mathrm{LT}(f_i)}{\gcd(\mathrm{LT}(f_i), \mathrm{LT}(f_j))}$$

and define the *S-polynomial*

$$S_{i,j} = s_{j,i} \cdot f_i - s_{i,j} \cdot f_j$$

REMARK. Note that $\mathrm{LT}(s_{j,i} \cdot f_i) = \mathrm{LT}(s_{i,j} \cdot f_j)$ so that they cancel out in $S_{i,j}$.

Buchberger's Theorem states that the *S*-polynomials give a criterion for a basis being Gröbner. It quickly leads to an *algorithm* for computing Gröbner bases.

THEOREM 2.3.10. *Let $F = \{f_1, \ldots, f_t\} \in k[X_1, \ldots, X_n]$ be a set of polynomials and let $\succ$ be an ordering on the monomials of $k[X_1, \ldots, X_n]$. Then F is a Gröbner basis of the ideal $\mathfrak{a} = (f_1, \ldots, f_t)$ if and only if*

$$S_{i,j} \to_F 0$$

for every S-polynomial one can form from the polynomials in F.

PROOF. If F is a Gröbner basis, then the division property (proposition 2.3.7 on the previous page) implies

$$S_{i,j} \to_F 0$$

since $S_{i,j} \in \mathfrak{a}$.

On the other hand, suppose all S-polynomials reduce to 0. Then there exist expressions

(2.3.4)
$$S_{i,j} = \sum_{\ell=1}^{t} a_{i,j}^{\ell} f_{\ell}$$

for all $1 \leq i < j \leq t$ and $1 \leq \ell \leq t$ such that

$$\mathrm{LT}(S_{i,j}) \quad \succeq \quad \mathrm{LT}(a_{i,j}^{\ell} f_i)$$

(2.3.5)
$$\mathrm{LT}(s_{i,j} f_j) \quad \succ \quad \mathrm{LT}(a_{i,j}^{\ell} f_i)$$

Suppose that F is *not* a Gröbner basis — i.e. $\mathrm{LT}(\mathfrak{a}) \neq (\mathrm{LT}(f_1), \ldots, \mathrm{LT}(f_t))$. Then there exists an element $f \in \mathfrak{a}$ with

(2.3.6)
$$f = \sum_{i=1}^{t} b_i f_i$$

such that $\mathrm{LT}(f_i) \nmid \mathrm{LT}(f)$ for all $i = 1, \ldots, t$. The only way this can happen is if the leading terms of *two* of the terms in equation 2.3.6 *cancel.* Suppose m is the *highest* (in the ordering) monomial of $\{\mathrm{LT}(b_i f_i)\}$ for $i = 1, \ldots, t$, suppose f has been chosen to make m minimal, and so that m occurs a minimal number of times.

Without loss of generality, suppose that $\bar{b}_1 \mathrm{LT}(f_1) = \mathrm{LT}(b_1 f_1)$ and $\bar{b}_2 \mathrm{LT}(f_2) = \mathrm{LT}(b_2 f_2)$ are equal to m, up to multiplication by an element of k. If we divide both of these by $\gcd(\mathrm{LT}(f_1), \mathrm{LT}(f_2))$, we get

$$k_1 \bar{b}_1 s_{1,2} = \bar{b}_2 \cdot s_{2,1}$$

where $s_{i,j}$ is as in equation 2.3.3 on the previous page. Since the $s_{i,j}$ have no common factors, we conclude that $s_{2,1} | \bar{b}_1$ or $\bar{b}_1 = c \cdot s_{2,1}$, for some monomial c, so $\bar{b}_1 \mathrm{LT}(f_1) = c \cdot s_{2,1} \cdot \mathrm{LT}(f_1)$. Now form the quantity

$$f' = f - c \left(S_{1,2} - \sum_{\ell=1}^{t} a_{1,2}^{\ell} f_{\ell} \right)$$

(where $S_{1,2}$ is as in definition 2.3.9 on the preceding page).

Our hypothesis (equation 2.3.4) implies that $f' = f$. On the other hand, the term $-c \cdot s_{2,1} \cdot f_1$ in $-cS_{1,2}$ cancels out $\bar{b}_1 \mathrm{LT}(f_1)$ and the term $+c \cdot s_{1,2} \cdot f_2$ combines with the term $b_2 f_2$ so that the number of occurrences of the monomial m decreases by at least 1. Equation 2.3.5 on page 50 implies that the terms $\{a_{1,2}^{\ell} f_{\ell}\}$ cannot affect this outcome, so we have a contradiction to the fact that

m occurred a minimal number of times in f. We conclude that F must have been a Gröbner basis. □

This result immediately leads to an algorithm for computing a Gröbner basis:

ALGORITHM 2.3.11 (Buchberger's Algorithm). *Given a set of polynomials* $F = \{f_1, \dots, f_t\} \in k[X_1, \dots, X_n]$,

(1) *for each pair (i, j) with $1 \le i < j \le t$, compute $S_{i,j}$ as in definition 2.3.9 on page 49,*
(2) *compute $S_{i,j} \to_F h_{i,j}$, using the Division Algorithm (2.3.4 on page 47),*
(3) *if $h_{i,j} \ne 0$, set*

$$F = F \cup \{h_{i,j}\}$$

and return to step 1.

The algorithm terminates when all of the $\{h_{i,j}\}$ found in step 3 are 0.

REMARK. The Hilbert Basis theorem (A.1.50 on page 362) implies that this process will terminate in a finite number of steps (since we are appending generators of $\mathrm{LT}(\mathfrak{a})$).

To get a *minimal* Gröbner basis, simply throw away unnecessary elements. To get a *reduced* basis, apply the Division Algorithm to each member of the output of this algorithm with respect to the other members.

Unfortunately, Buchberger's algorithm can have *exponential* time-complexity — for graded-reverse lexicographic ordering — and *doubly-exponential* (e^{e^n}) complexity for lexicographic ordering (see [**108**]).

In practice it seems to have a reasonable running time. In special cases, we have:

(1) For a system of *linear* polynomials, Buchberger's Algorithm become *Gaussian Elimination* for putting a matrix in upper triangular form.
(2) For polynomials over a single variable, it becomes *Euclid's algorithm* for finding the greatest common divisor for two polynomials (see A.1.16 on page 347).

Here is an example:

EXAMPLE 2.3.12. Let $f_1 = XY + Y^2$ and $f_2 = X^2$ in $k[X, Y]$ and we compute a Gröbner basis using lexicographical ordering with

$$X \succ Y$$

We have $\mathrm{LT}(f_1) = XY$ and $\mathrm{LT}(f_2) = X^2$. Neither is a multiple of the other and their greatest common divisor is X. Our first S-polynomial is

$$S_{1,2} = \frac{\mathrm{LT}(f_2)}{X} f_1 - \frac{\mathrm{LT}(f_1)}{X} f_2 = XY^2$$

The remainder after applying the Division Algorithm is $-Y^3$ so we set $f_3 = Y^3$. We compute

$$S_{1,3} = \frac{\mathrm{LT}(f_3)}{Y} f_1 - \frac{\mathrm{LT}(f_1)}{Y} f_3 = Y^4$$

$$S_{2,3} = \frac{\mathrm{LT}(f_3)}{1} f_2 - \frac{\mathrm{LT}(f_2)}{1} f_3 = 0$$

Since both of these are in the ideal generated by $\{f_1, f_2, f_3\}$, we are done.

Gröbner bases have an interesting history. In 1899, Gordon gave a new proof of the Hilbert Basis theorem[2] (theorem A.1.50 on page 362) that demonstrated the existence of a finite Gröbner basis (with lexicographic ordering) but gave no algorithm for computing it. See [**58**].

In 1920, Janet (see [**83**]) gave an algorithm for computing "involutive bases" of linear systems of partial differential equations, that can be translated into Buchberger's algorithm in a certain case. Given a system of differential equations that are linear combinations of products of partial derivatives of $\psi(x_1, \ldots, x_n)$ (with constant coefficients), one can substitute

$$\psi = e^{\sum \alpha_i x_i}$$

and get systems of *polynomials* in the α_i whose solution leads to solutions of the differential equations.

In 1950, Gröbner published a paper ([**68**]) that explored an algorithm for computing Gröbner bases, but could not prove that it ever terminated. One of Buchberger's signal contributions were the introduction of S-polynomials and theorem 2.3.10 on page 50.

Teo Mora (see [**117, 118**]) extended much of the theory of Gröbner bases to some non-polynomial rings, including local rings and power series rings.

At this point, we can prove another interesting property of Gröbner bases, when computed in a certain way:

PROPOSITION 2.3.13 (Elimination Property). *Suppose* $\{g_1, \ldots, g_j\}$ *is a Gröbner basis for the ideal* $\mathfrak{a} \in k[X_1, \ldots, X_n]$, *computed using lexicographic ordering with*

$$X_1 \succ X_2 \succ \cdots \succ X_n$$

If $1 \leq t \leq n$, *then*

$$\mathfrak{a} \cap k[X_t, \ldots, X_n]$$

has a Gröbner basis that is

$$\{g_1, \ldots, g_j\} \cap k[X_t, \ldots, X_n]$$

REMARK. It follows that we can use Gröbner bases to save ourselves the trouble of doing multiple computations of the *resultant*.

PROOF. Suppose $f \in \mathfrak{a} \cap k[X_t, \ldots, X_n]$ and its expansion using the Division Algorithm (2.3.4 on page 47) is

$$f = \sum q_i \cdot g_i$$

with $\mathrm{LT}(q_i \cdot g_i) \preceq \mathrm{LT}(f)$ for all i (see equation 2.3.2). Lexicographic ordering implies that, if $X_1, \ldots, X_{t-1}$ occur *anywhere* in $q_i \cdot g_i$ then these variables will be in the *leading term* of $q_i \cdot g_i$ and $\mathrm{LT}(q_i \cdot g_1) \succ \mathrm{LT}(f)$ — a *contradiction*. It follows that, for all i such that g_i contains variables $X_1, \ldots, X_{t-1}$, the corresponding $q_i = 0$. Since f is a linear combination of polynomials $\{g_1, \ldots, g_j\} \cap k[X_t, \ldots, X_n]$, they *generate* $\mathfrak{a} \cap k[X_t, \ldots, X_n]$.

Since

$$S_{i,i'} \rightarrow_{G'} 0$$

---

[2]He felt that Hilbert's proof was too abstract and gave a *constructive* proof.

whenever $g_i, g_{i'} \in \{g_1, \ldots, g_j\} \cap k[X_t, \ldots, X_n]$, theorem 2.3.10 on page 50 implies that $G' = \{g_1, \ldots, g_j\} \cap k[X_t, \ldots, X_n]$ is a *Gröbner* basis. $\square$

We already know how to test *membership* of a polynomial in an ideal via the Division Algorithm and proposition 2.3.7 on page 49. This algorithm also tells us when one ideal is contained in another since $(f_1, \ldots, f_j) \subseteq (g_1, \ldots, g_\ell)$ if and only if $f_i \in (g_1, \ldots, g_\ell)$ for $i = 1, \ldots, j$.

We can use Gröbner bases to compute *intersections* of ideals:

PROPOSITION 2.3.14 (Intersections of ideals). *Let* $\mathfrak{a} = (f_1, \ldots, f_j), \mathfrak{b} = (g_1, \ldots, g_\ell)$ *be ideals in* $k[X_1, \ldots, X_n]$. *If we introduce a new variable, T, and compute the Gröbner basis of*

$$(Tf_1, \ldots, Tf_j, (1-T)g_1, \ldots (1-T)g_\ell)$$

using lexicographic ordering and ordering T higher than the other variables, the Gröbner basis elements that do not contain T will be a Gröbner basis for $\mathfrak{a} \cap \mathfrak{b}$.

REMARK. If $f, g \in k[X_1, \ldots, X_n]$, this allows us to compute the least common multiple, z, of f and g since

$$(z) = (f) \cap (g)$$

and the greatest common divisor — even if we don't know how to factor polynomials! If $n > 1$, $k[X_1, \ldots, X_n]$ is *not* a Euclidean domain.

PROOF. Let $\mathfrak{I} \in k[T, X_1, \ldots, X_n]$ denote the big ideal defined above. We claim that

$$\mathfrak{I} \cap k[X_1, \ldots, X_n] = \mathfrak{a} \cap \mathfrak{b}$$

Suppose $f \in \mathfrak{I} \cap k[X_1, \ldots, X_n]$. Then

$$
\begin{aligned}
f &= \sum a_i T f_i + \sum b_j (1-T) g_j \\
&= T\left(\sum a_i f_i - \sum b_j g_j\right) + \sum b_j g_j
\end{aligned}
$$

so

$$\sum a_i f_i - \sum b_j g_j = 0$$

It follows that $f = \sum b_j g_j$ and that $\sum a_i f_i = \sum b_j g_j$ so that $f \in \mathfrak{a} \cap \mathfrak{b}$. The conclusion now follows from the Elimination Property. $\square$

We get a criterion for a polynomial to be in the radical of an ideal:

PROPOSITION 2.3.15. *If* $\mathfrak{a} \in k[X_1, \ldots, X_n]$ *is an ideal and* $f \in k[X_1, \ldots, X_n]$ *is a polynomial*

$$f \in \sqrt{\mathfrak{a}}$$

if and only if a Gröbner basis for

$$\mathfrak{a} + (1 - T \cdot f) \subset k[T, X_1, \ldots, X_n]$$

where T is a new indeterminate, contains 1.

PROOF. The proof of Hillbert's Nullstellensatz (theorem 2.2.5 on page 41) implies that $f \in \sqrt{\mathfrak{a}}$ if and only if $1 \in \mathfrak{a} + (1 - T \cdot f)$. The conclusion follows from corollary 2.3.8 on page 49. $\square$

EXAMPLE 2.3.16. Let $\mathfrak{a} = (Y - X^2, Z^2 - 2XYZ + Y^3) \subset k[X, Y, Z]$. If $f = Z^2 - Y^3$ then the polynomials given are a Gröbner basis for $\mathfrak{a}$ in the ordering by graded reverse lexicographic (in Maple, specified as `tdeg(X,Y,Z)`).

$$f \to_{\mathfrak{a}} 2Z^2 - 2XYZ$$

so $f \notin \mathfrak{a}$. If we compute a Gröbner basis (with respect to *any* ordering) for

$$\{Y - X^2, Z^2 - 2XYZ + Y^3, 1 - T \cdot f\}$$

we get (1). It follows that $f \in \sqrt{\mathfrak{a}}$. In fact, we get

$$f^2 \to_{\mathfrak{a}} 0$$

so $f^2 \in \mathfrak{a}$.

2.3.3. Mathematical software. All of the more commonly used systems of mathematical software are able to compute Gröbner bases and implement the Division Algorithm. Among commercial systems, Maple and Mathematica have very nice user-interfaces. Free software that can do this includes Maxima and Macaulay 2, and CoCoa . See [**39**] for much more information.

To use Macaulay 2, start it (in Unix-type systems) by typing `M2` (the command is actually capitalized). The default output format is rather awful, so you should change it by typing

`compactMatrixForm = false`

Now define a polynomial ring over $\mathbb{Q}$

`R = QQ[a..f,MonomialOrder=>Lex]`

Note that ordering is also specified here. Now define an ideal:

```
i3 : I = ideal(a*b*c-d*e*f,a*c*e-b*d*f,
a*d*f-b*c*e)
o3 = ideal (a*b*c - d*e*f, a*c*e - b*d*f,
a*d*f - b*c*e)
o3 : Ideal of R
```

To get a Gröbner basis, type:

`gens gb I`

You need to make the window wide enough to contain the entire output expression. Subscripted variables can be defined via

`x_2=3`

The ring above could have been defined via

`R = QQ[x_1..x_6,MonomialOrder=>Lex]`

In Maple the procedure is somewhat different: First, load the library via `'with(Groebner);'`. The library `PolynomialIdeals` is also very useful. Enter an ideal by simply enclosing its generators in square brackets. The command `Basis` computes a Gröbner basis:

```
Basis([a*b*c-d*e*f, a*c*e-b*d*f,
a*d*f-b*c*e], plex(a, b, c, d, e, f))
```

The output is nicely formatted.

The expression `plex` implies lexicographic ordering, and you must explicitly give the order of variables. For instance `plex(a, b, c, d, e, f)`

means

$$a \succ b \succ c \succ d \succ e \succ f$$

Maple also supports graded lexicographic ordering with the command `grlex(a,b,c,d,e,f)` or graded reverse lexicographic order via `tdeg(a,b,c,d,e,f)`.

To reduce a polynomial using the Division Algorithm (2.3.4 on page 47), the Maple command is

`NormalForm(list_polys,basis,monomial_order)` where the basis need not be Gröbner. It returns a list of remainders of the polynomials in the list.

Maxima has a package that computes Gröbner bases using lexicographic ordering (at present, no other ordering is available). To load it, type `load(grobner)`. The main commands are `poly_grobner(poly-list,var-list)`, and `poly_reduced_grobner(poly-list,var-list)`. For example:

`poly_grobner([x^2+y^2,x^3-y^4],[x,y]);` returns

$$(x^2 + y^2, x^3 - y^4, x^4 + xy^2, y^6 + y^4)$$

— the Gröbner basis with lexicographic order: $x \succ y$.

Another very powerful and free system is called Sage (it aims to "take over the world" of computer algebra systems!). It is available for all common computer systems and can even be used online (i.e., without installing it on your computer) at `http://www.sagemath.org/`.

Here's a small example:

The command:

`R.<a,b,c,d> = PolynomialRing(QQ, 4, order='lex')`

defines a polynomial ring, R, over $\mathbb{Q}$ with 4 indeterminates: a, b, c, and d. The statement `order='lex'` defines lexicographic ordering on monomials. The command:

```
I = ideal(a+b+c+d, a*b+a*d+b*c+c*d,
    a*b*c+a*b*d+a*c*d+b*c*d, a*b*c*d-1);
```

defines an ideal in R. Now the command:

`B = I.groebner_basis()`

computes the Gröbner basis with respect to the given ordering. Just typing the name B prints out the basis:

```
[a + b + c + d,
b^2 + 2*b*d + d^2, b*c - b*d + c^2*d^4 + c*d - 2*d^2,
b*d^4 - b + d^5 - d, c^3*d^2 + c^2*d^3 - c - d,
c^2*d^6 - c^2*d^2 - d^4 + 1]
```

We begin with some of the intersections in chapter 1 on page 1:

EXAMPLE 2.3.17. Consider the intersection in figure 1.4.5 on page 23 defined by the equations

$$
\begin{aligned}
5X^2 + 6XY + 5Y^2 + 6Y - 5 &= 0 \\
X^2 + Y^2 - 1 &= 0
\end{aligned}
$$

The intersection is the algebraic set defined by the radical of the ideal

$$\mathfrak{N} = (5X^2 + 6XY + 5Y^2 + 6Y - 5, X^2 + Y^2 - 1) \subset k[X, Y]$$

If we compute a Gröbner basis for this using lexicographic ordering with $X \succ Y$, we get

(2.3.7) $\mathfrak{N} = (Y^3, XY + Y, X^2 + Y^2 - 1)$

(using the Maple command `Basis(N, plex(X, Y))`). Since $Y^3 \in \mathfrak{N}$, we must have $Y \in \sqrt{\mathfrak{N}}$ and

$$\begin{aligned} \sqrt{\mathfrak{N}} &= (Y, XY + Y, X^2 + Y^2 - 1) \\ &= (Y, X^2 - 1) \end{aligned}$$

so the coordinate ring of the intersection is

$$\frac{k[X, Y]}{(X^2 - 1, Y)} = k_{(Y, X-1)} \oplus k_{(Y, X+1)}$$

a vector-space of two dimensions over k, implying two intersection-points. In fact this even tells us *where* they are: $Y = 0$ and $X = \pm 1$.

It is interesting to compute the quotient

$$Q = \frac{k[X, Y]}{\mathfrak{N}}$$

without taking the radical. If the images of X and Y in Q are x and y, respectively, the Gröbner basis in equation 2.3.7 on page 56 implies

$$\begin{aligned} xy &= -y & &\text{because } XY + Y \in \mathfrak{N} \\ x^2 &= 1 - y^2 & &\text{because } X^2 + Y^2 - 1 \in \mathfrak{N} \\ y^3 &= 0 & &\text{because } Y^3 \in \mathfrak{N} \\ x^3 &= x - xy^2 & &\text{because } x^2 = 1 - y^2 \\ &= x + y^2 & &\text{because } xy = -y \\ &= 1 + x - x^2 & &\text{because } y^2 = 1 - x^2 \end{aligned}$$

so Q is a k-vector-space with basis $\{1, x, y, y^2\}$. This suggests that Q counts the intersections with *multiplicities,* giving 4 in all. This reasoning will be the basis for our more sophisticated definition of intersection-multiplicity in sections 3.3.4 on page 139 and 5.8 on page 274.

Here's an application of Gröbner bases to the robotics problem in section 1.6 on page 28:

EXAMPLE 2.3.18. We set the lengths of the robot arms to 1. The system of equations 1.6.3 on page 30 gives rise to the ideal

$$r = (a_1 a_2 - b_1 b_2 + a_1 - x, a_2 b_1 + a_1 b_2 + b_1 - y, a_1^2 + b_1^2 - 1, a_2^2 + b_2^2 - 1)$$

in $\mathbb{C}[a_1, a_2, b_1, b_2]$. If we set $x = 1$ and $y = 1/2$, the Gröbner basis of r (using the command 'Basis(r,plex(a_1,b_1,a_2,b_2))' in Maple) is

$$(-55 + 64 b_2^2, 8 a_2 + 3, 16 b_2 - 5 + 20 b_1, -5 - 4 b_2 + 10 a_1)$$
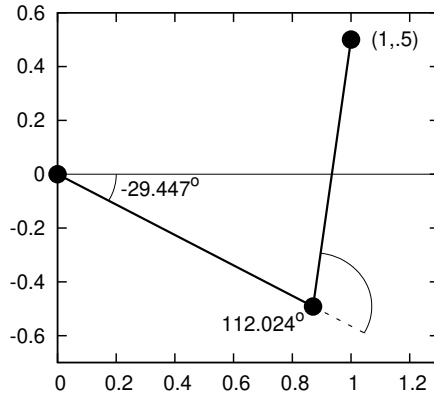
FIGURE 2.3.1.   Reaching a point

from which we deduce that $a_2 = -3/8$ and b_2 can be either $+\sqrt{55}/8$ in which case

$$a_1 = 1/2 + \sqrt{55}/20$$
$$b_1 = 1/4 - \sqrt{55}/10$$

or $-\sqrt{55}/8$ in which case

$$a_1 = 1/2 - \sqrt{55}/20$$
$$b_1 = 1/4 + \sqrt{55}/10$$

It follows that there are precisely *two* settings that allow the robot arm in figure 1.6.1 on page 28 to reach the point $(1, 1/2)$. It is straightforward to compute the angles involved in figure 1.6.1 on page 28: in the first case,

$$\theta = -29.44710523°$$
$$\phi = 112.024312°$$

as in figure 2.3.1 and in the second

$$\theta = 82.57720759°$$
$$\phi = -112.024312°$$

Another question we might ask is:

> *For what values of x are points on the line $y = 1 - 2x$ reachable?*

In this case, we start with the ideal

$$r = (a_1 a_2 - b_1 b_2 + a_1 - x, a_2 b_1 + a_1 b_2 + b_1 + 2x - 1,$$
$$a_1{}^2 + b_1{}^2 - 1, a_2{}^2 + b_2{}^2 - 1)$$

and   get   the   Gröbner   basis   (using   the   Maple   command
'Basis(r,plex(a_1,b_1,a_2,b_2,x))'

$$(-3 + 8x + 6x^2 + 4b_2^2 - 40x^3 + 25x^4, -5x^2 + 1 + 4x + 2a_2,$$
$$-1 + 6x - 13x^2 + 2xb_2 + 10x^3 + 2b_1 - 8xb_1 + 10x^2b_1,$$
$$3x - 2b_2 + 4x^2 + 4xb_2 - 5x^3 + 4b_1b_2,$$
$$-1 + 4x - b_2 - 5x^2 + 2b_1 - 5xb_1 + a_1)$$

The first monomial

$$-3 + 8x + 6x^2 + 4b_2^2 - 40x^3 + 25x^4$$

is significant: When all variables are real, $4b_2^2 \geq 0$, which requires

$$-3 + 8x + 6x^2 - 40x^3 + 25x^4 \leq 0$$

— since the basis elements are assumed to be set to 0. This only happens if

$$x \in \left[\frac{2 - \sqrt{19}}{5}, \frac{2 + \sqrt{19}}{5}\right]$$

— so those are the only points on the line $y = 1 - 2x$ that the robot-arm can
reach.

We can also analyze the Puma-type robot-arm in figure 1.6.2 on page 30:

EXAMPLE 2.3.19. If we set $\ell_1 = \ell_2 = 1$, equation 1.6.4 on page 32 implies
that the endpoint of the robot-arm are solutions to the system

$$\begin{aligned}
a_5a_4a_3 - a_5b_4b_3 + a_5a_4 - x &= 0 \\
b_5a_4a_3 - b_5b_4b_3 + b_5a_4 - y &= 0 \\
b_4a_3 + a_4b_3 + b_4 - z &= 0 \\
a_3^2 + b_3^2 - 1 &= 0 \\
a_4^2 + b_4^2 - 1 &= 0 \\
a_5^2 + b_5^2 - 1 &= 0
\end{aligned}$$

(2.3.8)

If we want to know which points it can reach with the hand pointing in the
direction

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

use equation 1.6.5 on page 32 to get

$$\begin{aligned}
(a_5a_4a_3 - a_5b_4b_3)a_2 + (-a_5a_4b_3 - a_5b_4a_3)b_2 - 1/\sqrt{3} &= 0 \\
(b_5a_4a_3 - b_5b_4b_3)a_2 + (-b_5a_4b_3 - b_5b_4a_3)b_2 - 1/\sqrt{3} &= 0 \\
(b_4a_3 + a_4b_3)a_2 + (a_4a_3 - b_4b_3)b_2 - 1/\sqrt{3} &= 0 \\
a_2^2 + b_2^2 - 1 &= 0
\end{aligned}$$

(2.3.9)

We regard these terms (in equations 2.3.8 and 2.3.9 as generators of an ideal,
$\mathfrak{P}$. The variety $\mathcal{V}(\mathfrak{P}) \subset \mathbb{R}^{10}$ is called the *variety of the movement problem*. Its
(real-valued) points correspond to possible configurations of the robot-arm.

To understand $\mathcal{V}(\mathfrak{P})$, we compute a Gröbner basis of $\mathfrak{P}$ with lexicographic ordering — giving the lowest weight to x, y, z — to get

$$(2.3.10) \quad \mathfrak{P} = (4\,y^2x^2 - 4\,z^2 + z^4 + 2\,z^2x^2 + 2\,y^2z^2,$$
$$-1 + 2\,b_5{}^2, -b_5 + a_5, 2\,zb_4 - z^2 - 2\,yx,$$
$$-4\,z + 4\,yb_4x + z^3 - 2\,xzy + 2\,y^2z + 2\,x^2z,$$
$$-2 - 2\,yx + 2\,b_4{}^2 + y^2 + x^2,$$
$$a_4 - b_5y + b_5x,$$
$$-b_5yz + b_5xz + b_3 + 2\,b_5yb_4,$$
$$2 + 2\,a_3 - 2\,y^2 - z^2,$$
$$2\,b_5z\sqrt{3} - \sqrt{3}b_5y - \sqrt{3}b_5x - 2\,b_4\sqrt{3}b_5 + 3\,b_2,$$
$$3\,a_2 - y\sqrt{3} - x\sqrt{3} - z\sqrt{3} + b_4\sqrt{3})$$

It follows that a point (x, y, z) is reachable (with the hand oriented as stated) *only* if it lies on the surface

$$4\,y^2x^2 - 4\,z^2 + z^4 + 2\,z^2x^2 + 2\,y^2z^2 = 0$$

Solving for z^2 gives

$$(2.3.11) \qquad z^2 = 2 - y^2 - x^2 \pm \sqrt{(2 - (x - y)^2)(2 - (x + y)^2)}$$

The fourth expression from the top in equation 2.3.10 is

$$\begin{aligned} -2 - 2\,yx + 2\,b_4{}^2 + y^2 + x^2 &= 0 \\ 2b_4^2 &= 2 - 2xy - x^2 - y^2 \end{aligned}$$

which implies that

$$2 - (x - y)^2 \geq 0$$

and gives the additional constraint on (x, y, z):

$$(x - y)^2 \leq 2$$

It follows that $2 - (x - y)^2 \geq 0$ so that the square root in equation 2.3.11 is only well-defined if $2 - (x + y)^2 \geq 0$ and we get an *additional* constraint on x and y.

The requirement that $z^2 \geq 0$ implies that the only case worth considering is

$$z^2 = 2 - y^2 - x^2 + \sqrt{(2 - (x - y)^2)(2 - (x + y)^2)}$$

and figure 2.3.2 on the following page shows the set of points that are reachable.

The Elimination Property of Gröbner bases is useful for algebraic computations:
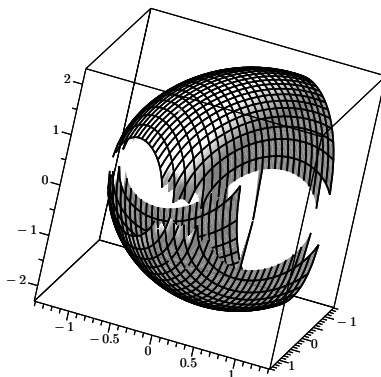
FIGURE 2.3.2.   Points reachable by the second robot arm

EXAMPLE 2.3.20. Suppose we have a field $F = \mathbb{Q}[\sqrt{2}, \sqrt{3}]$ (see lemma A.2.10 on page 389) and want to know the minimal polynomial (see A.2.9 on page 388) of $\alpha = \sqrt{2} + \sqrt{3}$. We regard F as a quotient

$$F = \mathbb{Q}[X, Y]/(X^2 - 2, Y^2 - 3)$$

Now form the ideal $\mathfrak{s} = (X^2 - 2, Y^2 - 3, A - X - Y) \subset \mathbb{Q}[X, Y, A]$ and *eliminate* X and Y by taking a Gröbner basis using lexicographic ordering with

$$X \succ Y \succ A$$

The result is

$$(1 - 10A^2 + A^4, -11A + A^3 + 2Y, 9A - A^3 + 2X)$$

and we claim that the minimal polynomial of α is

$$\alpha^4 - 10\alpha^2 + 1 = 0$$

It is a polynomial that α satisfies and generates $\mathfrak{s} \cap k[A]$, which is a principal ideal domain (see proposition A.2.4 on page 387), so any other polynomial that α satisfies must be a multiple of it. Since the degree of this minimal polynomial is the same as $[F:\mathbb{Q}] = 4$ it follows that $F = \mathbb{Q}[\alpha]$. Indeed, the second and third terms of the Gröbner basis imply that

$$\sqrt{2} = \frac{\alpha^3 - 9\alpha}{2}$$

$$\sqrt{3} = -\frac{\alpha^3 - 11\alpha}{2}$$

so $\mathbb{Q}[\alpha] = \mathbb{Q}[\sqrt{2}, \sqrt{3}]$. This is an example the Primitive Element Theorem (A.2.21 on page 394).

Here's a second example:

$F = \mathbb{Q}[2^{1/3}]$ and we want the minimal polynomial of

$$\alpha = \frac{1 + 2^{1/3}}{1 - 2^{1/3}}$$

We create the ideal $\mathfrak{b} = (X^3 - 2, (1 - X)A - 1 - X)$ (the second term is a polynomial that α satisfies) and take a Gröbner basis to get

$$\mathfrak{b} = (3 + 3A + 9A^2 + A^3, 1 - 8A - A^2 + 4X)$$

so the minimal polynomial of α is

$$\alpha^3 + 9\alpha^2 + 3\alpha + 3 = 0$$

The second term of the Gröbner basis implies that

$$2^{1/3} = \frac{\alpha^2 + 8\alpha - 1}{4}$$

so that $\mathbb{Q}[2^{1/3}] = \mathbb{Q}[\alpha]$.

EXERCISES.

1. Compare this to the solution of exercise 7 on page 39. Let V be the union of the 3 coordinate axes in $\mathbb{A}^3$. Find

$$\mathcal{I}(V) = (Y, Z) \cap (X, Y) \cap (X, Y)$$

2. The point $(1/2, 1/2, 1 + \sqrt{2}/2)$ lies on the "reachability surface" in example 2.3.19 on page 58 that can be reached by the robot arm with its hand pointed in the direction

$$\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Find the angles $\phi_1, \theta_1, \theta_2, \theta_3$ that accomplish this.

3. Find the reachability surface of the robot arm in example 2.3.19 on page 58 when we require the "hand" to point in the direction

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

4. Find the least common multiple of the polynomials

$$-X^3 - 2YX^2 - XY^2 + 2X$$

and

$$4 - 4X^2 - 4Y^2 + X^4 - 2Y^2X^2 + Y^4$$

in $k[X, Y]$.

5. Consider the ideal $\mathfrak{a} = (Y^3, X - Y)$.
Is $X + Y \in \mathfrak{a}$?

6. If $\mathfrak{a} = (Y^3, X - Y)$, is $X + Y \in \sqrt{\mathfrak{a}}$? If so, what power of it is in $\mathfrak{a}$?

7. If $\mathfrak{a} = (Y^6, -3Y^5 + 5XY^4, X^2 - 2XY + Y^2)$ is a Gröbner basis for an ideal, what is the lowest power of $X + Y$ that is contained in $\mathfrak{a}$?

8. If $F = \mathbb{Q}[2^{1/2}, 2^{1/3}]$ find the minimum polynomial of $\alpha = 2^{1/2} + 2^{1/3}$.

2.4. The coordinate ring

Now we return to geometry! The coordinate ring is one of the central concepts of algebraic geometry — particularly the theory of *affine* algebraic sets. It is the ring of algebraic functions on an algebraic set. and it determines *all* geometric properties.

DEFINITION 2.4.1. Let $V \subset \mathbb{A}^n$ be an algebraic set and let $\mathfrak{a} = \mathcal{I}(V)$. Then the *coordinate ring* of V is defined by

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

(where the X_i are indeterminates). It is the ring of polynomial functions of $\mathbb{A}^n$ *restricted* to V (or the algebraic functions *on* V).

REMARK. The coordinate ring is not only a ring, but also a structure called an *algebra over a field* (k in this case) — see definition A.2.5 on page 387. This gives it a few special properties — see exercise 6 on page 353 and section 2.5 on page 71.

EXAMPLE 2.4.2. Let $V \subset \mathbb{A}^2$ be the hyperbola defined by $XY = 1$ or $XY - 1 = 0$. It is easily checked that $\sqrt{(XY - 1)} = (XY - 1)$ so the defining ideal is $(XY - 1)$. The coordinate ring is

$$k[X, Y]/(XY - 1) = k[X, X^{-1}]$$

the ring of so-called Laurent polynomials.

PROPOSITION 2.4.3. *The coordinate ring, $k[V]$, of an algebraic set, V, has the following properties:*

   (1) *The points of V are in a 1-1 correspondence with the maximal ideals of $k[V]$.*
   (2) *The closed sets of V are in a 1-1 correspondence with the radical ideals of $k[V]$.*
   (3) *If $f \in k[V]$ and $p \in V$ with corresponding maximal ideal $\mathfrak{m}_p$, then the result of evaluating f at p is the same as the image of f under the canonical projection*
$$\pi\colon k[V] \to k[V]/\mathfrak{m}_p = k$$
   In other words, $f(p) = \pi(f)$.

PROOF. Let $V \subset \mathbb{A}^n$ be an algebraic set. If

$$\pi\colon k[X_1, \ldots, X_n] \to k[V]$$

is the canonical projection, and $\mathfrak{b} \subset k[V]$ is an ideal, then lemma A.1.25 on page 352 implies that

$$\mathfrak{b} \mapsto \pi^{-1}(\mathfrak{b})$$

is a bijection from the set of ideals of $k[V]$ to the set of ideals of $k[X_1, \ldots, X_n]$ containing $\mathfrak{a}$. Prime, and maximal ideals in $k[V]$ correspond to prime, and maximal ideals in $k[X_1, \ldots, X_n]$ containing $\mathfrak{a}$. The fact that radical ideals are intersections of maximal ideals (see corollary 2.2.7 on page 42) implies that this correspondence respects radical ideals too.

If $p = (a_1, \ldots, a_n) \in V \subset \mathbb{A}^n$ is a point, the maximal ideal of functions in $k[X_1, \ldots, X_n]$ that vanish at p is

$$\mathfrak{P} = (X_1 - a_1, \ldots, X_n - a_n) \subset k[X_1, \ldots, X_n]$$

and this gives rise to the maximal ideal $\pi(\mathfrak{P}) \subset k[V]$.

Clearly

$$\mathcal{V}\left(\pi^{-1}(\mathfrak{b})\right) = \mathcal{V}(\mathfrak{b}) \subset V$$

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

To see that $f(p) = \pi(f)$, note that the corresponding statement is true in $k[X_1, \ldots, X_n]$, i.e., the image of $f(X_1, \ldots, X_n)$ under the map

$$g: k[X_1, \ldots, X_n] \to k[X_1, \ldots, X_n]/\mathfrak{P} = k$$

is just $f(a_1 \ldots, a_n)$.

Let $h: k[X_1, \ldots, X_n] \to k[X_1, \ldots, X_n]/\mathfrak{a} = k[V]$ be the canonical projection. Then $\mathfrak{m}_p = h(\mathfrak{P})$ and the diagram

$$
\begin{array}{ccc}
k[X_1, \ldots, X_n] & \xrightarrow{\ h\ } & k[V] \\
{\scriptstyle g}\downarrow & & \downarrow{\scriptstyle \pi} \\
k & =\!\!=\!\!=\!\!= & k
\end{array}
$$

commutes by lemma A.1.26 on page 352. $\qquad\square$

We can define the Zariski Topology on an algebraic set:

If $V \subset \mathbb{A}^n$ is an algebraic set, the topology induced on V by the Zariski topology on $\mathbb{A}^n$ is given by:

- For each subset $W \subset \mathbb{A}^n$, $\mathcal{V}(\mathcal{I}W)$ is the closure in the Zariski topology of $\mathbb{A}^n$.
- There's a 1-1 correspondence between the closed subsets of $\mathbb{A}^n$ and the radical ideals of $k[X_1, \ldots, X_n]$.
- The closed subsets of an algebraic set V correspond to the radical ideals of $k[X_1, \ldots, X_n]$ that contain $\mathcal{I}(V)$.

DEFINITION 2.4.4. If $V \subset \mathbb{A}^n$ is an algebraic set, for every $h \in k[V]$, set

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

These are called the *principal open sets* of V.

REMARK. $D(h)$ is empty only if $h = 0$.

We will often want to regard open sets as algebraic sets in their own right:

PROPOSITION 2.4.5. *Let $V \subset \mathbb{A}^n$ be an affine algebraic set and let $D(h)$ be a principal open set. Then there is an algebraic set $D \subset \mathbb{A}^{n+1}$ homeomorphic to $D(h)$ (with respect to the Zariski topology). Consequently, $D(h)$, has the structure of an affine algebraic set with coordinate ring*

$$k[D(h)] = k[V]_h$$

(see definition A.1.91 on page 384).

PROOF. If $\mathfrak{a} = \mathcal{I}(V)$, set

$$D = \mathcal{V}(\mathfrak{a}, (hT - 1)) \subset \mathbb{A}^{n+1}$$

This is an algebraic set on which the polynomial h cannot vanish — since it has a multiplicative inverse. Projection onto the first n coordinates is a regular map (therefore continuous in the Zariski topology)

$$D \to D(h)$$

that is 1-1 and onto. We claim it maps closed sets to closed sets in the Zariski topology so that its *inverse* is also continuous. Let

$$g(X_1, \ldots, X_n, T) = 0$$

define a closed set in $K \subset \mathbb{A}^{n+1}$. Its intersection with D will be the set defined by

$$g(X_1, \ldots, X_n, h^{-1}) = \frac{p(X_1, \ldots, X_n)}{q(X_1, \ldots, X_n)} = 0$$

so the image of $K \cap D$ in $\mathbb{A}^n$ will be the closed set defined by

$$p(X_1, \ldots, X_n) = 0$$

It follows that the projection is a homeomorphism.

The final statement about the coordinate ring of $D(h)$ follows from lemma A.1.92 on page 384. $\square$

Here's an example of a principal open set that is an *algebraic group*:

EXAMPLE 2.4.6. The *general linear group*, $GL(n, k) \subset \mathbb{A}^{n^2+1}$. This is the group of $n \times n$ matrices with nonzero determinant. It is equal to $D(\det *) \subset \mathbb{A}^{n^2}$, where $\det *$ is the algebraic function on $\mathbb{A}^{n^2}$ that computes the determinant (regarding coordinates in $\mathbb{A}^{n^2}$ as entries of an $n \times n$ matrix). The coordinate ring is given by

(2.4.1) $k[GL(n, k)] =$

$$\frac{k[A_{1,1}, \ldots, A_{1,n}, A_{2,1}, \ldots, A_{2,n}, \ldots, A_{n,1}, \ldots, A_{n,n}, T]}{(T \cdot \det A - 1)}$$

The fact that $\det A$ has a multiplicative inverse prevents it from ever being 0.

PROPOSITION 2.4.7. *The principal open sets of an algebraic set form a basis for the Zariski topology.*

PROOF. Let K be a closed set in the Zariski topology so

$$\begin{aligned} \mathcal{I}(K) &= (x_1, \ldots, x_n) \\ &= \sum_{i=1}^{n} (x_i) \end{aligned}$$

is a radical ideal. By the same reasoning as was used in remark 5 on page 43, we have

$$K = \bigcap_{i=1}^{n} \mathcal{V}\left((x_i)\right)$$

and

$$\bar{K} = \bigcup_{i=1}^{n} \overline{\mathcal{V}(x_i)} = \bigcup_{i=1}^{n} D(x_i)$$

$\square$

PROPOSITION 2.4.8. *Let $V \in \mathbb{A}^n$ and $W \in \mathbb{A}^m$ be algebraic sets and let $f: V \rightarrow W$ be a regular map. Then f induces a homomorphism*

$$f^*: k[W] \rightarrow k[V]$$

of coordinate rings (as k-algebras).

PROOF. The fact that f is regular implies

$$f = \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix}$$

for $F_1, \ldots, F_m \in k[Y_1, \ldots, Y_n]$ and these polynomials induce a map

$$\begin{aligned} F^*: k[X_1, \ldots, X_m] &\rightarrow k[Y_1, \ldots, Y_n] \\ g(X_1, \ldots, X_m) &\mapsto g(F_1, \ldots, F_m) \end{aligned}$$

Since $f(V) \subset W$ we must have $F^*(\mathcal{I}(W)) \subset \mathcal{I}(V)$. But this means that F^* induces a homomorphism of k-algebras

$$f^*: k[X_1, \ldots, X_m]/\mathcal{I}(W) = k[W] \rightarrow k[Y_1, \ldots, Y_n]/\mathcal{I}(V) = k[V]$$

$\square$

EXAMPLE 2.4.9. Suppose $V \subset \mathbb{A}^2$ is the parabola $y = x^2$. Then projection to the x-axis

$$\begin{aligned} f: \mathbb{A}^2 &\rightarrow \mathbb{A}^1 \\ (x, y) &\mapsto x \end{aligned}$$

is a regular map. There is also a regular map $g: \mathbb{A}^1 \rightarrow V$

$$\begin{aligned} g: \mathbb{A}^1 &\rightarrow \mathbb{A}^2 \\ x &\mapsto (x, x^2) \end{aligned}$$

It is interesting that we have a converse to proposition 2.4.8:

PROPOSITION 2.4.10. *Let $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ be algebraic sets. Any homomorphism of k-algebras*

$$f: k[W] \rightarrow k[V]$$

induces a unique regular map

$$\bar{f}: V \rightarrow W$$

REMARK. This and proposition 2.4.8 imply that the coordinate ring is a contravariant functor (see definition A.5.7 on page 442) from the category of algebraic sets to that of k-algebras.

PROOF. We have a diagram

$$k[X_1,\ldots,X_m] \qquad\qquad k[Y_1,\ldots,Y_n]$$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

$$k[X_1,\ldots,X_m]/I(W) \qquad k[y_1,\ldots,y_n]/I(V)$$

$$\| \qquad\qquad\qquad\qquad \|$$

$$k[W] \xrightarrow{\quad f \quad} k[V]$$

and we can map each $X_i \in k[X_1,\ldots,X_m]$ to $k[Y_1,\ldots,Y_n]$ to make

(2.4.2) $k[X_1,\ldots,X_m] \dashrightarrow^{\,r\,} k[Y_1,\ldots,Y_n]$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

$$k[X_1,\ldots,X_m]/I(W) \qquad k[Y_1,\ldots,Y_n]/I(V)$$

$$\| \qquad\qquad\qquad\qquad \|$$

$$k[W] \xrightarrow{\quad f \quad} k[V]$$

commute as a diagram of k-algebras. Suppose $r(X_i) = g_i(Y_1,\ldots,Y_n)$. We claim that

$$\bar{f} = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix} : \mathbb{A}^n \to \mathbb{A}^m$$

is the required regular map. If $p = (k_1,\ldots,k_n) \in V \subset \mathbb{A}^n$ so

$$v(p) = 0$$

for any $v \in \mathcal{I}(V)$, then $w(f(p)) = f(w)(p) = 0$ for any $w \in \mathcal{I}(W)$ implying that $\bar{f}(V) \subset W$.

If we replace r in diagram 2.4.2 by a map r' that still makes it commute, the induced $\bar{f}'$ will differ from f by elements of $\mathcal{I}(V)$ so

$$\bar{f}|V = \bar{f}'|V$$

implying that the map $\bar{f}: V \to W$ is unique. $\square$

DEFINITION 2.4.11. Let $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ be algebraic sets. Then V and W are said to be *isomorphic* if there exist regular maps

$$\begin{aligned} f:V &\to W \\ g:W &\to V \end{aligned}$$

such that $f \circ g = 1: W \to W$ and $g \circ F = 1: V \to V$.

REMARK. We may regard isomorphic algebraic sets as equivalent in every way. Then example 2.4.9 shows that parabola $y = x^2 \subset \mathbb{A}^2$ is isomorphic to $\mathbb{A}^1$.

We have proved:

COROLLARY 2.4.12. *Algebraic sets $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ are isomorphic if and only if $k[V]$ and $k[W]$ are isomorphic as k-algebras.*

REMARK. This proves the claim made earlier: the coordinate ring defines *all* of the significant geometric properties of an algebraic set, including its isomorphism class. This is analogous to the Gelfand-Naimark theorem (in [**54**]) that locally compact Hausdorff spaces are determined up to *homeomorphism* by their C^*-algebras of continuous functions that vanish at infinity. Also see section 4.2.1 on page 167.

We can characterize the kinds of rings that can be coordinate rings of algebraic sets:

DEFINITION 2.4.13. Given an algebraically closed field k, an *affine k-algebra* is defined to be a finitely generated k-algebra that is reduced, i.e. $\sqrt{(0)} = (0)$. If A and B are affine k-algebras, the set of homomorphisms $f \colon A \to B$ is denoted

$$\hom_{k-\mathrm{alg}}(A, B)$$

REMARK. The requirement that the ring be *reduced* is equivalent to saying that it has no nilpotent elements. This is equivalent to saying that the intersection of its maximal ideals is 0 — see theorem A.1.47 on page 360.

If k is an algebraically closed field, Hilbert's Nullstellensatz (theorem 2.2.5 on page 41) implies that affine k-algebras (see definition 2.4.13) are Jacobson rings (see definition A.4.22 on page 428).

Dominating maps are, for most purposes, essentially surjective:

DEFINITION 2.4.14. A regular map $f \colon V \to W$ is said to be *dominating* if its image is dense in W.

REMARK. Recall that a subset of a space, $S \subset X$, is *dense* if the closure of S is *all* of X. Figure 2.1.2 on page 38 and the discussion surrounding it shows how easy it is for a map to have a dense image in the Zariski topology.

LEMMA 2.4.15. *A regular mapping $f \colon V \to W$ induces an injection*

$$f^* \colon k[W] \to k[V]$$

of coordinate rings if and only if it is dominating.

PROOF. If $g \in k[W]$ and the image of f is dense, $f(V)$ intersects all open sets of W, including the open set, $D(g)$. If $g \in \ker f^*$ then we must have $D(g) = \varnothing$, i.e. $g = 0$.

On the other hand, if f^* is injective, then $g \neq 0 \in k[W]$ implies that $f^*(g) \neq 0$ so $f(V) \cap D(g) \neq \varnothing$. □

Now we study *irreducible* algebraic sets — the building blocks of algebraic sets in general.

DEFINITION 2.4.16. A nonempty topological space is said to be *irreducible* if it is not the union of two proper closed subsets.

Irreducible spaces have a number of interesting properties. For instance, open sets in an irreducible space tend to be very "large":

PROPOSITION 2.4.17. *Let X be an irreducible space and let $O_1, \ldots, O_k \subset X$ be any finite collection of open sets. Then*

$$(2.4.3) \qquad\qquad \bigcap_{i=1}^{k} O_i \neq \varnothing$$

PROOF. Just take the complement of equation 2.4.3 to get

$$\bigcup_{i=1}^{k} \bar{O}_i = X$$

which contradicts the fact that X is irreducible. $\square$

It is easy to characterize irreducible algebraic sets via the coordinate ring:

PROPOSITION 2.4.18. *An algebraic set W is irreducible if and only if $\mathcal{I}(W)$ is prime. This happens if and only if $k[W]$ is an integral domain.*

PROOF. Suppose W is irreducible and $fg \in \mathcal{I}(W)$. At each point of W either f is zero or g is zero so $W \subset \mathcal{V}(f) \cup \mathcal{V}(g)$. It follows that

$$W = (W \cap \mathcal{V}(f)) \cup (W \cap \mathcal{V}(g))$$

Since W is irreducible, we must have

$$W = W \cap \mathcal{V}(f) \text{ or } W = W \cap \mathcal{V}(g)$$

so that either f or g are in $\mathcal{I}(W)$. Suppose $I(W)$ is prime and $W = \mathcal{V}(\mathfrak{a}) \cup \mathcal{V}(\mathfrak{b})$ with $\mathfrak{a}$ and $\mathfrak{b}$ radical ideals. We must show that $W = \mathcal{V}(\mathfrak{a})$ or $\mathcal{V}(\mathfrak{b})$. Recall that $\mathcal{V}(\mathfrak{a}) \cup \mathcal{V}(\mathfrak{b}) = \mathcal{V}(\mathfrak{a} \cap \mathfrak{b})$, and that $\mathfrak{a} \cap \mathfrak{b}$ is radical. It follows that $\mathcal{I}(W) = \mathfrak{a} \cap \mathfrak{b}$.
    If $W \neq \mathcal{V}(\mathfrak{a})$, then there exists an $f \in \mathfrak{a} \setminus \mathcal{I}(W)$. For all $g \in \mathfrak{b}$

$$fg \in \mathfrak{a} \cap \mathfrak{b} = \mathcal{I}(W)$$

Because $\mathcal{I}(W)$ is prime, this implies that $\mathfrak{b} \subset \mathcal{I}(W)$ so $W \subset \mathcal{V}(\mathfrak{b})$.
    The final statement about $k[V]$ being an integral domain follows from lemma A.1.30 on page 355. $\square$

DEFINITION 2.4.19. Let V be an algebraic set and let $U \subset V$ be an open subset. Then U is called an *open affine* of V if U has the structure of an affine algebraic set and the inclusions

$$\begin{aligned} U &\hookrightarrow V \\ D(h) &\hookrightarrow U \end{aligned}$$

are regular maps for all h such that $D(h) \subset U$.

REMARK. For instance any principal open set $D(h) \subset V$ is an open affine, as proposition 2.4.5 on page 63 shows.
    Exercise 7 on page 90 shows that there exist open sets that are *not* affine. There also exist open affines that are *not principal* — i.e., not of the form $D(h)$ for any regular function h (although they will be a union of such sets) — see exercise 13 on page 91.

The Zariski topology on an algebraic set gives it some interesting properties:

DEFINITION 2.4.20. A topological space, X, is *noetherian* if every descending chain of closed sets

$$X \supsetneq X_1 \supsetneq X_2 \supsetneq \cdots \supsetneq \varnothing$$

is finite.

REMARK. Clearly the topological spaces $\mathbb{R}$ and $\mathbb{C}$ (with their usual topologies) are far from being noetherian.

Under the Zariski topology, algebraic sets are noetherian:

PROPOSITION 2.4.21. *Algebraic sets with the Zariski topology are noetherian.*

PROOF. The descending chain condition on closed subspaces in definition 2.4.20 is equivalent to the ascending chain condition on ideals in definition A.1.48 on page 360. □

All noetherian spaces can be decomposed into a union of *irreducible subspaces:*

PROPOSITION 2.4.22. *Let V be a noetherian topological space. Then V is a finite union of irreducible closed subsets*

$$V = V_1 \cup \cdots \cup V_m$$

Moreover, if the decomposition is irredundant (no V_i is contained in any other), then the V_i are uniquely determined, up to order.

PROOF. Suppose V cannot be written as a *finite* union of irreducible closed subsets. Then, because V is noetherian, there will be a closed subset W of V that is *minimal* among those that cannot be written in this way. Since W cannot be irreducible, we have $W = W_1 \cup W_2$ with each W_i a proper closed subset of W.

From the *minimality* of W, we conclude that the W_i can be written as *finite* unions of irreducible close sets. This is a contradiction!

Suppose

$$V = V_1 \cup \cdots \cup V_m = W_1 \cup \cdots \cup W_t$$

are two irredundant decompositions. Then $V_i = \bigcup_j (V_i \cap W_j)$ and, because V_i is irreducible $V_i = V_i \cap W_j$ for some j. Consequently, there is some function

$$f \colon \{1, \ldots, m\} \to \{1, \ldots, t\}$$

such that $V_i \subset W_{f(i)}$.

Similar reasoning implies that there is a function

$$g \colon \{1, \ldots, t\} \to \{1, \ldots, m\}$$

such that the composites $f \circ g$ and $g \circ f$ are the identities. □

Recall that corollary 2.2.7 on page 42 showed that a radical ideal is equal to the intersection of all the maximal ideals containing it. Since an algebraic set is noetherian, it has a unique decomposition into irreducible closed sets, by proposition 2.4.22.

The correspondence between ideals and closed sets in corollary 2.2.6 on page 42 implies this variation on corollary 2.2.7 on page 42:

COROLLARY 2.4.23. *A radical ideal* $\mathfrak{a} \in k[X_1, \ldots, X_n]$ *is a finite intersection of prime ideals*

$$\mathfrak{a} = \mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_j$$

If there are no inclusions among the $\mathfrak{p}_i$ *then the* $\mathfrak{p}_i$ *are uniquely determined.*

PROOF. Write $\mathcal{V}(\mathfrak{a}) = \bigcup V_i$ (irreducible components) and set $\mathfrak{p}_i = \mathcal{I}(V_i)$.
$\square$

Recall that a topological space, X, is *connected* if the only sets that are both open and closed are the empty set and X itself. Equivalently, X is connected if and only if, whenever we have

(2.4.4) $X = X_1 \cup X_2$

with X_1, X_2 both closed and nonempty, we also have

(2.4.5) $X_1 \cap X_2 \neq \emptyset$

If X is an irreducible algebraic set, then equation 2.4.4 implies that $X_1 = X_2 = X$, so X is connected. The converse is not necessarily true.

For instance, $\mathcal{V}(X_1 X_2)$ is the union of the coordinate axes in $\mathbb{A}^2$, which is connected but not irreducible.

We can translate equations 2.4.4 and 2.4.5 into statements about ideals in the coordinate ring:

> An algebraic subset V of $\mathbb{A}^n$ is *connected* but not *irreducible* if and only if there exist ideals $\mathfrak{a}$ and $\mathfrak{b}$ such that 3 on page 38 $\mathfrak{a} \cap \mathfrak{b} = \mathcal{I}(V)$ but $\mathfrak{a} + \mathfrak{b} \neq k[X_1, \ldots, X_n]$.

EXERCISES.

1. Suppose the characteristic of the field k is $\neq 2$ and V is an algebraic set in $\mathbb{A}^3$ defined by the equations

$$X^2 + Y^2 + Z^2 = 0$$
$$X^2 - Y^2 - Z^2 + 1 = 0$$

Decompose V into its irreducible components.

2. Prove that the statements:
   a. X is connected if and only if the only subsets of X that are open and closed are $\emptyset$ and X,
   b. X is connected if, whenever $X = X_1 \cup X_2$ with X_1, X_2 closed nonempty subsets of X, then $X_1 \cap X_2 \neq \emptyset$.

are equivalent.

3. In an irreducible algebraic set, show that $D(h_1) \cap D(h_2) = D(h_1 h_2)$.

4. Show that $D(f) \subset D(g)$ if and only if $f^n \in (g)$ for some $n > 0$.

5. Show that the map

$$v\colon GL(n,k) \rightarrow GL(n,k)$$
$$A \mapsto A^{-1}$$

that sends a matrix to its inverse, is a regular map.

6. Suppose R is a ring that contains an infinite field with ideals $\mathfrak{I}_1, \ldots, \mathfrak{I}_n$ and $\mathfrak{J}$ such that

$$\mathfrak{J} \subset \bigcup_{i=1}^{n} \mathfrak{I}_i$$

Show that there exists an α such that $\mathfrak{J} \subset \mathfrak{I}_\alpha$.

This is one case of the oddly-named Prime Avoidance lemma.

7. Suppose $\mathfrak{I}_1, \ldots, \mathfrak{I}_n$ and $\mathfrak{J}$ are ideals in a ring R such that

$$\mathfrak{J} \subset \bigcup_{i=1}^{n} \mathfrak{I}_i$$

and all but *two* of the $\mathfrak{I}_j$ are *prime.* Show that there exists an α such that $\mathfrak{J} \subset \mathfrak{I}_\alpha$.

This is the second case of the Prime Avoidance lemma.

8. If $\mathfrak{a}, \mathfrak{b} \subset R = k[X_1, \ldots, X_n]$ are two ideals and $f \in R$ vanishes on $\mathcal{V}(\mathfrak{a}) \setminus \mathcal{V}(\mathfrak{b})$, show that, for any $b \in \mathfrak{b}$ that there exists a power n such $f \cdot b^n \in \mathfrak{a}$.

9. Suppose $\mathfrak{a}, \mathfrak{b} \subset R = k[X_1, \ldots, X_n]$ are two ideals and define the ideal[3]

$$(\mathfrak{a}\colon \mathfrak{b}^\infty) = \left\{ x \in R \,\middle|\, \exists_{n \in \mathbb{Z}^+} \text{ such that } x \cdot \mathfrak{b}^n \subset \mathfrak{a} \right\}$$

— *the saturation of* $\mathfrak{a}$ *with respect to* $\mathfrak{b}$. Show that

$$\mathcal{I}(\mathcal{V}(\mathfrak{a}) \setminus \mathcal{V}(\mathfrak{b})) = (\mathfrak{a}\colon \mathfrak{b}^\infty)$$

2.5. specm ∗

We are now in a position to define *affine varieties* — essentially affine algebraic sets with no preferred embedding in an affine space:

DEFINITION 2.5.1. If R is an affine k-algebra define the *maximal spectrum,* specm R, *of* R to be the set of all proper maximal ideals of R. In addition, define

(1) for all $r \in R$, $D(r) = \{\mathfrak{m} \in \text{specm } R \,|\, r \notin \mathfrak{m}\}$,
(2) the topology on specm R for which the sets $D(r)$, for all $r \in R$, form a base.
(3) Given $f \in R$ define $f(\mathfrak{m})$ to be the image of r under the canonical map

$$R \to R/\mathfrak{m} = k$$

the *value of* f at the point $\mathfrak{m} \in \text{specm } R$.

---

[3]The proof that this is an ideal is left as an exercise to the reader.

REMARK. The space $\operatorname{specm} *$ has been largely supplanted by the space $\operatorname{Spec} *$ whose points are *prime* ideals of a ring (see definition 4.2.1 on page 163) rather than maximal ideals. There are several reasons for this:

(1) the topology of $\operatorname{Spec} *$ is equivalent to that of $\operatorname{specm} *$
(2) $\operatorname{Spec} *$ is a well-defined functor (see definition A.5.7 on page 442) of rings, while $\operatorname{specm} *$ is not, except in the restricted case of affine k-algebras (proved in proposition 2.5.3).

We use $\operatorname{specm} *$ as a stepping stone to $\operatorname{Spec} *$ because it is somewhat easier to understand.

This gives an entirely "coordinate free" way of describing an affine algebraic set:

PROPOSITION 2.5.2. *Let $V \subset \mathbb{A}^n$ be an algebraic set with coordinate ring $k[V]$. Then there exists a canonical homeomorphism*

$$h: V \to \operatorname{specm} k[V]$$

of topological spaces.

PROOF. Proposition 2.4.3 on page 62 implies that every point $p \in V$ corresponds to a unique maximal ideal $\mathfrak{m}_p \subset k[V]$ of functions that vanish at p. We define h by

$$h(p) = \mathfrak{m}_p \in \operatorname{specm} k[V]$$

This is a one-to-one mapping and we claim that it is continuous. We must show that $h^{-1}(D(f))$ is open in V for all $f \in k[V]$.

In $\operatorname{specm} k[V]$, $D(f) = \{\mathfrak{m} \in \operatorname{specm} k[V] | f \notin \mathfrak{m}\}$, for $f \in k[V]$. Since evaluating a function, $f \in k[V]$ at a point $p \in V$ coincides with its image under the projection

$$\pi_p: k[V] \to k[V]/\mathfrak{m}_p = k$$

— see statement 3 in proposition 2.4.3 on page 62, it follows that

$$
\begin{aligned}
h^{-1}(D(f)) &= \{p \in V | \pi_p(f) \neq 0\} \\
&= \{p \in V | f(p) \neq 0\} \\
&= D(f) \subset V \text{ as in definition 2.4.4 on page 63}
\end{aligned}
$$

A similar argument shows that $h(D(f))$, where $D(f)$ is as defined in definition 2.4.4 is precisely $D(f)$ as defined in 2.5.1. This implies that the inverse of h is continuous. $\square$

Our $\operatorname{specm} *$ construction is functorial (see definition A.5.7 on page 442). This means that it mimics the behavior of algebraic sets under regular maps:

PROPOSITION 2.5.3. *Let $\alpha: A \to B$ be a homomorphism of affine k-algebras. Then α induces a continuous map of topological spaces*

$$\varphi: \operatorname{specm} B \to \operatorname{specm} A$$

with, for any maximal ideal $\mathfrak{n} \subset B$,

$$\varphi(\mathfrak{n}) = \alpha^{-1}(\mathfrak{n}) = \mathfrak{m}$$

a maximal ideal.

PROOF. For any maximal ideal $\mathfrak{n} \in B$, $\mathfrak{m} = \alpha^{-1}(\mathfrak{n})$ is maximal in A because $A/\mathfrak{m} \to B/\mathfrak{n} = k$ is an injective (see exercise 7 on page 353) map of k-algebras. Since quotients of k-algebras are also k-algebras (see exercise 6 on page 353, it follows that $k \subset A/\mathfrak{m}$ maps via the identity map to $B/\mathfrak{n} = k$.

Thus α defines a map

$$\varphi \colon \operatorname{specm} B \to \operatorname{specm} A, \quad \varphi(\mathfrak{n}) = \alpha^{-1}(\mathfrak{n}) = \mathfrak{m}$$

making the diagram

$$
\begin{array}{ccc}
A & \xrightarrow{\ \alpha\ } & B \\
\downarrow & & \downarrow \\
A/\mathfrak{m} & =\!\!=\!\!= & B/\mathfrak{n}
\end{array}
$$

commute.

This implies that

$$f(\varphi(\mathfrak{n})) = \alpha(f)(\mathfrak{n}) \quad \text{i.e., } f \circ \varphi = \alpha$$

for any $f \in A$, where $f(\mathfrak{m})$ (for $\mathfrak{m}$ a maximal ideal) is the image of f under the homomorphism

$$A \to A/\mathfrak{m} = k$$

Now note that α induces a map

$$\alpha_x \colon A_x \to B_{\alpha(x)}$$

for any $x \in A$.

Since the $D(x)$ for $x \in A$, are the principal open sets of $\operatorname{specm} A$ with respective coordinate rings A_x (see Proposition 2.4.5 on page 63), this implies that $\varphi^{-1}D(f) = D(f \circ \varphi)$, and

$$\varphi^{-1}(D(f)) = D(\alpha(f))$$

so φ is continuous. $\qquad\square$

We finally define:

DEFINITION 2.5.4. Let A be an affine k-algebra. Then the pair $(A, \operatorname{specm} A)$ is called an *affine variety*.

We finally summarize everything in the following result:

THEOREM 2.5.5. *The functor $A \mapsto \operatorname{specm} A$ is a contravariant equivalence (see definition A.5.13 on page 445) from the category of affine k-algebras to that of affine varieties with quasi-inverse $(V, \operatorname{specm} k[V]) \mapsto k[V]$. For any affine k-algebras A and B*

$$\operatorname{Hom}_{k-\mathrm{alg}}(A, B) \xrightarrow{\cong} \operatorname{hom}_{\mathrm{Var}}(\operatorname{specm} B, \operatorname{specm} A)$$

For any affine varieties V and W

$$\operatorname{hom}_{\mathrm{Var}}(V, W) \xrightarrow{\cong} \operatorname{Hom}_{k-\mathrm{alg}}(k[W], k[V])$$

REMARK. To state a result like this, we needed a formulation of algebraic sets that did not require their being embedded in an affine space.

An affine algebraic set is just an affine variety equipped with an embedding in an affine space. The following result implies that this is *always possible:*

LEMMA 2.5.6. *Let $(R, \operatorname{specm} R)$ be an affine variety. Then there exists an embedding*
$$f\colon \operatorname{specm} R \to \mathbb{A}^n$$
for some n.

PROOF. Since R is finitely generated over k, we have
$$R = k[x_1, \ldots, x_n]$$
and get a surjection
$$\begin{aligned} f\colon k[X_1, \ldots,, X_n] &\to R \\ X_i &\mapsto x_i \end{aligned}$$
(where the X_i are indeterminates) with a kernel, $\mathfrak{K}$. We claim that $\mathfrak{K}$ is a radical ideal. If not, let $x \in k[X_1, \ldots, X_n]$ be an element with the property that $x^n \in \mathfrak{K}$ but $x \notin \mathfrak{K}$. In that case, $f(x) \in R$ has the property that $f(x)^n = 0$ but $f(x) \neq 0$, contradicting the fact that R is a reduced ring.

If $V = \mathcal{V}(\mathfrak{K})$, then Hilbert's Nullstellensatz (see 2.2.5 on page 41) implies that $\mathfrak{K} = \mathcal{I}V$ with coordinate ring
$$k[X_1, \ldots, X_n]/\mathfrak{K} = R$$
Now map $\operatorname{specm} R$ to V via the inverse of the homeomorphism in proposition 2.5.2 on page 72. □

We can use theorem 2.5.5 on the preceding page to compute sets of *regular maps* between affine varieties:

(1) Theorem 2.5.5 on the previous page implies that
$$\hom_{\mathrm{Var}}(V, \mathbb{A}^1) = \operatorname{Hom}_{k-\mathrm{alg}}(k[X], k[V]) = k[V]$$
so the regular maps $V \to \mathbb{A}^1$ are just the regular functions on V.

(2) Define $\mathbb{A}^0$ to be the ringed space $(V_0, \mathcal{O}_{V_0})$ with V_0 consisting of a single point. Equivalently $\mathbb{A}^0 = \operatorname{specm} k$. Then, for any affine variety
$$\hom_{\mathrm{Var}}(\mathbb{A}^0, V) \simeq \hom_{k-\mathrm{alg}}(k[V], k) \simeq V$$
where the isomorphism
$$\hom_{k-\mathrm{alg}}(k[V], k) \simeq V$$
sends $\alpha \in \hom_{k-\mathrm{alg}}(k[V], k)$ to the point corresponding to the maximal ideal $\ker \alpha$.

(3) Consider $t \mapsto (t^2, t^3)\colon \mathbb{A}^1 \to \mathbb{A}^2$. This is bijective onto its image
$$V\colon y^2 = x^3$$
but is not an isomorphism onto its image — the inverse map is not regular. It suffices to show that $t \mapsto (t^2, t^3)$ doesn't induce an isomorphism of rings of regular functions. We have $k[\mathbb{A}^1] = k[T]$ and $k[V] = k[X, Y]/(Y^2 - X^3) = k[x, y]$, where $X \mapsto x$, $Y \mapsto y$.

The map on rings is
$$x \mapsto T^2, \quad y \mapsto T^3, \quad k[V] \to k[T]$$
which is injective, but its image is $k[T^2, T^3] \neq k[T]$, the ring of polynomials of degree > 1.
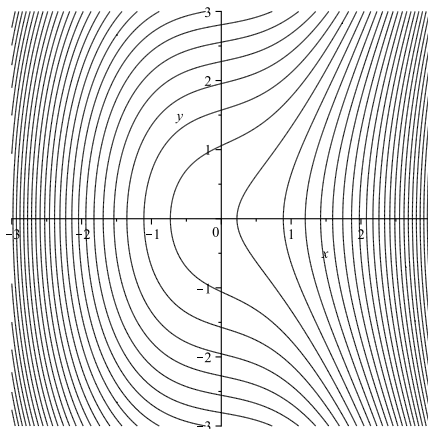
FIGURE 2.5.1.   Fibers of a regular map

In fact, unlike $k[T]$, $k[x,y]$ is not *integrally closed* (see definition A.4.10 on page 422):

$$(y/x)^2 - x = 0$$

so (y/x) is integral over $k[x,y]$ but $y/x \notin k[x,y]$ (it maps to T under the inclusion $k(x,y) \hookrightarrow k(T)$).

2.5.1. Fibers of regular maps. In algebraic geometry, inverse images of points (and subvarieties) under regular maps are called *fibers* of those maps.

Although this might seem to be an odd term for inverse images, figure 2.5.1 shows some fibers of a version of the elliptic map

$$\begin{array}{ccc} \mathbb{A}^2 & \to & \mathbb{A}^1 \\ (X,Y) & \mapsto & Y^2 - X^3 - X \end{array}$$

We can think of $\mathbb{A}^2$ as a union of an infinite number of these inverse images, which look somewhat like fibers. Another reason for calling inverse images fibers appears in section 4.3.3 on page 178 where *fibered products* are defined. It can be shown that fibers of maps occur as suitable fibered products — see example 9 on page 181.

DEFINITION 2.5.7. If $f\colon V \to W$ is a regular map and $X \subset W$ is a subvariety, we call $f^{-1}(X)$ the *fiber of f over X*.

One interesting feature of fibers over closed subvarieties is that they are closed subvarieties themselves:

LEMMA 2.5.8. *Let $f\colon V \to W$ is a regular map with induced homomorphism of coordinate rings*

$$f^*\colon k[W] \to k[V]$$

and let $X \subset W$ be a closed subvariety corresponding to the radical ideal $\mathfrak{x}$. Then $f^{-1}(X)$ is the closed subvariety of V corresponding to the ideal

$$\sqrt{f^*(\mathfrak{x}) \cdot k[V]} \subset k[V]$$

The coordinate ring of the fiber is the quotient

$$\frac{k[V]}{\sqrt{f^*(\mathfrak{x}) \cdot k[V]}}$$

PROOF. The subvariety X is defined by the fact that the functions $\{g_i\}$ that generate the ideal $\mathfrak{x}$ vanish on it. The set $f^{-1}(X)$ is defined by the fact that the composite of f with those functions vanish on $f^{-1}(X)$, i.e. that the functions $\{f \circ g_i\} \in k[V]$ vanish on it. But these are precisely $\{f^*(g_i)\}$ and the ideal they generate is $f^*(\mathfrak{x}) \cdot k[V]$. Since this might not be a radical ideal, we take the radical.   □

EXAMPLE 2.5.9. Suppose $f: \mathbb{A}^1 \to \mathbb{A}^2$ is the map that sends X to $(\bar{X}, \bar{Y}) = (X^2, X^3)$. The induced map of coordinate rings

$$\begin{array}{ccc} k[\bar{X}, \bar{Y}] & \xrightarrow{f^*} & k[X] \\ \bar{X} & \mapsto & X^2 \\ \bar{Y} & \mapsto & X^3 \end{array}$$

The point $(0, 0) \in \mathbb{A}^2$ corresponds to the maximal ideal $\mathfrak{m}_0 = (\bar{X}, \bar{Y})$, and its image under f^* is $(X^2, X^3) = (X^2) \subset k[X]$, which is not a radical ideal. After taking the radical, we get (X), the maximal ideal corresponding to the point $0 \in \mathbb{A}^1$, which is $f^{-1}(0, 0)$.

Consider $f^{-1}(2, 3)$. The maximal ideal is $(\bar{X} - 2, \bar{Y} - 3) \subset k[\bar{X}, \bar{Y}]$ and its image under f^* is

$$(X^2 - 2, X^3 - 3) \subset k[X]$$

and it is not hard to see that the greatest common divisor of $X^2 - 2$ and $X^3 - 3$ is 1. The Euclidean algorithm implies that $1 \in (X^2 - 2, X^3 - 3)$. This shows us what happens when we take the inverse image of a point that is not in the image of f — we get an ideal that contains 1 and represents the empty set.

We conclude this section with an two important classes of morphisms of affine varieties:

DEFINITION 2.5.10. A map $f: V \to W$ of affine varieties sets will be called *finite* if the induced map

$$f^*: k[W] \to k[V]$$

is injective and $k[V]$ is integral over $k[W] \subset k[V]$ (see definition A.4.3 on page 420).

REMARK. Proposition A.4.5 on page 421 implies that composites of finite maps are finite.

In order to understand the geometric properties of finite maps, we need the concept of *Artinian ring*, developed in section A.1.7 on page 381.

PROPOSITION 2.5.11. *If $f: V \to W$ is a finite map of irreducible affine varieties, then*

(1) f is surjective, and
(2) if $w \in W$ is any point, $f^{-1}(w)$ is a finite set of points.

REMARK. A dominating finite map is *almost* an isomorphism: every point in $f^{-1}(y)$ has a neighborhood that is mapped isomorphically to a neighborhood of y.

For instance the map $z \mapsto z^2$ defines a finite map

$$f: V = \mathbb{C} \to \mathbb{C} = W$$

and $f^{-1}(w)$ has two points in it unless $w = 0$.

PROOF. Since $f^*: k[W] \to k[V]$ is injective, f is dominating (see lemma 2.4.15 on page 67). Suppose $\mathfrak{a} \subset k[W]$ is a maximal ideal, representing a point, p, of W. Then the image of $\mathfrak{a}$ in $k[V]$ is just $\mathfrak{a} \cdot k[V]$, representing $f^{-1}(p)$ — see lemma 2.5.8 on page 75 and the surrounding discussion. If $\mathfrak{a} \cdot k[V] = k[V]$, then the inverse image of the point, p, is the empty set — so p is not in the image of f. Nakayama's Lemma A.1.80 on page 378 implies that this *never* happens:

If $\mathfrak{a} \cdot k[V] = k[V]$, then Nakayama's Lemma says that there exists an element $r \in k[W] \subset k[V]$ such that $r \equiv 1 \pmod{\mathfrak{a}}$ and $r \cdot k[V] = 0$. Since the rings in question are integral domains, $r = 0$ and $0 \equiv 1 \pmod{\mathfrak{a}}$, or $1 \in \mathfrak{a}$ so that $\mathfrak{a}$ also represents the empty set.

To prove the second statement, note that the points in $f^{-1}(p)$ correspond to the maximal ideals of $k[V]$ that contain $\mathfrak{a} \cdot k[V]$, i.e. the maximal ideals of the quotient

$$Q = k[V]/\mathfrak{a} \cdot k[V]$$

(see lemma A.1.25 on page 352). Since $k[V]$ is a finitely generated *module* over $k[W]$, Q will be a finitely generated module over the field, $F = k[W]/\mathfrak{a}$, i.e. a finite-dimensional vector space. We claim that Q is Artinian (see definition A.1.83 on page 381): every descending sequence

(2.5.1) $$\mathfrak{a}_1 \supseteq \mathfrak{a}_2 \supseteq \cdots$$

of ideals in Q must become constant after a finite number of terms. This is because each ideal is a sub-vector-space of Q and a proper subspace must be of a lower dimension. In fact the dimension of Q over F is the maximum number of distinct terms a sequence like 2.5.1 can have.

Statement 4 in lemma A.1.84 on page 381 implies that the number of maximal ideals in Q is *finite*. This completes the proof. □

Irreducible affine varieties have a number of interesting geometric properties that are not well defined for general affine varieties. For instance, they have a constant dimension. They are also key to defining general algebraic varieties.

The Noether Normalization Theorem (2.2.2 on page 40) implies the remarkable result:

THEOREM 2.5.12. *If V is an irreducible affine variety and the field of fractions of $k[V]$ has transcendence degree n, then there exists a dominating finite map*

$$f: V \to \mathbb{A}^n$$

REMARK. This is often called the *geometric* form of the Noether Normalization Theorem.

PROOF. The Noether Normalization Theorem implies that $k[V]$ is integral over $k[x_1, \dots, x_n]$ where $x_1, \dots, x_n \in k[V]$ are algebraically independent. The inclusion

$$k[x_1, \dots, x_n] \hookrightarrow k[V]$$

induces f via proposition 2.4.10 on page 65. $\qquad\square$

We conclude this section with the topic of flat morphisms:

DEFINITION 2.5.13. A ring-homomorphism $f\colon R \to S$ will be called *flat* if it makes S a flat module over R. A morphism $f\colon V \to W$ of affine varieties will be called flat if the corresponding morphism $f^*\colon k[W] \to k[V]$ of coordinate rings is flat.

Flat morphisms of rings have an important algebraic property, called *Going Down:*

PROPOSITION 2.5.14. *Let $f\colon R \to S$ be a flat morphism of rings and let $\mathfrak{p}_1 \subset \mathfrak{p} \subset R$ be an inclusion of prime ideals. If $\mathfrak{q} \subset S$ is a prime ideal such that $f^{-1}(\mathfrak{q}) = \mathfrak{p}$, then there exists an ideal $\mathfrak{q}_1 \subset \mathfrak{q}$ such that $f^{-1}(\mathfrak{q}_1) = \mathfrak{p}_1$.*

PROOF. Let $\mathfrak{q}_1$ be a minimal prime ideal containing $f(\mathfrak{p}_1) \cdot S$ — this exists because the intersection of a descending chain of prime ideals is prime (see exercise 12 on page 354). Exercise 22 on page 473 shows that

$$\frac{S}{f(\mathfrak{p}_1) \cdot S} = S \otimes_R \left(\frac{R}{\mathfrak{p}_1} \right)$$

is a flat module over

$$\frac{R}{\mathfrak{p}_1}$$

so there is no loss of generality if we assume that $\mathfrak{p}_1 = 0$ and $\mathfrak{q}_1$ is minimal. Then exercise 31 on page 364 implies that all of the elements of $\mathfrak{q}_1$ are zero-divisors. On the other hand, $R/\mathfrak{p}_1$ is an integral domain and non-zero divisors of $R/\mathfrak{p}_1$ map to non-zero divisors of $S/f(\mathfrak{p}_1) \cdot S$ (by exercise 21 on page 473) so

$$\operatorname{im} \bar{f}(R/\mathfrak{p}_1) \cap \operatorname{im} \mathfrak{q}_1 = 0$$

and $f^{-1}(\mathfrak{q}_1) = 0 = \mathfrak{p}_1$. $\qquad\square$

This has an interesting geometric interpretation:

If $f\colon V \to W$ is a flat morphism of affine varieties and $V' \subset V$ has the property that $f(V') \subset W' \subset W$ are subvarieties, then there exists a subvariety $V'' \supset V'$ such that $f(V'')$ is dense in W' (i.e., $f|V''\colon V'' \to W'$ is effectively surjective). See figure 2.5.2 on the facing page.
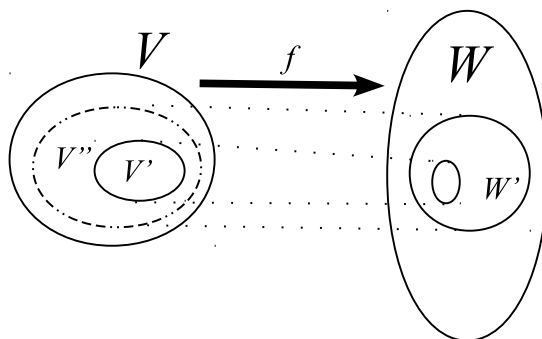
FIGURE 2.5.2.   Flat morphism

EXERCISES.

1. If

$$f\colon V \to W$$

is a regular map represented by

$$F\colon k[W] \to k[V]$$

and with $p \notin \operatorname{im} f$ represented by the maximal ideal $\mathfrak{m}_p \subset k[W]$, what can one say about $F(\mathfrak{m}_p) \subset k[V]$?

2. Consider the regular map

$$\begin{aligned} f\colon \mathbb{A}^2 &\to \mathbb{A}^2 \\ (X, Y) &\mapsto (X, XY) \end{aligned}$$

Determine the induced map of coordinate rings and the fiber over the point $(a, b) \in \mathbb{A}^2$.

3. Let H be the hyperbola $XY = 1$ in $\mathbb{A}^2$. Construct a surjective map

$$f\colon H \to \mathbb{A}^1$$

with $f^{-1}(\text{point})$ a finite set of points (as per theorem 2.5.12 on page 77). Represent the coordinate ring of V as a finitely-generated module over a polynomial algebra.

4. Suppose M is a finitely generated module over $R = k[X_1, \ldots, X_n]$ with a prime filtration (see theorem A.1.77 on page 375)

$$0 = M_0 \subset \cdots \subset M_j = M$$

with

$$\frac{M_{i+1}}{M_i} \cong \frac{R}{\mathfrak{p}_i}$$

If $\operatorname{Ann}(M) \subset R$ is a radical ideal, show that

$$V = \mathcal{V}\left(\operatorname{Ann}(M)\right) = \bigcup_{i=1}^{j-1} \mathcal{V}\left(\mathfrak{p}_i\right)$$

is a decomposition of V into its irreducible components. This gives a geometric proof of the statement:

If $\mathfrak{k}$ is a *radical* ideal, then the prime ideals in the prime filtration
of $R/\mathfrak{k}$ are all *distinct* and *uniquely* determined by it.

This is *not* true in general (i.e., when $\mathfrak{k}$ is not a radical ideal).

5. If $f\colon V \to W$ is a regular map and $x \in k[W]$, what is $k[f^{-1}(D(x))]$?

2.5.2. Elimination theory. In this section, we consider the question

If $V \subset \mathbb{A}^n$ is an algebraic set and $\mathbb{A}^n \to \mathbb{A}^{n-d}$ is projection to
the last $n - d$ coordinates, what can we say about the image of
V in $\mathbb{A}^{n-d}$?

To answer this, we need

DEFINITION 2.5.15. Let $\mathfrak{a} \subset k[X_1, \ldots, X_n]$ be an ideal. Its dth *elimination
ideal*, $\mathfrak{a}_d$, is defined by

$$\mathfrak{a}_d = \mathfrak{a} \cap k[X_{d+1}, \ldots, X_n]$$

REMARK. In other words, $\mathfrak{a}_d$ "eliminates" the first d coordinates.

And the answer to the question posed above is

PROPOSITION 2.5.16. *If $\mathfrak{a} \in k[X_1, \ldots, X_n]$ is a radical ideal generating an alge-
braic set $\mathcal{V}(\mathfrak{a}) \subset \mathbb{A}^n$ and $p\colon \mathbb{A}^n \to \mathbb{A}^{n-d}$ is projection to the last $n - d$ coordinates,
then*

$$\mathcal{I}(p(\mathcal{V}(\mathfrak{a}))) = \mathfrak{a}_d$$

in the notation of 2.1.4 on page 38.

REMARK. Note that this does *not* compute $p(\mathcal{V}(\mathfrak{a}))$ — at best it computes
its Zariski closure, $\overline{p(\mathcal{V}(\mathfrak{a}))}$. For instance, the hyperbola

$$XY = 1$$

in $\mathbb{A}^2$ projects to $\mathbb{A}^1 \setminus \{0\}$ in the Y-axis, but its elimination ideal is $(XY - 1) \cap$
$k[Y] = (0)$, which defines the *entire* Y-axis.

PROOF. The projection

$$p\colon \mathbb{A}^n \to \mathbb{A}^{n-d}$$

induces the map

$$p^*\colon k[X_{d+1}, \ldots, X_n] \hookrightarrow k[X_1, \ldots, X_n]$$

of coordinate rings and a function $f \in k[X_{d+1}, \ldots, X_n]$ vanishes on $p(\mathcal{V}(\mathfrak{a}))$
if and only if $f \circ p$ vanishes on $\mathcal{V}(\mathfrak{a})$. This, in turn, happens if and only if
$p^*(f) \in \mathfrak{a}$. □

Computing elimination ideals is usually done using proposition 2.3.13 on
page 52, which states that, if $\mathfrak{a} \subset k[X_1, \ldots, X_n]$ is an ideal and

$$G = \{g_1, \ldots, g_s\}$$

is a Gröbner basis for $\mathfrak{a}$ computed with respect to a lexicographical ordering of
monomials using $X_1 \succ X_2 \succ \cdots \succ X_n$, then

$$G \cap k[X_{d+1}, \ldots, X_n]$$

is a Gröbner basis for $\mathfrak{a}_d$.

In the example above, Maple computes the Gröbner basis of $\mathfrak{a}$ with respect to a lexicographical ordering of monomials with $X \succ Y \succ Z$ as

$$\mathfrak{a} = (-Y^2 + Z^2 + 4\,Z^2Y^2, -4\,Z^2Y + 2\,Z^2X + Y - Z,$$
$$- 2\,ZY + ZX + YX, -X + Z + ZX^2)$$

The only one of these four terms lacking the variable X is the *first*. The fact that none of these terms only contains the variable Z implies that $\mathfrak{a}_2 = (0)$.

2.5.3. Rational functions. If V is an irreducible affine variety, its coordinate ring $k[V]$ is an integral domain by proposition 2.4.18 on page 68, so we can form its field of fractions.

DEFINITION 2.5.17. Let V be an irreducible affine variety with coordinate ring $k[V]$. If we form the field of fractions, $k(V)$, we get an extension of the field k: The *field of rational functions* on V.

Unlike regular functions, rational functions on an algebraic set $V \subset \mathbb{A}^n$ do not generally extend in a simple way to all of $\mathbb{A}^n$.

PROPOSITION 2.5.18. *Let $V \subset \mathbb{A}^n$ be an irreducible algebraic set with rational function field $k(V)$. If $\varphi \in k(V)$ then there exist representations of*

$$\frac{f_i}{g_i} \in k(\mathbb{A}^n)$$

$i = 1, \ldots, k$, *with the property that*

(1) for each i

$$\frac{f_i}{g_i}$$

is well-defined on an open set, U_i where $g_i \neq 0$

(2) for each i, j

(2.5.2) $$\frac{f_i}{g_i}|U_i \cap U_j \cap V = \frac{f_j}{g_j}|U_i \cap U_j \cap V$$

The open sets $U_i \cap V$ will be open affines as in definition 2.4.19 on page 68. The union of all of these open affines is called the open set of regular points, $R(\varphi)$, of φ.

PROOF. Equation 2.5.2 simply says that

$$f_i g_j - f_j g_i \in \mathcal{I}(V)$$

The only thing to be proved is that the $U_i \cap V$ are affine algebraic sets. This follows from proposition 2.4.5 on page 63 and the fact that $U_i \cap V = D(\bar{g}_i)$, a principal open set, where $\bar{g}_i$ is the image of g_i in $k[V]$.

The open set $R(\varphi)$ might not have an affine structure (unions of principal open sets do not always have such a structure) but the *intersection* of the $D(\bar{g}_i)$ *will* have such a structure by proposition 2.4.5 on page 63. It will still be a dense open set in V. $\qquad\square$

Here is an example:
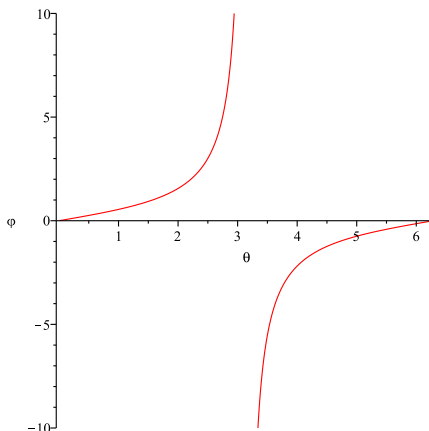
FIGURE 2.5.3.   A rational function on a unit circle

EXAMPLE 2.5.19. Suppose C is the unit circle with coordinate ring

$$k[C] = k[x,y] = k[X,Y]/(X^2 + Y^2 - 1)$$

where $X \mapsto x$ and $Y \mapsto y$ and consider the function $\varphi \in k(C)$ represented by

$$\frac{f_1}{g_1} = \frac{1-x}{y} \in k(x,y)$$

on the open set U_1 defined by $y \neq 0$. It would seem that this function fails to be regular when $y = 0$ but we can find a second representative

$$\frac{1-x}{y} \cdot \frac{1+x}{1+x} = \frac{1-x^2}{y(1+x)} \sim \frac{y^2}{y(1+x)} = \frac{y}{1+x} = \frac{f_2}{g_2}$$

valid on the open set U_2, defined by $1 + x \neq 0$. Here $\sim$ means "apply the identities valid in $k[C]$." The set of regular points of this function is $U_1 \cup U_2 \cap C$, or $x \neq -1$.

   If we parametrize this circle by angle, figure 2.5.3 shows that this function does blow up at the one point represented by $\theta = \pi$ or $x = -1$.

   Since the coordinate ring determines all geometric properties of an affine variety, one might ask what its field of fractions determines. The following is a partial answer:

LEMMA 2.5.20. *Let V be an irreducible affine variety with open affine, U. Then*

   (1) the coordinate ring, $k[U]$, is an integral domain.
   (2) the homomorphism of coordinate rings

$$k[V] \to k[U]$$

   induced by the inclusion $U \to V$ is injective.
   (3) the induced map of rational function fields

$$k(V) \to k(U)$$

   is an isomorphism.

PROOF. Since the $D(h)$ form a basis for the Zariski topology of V, we will have $D(h) \subset U$ for some h. The coordinate ring of $D(h)$ is $k[V]_h$ — see proposition 2.4.5 on page 63. This is an integral domain since $k[V]$ is, and we get a sequence of maps

(2.5.3) $$k[V] \to k[U] \to k[D(h)] = k[V]_h$$

We claim these are all inclusions. Suppose $f \in k[V]$ maps to 0 in $k[U]$ or $k[D(h)]$. Then $f|U = 0$ or $f|D(h) = 0$. This is impossible since the set of points where f is *nonzero* is *open* (it is $D(f)$) and all open sets in V have nonempty intersection — see proposition 2.4.17 on page 68.

We conclude that $k[U]$ is an integral domain, since it is a subring of the integral domain, $k[V]_h$. The sequence of inclusions in equation 2.5.3 induces inclusions of rational function fields

$$k(V) \to k(U) \to k(D(h))$$

and the conclusion follows from the fact that the field of fractions of $k[V]_h$ must be the same as that of $k[V]$ so $k(D(h)) = k(V)$. $\square$

It is encouraging that a rational function that is regular *everywhere* is actually regular:

PROPOSITION 2.5.21. *Suppose V is an irreducible affine variety and suppose $\varphi \in k(V)$ is regular everywhere on V. Then φ is in the image of $k[V]$ under the standard inclusion*

$$k[V] \hookrightarrow k(V)$$

PROOF. Embed V in $\mathbb{A}^n$, making it an irreducible algebraic set (see lemma 2.5.6 on page 74). Suppose we have a representation of φ as in proposition 2.5.18 on page 81. Then, since φ is regular everywhere on V we must have

$$\mathcal{I}(V) + \sum_{i=1}^{k}(g_i)$$

has *no common zero* in $\mathbb{A}^n$ (i.e., the g_i might have a common zero *off* of V, which is compensated for by adding $\mathcal{I}(V)$).

Hilbert's Nullstellensatz implies that

$$\mathcal{I}(V) + \sum_{i=1}^{k}(g_i) = (1)$$

or there exist $u_i \in k[\mathbb{A}^n]$ such that

$$1 = J + \sum_{i=1}^{k} u_i g_i$$

with $J \in \mathcal{I}(V)$. Reducing modulo $\mathcal{I}(V)$ gives

$$1 = \sum_{i=1}^{k} \bar{u}_i \bar{g}_i$$

where $\bar{u}_i, \bar{g}_i \in k[V]$ denote images in $k[V]$. If we multiply both sides of this equation by

$$\varphi = \frac{\bar{f}_1}{\bar{g}_1} = \cdots = \frac{\bar{f}_k}{\bar{g}_k}$$

(multiply the first term of the sum by the first representation of φ, and the second by the second, and so on) we get

$$\varphi = \sum_{i=1}^{k} \bar{u}_i \bar{f}_i \in k[V]$$

$\square$

We can also define a more general class of maps:

DEFINITION 2.5.22. If $V \subset \mathbb{A}^n$ is an irreducible algebraic set, a *rational map* $f: V \to \mathbb{A}^m$ is one of the form

$$f = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_m \end{bmatrix}$$

with $\varphi_1, \ldots, \varphi_m \in k(V)$. The set of points where all of the φ_i are regular is called the *regular set* of the map, denoted $R(f)$. This is an open set that contains an affine open set.

If $W \in \mathbb{A}^m$ and $f(V) \subset W$ then f is called a *rational map from V to W*.

REMARK. The set of regular points contains the open set that is the intersection of the open sets $U(\varphi_i)$ defined in proposition 2.5.18 on page 81, so it is an affine algebraic set.

EXAMPLE 2.5.23. For instance, let H be the hyperbola $xy = 1$ in $\mathbb{A}^2$. Then the map

$$\begin{aligned} f: \mathbb{A}^1 &\to H \\ x &\mapsto (x, 1/x) \end{aligned}$$

is rational where $R(f) = \{x \in \mathbb{A}^1 | x \neq 0\}$.

We also can classify algebraic sets in a coarser manner than isomorphism:

DEFINITION 2.5.24. Irreducible algebraic sets V and W are *birationally equivalent* if there exist rational maps

$$\begin{aligned} f: V &\to W \\ g: W &\to V \end{aligned}$$

such that $f \circ g | R = 1: W \to W$ and $g \circ f | R' = 1: V \to V$ where R is the set of points where $f \circ g$ is regular and R' is the set of points where $g \circ f$ is regular.

EXAMPLE. Returning to example 2.5.23, we see that the hyperbola, H is birationally equivalent to $\mathbb{A}^1$ — the inverse mapping is just the projection $(x, y) \mapsto x$. The two algebraic sets are not isomorphic, however, since their coordinate rings are not isomorphic:

The coordinate ring of $\mathbb{A}^1$ is $k[X]$ and that of H is $k[T, T^{-1}]$ — see example 2.4.2 on page 62. The only units in $k[X]$ are elements of k itself (degrees of polynomials add when you multiply them). Any homomorphism $f: k[T, T^{-1}] \to k[X]$ would, therefore, have to map T to an element of k forcing $f(k[T, T^{-1}]) \subset k \subset k[X]$.

It follows that birational equivalence as a system of classifying algebraic sets is *strictly weaker* than isomorphism.

PROPOSITION 2.5.25. *Let $f: V \to W$ be a rational map of irreducible varieties whose image is dense in W. Then f induces a homomorphism of fields over k*

(2.5.4) $f^*: k(W) \to k(V)$

Conversely, any such homomorphism is induced by a rational map whose image is dense.

REMARK. See figure 2.1.2 on page 38 and the discussion surrounding it for the significance of the image of a map being dense.

PROOF. By lemma 2.4.15 on page 67, the restriction of this to an open affine, $D(g) \subset V$, where f is regular induces an injection

$$k[W] \xrightarrow{(f|D(g))^*} k[D(g)] \subset k(V)$$

(see proposition 2.4.5 on page 63). This induces a homomorphism of fields of fractions

$$k(W) \to k(V)$$

Conversely, if

$$g: k(W) \hookrightarrow k(V)$$

is a homomorphism of fields (and homomorphisms of fields are *always* injective), consider the composite

$$\bar{g}: k[W] \hookrightarrow k(W) \hookrightarrow k(V)$$

Let $x_1, \ldots, x_n$ be generators of $k[W]$ as an algebra over k, and let their images be

$$\left\{ \frac{p_1}{q_1}, \ldots, \frac{p_n}{q_n} \right\} \in k(V)$$

with $p_i, q_i \in k[V]$ for $i = 1, \ldots, n$. It is not hard to see that

$$\left\{ \frac{p_1}{q_1}, \ldots, \frac{p_n}{q_n} \right\} \in k[U] \subset k(V)$$

(see proposition 2.4.5 on page 63) where

$$U = \bigcap_{i=1}^n D(q_i)$$

so we get an injective map

$$k[W] \hookrightarrow k[U]$$

induced by a regular map

$$U \to W$$

whose image is dense (since the map of coordinate rings was injective — see lemma 2.4.15 on page 67). $\square$

COROLLARY 2.5.26. *Two irreducible affine varieties are birationally equivalent if and only if their fields of rational functions are isomorphic.*

The following result clarifies the relationship between isomorphism and birational equivalence:

THEOREM 2.5.27. *Two irreducible affine varieties, V and W are birationally equivalent if and only if there exist open affine subsets $A \subset V$ and $B \subset W$ that are isomorphic.*

REMARK. This theorem is frequently stated using the wording "open sets that are isomorphic". Unfortunately, the term "isomorphic" is meaningless unless the open sets have some sort of structure. And exercise 7 on page 90 shows that there are open sets that have no affine structure.
It won't be possible to define isomorphisms of *arbitrary* open sets until we define *general* (i.e., non-affine) algebraic varieties in section 4.6 on page 205. It turns out that an arbitrary open set in an algebraic set (or even an algebraic variety) is a *general* algebraic variety (see corollary 4.6.9 on page 207).

PROOF. Clearly, if these open affines exist and are isomorphic, then

$$k(A) \cong k(V) \cong k(W) \cong k(B)$$

by lemma 2.5.20 on page 82.
Conversely, suppose $f \colon k(V) \to k(W)$ is an isomorphism with inverse $g \colon k(W) \to k(V)$.
Suppose

$$k[V] = k[x_1, \ldots, x_t]$$

where the $x_i \in k[V]$ generate it as a ring.
If

$$f(x_i) = \frac{a_i}{b_i} \in k(W)$$

then $f(k[V]) \subseteq k[W]_h$ where $h = b_1 \cdots b_t \in k[W]$. Then $k[W]_h$ is an open set, $C \subset W$, such that $f|f^{-1}(C)$ is regular.
It follows that $f|k[V] \colon k[V] \to k[W]_h = k[C]$ is *injective*, and we identify $k[V]$ with its image under f. We get

(2.5.5) $$k[V] \subset k[W]_h \subset k(V) = k(W)$$

Let

$$k[W]_h = k[V][y_1, \ldots, y_s]$$

where $y_i \in k(V)$ so that $y_i = c_i/d_i$, with $c_i, d_i \in k[V]$. If $\ell = d_1 \cdots d_s$, then

$$\ell \cdot y_i \in k[V]$$

for all i, which implies that $k[W]_h \subset k[V]_\ell$, and even $k[W]_{h \cdot \ell} \subset k[V]_\ell$. Since equation 2.5.5 implies the *reverse* inclusion, it follows that $k[V]_\ell = k[W]_{h \cdot \ell}$ and these define the isomorphic open sets.   □

Here's an application of birational equivalence:

DEFINITION 2.5.28. An affine variety will be called *rational* if it is birationally equivalent to an affine space.

One *application* of rationality is in computing certain indefinite integrals:

EXAMPLE 2.5.29. Suppose

$$f(x,y) = 0$$

defines y implicitly as a function of x and we want to compute an integral

$$\int g(x, y(x))\, dx$$

where $g(x, y)$ is a rational function.

Rationality of the curve $f(x,y) = 0$ implies that we can find a rational parametrization

$$x(t) = \varphi_1(t)$$
$$y(t) = \varphi_2(t)$$

converting the integral into

$$\int g(\varphi_1(t), \varphi_2(t)) \cdot \varphi_1'(t)\, dt$$

which can be computed using elementary functions. This process is called Euler Substitution. For instance, suppose we want to compute the integral

$$\int \frac{\sqrt{x^2 + x + 1}}{x + 1}\, dx$$

We rewrite this as the integral

$$\int \frac{y}{x+1}\, dx$$

where

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

The point $(0, 1)$ is on this curve that this defines, so we set $y - 1 = tx$ or $y = tx + 1$ so get

$$t^2 x^2 + 2tx + 1 = x^2 + x + 1$$

which gives

(2.5.6)
$$x = \frac{1 - 2t}{t^2 - 1}$$
$$y = tx + 1$$
$$= \frac{-t^2 + t - 1}{t^2 - 1}$$

and our integral becomes

(2.5.7) $$\int \frac{-t^2 + t - 1}{t(t - 2)} \cdot \frac{2(-t^2 + t - 1)}{(t^2 - 1)^2}\, dt =$$
$$-1/2 \ln (t - 1) - \frac{3}{2(t + 1)} + 1/2 \ln (t + 1)$$
$$+ \ln (t - 2) - \ln (t) + \frac{1}{2(t - 1)}$$

If we solve equation 2.5.6 for t we get

$$t = \frac{-1 \pm \sqrt{x^2 + x + 1}}{x}$$
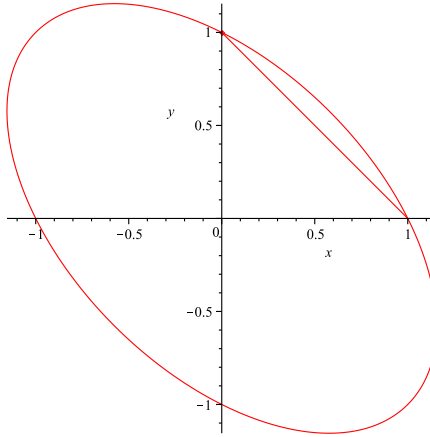
FIGURE 2.5.4.   Parametrization of an ellipse

which we can plug into equation 2.5.7 on the preceding page to get our indefinite integral.

EXAMPLE 2.5.30. Consider the ellipse defined by

$$(2.5.8) \qquad\qquad x^2 + xy + y^2 = 1$$

We can find a parametrization by taking any rational point of the curve, like $(1,0)$ and considering where the line of slope t passing through the curve intersects it. See figure 2.5.4

Set

$$(2.5.9) \qquad\qquad y = t(x-1)$$

and substitute this into equation 2.5.8 to get

$$x^2 + xt\,(x-1) + t^2\,(x-1)^2 = 1$$

which we solve for x to get

$$(2.5.10) \qquad\qquad \begin{aligned} x &= \frac{t^2-1}{t^2+t+1} \\ y &= t(x-1) \\ &= -\frac{t\,(t+2)}{t^2+t+1} \end{aligned}$$

This is a rational map from $\mathbb{A}^1$ to the curve. The inverse comes from how we defined t in the first place

$$\begin{aligned} y &= t(x-1) \\ t &= \frac{y}{x-1} \end{aligned}$$

We can also parametrize the 2-sphere:
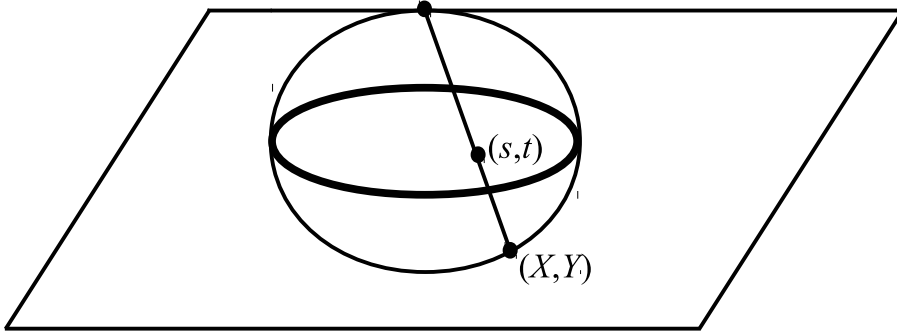
FIGURE 2.5.5.   Rationalization of the two-sphere

EXAMPLE 2.5.31. To find an implicit equation for this, we rationally parametrize the unit sphere by setting

$$\begin{aligned} s &= X/(1-Z) \\ t &= Y/(1-Z) \end{aligned}$$

and plugging into

$$X^2 + Y^2 + Z^2 = 1$$

to get

$$s^2(1-Z)^2 + t^2(1-Z)^2 + Z^2 = 1$$

which we can solve for Z to get

$$Z = \frac{s^2 + t^2 - 1}{s^2 + t^2 + 1}$$

and

$$\begin{aligned} X &= 2s/(1+s^2+t^2) \\ Y &= 2t/(1+s^2+t^2) \end{aligned}$$

Geometrically, this is depicted in figure 2.5.5. It follows that the 2-sphere, S^2, is *birationally equivalent* to $\mathbb{A}^2$! This shows how *coarse* the relation of birational equivalence is.

Theorem 2.5.27 on page 86 implies that an affine open set of S^2 must be isomorphic to one of $\mathbb{A}^2$. We show that the open set $D_{Z-1}(S^2) = k[S^2 \setminus (0,0,1)]$ is *isomorphic* to the open set $P = \mathbb{A}^2 \setminus S'$, where S' is the circle of radius i defined by

$$1 + s^2 + t^2 = 0$$

The coordinate ring of P is

$$A = k[P] = k[S, T, W]/(W(1 + S^2 + T^2) - 1)$$

and that of $D_{Z-1}(S^2)$ is

$$B = k[X, Y, Z, U]/(X^2 + Y^2 + Z^2 - 1, U(1 - Z) - 1)$$

(see definition 2.4.4 on page 63 and proposition 2.4.5 on page 63). The following maps are easily verified to be isomorphisms:

$$
\begin{aligned}
f\colon A &\to B \\
S &\mapsto U \cdot X \\
T &\mapsto U \cdot Y \\
W &\mapsto (1 - Z)/2
\end{aligned}
$$

and

$$
\begin{aligned}
g\colon B &\to A \\
X &\mapsto 2S \cdot W \\
Y &\mapsto 2T \cdot W \\
Z &\mapsto (S^2 + T^2 - 1) \cdot W \\
U &\mapsto (1 + S^2 + T^2)/2
\end{aligned}
$$

By corollary 2.5.26 on page 86, the fraction-fields of

$$k[X, Y, Z]/(X^2 + Y^2 + Z^2 - 1]$$

and $k[S, T]$ must be isomorphic.

Another interesting application of rationality is to *number theory.* Suppose we want to know all of the *rational* solutions of the equation in example 2.5.30 on page 88. Since any rational values for x and y give a rational value for t in equation 2.5.9 on page 88 and rational values for t give rational values for x and y, it follows that the rational solutions of equation 2.5.8 on page 88 are precisely those given by equation 2.5.10 on page 88 with *rational* values of t.

EXERCISES.

6. If V is an irreducible affine variety and $f \in k(V)$ is a rational function, show that there exists an open set $U \subset V$ where f is regular — i.e., $f \in k[U]$.

7. Show that $U = \mathbb{A}^2 \setminus \{(0,0)\} = D(X) \cup D(Y) \subset \mathbb{A}^2$ is an open set that is not affine.

8. Determine the rationality of :
   a. The curve $Y^2 = X^3$.
   b. The circle $X^2 + Y^2 = 1$
   c. The curve $Y^2 = X^3 + X^2$

9. At which points of the affine variety $Y^2 = X^2 + X^3$ is the rational function $\frac{Y}{X}$ defined?

10. Compute all *rational* solutions to the equation

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

11. Why can't we use the rationalization in example 2.5.30 on page 88 to find solutions of

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

in the field $\mathbb{F}_5$?

12. If k is a field of characteristic 0, show that the polynomial $XY + X^2U + Y^2V \in k[X, Y, U, V]$ is irreducible.

13. Here is an example of an open affine that is not principal. The author is indebted to Hailong Dao on the online discussion group, `http://mathoverflow.net` for it.

Let V be the algebraic set with coordinate ring

$$R = k[x, y, u, v] = k[X, Y, U, V]/(XY + X^2U + Y^2V)$$

where the lowercase letters are the images of the corresponding uppercase indeterminates and we know that the denominator is irreducible by exercise 12. Suppose the field of fractions of R is F and consider the open set

$$D = D(x) \cup D(y)$$

By the same reasoning in the solution of exercise 7, we conclude that — *if D is affine*

$$k[D] = R_x \cap R_y$$

In exercise 7, we got $k[D] = R$, which led to a contradiction. In the present case, set

$$
\begin{aligned}
f &= -\frac{u}{y} \\
&= \frac{x + yv}{x^2}
\end{aligned}
$$

(by the identities that hold in R). This is clearly contained in $R_x \cap R_y \setminus R$. Now we also define

$$
\begin{aligned}
g &= -\frac{v}{x} \\
&= \frac{y + xu}{y^2}
\end{aligned}
$$

which is also contained in $R_x \cap R_y \setminus R$. An easy computation shows

$$x \cdot f + y \cdot g = 1 \in F$$

Why does this imply that $D(x) \cup D(y)$ is affine?

2.6. Applications to optimization theory

This is a brief sketch of material from the paper [**34**], which applies algebraic geometry to problems that arise from linear programming.

Recall that linear programming minimizes or maximizes a linear function (called the *objective function*) over a polygonal region (called the *feasible region*) defined by a finite number of inequalities. Basic calculus tells us that the extrema will lie on the vertices of this feasible region. The traditional Simplex Method involves "crawling" around the boundary of the polygon searching for this extremum. While the simplex method is fairly efficient in general, it

has the flaw that the number of vertices to be checked is an *exponential* function of the dimension of the polygon (for instance an n-cube has 2^n vertices).

John von Neumann suggested the so called *interior point method* which was perfected by Narendra Karmarkar(see [**86**]) in 1984. The idea here is that we modify the objective function by making it nonlinear so that ordinary calculus techniques work.

Suppose we have put our linear programming problem into normal form:

Maximize

$$(2.6.1) \qquad\qquad f(x) = c^T x$$

subject to the conditions that

$$(2.6.2) \qquad\qquad \begin{aligned} x &\geq 0 \\ Ax &= b \end{aligned}$$

where

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \, c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \, b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

and A is an $m \times n$ matrix. In this problem, the feasible region is the set of solutions to equation 2.6.2 and its boundary consists of points where at least one of the $x_i = 0$.

In interior point methods, we replace the linear objective function in equation 2.6.1 by

$$(2.6.3) \qquad\qquad f_\lambda(x) = c^t x + \lambda \cdot \sum_{i=1}^n \ln(x_i)$$

where $\lambda > 0$ is a real parameter. Note that this coincides with equation 2.6.1 when $\lambda = 0$. Over $\mathbb{R}$, $f_\lambda(x)$ is strictly concave[4] so it has a unique maximum, $x^*(\lambda)$, inside the feasible region. As we vary λ, $x^*(\lambda)$ traces out a curve in the feasible region, called the *primal central path*. Since $f_\lambda(x) \to -\infty$ as x approaches the boundary of the feasible region, we constrain the x_i to be nonzero by setting

$$x_i \cdot s_i = \lambda$$

for $i = 1, \ldots n$ — recall how we forced a variable to be nonzero in example 2.4.6 on page 64. The variables s_i have a meaning in linear programming (they are called *slack variables*) but we will not be concerned with them here.

Intuition suggests that x_λ^* approaches a solution to the original linear programming problem as $\lambda \to 0$, and this is indeed the case — see [**167**]. To compute x_λ^*, we form the partial derivatives of $f_\lambda(x)$ in equation 2.6.3 with respect to the x_i and set them to 0 — as with conventional optimization.

DEFINITION 2.6.1. If we define the *central sheet of the optimization* problem, $\mathscr{L}_{A,c}^{-1}$, to be the Zariski closure of

$$\left\{ \frac{1}{u_1}, \ldots, \frac{1}{u_n} \in \mathbb{C}^n \,|\, u_1, \ldots u_n \in \mathrm{Span}(A_i, *, c) \text{ and } \neq 0 \right\}$$

---

[4]This means that $f_\lambda(t \cdot x_1 + (1-t) \cdot x_2) > t \cdot f_\lambda(x_1 + (1-t) \cdot f_\lambda(x_2)$ for $t \in (0,1)$.

Given this, Sturmfels et al, show that the Zariski closure of $\{x^*_\lambda, \lambda \geq 0\}$ is the intersection of $\mathscr{L}^{-1}_{\mathbf{A},\mathbf{c}}$ with the affine subspace defined by $\mathbf{A}\mathbf{x} = \mathbf{b}$ and they compute its defining ideal (using a Gröbner basis provided by [**138**]). The goal of these computations is to estimate the *total curvature* of $\{x^*_\lambda, \lambda \geq 0\}$, giving an estimate of the running time of the optimization problem. This involves using Bézout's Theorem, and the theory of divisors (presented in section 5.9 on page 284). See [**34**] for more details.

2.7. Products

There are many ways we could define products of affine varieties. Simply taking the topological product of the underlying spaces is probably not the correct one. For instance, using this definition of product would imply that

$$\mathbb{A}^n \times \mathbb{A}^m \neq \mathbb{A}^{n+m}$$

(recall exercise 1 on page 38) in the Zariski topology. We have already seen the category-theoretic concept of product — see definition A.5.1 on page 439. If Z is the category-theoretic product of A and B, we get

$$\hom(W, Z) \leftrightarrow \hom(W, A) \times \hom(W, B)$$

is a *bijection*. Since the underlying set of points of an affine variety, V, is $\hom_{\mathrm{Var}}(\mathbb{A}^0, V)$, we claim that

$$
\begin{aligned}
(\text{points of } V \times W) \quad &= \quad \hom_{\mathrm{Var}}(\mathbb{A}^0, V \times W) \\
&= \quad \hom_{\mathrm{Var}}(\mathbb{A}^0, V) \times \hom_{\mathrm{Var}}(\mathbb{A}^0, W) \\
&= \quad (\text{points of } V) \times (\text{points of } W)
\end{aligned}
$$

so the underlying set of points is a Cartesian product (although its topology is different).

2.7.1. Tensor products and affine varieties. If A and B are k-algebras, we can define the structure of a k-algebra on $A \otimes_k B$ by defining

$$(a \otimes b) \cdot (c \otimes d) = (a \cdot c) \otimes (b \cdot d)$$

PROPOSITION 2.7.1. *If $A = k[X_1, \ldots, X_n]$ and $B = k[Y_1, \ldots, Y_m]$, then there is an isomorphism of k-algebras*

$$A \otimes_k B \xrightarrow{\cong} k[X_1, \ldots, X_n, Y_1, \ldots, Y_m]$$

REMARK. So $k[\mathbb{A}^n] \otimes_k k[\mathbb{A}^m] = k[\mathbb{A}^{n+m}]$.

PROOF. As vector-spaces A has a basis $\{m_\alpha(X_i)\}$ of monomials in the X_i and B has a basis $\{m_\beta(Y_j)\}$ of monomials in the Y_j. The set

$$\{m_\alpha(X_i) \otimes m_\beta(Y_j)\}$$

are a basis for $A \otimes_k B$ (see exercise 15 on page 472) and this set is in a one-to-one correspondence with monomials in the X_i and Y_j. The isomorphism sends $m_\alpha(X_i) \otimes m_\beta(Y_j)$ to $m_\alpha(X_i) \cdot m_\beta(Y_j)$. $\qquad\square$

COROLLARY 2.7.2. *If $A = k[X_1, \ldots, X_n]/\mathfrak{a}$ and $B = k[Y_1, \ldots, Y_m]/\mathfrak{b}$, then there is an isomorphism of k-algebras*

$$A \otimes_k B \xrightarrow{\cong} k[X_1, \ldots, X_n, Y_1, \ldots, Y_m]/(\mathfrak{a}, \mathfrak{b})$$

If A and B are affine k-algebras, so is $A \otimes_k B$.

PROOF. The tensor product of the projections (see proposition A.5.47)

$$(2.7.1) \qquad\qquad\qquad \begin{aligned} p : k[X_1, \ldots, X_n] &\rightarrow A \\ q : k[Y_1, \ldots, Y_m] &\rightarrow B \end{aligned}$$

gives a surjective homomorphism of rings

$$k[X_1, \ldots, X_n, Y_1, \ldots, Y_m] \rightarrow A \otimes_k B$$

We claim that the kernel is

$$\mathfrak{a} \cdot k[Y_1, \ldots, Y_m] + k[X_1, \ldots, X_n] \cdot \mathfrak{b} = (\mathfrak{a}, \mathfrak{b})$$

Certainly,

$$\mathfrak{a} \cdot k[Y_1, \ldots, Y_m] \subset \ker p \otimes q$$

since $p(\mathfrak{a}) = 0$ so $\mathfrak{a} \otimes_k k[Y_1, \ldots, Y_m]$ maps to $0 \otimes_k B = 0$. The same reasoning applies to the second term, so we conclude that

$$(\mathfrak{a}, \mathfrak{b}) \subset \ker p \otimes q$$

Temporarily forget the ring structures here and regard all objects as vector spaces. As vector spaces

$$\begin{aligned} k[X_1, \ldots, X_n] &= A \oplus \mathfrak{a} \\ k[Y_1, \ldots, Y_m] &= B \oplus \mathfrak{b} \end{aligned}$$

Proposition A.5.48 on page 464 implies, that, as *vector spaces*

$$\begin{aligned} k[X_1, \ldots, X_n] \otimes_k k[Y_1, \ldots, Y_m] &= A \otimes_k B \\ &\oplus A \otimes_k \mathfrak{b} \\ &\oplus \mathfrak{a} \otimes_k B \\ &\oplus \mathfrak{a} \otimes_k \mathfrak{b} \end{aligned}$$

and the kernel of the projection to the subspace $A \otimes_k B$ consists of

$$A \otimes_k \mathfrak{b} \oplus \mathfrak{a} \otimes_k B \oplus \mathfrak{a} \otimes_k \mathfrak{b} \subset (\mathfrak{a}, \mathfrak{b})$$

The only thing needed now is to show that $A \otimes_k B$ is an affine *k*-algebra, i.e., that it is *reduced*. Suppose $x \in A \otimes_k B$ is nilpotent. Then

$$x = \sum_{j=1}^{t} a_j \otimes b_j$$

with the a_j in A and the b_j in B. We may assume the b_j are all linearly independent because, if

$$b_1 = \sum_{i=2}^{t} f_i b_i$$

we could fold that into the tensor product

$$x = \sum_{j=2}^{t} (a_j + f_j a_1) \otimes b_j$$

For every maximal ideal, $\mathfrak{m}$, of A we have a homomorphism

$$A \to A/\mathfrak{m} = k$$

inducing

$$
\begin{aligned}
A \otimes_k B &\to k \otimes_k B = B \\
a_j \otimes b_j &\mapsto \bar{a}_j \cdot b_j \\
x &\mapsto \sum_{j=1}^{t} \bar{a}_j \cdot b_j
\end{aligned}
$$

where $\bar{a}_j \in k$ is the image of a_j under the projection $A \to A/\mathfrak{m} = k$.

The image of x under this map is nilpotent, hence 0 since B is an affine k-algebra. Since the $\{b_j\}$ are linearly independent, it follows that all of the $\bar{a}_j = 0$, and this is true for *any* maximal ideal of A. This implies that the a_j are all 0 (they are regular functions that vanish at *all points* of a variety), so $x = 0$. □

PROPOSITION 2.7.3. *Let A_1, A_2 be affine k-algebras. Then*

$$A_1 \otimes_k A_2$$

is a coproduct in the category $\mathscr{A}_k$ of affine k-algebras (see definition A.5.4 on page 441).

REMARK. The ring-structure is very important here. Tensor products are neither products nor coproducts in the category of vector-spaces.

PROOF. Corollary 2.7.2 on the preceding page implies that $A_1 \otimes_k A_2$ is an affine k-algebra. Recall that coproducts have canonical maps from their factors to themselves. We use the ring structure to define these:

$$
\begin{aligned}
A_1 &\to A_1 \otimes_k A_2 \\
r_1 &\mapsto r_1 \otimes 1
\end{aligned}
$$
(2.7.2)

and

$$
\begin{aligned}
A_2 &\to A_1 \otimes_k A_2 \\
r_2 &\mapsto 1 \otimes r_2
\end{aligned}
$$
(2.7.3)

Given homomorphisms of affine k-algebras $f_1 \colon A_1 \to W$ and $f_2 \colon A_2 \to W$, we can define

$$
\begin{aligned}
F \colon A_1 \otimes_k A_2 &\to W \\
r_1 \otimes r_2 &\mapsto f_1(r_1) \cdot f_2(r_2)
\end{aligned}
$$

and this is the *unique* map that is compatible with f_1, f_2 and the maps in equations 2.7.2 and 2.7.3.

Since the coordinate ring and specm $*$ are inverse contravariant functors, we define □

DEFINITION 2.7.4. If V and W are affine varieties, then

$$V \times W = \text{specm}(k[V] \otimes_k k[W])$$

REMARK. This construct will have the proper functorial properties of a product, as discussed in definition A.5.1 on page 439. And it is reassuring that

$$\mathbb{A}^n \times \mathbb{A}^m = \mathbb{A}^{n+m}$$

due to proposition 2.7.1 on page 93.

LEMMA 2.7.5. *Let V and W be irreducible affine varieties. Then $V \times W$ is also irreducible.*

REMARK. This implies that, if A and B are affine k-algebras that are integral domains, so is $A \otimes_k B$. Note that this algebraic fact is not true for arbitrary rings. For instance

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}$$

is *not* an integral domain.

PROOF. We will prove this by contradiction. Suppose $V \times W = Z_1 \cup Z_2$. For any point $x \in X$, the product $x \times Y$ is isomorphic to Y and, therefore, irreducible.
Since

(2.7.4) $x \times Y = ((x \times Y) \cap Z_1 \cup ((x \times Y) \cap Z_2)$

it follows that $x \times Y \subset Z_1$ or $x \times Y \subset Z_2$. We claim that the set, X_y, of points $x \in X$ with the property that $x \times y \subset Z_1$ is closed. This is because it is given by

$$X_y \times y = (X \times y) \cap Z_1$$

which is an intersection of closed sets. Then the set

$$X_1 = \bigcap_{y \in Y} X_y$$

is also closed since it is the intersection of closed sets. This set, X_1, is the set of $x \in X$ such that

$$x \times Y \subset Z_1$$

In like fashion, we can define a closed set X_2 and equation and the reasoning after it implies that $X = X_1 \cup X_2$. The irreducibility of X implies that $X = X_1$ or $X = X_2$ so $X \times Y = Z_1$ or Z_2.

One type of product that will interest us is the product of a variety by itself and the diagonal   □

DEFINITION 2.7.6. Let V be an affine variety and let $V \times V$ be its product with itself. The *diagonal map*

$$\Delta: V \to V \times V$$

is the unique morphism that makes the diagrams

(2.7.5)

commute, where $p_i: V \times V \to V$ with $i = 1, 2$, is projection onto the first and second factors.

REMARK. This map is unique in the category-theoretic sense, where a map to a product is determined by its composites with the maps to the factors. To actually get an idea of what this maps does, it is necessary to compute the induced map of coordinate rings.

It is a map that makes the diagrams

$$
\begin{array}{ccc}
& k[V] \otimes_k k[V] & \\
{\scriptstyle \Delta^*} \nearrow & & \uparrow {\scriptstyle p_i^*} \\
k[V] & \!\!=\!\!=\!\!=\!\!=\!\! & k[V]
\end{array}
$$

(gotten by reversing the arrows in the previous diagram) commute. Here $p_i^*: k[V] \to k[V] \otimes_k k[V]$ includes $k[V]$ as $k[V] \otimes 1$ or $1 \otimes k[V]$. It is not hard to see that the only algebra-homomorphism compatible with these is

$$
\begin{array}{rcl}
\Delta^*: k[V] \otimes_k k[V] & \to & k[V] \\
v_1 \otimes v_2 & \mapsto & v_1 \cdot v_2
\end{array}
$$

We will sometimes be interested in varieties equipped with extra structure.

DEFINITION 2.7.7. An affine variety, V, is called an *algebraic group* if it is equipped with a point p, called the *identity element*, and regular maps

$$
\begin{array}{rcl}
\mu: V \times V & \to & V \\
\iota: V & \xrightarrow{\cong} & V
\end{array}
$$

such that:

(1) the diagram

$$
\begin{array}{ccc}
& V \times \{p\} & \\
{\scriptstyle =} \nearrow & & \searrow {\scriptstyle \mu} \\
V \!\!=\!\!=\!\!=\!\!=\!\!=\!\!=\!\!=\!\! & & V \\
{\scriptstyle =} \searrow & & \nearrow {\scriptstyle \mu} \\
& \{p\} \times V &
\end{array}
$$

commutes, implying that $\{p\} \in V$ acts as the identity element.
(2) $\mu \circ (\mu \times 1) = \mu \circ (1 \times \mu): V \times V \times V \to V$ (associativity).
(3) if $\Delta: V \to V \times V$ is the diagonal map, the diagram

$$
\begin{array}{ccc}
V \times V \xrightarrow{1 \times \iota} V \times V & \\
{\scriptstyle \Delta} \nearrow & & \searrow {\scriptstyle \mu} \\
V & & p \in V \\
{\scriptstyle \Delta} \searrow & & \nearrow {\scriptstyle \mu} \\
V \times V \xrightarrow[\iota \times 1]{} V \times V &
\end{array}
$$

commutes, implying that $\iota: V \to V$ maps each point to its multiplicative inverse.

REMARK. The group, $SL(n,k) \subset \mathbb{A}^{n^2}$ defined in statement 6 on page 36 and $GL(n,k) \subset \mathbb{A}^{n^2+1}$ defined in example 2.4.6 on page 64 are both examples of algebraic groups. Their multiplication-maps are matrix-multiplication, hence regular maps, and their inversion maps are proved to be regular in exercise 5 on page 71. These are *affine* algebraic groups.

We will study *projective* algebraic groups (called abelian varieties) in section 6.2 on page 313.

EXERCISES.

1. If $W \subset V$ show that the diagonal map

$$\Delta: V \to V \times V$$

induces an isomorphism

$$\Delta: W \to \Delta(W) \subset V \times V$$

2. If V is an algebraic group and $a, b \in V$ are two points, show that there exists an isomorphism of affine varieties

$$f_{a,b}: V \to V$$

such that $f_{a,b}(a) = b$.

2.8. Dimension

2.8.1. Introduction. For centuries, concept of dimension has been used to define the "size" of a geometric object or the number of independent parameters needed to specify its points. In 1890, this simple notion of dimension was complicated by Peano's proof of the existence of a *continuous surjective* function

$$(2.8.1) \hspace{3cm} f: [0,1] \to [0,1] \times [0,1]$$

The early 20th century saw an effort to give a rigorous topological definition of dimension. The book [81] gives an excellent treatment of the topological theory of dimension that developed.

In algebraic geometry, we can fall back on the number of parameters needed to specify a point since it is hard to imagine a "space filling" curve that is algebraic. This leads to a concept of dimension that is not always compatible with the topological one. For instance, it is natural to regard $\mathbb{C}^n$ as an n-dimensional affine variety over $\mathbb{C}$, but its *topological* dimension is $2n$.

If an affine variety is over a *discrete* field like $\bar{\mathbb{Q}}$, it is topologically zero-dimensional, whereas we would like to think of $\mathbb{A}^n = \bar{\mathbb{Q}}^n$ as n-dimensional. Consequently, the dimension of an irreducible affine variety was generally defined to be the transcendence degree of its rational function-field (see [95]).

Unfortunately, using transcendence degree of the function field over the ground field has several drawbacks:

- It is undefined for reducible varieties
- It is undefined for rings that are not algebras over fields (which we will consider in chapter 4).

Consequently, we will want follow a different course in defining dimension. Consider the sequence of affine spaces

$$\mathbb{A}^0 \subsetneq \mathbb{A}^1 \subsetneq \cdots \subsetneq \mathbb{A}^n$$

each of which is an irreducible variety within the next. The fact that $\mathbb{A}^{i+1}$ is an irreducible implies that $\mathbb{A}^i \subset \mathbb{A}^{i+1}$ is not simply "smaller" but "thinner" — it is defined via a smaller set of free parameters.

We base our definition of dimension on this:

DEFINITION 2.8.1. If V is an affine variety, the *dimension* of V is defined to be the largest n that occurs in a sequence

$$V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$$

where each term is an irreducible subset of the next.

It is immediately clear that

PROPOSITION 2.8.2. *The dimension of an affine variety is the maximum of the dimensions of its irreducible components.*

We prove a few other seemingly-trivial properties of dimension:

PROPOSITION 2.8.3. *If V is an irreducible algebraic set and Z is a proper irreducible algebraic subset of V, then $\dim Z < \dim V$.*

PROOF. Suppose $\dim V = n$ and $\dim Z = m \geq n$. Then Z has a chain of irreducible subspaces

$$Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_m = Z$$

and this can be extended to

$$Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_m = Z \subsetneq V$$

which shows that $m + 1 \leq n$. □

DEFINITION 2.8.4. Let $W \subset V$ be an inclusion of irreducible affine varieties. Then the *codimension of W in V* is defined to be

$$\dim V - \dim W$$

PROPOSITION 2.8.5. *If V is a 0-dimensional irreducible affine variety, then V is a single point.*

PROOF. If V is not a point, it *contains* a point, p, so we have the following sequence of irreducibles (at least)

$$p \subsetneq V$$

□

PROPOSITION 2.8.6. *If $f\colon V \to W$ is a surjective regular map of irreducible affine varieties, then $\dim V \geq \dim W$.*

REMARK. This shows that a Peano-type curve like that in equation 2.8.1 on the preceding page is impossible in algebraic geometry.

PROOF. We do induction on $\dim V$. If $\dim V = 0$ then V and, therefore, W must be a single point by proposition 2.8.5 on the preceding page so the result is true.

Assume the result true whenever the dimension of V is $\leq n - 1$, suppose the dimension of V is n, and $\dim V = m$. Let

$$W_0 \subsetneq W_1 \subsetneq \cdots \subsetneq W_m = W$$

be a maximal irreducible sequence in W and let

$$A = f^{-1}(W_{m-1})$$

If

$$A = \bigcup_{i=1}^{k} Z_i$$

is a decomposition into irreducible components, then

$$W_{m-1} = \bigcup_{i=1}^{k} f(Z_i)$$

Since W_{m-1} is irreducible, it cannot be a union of proper irreducible subsets, so there exists an α such that

$$W_{m-1} = f(Z_\alpha)$$

The inductive hypothesis implies that

$$\dim Z_\alpha \geq \dim W_{m-1} = m - 1$$

and proposition 2.8.3 on the previous page implies that

$$\dim V > \dim Z_\alpha \geq m - 1$$

so $\dim V \geq \dim W$. $\square$

2.8.2. The Krull dimension of a ring. In commutative algebra, we have another definition of dimension:

DEFINITION 2.8.7. If R is a commutative ring, it is said to have *Krull dimension* $\dim R = d$ if the maximal length of a chain of distinct prime ideals

$$\mathfrak{p}_d \supsetneq \cdots \supsetneq \mathfrak{p}_0$$

is $d + 1$. We define the Krull dimension of the trivial ring to be -1.

Given a prime ideal $\mathfrak{p} \subset R$, its *height*, $\mathrm{ht}(\mathfrak{p})$ is defined as the maximum d that occurs in a chain of prime ideals

$$\mathfrak{p} = \mathfrak{p}_d \supsetneq \cdots \supsetneq \mathfrak{p}_0$$

REMARK. A field has Krull dimension 0 and a principal ideal domain or discrete valuation ring has dimension 1.

Although the possibility of a ring having dimension -1 seems odd, it is standard in inductive dimension theory for the empty set to have dimension -1 — see [**40**] or [**81**].

Theorem A.1.87 on page 382 could be restated as

> "A ring is Artinian if and only if it is noetherian of dimension 0."

Wolfgang Krull (1899–1971) was a German mathematician who made great contributions to commutative algebra. In 1938, he proposed the definition of dimension given above and proved that it coincides with transcendence degree for polynomial rings over a field (see[**98**]). He was responsible for many other notable results that can be applied to algebraic geometry, including the Krull Principal Ideal Theorem (2.8.29 on page 109).

In general, we have

PROPOSITION 2.8.8. *If R is a ring and $\mathfrak{p} \subset R$ is a prime ideal, the inequality*

$$\mathrm{ht}(\mathfrak{p}) + \dim R/\mathfrak{p} \leq \dim R$$

REMARK. This is usually the best one can say unless *R catenary* — see definition 2.8.17 on page 105.

PROOF. Lemma A.1.25 on page 352 implies that a sequence of distinct prime ideals in $R/\mathfrak{p}$ corresponds to a sequence of distinct prime ideals in R that contain $\mathfrak{p}$. If we splice this together with the maximal sequence

$$\mathfrak{p} = \mathfrak{p}_d \supsetneq \cdots \supsetneq \mathfrak{p}_0$$

we get an increasing sequence of prime ideals in R and the conclusion follows from definition 2.8.7 on the facing page. □

PROPOSITION 2.8.9. *The dimension of an affine variety, V, is equal to the Krull dimension of its coordinate ring.*

PROOF. This follows from proposition 2.4.18 on page 68, proposition 2.4.3 on page 62 and the fact that the correspondence

$$\mathfrak{a} \leftrightarrow \mathcal{V}(\mathfrak{a})$$

between radical ideals and subvarieties is order-inverting. Any sequence of irreducibles

$$V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$$

gives rise to a sequence of prime ideals

$$\mathcal{I}(V_0) \supsetneq \mathcal{I}(V_1) \supsetneq \cdots \subsetneq \mathcal{I}(V_n) = (0)$$

(in the notation of definition 2.1.4 on page 38) and vice-versa. □

The rest of this section will be devoted to proving that

$$\dim k[X_1, \ldots, X_n] = n$$

and some related results.

The sequence of prime ideals

$$0 \subset (X_1) \subset \cdots \subset (X_1, \ldots, X_n)$$

shows that

$$\dim k[X_1, \ldots, X_n] \geq n$$

The reverse inequality is surprising difficult to prove. We will follow the shortest and slickest treatment the author has ever seen in [**33**]. Coquand and Lombardi's ingenious argument shows that the Krull dimension of an integral domain is always bounded from above by the transcendence degree if its field of fractions. It is also interesting in that it doesn't require the rings in question to be noetherian.

DEFINITION 2.8.10. Let R be a commutative ring and let $x \in R$ be an element. Define the *boundary*, $R_{\{x\}}$ of x to be the ring of fractions $S_{\{x\}}^{-1} R$ where

$$S_{\{x\}} = \{z \in R | n \in \mathbb{Z}, n > 0, a \in R, z = x^n (1 + ax)\}$$

where n runs over all positive integers and a runs over all elements of R.

REMARK. Note that $R_{\{x\}} = 0$ if x is invertible or nilpotent because $0 \in S_{\{x\}}$ in either of these cases (see proposition A.1.90 on page 384).

This construct has the interesting property that it lowers Krull dimension:

THEOREM 2.8.11. *Let R be a commutative ring and let ℓ be a nonnegative integer. The following statements are equivalent:*
   (1) the Krull dimension of R is $\leq \ell$.
   (2) the Krull dimension of $R_{\{x\}}$ is $\leq \ell - 1$ for all $x \in R$.

PROOF. This follows from:
Claim 1: If $\mathfrak{m}$ is maximal, then $\mathfrak{m} \cap S_{\{x\}} \neq \emptyset$.
If $x \in \mathfrak{m}$, this is immediate. Otherwise x is invertible modulo $\mathfrak{m}$ so $1 + ax \in \mathfrak{m}$ for some $a \in R$.
Claim 2: If $\mathfrak{m} \subset R$ is a maximal ideal and $\mathfrak{p} \subset \mathfrak{m}$ is a prime ideal and $x \in \mathfrak{m} \setminus \mathfrak{p}$, then $S_{\{x\}} \cap \mathfrak{p} = \emptyset$.
Since $x \notin \mathfrak{p}$, the only way $x^k(1 + xy) \in \mathfrak{p}$ is for $1 + xy \in \mathfrak{p} \subset \mathfrak{m}$. This implies that $1 + xy \in \mathfrak{m}$, and $x \in \mathfrak{m}$ implies that $1 \in \mathfrak{m}$, a contradiction.
The first claim shows that forming the boundary lowers Krull dimension because the maximal ideals in R get killed off in $R_{\{x\}}$ (see corollary A.1.94 on page 385). The second shows that forming it with respect to a suitable element of R *only* kills the maximal ideals, so the Krull dimension is lowered by *precisely* 1. $\qquad \square$

EXAMPLE 2.8.12. If R is a ring and, for any $x \in R$, we can find $n > 0$ and $a \in R$ such that
$$x^n(1 + ax) = 0 \in S_{\{x\}}$$
then R has dimension ≤ 0. In this case $R_{\{x\}} = 0$, the trivial ring. Every element that is not a zero-divisor is a unit.
If R is required to be an integral domain, then $1 + ax = 0$ so x has a multiplicative inverse and R is a field.

The next example will be useful in characterizing Krull dimension:

EXAMPLE 2.8.13. If $S^{-1}R$ is a localized ring, then $S^{-1}R$ has Krull dimension ≤ 0 if and only if, for any $s_0^{-1}x \in S^{-1}R$, there exists an $n > 0$ and an $s_1^{-1}a \in S^{-1}R$ such that
$$(s_0^{-1}x)^n(1 + (s_1^{-1}a)(s_0^{-1}x)) = 0 \in S_{\{s^{-1}x\}}$$
or, equivalently
$$x^n(s_0 s_1 + ax) = 0 \in R$$
In other words, $S^{-1}R$ has Krull dimension ≤ 0 if and only if, for any $x_0 \in R$, there exists an $n > 0$ and an $a_0 \in R$ and an $s \in S$ such that
$$x_0^n(s + a_0 x_0) = 0 \in R$$

If we let $s = x_1^m(1 + a_1x_1) \in S_{\{x_1\}}$, theorem 2.8.11 on the facing page implies: The ring, R, has Krull dimension ≤ 1 if and only if for any $x_0, x_1 \in R$, there exist $a_0, a_1 \in R$ and exponents $n, m > 0$ such that

$$(2.8.2) \qquad x_0^n(x_1^m(1 + a_1x_1) + a_0x_0) = 0 \in R$$

THEOREM 2.8.14. *Let $\ell > 0$ be some integer and let R be a ring. Then R has Krull dimension $\leq \ell$ if and only if, for any $x_0, \dots, x_\ell \in R$, there exist positive integers $n_0, \dots, n_\ell$ and elements $a_0, \dots, a_\ell \in R$ such that*

$$(2.8.3) \quad x_0^{n_0}\left(\cdots x_\ell^{n_\ell}(1 + a_\ell x_\ell) + a_{\ell-1}x_{\ell-1}\right)\cdots + a_0x_0\right) =$$
$$x_0^{n_0} \cdots x_\ell^{n_\ell} + a_\ell x_0^{n_0} \cdots x_\ell^{n_\ell+1} + a_{\ell-1}x_0^{n_0} \cdots x_{\ell-1}^{n_{\ell-1}+1}$$
$$+ \cdots + a_0 x_0^{n_0+1} = 0$$

REMARK. The equation on the top line is the result of replacing the unique term equal to 1 in equation 2.8.2, by an element of $S_{\{x_2\}}$, resulting in an equation with a unique term equal to 1, replacing *that* 1 with an element of $S_{\{x_3\}}$ and so on.

This and the following theorem may be called an *equational characterization of Krull dimension*.

PROOF. The statement about Krull dimension follows by induction, using theorem 2.8.11 on the facing page — induction that was begun by equation 2.8.2.

The second statement about the form of the equation follows from the fact that the leading monomial (with coefficient 1) corresponds to the unique 1 in the top line of equation 2.8.3. An iteration of the construction multiplies this by $x_i^{n_i}(1 + a_ix_i)$ and leaves all other terms alone.   □

THEOREM 2.8.15. *Let k be a field and let R be an algebra over k. If all sequences $x_0, \dots, x_\ell \in R$ are algebraically dependent then R has Krull dimension $\leq \ell$.*

REMARK. If we want to drop the requirement that R be an algebra over k, we must replace the statement by

> "If all sequences $x_0, \dots, x_\ell \in R$ are algebraically dependent via a relation
> $$Q(x_0, \dots, x_\ell) = \sum \alpha_{n_0 \cdots n_\ell} x_0^{n_0} \cdots x_\ell^{n_\ell}$$
> whose lowest term (in lexicographic order in the exponents of the x_i) has a coefficient of 1, then R has Krull dimension $\leq \ell$."

PROOF. we will show that any algebraic dependence relation can be written in the form of equation 2.8.3.

By the hypothesis, given *any* set of elements, $x_0, \dots, x_\ell$, there exists an algebraic dependence relation, $Q(x_0, \dots, x_k) = 0$. Order the monomials, $\alpha_{n_0, \dots, n_\ell} x_0^{n_0} \cdots x_\ell^{n_\ell}$, of Q lexicographically on the strings $(n_0, \dots, n_\ell)$ and pick the one of *lowest* degree, say

$$x_0^{m_0} \cdots x_k^{m_k}$$

and make its coefficient 1 — this is the *only* place where we use the fact that k is a field.

Set $Q_{\ell+1} = Q - x_0^{m_0} \cdots x_\ell^{m_\ell}$ and write

$$Q_{\ell+1} = x_0^{m_0} \cdots x_\ell^{m_\ell+1} \cdot a_\ell + Q_\ell$$

where Q_ℓ consists of all terms, t, of $Q_{\ell+1}$ such that $x_0^{m_0} \cdots x_k^{m_k+1} \nmid t$ and

$$a_\ell = \frac{Q_{\ell+1} - Q_\ell}{x_0^{m_0} \cdots x_k^{m_k+1}}$$

We continue, writing

$$Q_{i+1} = x_0^{m_0} \cdots x_i^{m_i+1} \cdot a_i + Q_i$$

for $i = 1, \ldots, \ell$, where Q_i consists of terms, t, of Q_{i+1} such that $x_0^{m_0} \cdots x_i^{m_i+1} \nmid t$.
   We have

$$
\begin{aligned}
Q &= x_0^{m_0} \cdots x_\ell^{m_\ell} + Q_{\ell+1} \\
&= x_0^{m_0} \cdots x_\ell^{m_\ell} + x_0^{m_0} \cdots x_\ell^{m_\ell+1} \cdot a_\ell + Q_\ell \\
&\;\;\vdots \\
&= x_0^{m_0} \cdots x_\ell^{m_\ell} + x_0^{m_0} \cdots x_\ell^{m_\ell+1} \cdot a_\ell + \cdots + a_0 x_0^{m_0+1} \\
&= 0
\end{aligned}
$$

The conclusion follows from theorem 2.8.14 on the preceding page. $\qquad \square$

At this point, it is reassuring to see that:

COROLLARY 2.8.16. *If k be a field and $R = k[X_1, \ldots, X_n]$ is the ring of polynomials over k, then the Krull dimension of R is n. This implies that the dimension of $\mathbb{A}^n$ is n.*

REMARK. The *second* shortest proof of this appears in chapter 10 of [**38**].

PROOF. It is easy to construct a sequence of prime ideals of length $n + 1$:

$$0 \subsetneq (X_1) \subsetneq (X_1, X_2) \subsetneq \cdots \subsetneq (X_1, \ldots, X_n)$$

but this only proves that the Krull dimension of R is $\geq n$. Theorem 2.8.15 on the previous page shows that it is exactly n since the transcendence degree of the field of fractions of R is n (see A.2.52 on page 409). Proposition 2.8.9 on page 101 implies the conclusion. $\qquad \square$

Examples of dimension:

- If V is a linear subspace of $\mathbb{A}^n$ (or the translate of such a subspace) it is easy to see that $k[V]$ 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 equations defining V — so its Krull dimension coincides with the linear-algebra dimension.
- Consider the direct sum $R = k[X] \oplus k[Y_1, Y_2]$ consisting of all linear combinations $k_1 p(X) + k_2 q(Y_1, Y_2)$ with multiplication defined so that $X \cdot Y_i = 0$. This is an affine k-algebra but not an integral domain, so

it doesn't have a field of fractions. We can compute its realization by mapping

$$
\begin{aligned}
k[T_1, T_2, T_3] &\;\rightarrow\; R \\
T_1 &\;\mapsto\; X \\
T_2 &\;\mapsto\; Y_1 \\
T_3 &\;\mapsto\; Y_2
\end{aligned}
$$

The kernel of this mapping is the ideal (XY_1, XY_2) so the realization of R is the union

$$
\mathbb{A}^1 \cup \mathbb{A}^2
$$

where they only share the *point* $(0,0,0)$. The dimensions of the irreducible components are different, so dimension is not necessarily well-defined for varieties that are not irreducible.

In the sequel, we'll need a refinement of this result. We begin by defining a particularly well-behaved type of ring:

DEFINITION 2.8.17. Let R be a ring of Krull dimension n. We say that the ring R is *catenary* if *any* increasing sequence of primes

$$
\mathfrak{p}_0 \subsetneqq \mathfrak{p}_1 \subsetneqq \cdots \subsetneqq \mathfrak{p}_t \subsetneqq R
$$

that is maximal in the sense that no prime ideals can be inserted between successive terms or added to either end must have $t = n$.

Here's an example of a ring of *a non*-catenary ring:

EXAMPLE 2.8.18. It is an algebraic set that is the union of a line and a plane. The line has coordinate ring $k[X, Y, Z]/(X, Y)$ and the plane has $R = k[X, Y, Z]/(Z)$. Their *union* has coordinate ring $k[X, Y, Z]/\mathfrak{a}$ where

$$
\mathfrak{a} = (X, Y) \cap (Z) = (XZ, YZ)
$$

where the intersection can be computed via proposition 2.3.14 on page 53.

Note that (0) is *not* a prime ideal in R because it has zero-divisors. Note that (X) and (Y) *also* fail to be prime ideals: they contain 0 and $XZ = 0 = YZ$ in the quotient. The only way to guarantee that

$$
u \cdot v \in \mathfrak{f} \implies u \in \mathfrak{f} \text{ or } v \in \mathfrak{f}
$$

is to include *both* X and Y *or* Z in $\mathfrak{f}$.

We get two ascending chains of prime ideals of different lengths:

(1) $(X, Y) \subsetneqq (X, Y, Z - 1)$ and
(2) $(Z) \subsetneqq (Y - 1, Z) \subsetneqq (X - 1, Y - 1, Z)$

Our coordinate rings of affine spaces have this property:

THEOREM 2.8.19. *If k is an infinite field and n is a positive integer, then $k[X_1, \ldots, X_n]$ is catenary.*

REMARK. In other words, *all* maximal sequences of prime ideals have *exactly* length n, not only the *longest* such sequence.

PROOF. We will prove this by induction on n. If $n = 1$, proposition A.2.4 on page 387 implies that $k[X_1]$ is a principal ideal domain, so the conclusion is clear.

If $n > 1$, theorem A.3.7 on page 413 and a simple induction imply that $k[X_1, \ldots, X_n]$ is a unique factorization domain. We claim that the minimal prime $\mathfrak{p}_1 = (f)$ for some irreducible polynomial f. If

$$\mathfrak{p}_1 = (f_1, \ldots, f_s)$$

the f_i must be irreducible and each $(f_i) \subsetneq \mathfrak{p}_1$.

If f is of degree d, apply the argument of lemma 2.2.1 on page 39 to change coordinates so that the coefficient of X_n^d is a nonzero constant — which we can take as 1. It follows that

$$(2.8.4) \qquad k[X_1, \ldots, X_{n-1}] \to \frac{k[X_1, \ldots, X_n]}{(f)}$$

is an *integral extension* of rings (see definition A.4.1 on page 419 and corollary A.4.6 on page 421). Giver our ascending sequence of prime ideals

$$0 \subsetneq (f) = \mathfrak{p}_1 \subsetneq \cdots \subsetneq \mathfrak{p}_t \subsetneq k[X_1, \ldots, X_n]$$

we factor out by (f) to get a sequence

$$0 \subsetneq \frac{\mathfrak{p}_2}{(f)} \subsetneq \cdots \subsetneq \frac{\mathfrak{p}_t}{(f)} \subsetneq \frac{k[X_1, \ldots, X_n]}{(f)}$$

This preserves maximal sequences of prime ideals, by lemma A.1.25 on page 352. Our sequence of length $t + 1$ becomes one of length t. Now we intersect it with

$$k[X_1, \ldots, X_{n-1}] \subset \frac{k[X_1, \ldots, X_n]}{(f)}$$

The sequence

$$\cdots \subset \frac{\mathfrak{p}_i}{(f)} \cap k[X_1, \ldots, X_{n-1}] \subset \frac{\mathfrak{p}_{i+1}}{(f)} \cap k[X_1, \ldots, X_{n-1}] \subset \cdots$$

consists of prime ideals. Since the extension in equation 2.8.4 is *integral*, proposition A.4.8 on page 422 implies that

$$(2.8.5) \qquad \cdots \subsetneq \frac{\mathfrak{p}_i}{(f)} \cap k[X_1, \ldots, X_{n-1}] \subsetneq \frac{\mathfrak{p}_{i+1}}{(f)} \cap k[X_1, \ldots, X_{n-1}] \subsetneq \cdots$$

for all i (i.e., that successive terms are *distinct*). The induction hypothesis implies that the sequence in equation 2.8.5 is of length $t = n$ and the conclusion follows. $\qquad \square$

It turns out that certain quotients of these rings also have these nice properties:

COROLLARY 2.8.20. *Let $\mathfrak{p} \subset R$ be a prime ideal in a ring that is catenary. Then the quotient, $R/\mathfrak{p}$, is also catenary.*

If k is an infinite field and R is a finitely generated k-algebra that is an integral domain, then R is catenary.

PROOF. The first statement follows immediately from lemma A.1.25 on page 352. The second follows from the fact that

$$R = \frac{k[X_1, \ldots, X_t]}{\mathfrak{p}}$$

for some prime ideal $\mathfrak{p} \subset k[X_1, \ldots, X_t]$. $\qquad \square$

Compare this to proposition 2.8.8 on page 101

COROLLARY 2.8.21. *If R is a catenary ring and $\mathfrak{p} \subset R$ is a prime ideal*

$$\mathrm{ht}(\mathfrak{p}) + \dim R/\mathfrak{p} = \dim R$$

PROOF. We follow the same reasoning as in proposition 2.8.8 on page 101, except that the maximal chain of ideals must have a length exactly equal to $\dim R$. $\qquad \square$

EXERCISES.

1. If R is a ring and $S \subset R$ is a multiplicative set, show that

$$\dim S^{-1}R \le \dim R$$

2. If R is a ring, show that

$$\dim R = \max_{\mathfrak{m}} \dim R_{\mathfrak{m}}$$

where $\mathfrak{m} \subset R$ runs over all of the maximal ideals.

2.8.3. Dimension and transcendence degree. The fact that the ring $k[X_1, \dots, X_n]$ has Krull dimension n suggests a possible connection between Krull dimension and the transcendence degree of the field of fractions. This, indeed, turns out to be the case.

THEOREM 2.8.22. *Suppose V is an irreducible affine variety such that $k(V)$ has transcendence degree n. Then the dimension of V is also n.*

PROOF. The Noether normalization theorem (2.2.2 on page 40) implies that there exists a finite map

$$f \colon V \to \mathbb{A}^n$$

which implies that $\dim V \ge n$ (see proposition 2.8.6 on page 99). At this point theorem 2.8.15 on page 103 implies that $\dim V \le n$ and the conclusion follows. $\square$

This immediately implies:

COROLLARY 2.8.23. *If $f \colon X \to Y$ is a finite map or birational equivalence of irreducible affine varieties, then $\dim X = \dim Y$.*

PROOF. If f is a finite map, then $k[X]$ is a finitely-generated module over $k[Y]$, and this implies that $k(X)$ is a finite extension of $k(Y)$, i.e. an *algebraic* extension (see proposition A.2.12 on page 390), so the fields have the same transcendence degree. $\square$

We have a partial converse:

PROPOSITION 2.8.24. *Let $f \colon X \to Y$ be a morphism of irreducible affine varieties such that $f(X) \subset Y$ is dense. If $\dim X = \dim Y$, then*

$$f^* \colon k(Y) \hookrightarrow k(X)$$

and $k(X)$ is a finite extension of $k(Y)$.

PROOF. Lemma 2.4.15 on page 67 implies that $k[Y] \to k[X]$ is injective so we get an injection $k(Y) \to k(X)$. The conclusion follows from the fact that $k(X)$ and $k(Y)$ both have the same transcendence degree, so $k(X)$ is an algebraic extension of $k(Y)$. We claim that this must be finite. Let $x_1, \dots x_n \in k[X]$ be generators of $k[X]$ as an algebra over k. Then each of the x_i is algebraic over $k[Y]$ within a finite extension of $k[Y]$ so the set of all must be contained in a finite extension. $\square$

COROLLARY 2.8.25. *Let V and W be irreducible affine varieties of dimension n and m, respectively. Then*

$$\dim V \times W = n + m$$

REMARK. We know that $V \times W$ is irreducible by lemma 2.7.5 on page 96.

PROOF. The finite maps

$$\begin{aligned} f &: V &\to& \quad \mathbb{A}^n \\ g &: W &\to& \quad \mathbb{A}^m \end{aligned}$$

induce a surjective map

$$f \times g : V \times W \to \mathbb{A}^n \times \mathbb{A}^m = \mathbb{A}^{n+m}$$

We claim that this is a finite map. This follows from the induced maps of coordinate rings:

$$\begin{aligned} f^* &: k[\mathbb{A}^n] &\to& \quad k[V] \\ g^* &: k[\mathbb{A}^m] &\to& \quad k[W] \end{aligned}$$

where $k[V]$ is a finitely generated module over $f^*(k[\mathbb{A}^n])$ — say with generators $\{e_1, \ldots, e_r\}$ and $k[W]$ is a finitely generated module over $g^*(k[\mathbb{A}^m])$ — with generators $\{f_1, \ldots, f_t\}$. We claim that

$$\{e_i \otimes f_j\}$$

generate $k[V] \otimes_k k[W]$ over $k[\mathbb{A}^n] \otimes_k k[\mathbb{A}^m] = k[\mathbb{A}^{n+m}]$. This follows from the same reasoning as that used in exercise 15 on page 472 except that we cannot make any claims of linear independence of the $\{e_i \otimes f_j\}$. The conclusion follows from corollary 2.8.23 on the preceding page.   □

We can use dimension to define the important concept of the degree of a map:

DEFINITION 2.8.26. Under the hypotheses of proposition 2.8.24 on the previous page, above, the degree of the extension

$$f^* : k(Y) \hookrightarrow k(X)$$

is called the *degree of the map, f*.

REMARK. Finite maps satisfy all of the conditions listed above, so they have degrees.

Here's an example of degree of a map:

EXAMPLE 2.8.27. Consider the projection of a parabola $y = x^2$ onto the y-axis. Its image over $\mathbb{C}$ is the entire y-axis and the extension of fields is given by

$$\mathbb{C}(X^2) \to \mathbb{C}(X)$$

which is an extension of degree 2 (see exercise 2 on page 395). It follows that this projection is of degree 2.

Note that the inverse image of most points in this example (i.e., points other than the origin) consists of two points.

Now we consider affine varieties generated by single elements:

DEFINITION 2.8.28. The zero-set of a nonconstant polynomial $f(x_1, \ldots, x_n)$ is called a *hypersurface* in $\mathbb{A}^n$.

Now we will analyze the zero-set of an element, $f \in k[V]$ of the coordinate ring of an affine variety, V. The irreducible components of this set will be the minimal primes that contain f. Intuition tells us that the components of the set $f = 0$ will be one dimension lower than that of V. Krull's Principal Ideal Theorem (originally published in [93]) confirms this:

THEOREM 2.8.29 (Krull's Principal Ideal Theorem). *If R is a noetherian ring, $x \in R$ is a nonzero element that is not a unit, and $\mathfrak{p}$ is a minimal prime such that $(x) \subset \mathfrak{p}$,*
  (1) $\mathrm{ht}(\mathfrak{p}) = 0$ *if and only if x is a zero-divisor*
  (2) if x is not a zero-divisor $\mathrm{ht}(\mathfrak{p}) = 1$,

REMARK. Exercise 31 on page 364 shows that prime ideals that are *minimal* consist of zero-divisors — or are the ideal (0). This theorem considers what happens when a prime ideal contains a non-zero divisor.

In [85], Kaplansky called this "probably the most important single theorem in the theory of Noetherian rings." Although it makes intuitive sense[5], it is surprisingly hard to prove.

PROOF. The proof is divided into two cases:

Statement 1: Suppose x is a zero divisor.

If x is *nilpotent,* then theorem A.1.47 on page 360 implies that

$$x \in \mathfrak{N}(R) = \bigcap_{\text{all prime ideals}} \mathfrak{p}_i = \bigcap_{\text{minimal prime ideals}} \mathfrak{p}_j$$

so a minimal prime ideal containing x is a minimal prime ideal (i.e., not simply a minimal prime *containing x*) — therefore of height zero. Without loss of generality, we can form the quotient by the nilradical — lemma A.1.25 on page 352 implies that this preserves the structure of prime ideals. *We assume that R is reduced.*

If x is a (non-nilpotent) *zero-divisor,* then $x \cdot y = 0$ for some $y \in R$ with $y \neq 0$. The nilradical of R is 0 (since R is reduced) and

$$0 = \bigcap_{\text{all prime ideals}} \mathfrak{p}_i = \bigcap_{\text{minimal prime ideals}} \mathfrak{p}_j$$

Since $y \neq 0$, there must exist a minimal prime $\bar{\mathfrak{p}}$ with $y \notin \bar{\mathfrak{p}}$. Then $x \cdot y = 0 \in \bar{\mathfrak{p}}$ implies that $x \in \bar{\mathfrak{p}}$ so that $\bar{\mathfrak{p}}$ has height 0.

Conversely, if $\mathrm{ht}(\mathfrak{p}) = 0$, so that x is contained in a minimal prime ideal, exercise 31 on page 364 shows that x is a zero-divisor.

Statement 2: Assume that x is *not* a zero-divisor *and* $\mathrm{ht}(\mathfrak{p}) > 0$.

If $\mathfrak{p} \subset R$ is a prime ideal, define $\mathfrak{p}^{(n)}$, the nth *symbolic power* of $\mathfrak{p}$ to be $R \cap (\mathfrak{p} \cdot R_{\mathfrak{p}})^n$ or $\{x \in R | x \cdot y \in \mathfrak{p}^n, y \notin \mathfrak{p}\}$.

---

[5]If $\mathfrak{p}$ is a minimal prime containing x, it is hard to imagine it containing a *smaller* prime (other than (0)).
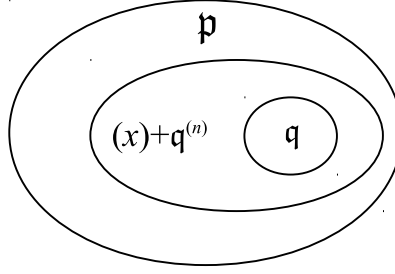
FIGURE 2.8.1.   Minimal prime ideal containing (x)

Suppose $\mathfrak{q} \subsetneq \mathfrak{p}$ is another prime ideal with $x \notin \mathfrak{q}$ (if $x \in \mathfrak{q}$, it would contradict the minimality of $\mathfrak{p}$). We will show that the ring $R_{\mathfrak{q}}$ has dimension 0 so $\operatorname{ht}(\mathfrak{q}) = 0$ and $\operatorname{ht}(\mathfrak{p}) \leq 1$ — see figure 2.8.1 on page 110. Without loss of generality, we may assume that $\mathfrak{p}$ is maximal: if not, simply form the localization $R_{\mathfrak{p}}$; this process will preserve all of the ideals contained in $\mathfrak{p}$ (see corollary A.1.94 on page 385).

Since $\mathfrak{p} \subset R$ is maximal and the minimal prime ideal that contains (x), we conclude that $R/(x)$ is Artinian (see theorem A.1.87 on page 382). It follows that the sequence

$$\mathfrak{q}^{(1)} + (x) \supset \mathfrak{q}^{(2)} + (x) \supset \cdots$$

becomes constant from some finite point on, i.e. $\mathfrak{q}^{(n+1)} + (x) = \mathfrak{q}^{(n)} + (x)$ or $\mathfrak{q}^{(n)} \subset \mathfrak{q}^{(n+1)} + (x)$. If $y \in \mathfrak{q}^{(n)}$, then there exist $a \in R$ and $q \in \mathfrak{q}^{(n+1)}$ such that

(2.8.6) $$y = ax + q$$

Since $\mathfrak{q}^{(n+1)} \subset \mathfrak{q}^{(n)}$, we have $q \in \mathfrak{q}^{(n)}$ and it follows that $ax \in \mathfrak{q}^{(n)}$.

Since $x \notin \mathfrak{q}$, $ax \in \mathfrak{q}^{(n)}$ implies that $a \in \mathfrak{q}^{(n)}$. Equation 2.8.6 implies that

(2.8.7) $$\mathfrak{q}^{(n)} = (x) \cdot \mathfrak{q}^{(n)} + \mathfrak{q}^{(n+1)}$$

At this point, Nakayama's Lemma (A.1.80 on page 378) applied to

$$\frac{\mathfrak{q}^{(n)}}{\mathfrak{q}^{(n+1)}}$$

(which is unchanged when multiplied by (x)) implies that

(2.8.8) $$\mathfrak{q}^{(n)} = \mathfrak{q}^{(n+1)}$$

which implies that $(\mathfrak{q} \cdot R_{\mathfrak{q}})^{n+1} = (\mathfrak{q} \cdot R_{\mathfrak{q}})^{n}$. A second application of Nakayama's Lemma then implies that $(\mathfrak{q} \cdot R_{\mathfrak{q}})^{n} = 0$. Since the nilpotent elements are the intersection of all prime ideals (see theorem A.1.47 on page 360), it follows that $\mathfrak{q} \cdot R_{\mathfrak{q}}$ does *not* properly contain any other prime ideals. Since $\mathfrak{q} \cdot R_{\mathfrak{q}}$ is maximal, it follows that $\dim R_{\mathfrak{q}} = 0$, and $\operatorname{ht}(\mathfrak{q}) = 0$ in R.

□

Krull's Principal Ideal theorem immediately implies that

COROLLARY 2.8.30. *Let V be an n-dimensional irreducible affine variety and let* $f \neq 0 \in k[V]$ *be a regular function that has a 0 on V. Then the zero-set*

$$W = \{x \in V | f(x) = 0\}$$

is composed of irreducible components of dimension $n - 1$.

PROOF. Since V is irreducible its coordinate ring is an integral domain, hence catenary, by corollary 2.8.20 on page 106. The irreducible components of W correspond to the minimal primes, $\{\mathfrak{p}_i\}$ that contain (f), and these are of height 1, by theorem 2.8.29 on page 109. Corollary 2.8.21 on page 106 implies that the dimension of $k[V]/\mathfrak{p}_i$ is $n - 1$ for all i. $\square$

We also get a partial converse:

PROPOSITION 2.8.31. *Let V be an irreducible variety such that* $k[V]$ *is a unique factorization domain (for example,* $V = \mathbb{A}^d$).
If $W \subset V$ *is a closed variety of dimension* $\dim V - 1$, *then* $\mathcal{I}(W) = (f)$ *for some* $f \in k[V]$.

PROOF. If W_i are the irreducible components of W, we have

$$\mathcal{I}(W) = \bigcap \mathcal{I}(W_i)$$

and if $\mathcal{I}(W_i) = (f_i)$, then $\mathcal{I}(W) = (\prod f_i)$. It follows that it *suffices* to prove the result for W *irreducible*. Since $W \neq V$, there is some polynomial, H, that vanishes on it. Since it is irreducible, some irreducible factor, f, of H vanishes on it as well. Let $\mathfrak{p} = \mathcal{I}(W)$. It is a prime ideal and is nonzero since otherwise $\dim W = \dim V$.

It contains the irreducible factor, f, of H and (f) is prime (because $k[V]$ is a unique factorization domain). If $(f) \neq \mathfrak{p}$ then

$$W = \mathcal{V}(\mathfrak{p}) \not\subseteq \mathcal{V}((f)) \not\subseteq V$$

so

$$\dim W < \dim \mathcal{V}((f)) < \dim V$$

which contradicts the hypothesis. $\square$

We get a straightforward generalization:

PROPOSITION 2.8.32. *Let V be an irreducible affine variety and let* $f_1, \ldots, f_t \in k[V]$ *be regular functions and let* W_j *be the zero-set of* $f_1, \ldots, f_j$. *If*

$$W_j \subset V$$

is nonempty, it is of codimension $\leq j$.

We also get the corresponding algebraic statement:

THEOREM 2.8.33 (Krull's Height Theorem). *If R is a noetherian ring with elements* $x_1, \ldots, x_t \in R$, *and* $\mathfrak{p}$ *is a minimal prime ideal containing* $(x_1, \ldots, x_t)$ *then* $\mathrm{ht}(\mathfrak{p}) \leq t$.

REMARK. Proposition 2.8.32 already proves this for R an affine k-algebra.

PROOF. We will do induction on t, using theorem 2.8.29 on page 109 as the ground case. Without loss of generality, we may assume that $\mathfrak{p}$ is maximal: localizing with respect to $\mathfrak{p}$ does not change the inclusion-relations between ideals contained in $\mathfrak{p}$. It follows that

- $\operatorname{im}\mathfrak{p} \subset R' = R/(x_1, \ldots, x_t)$ is the *only* prime ideal of R' so
- R' is Artinian (see theorem A.1.87 on page 382) and $\operatorname{im}\mathfrak{p} = \mathfrak{N}(R')$ is nilpotent (by proposition A.1.86 on page 382).

So $\mathfrak{p}$ is nilpotent modulo the ideal $(x_1, \ldots, x_t)$ — let $\mathfrak{q}$ be a maximal prime with $\mathfrak{q} \subset \mathfrak{p}$. We will show that $\mathfrak{q}$ is minimal over an ideal with $t-1$ generators and the result will follow by induction. The hypothesis that $\mathfrak{p}$ is minimal over $(x_1, \ldots, x_t)$ implies that one of the $x_i \notin \mathfrak{q}$ — say x_1. Then $\mathfrak{p}$ is minimal over $\mathfrak{q} + (x_1)$. The same reasoning as that used above to show that $\mathfrak{p}$ is nilpotent over $(x_1, \ldots, x_t)$ implies that the x_i are nilpotent modulo $\mathfrak{q} + (x_1)$ so that

$$x_i^n = a_i x_1 + y_i$$

for $i = 2, \ldots, t$ with $y_i \in \mathfrak{q}$. The fact that $\mathfrak{p}$ is nilpotent modulo $(x_1, y_2, \ldots, y_t)$ and theorem 2.8.29 on page 109 implies that the height of the image of $\mathfrak{p}$ in $R/(y_2, \ldots, y_t)$ is ≤ 1. The definition of $\mathfrak{q}$ implies that its height is 0 so $\mathfrak{q}$ is minimal over $(y_2, \ldots, y_n)$. This completes the induction. $\square$

This simple generalization of theorem 2.8.29 on page 109 gives us an interesting result involving intersections of varieties:

COROLLARY 2.8.34. *Let $V, W \subset \mathbb{A}^n$ be irreducible algebraic sets. If $V \cap W \neq \varnothing$, then*

$$\dim(V \cap W) \geq \dim V + \dim W - n$$

PROOF. This follows immediately from proposition 2.8.32 on the preceding page and corollary 2.8.25 on page 108 on realizing

$$V \cap W \cong (V \times W) \cap \Delta(\mathbb{A}^n)$$

(see exercise 1 on page 98). The intersection with $\Delta(\mathbb{A}^n)$ is equivalent to imposing n equations

$$f_i = X_i - Y_i$$

where the X_i and the Y_i are the coordinates in the two factors of $\mathbb{A}^n$ in $\mathbb{A}^n \times \mathbb{A}^n$. $\square$

EXAMPLE. Let $F(X, Y)$ and $G(X, Y)$ be nonconstant polynomials with no common factor. Then $\mathcal{V}(F(X, Y)) \subset \mathbb{A}^2$ has dimension 1 and

$$\mathcal{V}(F(X, Y)) \cap \mathcal{V}((G(X, Y))$$

has dimension zero — it is a finite set of points.

EXAMPLE 2.8.35. We can classify the irreducible closed subsets V of $\mathbb{A}^2$:

- If V has dimension 2, then it can't be a proper subset of $\mathbb{A}^2$ (since $\mathbb{A}^2$ is irreducible and proper subsets must have a lower dimension).
- If V has dimension 1, then $V \neq \mathbb{A}^2$ so $\mathcal{I}(V)$ contains a nonzero polynomial, hence a nonzero irreducible polynomial f. Then $V \supset \mathcal{V}(f)$ so it must equal $\mathcal{V}(f)$.
- Finally, if $\dim V = 0$, V is a point.

- We can list all of the prime ideals of $k[X, Y]$: They are: (0), (f) with f irreducible, and $(X - a, X - b)$ for $\{a, b\} \in \mathbb{A}^2$.

We can also say something about the dimension of *fibers of maps*:

COROLLARY 2.8.36. *If $f \colon V \to W$ is a surjective regular map of irreducible affine varieties*

(1) $\dim F \geq \dim V - \dim W$ *for any component, F, of $f^{-1}(p)$, for any $p \in W$,*
(2) *there exists an open set $U \subset W$ such that $\dim f^{-1}(p) = \dim V - \dim W$ for $p \in U$.*

REMARK. Exercise 2 on page 79 gives a degree-1 mapping $f \colon \mathbb{A}^2 \to \mathbb{A}^2$ with $\dim F$ can be > 0 at some points.

PROOF. The induced map

$$f^* \colon k[W] \to k[V]$$

will be injective. Let g_1 be a regular function on W with p in a component, W_1, of its zero-set. Likewise, let g_2 be a regular function on W_1 with p in a component, W_2, of its zero-set. In this fashion, we construct a chain of irreducible subvarieties

$$W \supsetneq W_1 \supsetneq \cdots \supsetneq W_m = p$$

and p is the only common zero of m regular functions $g_1, \ldots, g_m \in k[W] \subset k[V]$, where the definition of dimension (definition 2.8.1 on page 99) implies that $m \leq \dim W$.

The subvariety $f^{-1}(p)$ is defined by

$$f^*(g_1) = \cdots = f^*(g_m) = 0$$

Proposition 2.8.32 on page 111 implies that the dimension of this zero-set will be $\geq \dim V - m$. The conclusion follows.

To prove the second statement, let $k[W] = k[w_1, \ldots, w_s]$ and $k[V] = k[v_1, \ldots, v_t]$ so f induces an inclusion $f^* \colon k[W] \to k[V]$ that induces an extension $k(W) \to k(V)$. Assume, without loss of generality, that $v_1, \ldots, v_j$ is algebraically independent over $k(W)$, where $j = \dim V - \dim W$. If there is an algebraic relation

$$F_\beta(w_1, \ldots w_s, v_1, \ldots, v_t) = 0$$

is the set of all algebraic relations between the v_i over $k(W)$, we can regard these as polynomials in the v_i coefficients that are polynomials in the w_α. Let J_β be the zero-set of these coefficient-polynomials in F_β. Then $U = W \setminus \bigcup_\beta J_\beta$ is an open set and if $p \in U$, F_β maps to the nonzero polynomial

$$\bar{F}_\beta(v_1, \ldots, v_t) = 0$$

when evaluated at p (or under the map $k[V] \to k[V]/k[V] \cdot \mathfrak{m}_p$, where $\mathfrak{m}_p \subset k[W]$ is the maximal ideal corresponding to p). The $\bar{F}_\beta$ will be polynomials with coefficients in $k[W]/\mathfrak{m}_p = k$. Since the v_i are algebraically dependent on the $v_1, \ldots, v_j$ it follows that $\dim f^{-1}(p) \leq j$. The conclusion follows from statement 1.   □

The following result shows just how coarse birational equivalence is as an equivalence relation. Every irreducible affine variety is birationally equivalent to a hypersurface:

THEOREM 2.8.37. *If k is of characteristic zero and V is an n-dimensional irreducible affine variety, then there exists an irreducible hypersurface $W \subset \mathbb{A}^{n+1}$ and a birational equivalence*

$$f: V \to W$$

REMARK. In the next section, this result will be used to establish various properties of the tangent space of a variety.

PROOF. Since V is n-dimensional, theorem A.2.52 on page 409 implies that

$$k(V) \cong k(X_1, \ldots, X_n)[\alpha_1, \ldots, \alpha_t]$$

where the α_i are algebraic over $k(X_1, \ldots, X_n)$. The Primitive Element Theorem (A.2.21 on page 394) implies that there exists a single element β that is algebraic over $k(X_1, \ldots, X_n)$ such that

$$k(V) \cong k(X_1, \ldots, X_n)[\beta]$$

If we map

$$\begin{aligned} k(X_1, \ldots, X_n)[X_{n+1}] &\to& k(X_1, \ldots, X_n)[\beta] \\ X_{n+1} &\mapsto& \beta \end{aligned}$$

the kernel is (f), where $f \in k(X_1, \ldots, X_n)[X_{n+1}]$ is the minimal polynomial of β and

$$k(V) \cong k(X_1, \ldots, X_n)[X_{n+1}]/(f)$$

Note that f is a polynomial in X_{n+1} whose coefficients are rational functions in $X_1, \ldots, X_n$. Let $g(X_1, \ldots, X_{n+1}) \in k[X_1, \ldots, X_{n+1}]$ be the result of clearing out all of the denominators of the coefficients of f. Then

$$g(X_1, \ldots, X_{n+1}) = 0$$

defines a hypersurface, W, in $\mathbb{A}^{n+1}$ with

$$k(W) \cong k(X_1, \ldots, X_n)[X_{n+1}]/(f) \cong k(V)$$

so the conclusion follows from corollary 2.5.26 on page 86.   □

COROLLARY 2.8.38. *Let $f: V \to W$ be a rational map of algebraic sets, with V irreducible. Then the set of points where f is not regular is a subvariety, $Z \subset V$, with $\dim Z < \dim V$. Consequently, the set of points where f is regular is open.*

Restricted to the open set, U, of regular points, f is a regular map and induces a homomorphism

$$f^*: k[V] \to k[U]$$

REMARK. Since open sets are dense, this means that f is regular "almost everywhere."

PROOF. If $W \subset \mathbb{A}^m$, the rational map, f, is given by

$$f = \begin{bmatrix} \varphi_i \\ \vdots \\ \varphi_m \end{bmatrix} : V \to \mathbb{A}^m$$

with $\varphi_i \in k(V)$, and the set of points where it is regular is just

$$\bigcap_{i=1}^{m} U(\varphi_i) = U$$

a nonempty open set, by proposition 2.4.17 on page 68. The complement $Z = V \setminus U$ is a closed set, hence (in the Zariski topology) a subvariety. Since V is irreducible, Z must be of a strictly lower dimension, by proposition 2.8.3 on page 99.

If we restrict f to U, proposition 2.5.21 on page 83 implies that this restriction is regular and, therefore, induces a homomorphism of coordinate rings. $\square$

EXERCISES.

3. Use theorem 2.8.29 on page 109 to prove:

LEMMA 2.8.39. *Suppose A is a noetherian domain. Then A is a unique factorization domain if and only if every prime of height* 1 *is principal.*

4. If $f\colon V \to W$ is a dominating (see definition 2.4.14 on page 67) regular map of affine varieties and $U \subset V$ is a dense subset, show that $f(U) \subset W$ is dense.

5. Let $f\colon V \to W$ be a degree-d (see definition 2.8.26 on page 108) regular map of irreducible affine varieties. If $U \subset V$ is an open set, show that $f(U)$ contains a nonempty open set of W. Hint: use the norm-construction for a finite field-extension (see definition A.2.22 on page 396).

6. If k is of characteristic 0, $V \subset \mathbb{A}^n$ with $k[V] = k[X_1, \dots, X_n]/\mathfrak{A}$ and $f\colon V \to W$ is a surjective regular map of irreducible affine varieties of degree d show that

$$k(V) = k(W)[\alpha]$$

where

$$\alpha = \sum_{j=1}^{n} \beta_j x_j$$

with $\beta_j \in k$, and x_j is the image of X_j in $k[W]$.

7. Under if the hypotheses of exercise 6, show that f factors as

$$V \xrightarrow{\varphi} \bar{V} \subset W \times \mathbb{A}^1 \xrightarrow{\pi} W$$

where φ is a birational equivalence and $\pi\colon W \times \mathbb{A}^1 \to W$ is the projection. This is a *relative* version of theorem 2.8.37 on the preceding page.

8. Why do exercise 6 and 7 have the hypothesis that the characteristic of k is 0?

9. Under the hypotheses of exercises 7 and 6, show that there exists a nonempty open set $W_1 \subset W$ such that $f^{-1}(w)$ has precisely d points for $w \in W_1$.

CHAPTER 3

Local properties of affine varieties

> "The introduction of the digit 0 or the group concept was general non-
> sense too, and mathematics was more or less stagnating for thousands
> of years because nobody was around to take such childish steps..."
> —Alexander Grothendieck, writing to Ronald Brown.

3.1. Introduction

The local properties of an affine variety are widely used in more general settings. All algebraic varieties (and schemes, for that matter) are *locally* affine and arguments of a local nature really only use the properties of affine varieties.

3.2. The coordinate ring at a point

We define a structure that determines the properties of an algebraic set in the neighborhood of a point.

DEFINITION 3.2.1. Let $V \subset \mathbb{A}^n$ be an irreducible algebraic set with coordinate ring $k[V]$. If $p \in V$ is a point, then p corresponds to a maximal ideal, $\mathfrak{m}_p \subset k[V]$ (see proposition 2.4.3 on page 62) and $S = k[V] \setminus \mathfrak{m}_p$ is a multiplicative set. Define

$$\mathcal{O}_{V,p} = S^{-1}k[V]$$

to be the *coordinate ring of V at the point p*.

REMARK. Basically, $\mathcal{O}_{V,p} \subset k(V)$ is the subring of the rational function field generated by functions

$$\frac{f(x_i)}{g(x_i)} \in k(V)$$

such that $g(p) \neq 0$, i.e., with $g(x_i) \notin \mathfrak{m}_p$.

Since we have inverted all elements not in $\mathfrak{m}_p$, it follows that $\mathfrak{m}_p \subset k[V] \subset \mathcal{O}_{V,p}$ is the *only* maximal ideal in $\mathcal{O}_{V,p}$ (see corollary A.1.94 on page 385) — making $\mathcal{O}_{V,p}$ a *local ring* (see definition A.1.24 on page 351).

In a manner of speaking, $\mathcal{O}_{V,p}$ is the "coordinate ring of the intersection of all open sets that contain p." This is made more precise as follows:

PROPOSITION 3.2.2. *Let $V \subset \mathbb{A}^n$ be an irreducible algebraic set with coordinate ring $k[V]$. If $p \in V$, let $\{D(h_i)\}$ be the set of all principal open sets such that $p \in D(h_i)$. These are ordered by inclusion, and any inclusion*

$$D(h) \subset D(j)$$

induces a restriction map

$$k[D(j)] = k[V]_j \to k[D(h)] = k[V]_h$$

and

$$\mathcal{O}_{V,p} = \varinjlim k[D(h_i)]$$

PROOF. This follows from corollary A.5.25 on page 452 and the fact that we have inverted precisely the functions that do not vanish at p. $\qquad\square$

It is not hard to see that

PROPOSITION 3.2.3. *A regular map of affine varieties, $f: V \to W$, induces a homomorphism*

$$f_p: \mathcal{O}_{W,f(p)} \to \mathcal{O}_{V,p}$$

for all points, $p \in V$. If f is an isomorphism, then so is f_p for all $p \in V$.

What do we know about $\mathcal{O}_{V,p}$, algebraically? Well, exercise 2 on page 107 implies that:

PROPOSITION 3.2.4. *If V is an irreducible affine variety, then*

$$\dim V = \max_{p \in V} \dim \mathcal{O}_{V,p}$$

To investigate other properties of $\mathcal{O}_{V,p}$, we need a partial converse to theorem 2.8.33 on page 111:

COROLLARY 3.2.5. *Any prime ideal $\mathfrak{p}$ in a noetherian ring, R, that is of height t is minimal over an ideal generated by t elements.*

PROOF. We use induction over t: the ground case follows from theorem 2.8.29 on page 109. If $\mathfrak{p}$ is of height t, let

$$\mathfrak{p}_0 \subset \cdots \subset \mathfrak{p}_{t-1} \subset \mathfrak{p}$$

be a chain of prime ideals defining this fact. By induction, $\mathfrak{p}_{t-1}$ is minimal over $(x_1, \ldots, x_{t-1})$ and we pick an element $x_t \in \mathfrak{p} \setminus \mathfrak{p}_{t-1}$ that is not contained in any of the finite number of primes minimal over $(x_1, \ldots, x_{t-1})$. This is possible by prime avoidance — see exercises 6 on page 71 and 7 on page 71. $\qquad\square$

In one case, we can characterize the dimension of a ring in terms of the number of elements needed to generate an ideal:

THEOREM 3.2.6. *Let R be a noetherian local ring with unique maximal ideal $\mathfrak{m}$. Then the dimension of R is the minimum value of t such that there exist t elements $\{x_1, \ldots, x_t\} \in R$ with $\mathfrak{m}^k \subset (x_1, \ldots, x_t)$ for some value of k.*

REMARK. The hypotheses are equivalent to saying that $\mathfrak{m}$ is minimal over $(x_1, \ldots, x_t)$:
  (1) If $\mathfrak{m}$ is minimal over $(x_1, \ldots, x_t)$ then the image of $\mathfrak{m}$ in $R/(x_1, \ldots, x_t)$ is its *only* prime ideal so it must be nilpotent (by theorem A.1.47 on page 360 applied to $\sqrt{(0)}$ and exercise 30 on page 364). This means that $\mathfrak{m}^k \subset (x_1, \ldots, x_t)$ for some value of k.
  (2) On the other hand, if $\mathfrak{m}^k \subset (x_1, \ldots, x_t)$ then the image, $p(\mathfrak{m})$, of $\mathfrak{m}$ in $R/(x_1, \ldots, x_t)$ is nilpotent, and theorem A.1.47 on page 360 implies that $p(\mathfrak{m})$ must be a minimal prime. This implies that $\mathfrak{m}$ is minimal over $(x_1, \ldots, x_t)$.

PROOF. If $\mathfrak{m}$ is minimal over $(x_1, \ldots, x_t)$, then its height is $\leq t$, by corollary 2.8.33 on page 111. On the other hand, corollary 3.2.5 on the preceding page implies that if the dimension of R is t, we can find elements $x_1, \ldots, x_t$ such that $\mathfrak{m}$ is minimal over $(x_1, \ldots, x_t)$. $\qquad\square$

3.3. The tangent space

One nice thing about algebraic geometry is that it enables us to do *calculus* in situations where the usual concepts of limits might not be well-defined.

For instance, we can simply define:

$$\frac{\partial x_i^n}{\partial x_j} = \begin{cases} 0 & \text{if } i \neq j \\ nx_i^{n-1} & \text{otherwise} \end{cases}$$

We will use this to develop the important concept of the tangent space to a variety at a point.

Since this is a local property of a variety, we may assume that our variety is affine and even embedded in an affine space (i.e., making it into an algebraic set). Afterward, we will give a "coordinate-free" construction.

DEFINITION 3.3.1. Let $f \in k[X_1, \ldots, X_n]$ be a function defined over $\mathbb{A}^n$. If $p = (\alpha_1, \ldots, \alpha_n) \in \mathbb{A}^n$ is a point, we define the *differential of f at p* to be the linear function

$$d_p f = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial X_i} \right)_p (X_i - \alpha_i)$$

where $\left(\frac{\partial f}{\partial X_i} \right)_p$ means "compute the derivative and then plug in the coordinates of p" — so $\left(\frac{\partial f}{\partial X_i} \right)_p \in k$.

REMARK. If we were expanding f in a Taylor series centered at p, this would be the linear term.

Now we give our first definition of the tangent space — identical to that used in analytic geometry or vector calculus:

DEFINITION 3.3.2. Let $V \subset \mathbb{A}^n$ be an algebraic set with $\mathcal{I}(V) = (F_1, \ldots, F_t) \subset k[X_1, \ldots, X_n]$ and let $p \in V$ be a point. The *tangent space of V at p* is the linear subspace, T_p, of $\mathbb{A}^n$ defined by the linear equations

(3.3.1)
$$\begin{aligned} d_p F_1 &= 0 \\ &\vdots \\ d_p F_t &= 0 \end{aligned}$$

These equations simply say that the tangent plane is perpendicular to the gradients of the defining functions.

REMARK. If we think of the functions $F_i \in k[X_1, \ldots, X_n]$ as defining a map

$$\mathbb{A}^n \to \mathbb{A}^t$$
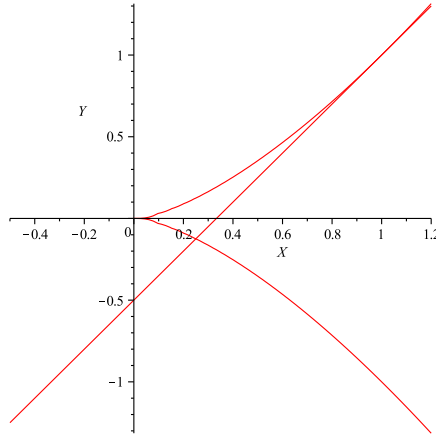
FIGURE 3.3.1.   Tangent space

with a Jacobian matrix

$$(3.3.2) \qquad\qquad \mathcal{J} = \left(\frac{\partial F_i}{\partial X_j} \right)_p$$

then equations 3.3.1 on the preceding page imply that the tangent space is the kernel of this Jacobi map and

$$(3.3.3) \qquad\qquad \dim T_p = t - \operatorname{rank} \mathcal{J}$$

EXAMPLE 3.3.3. For instance, let $V \subset \mathbb{A}^2$ be defined by $Y^2 - X^3 = 0$. At the point $p = (1, 1)$, the derivatives are

$$\left(\frac{\partial(Y^2 - X^3)}{\partial X} \right)_{(1,1)} = -3$$

$$\left(\frac{\partial(Y^2 - X^3)}{\partial Y} \right)_{(1,1)} = 2$$

so the tangent space, T_p, at p is

$$-3(X - 1) + 2(Y - 1) = 0$$

as depicted in figure 3.3.1.

Note that the tangent space is 1-dimensional, as is V. The point $(0, 0) \in V$ is "suspicious looking" — the curve has a "crease" in it. If we compute the tangent space at this point, we get

$$\left(\frac{\partial(Y^2 - X^3)}{\partial X} \right)_{(0,0)} = 0$$

$$\left(\frac{\partial(Y^2 - X^3)}{\partial Y} \right)_{(0,0)} = 0$$

and the equation for the tangent space is

$$0 \cdot X + 0 \cdot Y = 0$$

so that all points of $\mathbb{A}^2$ satisfy it and the dimension of *this* tangent space is 2. The point $(0,0)$ is an example of a *singular point* — defined rigorously in definition 3.3.11 on page 125.

PROPOSITION 3.3.4. *If $V \subset \mathbb{A}^n$ is an affine algebraic set with a point $p \in V$, the differential of a function $g \in k[V]$ defines a linear form on the tangent space $T_{V,p}$.*

REMARK. A linear form on $T_{V,p}$ is an element of the *dual vector* space, $T_{V,p}^*$.

PROOF. As defined above, the differential is a linear form. The only question is whether it is *well-defined* on the tangent space: a function $g \in k[V]$ must be lifted up to $k[X_1, \ldots, X_n]$ before it can be differentiated. Suppose $G_1, G_2 \in k[X_1, \ldots, X_n]$ both map to g under the projection

$$k[X_1, \ldots, X_n] \to k[V]$$

Then $G_1 - G_2 \in (F_1, \ldots, F_r)$ where V is defined by $F_i = 0$, $i = 1, \ldots, r$. If

$$G_1 - G_2 = \sum_{i=1}^{r} f_i \cdot F_i$$

the product-rule for differentiation gives

$$d_p(G_1 - G_2) = \sum_{i=1}^{r} d_p f_i \cdot F_i + f_i \cdot d_p F_i$$

Since V is *defined* by $F_i = 0$, restricting this to $V \subset \mathbb{A}^n$ makes the first set of terms vanish and gives

$$d_p(G_1 - G_2) = \sum_{i=1}^{r} f_i \cdot d_p F_i$$

Since the tangent plane is *defined* by $d_p F_i = 0$, we see that $d_p(G_1 - G_2) = 0$, there. It follows that all representatives of g in $k[X_1, \ldots, X_n]$ give rise to the same linear form on the tangent plane. □

In [**176**], Zariski developed a "coordinate free" description of the tangent space in keeping with our coordinate free algebraic sets.

LEMMA 3.3.5. *If $\mathfrak{m} \subset k[V]$ is the unique maximal ideal of functions that vanish at $p \in V$, then differentiation defines a surjective homomorphism*

$$\mathfrak{m} \to T_{V,p}^*$$

whose kernel is $\mathfrak{m}^2$.

PROOF. We essentially prove the statement for $\mathbb{A}^n$ and show that this makes it true for V.

Embed V in $\mathbb{A}^n$ (see lemma 2.5.6 on page 74) and, without loss of generality, assume that $p = (0, \ldots, 0)$. Let

$$g \colon T_{V,p} \to \mathbb{A}^n$$

be the corresponding embedding of $T_{V,p}$. Let $\{e_1, \ldots, e_r\}$ be a basis for $T_{V,p}$ and extend it to a basis for all of $\mathbb{A}^n$: $\{e_1, \ldots, e_n\}$.

Also assume $V = I(\mathfrak{F})$ for an ideal

$$\mathfrak{F} = (F_1, \ldots, F_m) \subset k[X_1, \ldots, X_n]$$

in this basis. In addition, let $\{e^1, \ldots, e^n\}$ be a dual basis for $(\mathbb{A}^n)^*$ with $\{e^1, \ldots, e^r\}$ a basis of $T^*_{V,p}$ so

$$e^i(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Let $\mathfrak{M} = (X_1, \ldots, X_n) \subset k[X_1, \ldots, X_n]$, a maximal ideal by the Nullstellensatz (see theorem 2.2.3 on page 40). The composite

(3.3.4) $$k[X_1, \ldots, X_n] \xrightarrow{\pi} k[X_1, \ldots, X_n]/\mathfrak{F} = k[V]$$

defines $\mathfrak{m}$ in terms of $\mathfrak{M}$ via

(3.3.5) $$\mathfrak{m} = \iota \circ \pi(\mathfrak{M})$$

— see lemma A.1.30 on page 355 and proposition A.1.93 on page 385.

If

$$\ell = \sum_{i=1}^{r} c_i e^i$$

with $c_i \in k$, is any linear functional on T_p, it extends to a linear functional on $\mathbb{A}^n$ and the function

$$v(\ell) = \sum_{i=1}^{r} c_i X_i \in \mathfrak{M}$$

has the required differential. Its image under the maps in equation 3.3.4 is a function in $\mathfrak{m}$ whose differential is precisely ℓ.

If $g \in \mathfrak{m}$ has a differential in T_p that is 0, let $g = \pi(G)$ where $G \in k[X_1, \ldots, X_n]$ vanishes at p (i.e., $G \in \mathfrak{M}$) and has a differential on T_p that is 0. Since this differential vanishes on T_p, we have

$$d_p G = \sum_{j=1}^{m} \lambda_j d_p F_j$$

for some $\lambda_i \in k$. Let

$$\bar{G} = G - \sum_{j=1}^{m} \lambda_j F_j \in k[X_1, \ldots, X_n]$$

defines the same element $g \in k[V]$ and $\mathfrak{m} \subset k[V]$ as G but its differential in $\mathbb{A}^n$ vanishes at p. This means that its constant and linear terms vanish, so

$$\bar{G} \in \mathfrak{M}^2$$

and equation 3.3.5 implies that $g \in \mathfrak{m}^2$.   □

Suppose R is a ring with a maximal deal $\mathfrak{m}$. Then

$$\begin{aligned} R \cdot \mathfrak{m} &\subset \mathfrak{m} \\ R \cdot \mathfrak{m}^2 &\subset \mathfrak{m}^2 \end{aligned}$$

so we may regard $\mathfrak{m}$ and the quotient

$$\mathfrak{m}/\mathfrak{m}^2$$

as *modules* over R (see definition A.1.57 on page 367). Under this module-structure, multiplying an element of

$$\mathfrak{m}/\mathfrak{m}^2$$

by an element of $\mathfrak{m} \subset R$ gives 0. Consequently, we can regard

$$\mathfrak{m}/\mathfrak{m}^2$$

as a module over the field $R/\mathfrak{m} = F$ — i.e., a vector space over F.

We have proved that

LEMMA 3.3.6. *If R is a ring with maximal ideal $\mathfrak{m}$, then*

$$\mathfrak{m}/\mathfrak{m}^2$$

is naturally a vector-space over the field $R/\mathfrak{m}$.

We are finally ready to give Zariski's coordinate-free definition of the tangent space:

THEOREM 3.3.7. *If V is an affine variety, $p \in V$ is a point, and $\mathfrak{m} \subset k[V]$ is the maximal ideal of functions that vanish at p. Then there exists an isomorphism of vector spaces over k*

$$(\mathfrak{m}_p/\mathfrak{m}_p^2)^* \cong (\mathfrak{m}/\mathfrak{m}^2)^* \cong T_{V,p}$$

where $\mathfrak{m}_p = \mathfrak{m} \cdot \mathcal{O}_{V,p}$.

REMARK. This is sometimes called the Zariski tangent space. We have implicitly used the fact that $k[V]/\mathfrak{m} = k$. The isomorphism

$$(\mathfrak{m}_p/\mathfrak{m}_p^2)^* \cong (\mathfrak{m}/\mathfrak{m}^2)^*$$

implies that the tangent space is a *local* property of V.

See section A.7 on page 482 for an alternate handling of cotangent spaces.

PROOF. If $\mathfrak{m} \subset k[V]$ is the maximal ideal of functions that vanish at p, lemma 3.3.5 on page 121 implies that

$$(\mathfrak{m}/\mathfrak{m}^2)^* \cong T_{V,p}$$

To prove the second isomorphism note that there is a natural injection

$$\mathfrak{m} \hookrightarrow \mathfrak{m}_p$$

that induces an injection

$$\mathfrak{m}/\mathfrak{m}^2 \hookrightarrow \mathfrak{m}_p/\mathfrak{m}_p^2$$

We claim that this is an *isomorphism* of vector spaces. Elements of $\mathfrak{m}_p$ are of the form

$$F/G$$

where $F \in \mathfrak{m}$ and $G \notin \mathfrak{m}$. Since $\mathfrak{m}$ is maximal, there exists an $H \in k[V]$ such that

$$G \cdot H - 1 \in \mathfrak{m}$$

Now multiply this by F to get

$$F \cdot G \cdot H - F \in \mathfrak{m}^2$$

(since $F \in \mathfrak{m}$). We have proved that (see definition A.1.89 on page 383)

$$\frac{F}{G} \sim \frac{F \cdot H}{1} \quad (\text{mod } \mathfrak{m}^2)$$

So every element of $\mathfrak{m}_p/\mathfrak{m}_p^2$ is equivalent to one from $\mathfrak{m}/\mathfrak{m}^2$.   □

This result is extremely useful because it implies:

COROLLARY 3.3.8. *Let $f: V \to W$ be a regular map of affine varieties and let $p \in V$. Then f induces a natural homomorphism of tangent spaces*

$$d_p f: T_{V,p} \to T_{W,f(p)}$$

If f is an isomorphism, then so is $d_p f$.

PROOF. The map f induces a homomorphism of rings

$$g: k[W] \to k[V]$$

If $\mathfrak{m} \subset k[V]$ is the maximal ideal corresponding to the point p, then

$$g^{-1}(\mathfrak{m}) = \mathfrak{n}$$

is the maximal ideal in $k[W]$ corresponding to $f(p)$ and we get a natural homomorphism of vector spaces

$$\mathfrak{n}/\mathfrak{n}^2 \to \mathfrak{m}/\mathfrak{m}^2$$

that induces a natural homomorphism of dual vector-spaces

$$T_{V,p} = (\mathfrak{m}/\mathfrak{m}^2)^* \to (\mathfrak{n}/\mathfrak{n}^2)^* = T_{W,f(p)}$$

□

We can use theorem 3.3.7 on the previous page to get a coordinate free version of the differential:

LEMMA 3.3.9. *Let V be an irreducible affine variety and let p be a point. If $f \in k[V]$ then $f - f(p) \in \mathfrak{m}$, where $\mathfrak{m}$ is the maximal ideal of regular functions that vanish at p and*

$$d_p f = im\,(f - f(p)) \in \mathfrak{m}/\mathfrak{m}^2$$

PROOF. The proof is very similar to that of lemma 3.3.5 on page 121: we lift f to a function $F \in k[X_1, \ldots, X_n]$ and note that the image of $F - F(p)$ in

$$\mathfrak{M}/\mathfrak{M}^2$$

consists of the *linear terms* of F — i.e., taking the quotient by $\mathfrak{M}^2$ kills off the quadratic and higher-degree terms.   □

We can say something about the tangent spaces of varieties:

THEOREM 3.3.10. *Let V be an n-dimensional irreducible affine variety. Then there exists an open set $U_V \subset V$ such that*

$$\dim T_{V,p} = n$$

for all points $p \in U_V$.

REMARK. Since open sets in an irreducible variety are "large," this means the tangent spaces are n-dimensional "almost everywhere," and the set of points where the dimension is $\neq n$ forms a subvariety of strictly lower dimension.

PROOF. First, suppose that $V \subset \mathbb{A}^{n+1}$ is a hypersurface defined by

$$F(X_1, \ldots, X_{n+1}) = 0$$

for some $F \in k[X_1, \ldots, X_{n+1}]$.

From definition 3.3.2 on page 119, the tangent space will be n-dimensional wherever all of the partial derivatives

$$\frac{\partial F}{\partial X_i} \neq 0$$

so the open set

$$U_V = V \cap \bigcap_{i=1}^{n+1} D\left(\frac{\partial F}{\partial X_i}\right)$$

has the required properties.

If V is not a hypersurface, there theorem 2.8.37 on page 114 implies the existence of a birational equivalence

$$f: V \to W$$

where $W \subset \mathbb{A}^{n+1}$ is an irreducible hypersurface. Furthermore, theorem 2.5.27 on page 86 shows that there exists an affine open set $R \subset V$ mapped isomorphically onto an open set in $R' \subset W$, so we define

$$U_V = f^{-1}(U_W \cap R') \cap R$$

and the conclusion follows from theorem 3.3.7 on page 123. $\square$

We can now define simple and singular points of a variety:

DEFINITION 3.3.11. If V is an n-dimensional irreducible affine variety, a point $p \in V$ is

(1) *simple* or *smooth* if the tangent space, $T_{V,p}$, is n-dimensional
(2) *singular* otherwise.

If all points of V are simple, V is said to be *smooth*.

REMARK. We have seen that the set of singular points of an irreducible variety form a subvariety of strictly lower dimension.

EXAMPLE 3.3.12. For instance, let $V \subset \mathbb{A}^2$ be defined by $Y^2 - X^3 = 0$ — a one-dimensional affine variety (see 3.3.1 on page 120). Let the images of X and Y in $k[V]$ be x and y, respectively.

At the point $(1, 1)$, the maximal ideal is $\mathfrak{m} = (x - 1, y - 1)$ and

$$\begin{aligned}
\mathfrak{m}^2 &= (x^2 - 2x + 1, (x - 1)(y - 1), y^2 - 2y + 1) \\
&= (x^2 - 2x + 1, (x - 1)(y - 1), x^3 - 2y + 1)
\end{aligned}$$

due to the relation $y^2 = x^3$ in $k[V]$. This implies that

$$\begin{aligned}
y &= (x^3 + 1)/2 \\
&= (x \cdot (2x - 1) + 1)/2 \\
&= (2x^2 - x + 1)/2 \\
&= (3x - 1)/2
\end{aligned}$$

in the quotient $q = \mathfrak{m}/\mathfrak{m}^2$. In addition, $x^2 = 2x - 1$ and $xy = x + y - 1 = 5x/2 - 3/2$. Now, we consider the point $(0,0)$ — where

$$\mathfrak{m} = (x,y)$$

and

$$\mathfrak{m}^2 = (x^2, xy, y^2) = (x^2, xy)$$

so that $\mathfrak{m}/\mathfrak{m}^2$ is the vector space generated by x and y. Since this is two dimensional, $(0,0)$ is a *singular* point of V.

THEOREM 3.3.13 (Jacobi Criterion). *Let $V \subset \mathbb{A}^n$ be an irreducible affine algebraic set with $I(V) = (F_1, \ldots, F_t) \subset k[X_1, \ldots, X_n]$ and let $p \in V$ be a point. Then the point p is smooth if and only if the Jacobian, $\mathcal{J}$ (see equation 3.3.2 on page 120), of the map*

$$\begin{bmatrix} F_1 \\ \vdots \\ F_t \end{bmatrix} : \mathbb{A}^n \to \mathbb{A}^t$$

evaluated at p, has rank $n - \dim V$.

REMARK. We can use this to *compute* the dimension of V — it is $n - r$, where r is the *maximal rank* of $\mathcal{J}$ as p runs over all of the points of V. Since the set of singular points of a variety is "small", we get a "statistical" algorithm for computing the dimension of V:

> Plug *random* points of V into the variables of $\mathcal{J}$ and compute the rank of the resulting matrix. There is a high probability that this will equal the maximal rank, r.

In many cases, one can carry this out by plugging random points of $\mathbb{A}^n$ into $\mathcal{J}$: The maximal rank of $\mathcal{J}$ is the smallest number r such that $\mathcal{R}(\mathcal{J}, r) = \mathbb{A}^n$ and we get a chain of algebraic sets (see example 7 on page 37):

$$\mathcal{R}(\mathcal{J}, 0) \subset \cdots \subset \mathcal{R}(\mathcal{J}, r-1) \subset \mathcal{R}(\mathcal{J}, r) = \mathbb{A}^n$$

If $V \not\subset \mathcal{R}(\mathcal{J}, r-1)$, then this r will be the maximal rank of $\mathcal{J}$ on V as well.

> Jakob Steiner (1796–1863) was a Swiss mathematician who specialized in synthetic geometry (geometry that uses axioms and logical reasoning rather than analytic methods). Despite his hostility to analytic methods, he is responsible for several interesting examples of surfaces in algebraic geometry — all *attempts* to embed $\mathbb{R}P^2$ in $\mathbb{R}^3$. Twentieth-century topology has proved that no such embedding is possible — see [**115**].

EXAMPLE 3.3.14. Steiner's Cross-cap
Consider the mapping

$$\begin{aligned} f: \mathbb{A}^3 &\to \mathbb{A}^3 \\ (X, Y, Z) &\mapsto (YZ, 2XY, X^2 - Y^2) \end{aligned}$$

Note that this map sends diametrically opposite points to the *same* point. It follows that the image of the unit sphere under this map is the same as that of $\mathbb{R}P^2$. Figure 3.3.2 on the facing page shows that it is not an embedding of $\mathbb{R}P^2$ in $\mathbb{R}^3$.
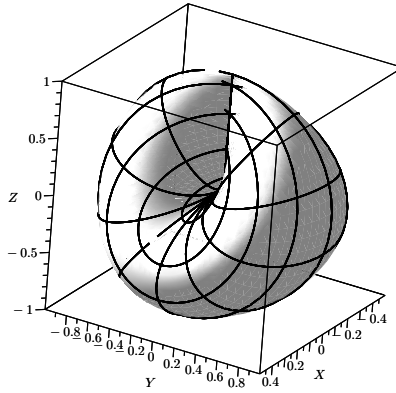
FIGURE 3.3.2.   Steiner's crosscap

Recall the rational parametrization of the two-sphere in example 2.5.31 on page 89. Plugging it into f gives a parametric representation of Steiner's crosscap

$$
\begin{aligned}
X &= 2t(s^2 + t^2 - 1/(1 + s^2 + t^2)^2 \\
Y &= 8st/(1 + s^2 + t^2)^2 \\
Z &= 4(s^2 - t^2)/(1 + s^2 + t^2)^2
\end{aligned}
$$
(3.3.6)

To get an implicit equation, use techniques like those in section 1.3 on page 10 — *eliminate s* and *t* from equations 3.3.6. We can accomplish this by either performing several resultant calculations or using Gröbner bases: form a Gröbner basis for the ideal

$$
\begin{aligned}
\mathfrak{z} = (X(1 + s^2 + t^2)^2 - 2t(s^2 + t^2 - 1), \\
Y(1 + s^2 + t^2)^2 - 8st, \\
Z(1 + s^2 + t^2)^2 - 4(s^2 - t^2))
\end{aligned}
$$

with lexicographic ordering, and with s and t ordered *higher* than X, Y, and Z (see proposition 2.3.13 on page 52). Only one term in this Gröbner basis does *not* contain s or t. We set it to 0 to get our implicit equation:

$$
F(X, Y, Z) = 4X^2 \left(X^2 + Y^2 + Z^2 + Z \right) + Y^2 \left(Y^2 + Z^2 - 1 \right) = 0
$$

It is two-dimensional by corollary 2.8.30 on page 111. The Jacobian is

$$
\mathcal{J} = \left[\begin{array}{c}
8X \left(X^2 + Y^2 + Z^2 + Z \right) + 8X^3 \\
8X^2Y + 2Y \left(Y^2 + Z^2 - 1 \right) + 2Y^3 \\
4X^2(2Z + 1) + 2Y^2Z
\end{array} \right]
$$

and smooth points are points where $3 - \text{rank}(\mathcal{J}) = 2$ or $\text{rank}(\mathcal{J}) = 1$. Singular points are where the rank is less than 1 or 0, i.e. points where *all* of these derivatives vanish.

It is not hard to see that F and its derivatives vanish if $X = Y = 0$ so we get a singular line. To see what else can happen, find a Gröbner basis for the terms in $\mathcal{J}$ with lexicographic ordering: $X \succ Y \succ Z$. The first term in the basis is

$$-ZY - YZ^2 + YZ^3 + YZ^4$$

If $Y \neq 0$ and $Z \neq 0$, we can divide by YZ to get

$$-1 - Z + Z^2 + Z^3 = 0$$

which only has $Z = \pm 1$ as roots. If $Z = -1$, add $Z + 1$ to the original list of terms in $\mathcal{J}$ and take another Gröbner basis. We get

$$(Z + 1, 2\,X^2 + Y^2)$$

so we get $Y = \pm X\sqrt{-2}$ as solutions. When we set $Z = +1$ and adjoin $Z - 1$ to the terms in $\mathcal{J}$ and take a Gröbner basis, we get

$$(Z - 1, Y^2, X)$$

so that the *only* singularities are

(1) $X = Y = 0$, Z *arbitrary,* and
(2) $Z = -1$, $Y = \pm X\sqrt{-2}$.

It is possible to embed real projective space in *higher* dimensional affine spaces — see section 5.2.4 on page 232.

The following result (from [**57**]) is useful in calculating ranks of matrices:

LEMMA 3.3.15. *If A is an $m \times n$ matrix with $A_{1,1} \neq 0$ define*

$$d_{i,j} = \det \begin{bmatrix} A_{1,1} & A_{1,j} \\ A_{i,1} & A_{i,j} \end{bmatrix}$$

for all $i = 2, \ldots m$ and $j = 2, \ldots, n$. Then

$$\text{rank}(A) = 1 + \text{rank} \begin{bmatrix} d_{2,2} & \cdots & d_{2,n} \\ \vdots & \ddots & \vdots \\ d_{m,2} & \cdots & d_{m,n} \end{bmatrix}$$

REMARK. This suits our purposes somewhat better than Gaussian Elimination (the usual way of calculating rank) because it does not require division.

EXAMPLE 3.3.16. Let $V \subset \mathbb{C}^4$ be the complex variety defined by the equations

$$\begin{aligned} XY + YZ + ZW - 1 &= 0 \\ X + Y + Z + W - 2 &= 0 \end{aligned}$$

with Jacobian

$$\mathcal{J} = \begin{bmatrix} Y & 1 \\ X + Z & 1 \\ Y + W & 1 \\ Z & 1 \end{bmatrix}$$

To determine its rank, we can interchange its columns to get

$$\begin{bmatrix} 1 & Y \\ 1 & X + Z \\ 1 & Y + W \\ 1 & Z \end{bmatrix}$$

Using lemma 3.3.15 on the preceding page, we get

$$\text{rank}(\mathcal{J}) = 1 + \text{rank} \begin{bmatrix} X + Z - Y \\ W \\ Z - Y \end{bmatrix}$$

This means that the rank of $\mathcal{J}$ is 2 unless $X = W = 0$ and $Y = Z$, when its rank is 1. This defines a line that intersects V in the single point $(0, 1, 1, 0)$. It follows that V is two-dimensional and smooth everywhere except for the singular point $(0, 1, 1, 0)$.

We conclude this section with an important class of smooth varieties:

LEMMA 3.3.17. *An algebraic group (see definition 2.7.7 on page 97) is smooth.*

PROOF. Suppose that an algebraic group, V, has a singular point, s. Then there exists an isomorphism

$$f_{s,a} \colon V \to V$$

mapping s into any other point, $a \in V$. It follows that *all* points of V are singular, by corollary 3.3.8 on page 124, which contradicts the fact that singular points form a subvariety of lower dimension than V. □

DEFINITION 3.3.18. Let $f \colon V \to W$ be a regular map of affine varieties and let $p \in V$ be a point. Then f is said to be *étale* at p if the induced map of tangent spaces

$$T_{V,p} \to T_{W,f(p)}$$

is an isomorphism.

REMARK. In differential geometry, étale maps are *local isomorphisms,* i.e. if a map is étale at p then there exists a neighborhood $p \in U$ that is mapped isomorphically by f. The idea is that in a closeup of figure 3.3.1 on page 120, we get a 1-1 mapping between V and its tangent space, as in figure 3.3.3 on the next page

This is a useful property of étale maps that unfortunately does not apply in algebraic geometry. The problems is that open sets (=neighborhoods) are very large. In this example, every open set that contains the point p includes *both* branches of the function $Y^2 = X^3$.

This problem is solved by redefining the concept of neighborhood — replacing them with so-called *étale* neighborhoods. This is the basis of a deep field of mathematics called *étale cohomology* (see [**164**] or [**112**]).
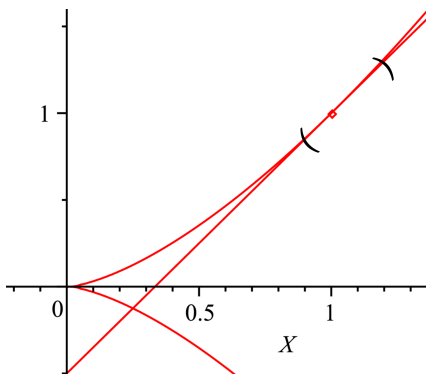
FIGURE 3.3.3.   A neighborhood

EXERCISES.

1. Determine the simple and singular points of the algebraic set

$$Y^2 = X^3 + X^2 \subset \mathbb{A}^2$$

2. For what values of $a \in \mathbb{C}$ is the variety $V \subset \mathbb{A}^4$ defined by

$$
\begin{aligned}
X^3 + Y^3 + Z^3 + W^3 &= a \\
X^2 + Y^2 + Z^2 + W^2 &= 1
\end{aligned}
$$

(3.3.7)

smooth? What is its dimension in these cases?

3. Steiner's *Roman surface* (called this because he was in Rome at the time) is defined to be the image of the unit sphere under the map

(3.3.8) $$(X, Y, Z) \mapsto (YZ, XZ, XY)$$

This clearly maps diametrically opposite points to the same point, so the image of the sphere is a representation of $\mathbb{R}P^2$. Find an implicit equation for this surface and determine its singularities (see example 3.3.14 on page 126).

4. Show that the image of the unit-sphere under a slight modification of the map in equation 3.3.8

$$f \colon \mathbb{R}^3 \to \mathbb{R}^4$$

defined by

$$f(X, Y, Z) = (YZ, XZ, XY, Y^2 - Z^2)$$

is an embedded copy of $\mathbb{R}P^2$ in $\mathbb{R}^4$.

5. If $W \subset \mathbb{A}^2$ is a smooth variety defined by a single equation $f(X, Y) = 0$, show that there exist functions $r, s \in k[X, Y]$ such that

$$r \cdot \frac{\partial f}{\partial X} + s \cdot \frac{\partial f}{\partial Y} \equiv 1 \quad (\text{mod } (f(X, Y)))$$
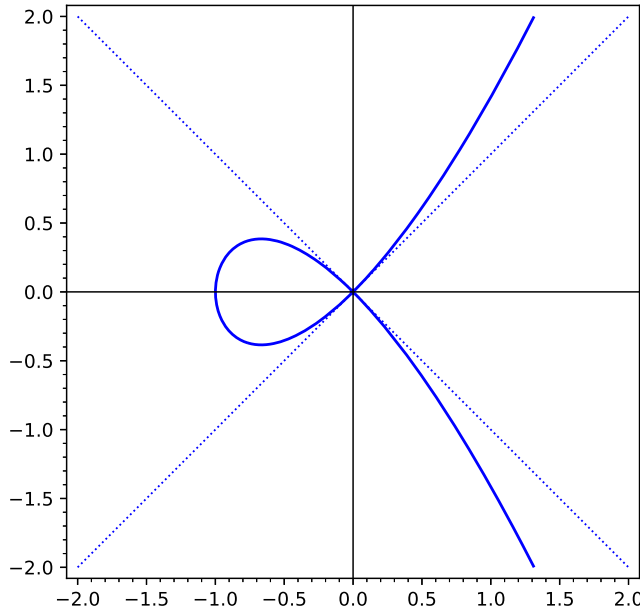
FIGURE 3.3.4.   Tangent cone

3.3.1. The tangent cone . At singular points, one can devise a similar invariant called the *tangent cone* — and it directly generalizes the tangent space.

If $V \subset \mathbb{A}^n$ is an affine variety defined by the ideal $\mathfrak{I} \subset k[X_1, \ldots, X_n]$ that passes through the origin, let $\mathrm{IT}(f)$ be the lowest-degree term of $f \in k[X_1, \ldots, X_n]$, and let $\mathrm{IT}(\mathfrak{I})$ be the homogeneous ideal generated by $\mathrm{IT}(f)$ for all $f \in \mathfrak{I}$.

DEFINITION 3.3.19. The *tangent cone* of V as defined above *at the origin* is the Zariski closed subset of $\mathbb{A}^n$ defined by $\mathrm{IT}(\mathfrak{I})$.

REMARK. We compute tangent cones at other points by displacing them to the origin.

EXAMPLE 3.3.20. Consider $\mathfrak{I} = (Y^2 - X^2 - X^3)$. Then $\mathrm{IT}(\mathfrak{I}) = (Y^2 - X^2)$ and the tangent cone at the origin is given by $Y^2 - X^2 = 0$ or $Y = \pm X$. In figure 3.3.4, the dotted lines are the tangent cone.

Note that this is an actual generalization of the tangent space. For instance, we can compute the tangent cone at $(1, \sqrt{2})$ by displacing this point to the origin and replacing $\mathfrak{I} = (Y^2 - X^2 - X^3)$ by

$$\mathfrak{I} = ((Y + \sqrt{2})^2 - (X + 1)^2 - (X + 1)^3)$$
$$= (-X^3 - 4X^2 + Y^2 + 2\sqrt{2}Y - 5X)$$

In this case, $\mathrm{IT}(\mathfrak{I}) = (+2\sqrt{2}Y - 5X)$, a linear term, and the tangent cone becomes a tangent line, as in figure 3.3.5 on the following page.

This is to be expected, though. The coefficients of leading *linear* terms are precisely the derivatives of the defining polynomials with variables set to zero.
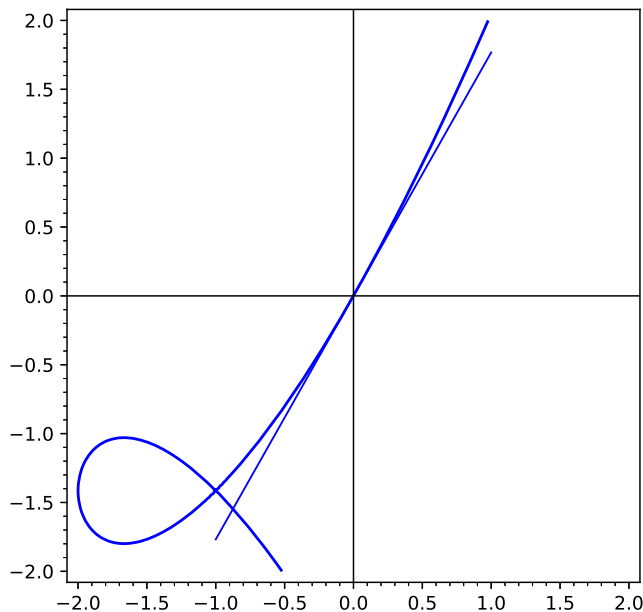
FIGURE 3.3.5.   Tangent cone becomes the tangent space

If the ideal, $\mathfrak{I}$, is complex (i.e., with many generators), computing $\mathrm{IT}(\mathfrak{I})$ can be daunting. See [**117**] for an algorithm based on the one for computing Gröbner bases in section 2.3.2 on page 49.

For a "coordinate free" treatment of this subject, see section 4.4.2 on page 191.

EXERCISES.

6. Suppose we displace a point that is *not* on the affine variety to the origin and compute the tangent cone (for instance, displace $(3, 5)$ to the origin in example 3.3.20 on the previous page)?

3.3.2. Local parameters. We can use Zariski's construction to develop a "local coordinate system" on a variety:

DEFINITION 3.3.21. Let $p \in V$ be a point in an irreducible affine variety with corresponding maximal ideal $\mathfrak{m}$. Elements $u_1, \ldots, u_r \in \mathfrak{m}$ that map to basis elements of $\mathfrak{m}/\mathfrak{m}^2$ are called *local parameters* of V at p.

REMARK. What we call local parameters here are sometimes called *local uniformizing parameters* or *uniformizers*.

Local parameters determine the local behavior of a variety to the extent that every regular function can be expanded into a power series in local parameters.

EXAMPLE 3.3.22. Let $V \subset \mathbb{A}^2$ be defined by $X^2 + Y^2 = 1$. Then $k[V] = k[X, Y]/(X^2 + Y^2 - 1)$.

Let x and y be the images of X and Y, respectively, under the projection

$$k[X, Y] \to k[V]$$

At the point $(0, 1)$ the maximal ideal $\mathfrak{m} = (x, y - 1)$, and

$$
\begin{aligned}
(y - 1)^2 &= y^2 - 2y + 1 \\
&= 1 - x^2 - 2y + 1 \\
&= 2(1 - y) + x
\end{aligned}
$$

so

$$y - 1 \equiv \frac{1}{2}\left(x^2 - (y-1)^2\right) \in \mathfrak{m}^2 \subset k[V]$$

and $\mathfrak{m}/\mathfrak{m}^2$ is one-dimensional over k, and x is a local parameter at the point $(0, 1)$.

A simple application of Nakayama's lemma shows that

LEMMA 3.3.23. *Let $p \in V$ and $\mathfrak{m}$ be as in definition 3.3.21 on the facing page. If $u = \{u_1, \dots, u_r\} \in \mathfrak{m}$ are local parameters, then the set u generates*

$$\mathfrak{m}_p = \mathfrak{m} \cdot \mathcal{O}_{V,p}$$

as an ideal in $\mathcal{O}_{V,p}$.

REMARK. For instance, in example 3.3.22, we have

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

since functions that do not vanish at $(0, 1)$ are invertible in $\mathcal{O}_{V,p}$.

Corollary 3.2.5 on page 118 shows that at *least* dim R generators are needed to generate $\mathfrak{m}$.

At singular points, more generators are needed but theorem 3.2.6 on page 118 still applies. For instance, let $\mathfrak{I} = (X^3 - Y^2) \subset k[X, Y]$, let x and y be the images of X and Y, respectively, in the quotient, and let $V = \mathcal{V}(\mathfrak{I})$. The point $(0, 0) \in V$ is singular (see the example 3.3.3 on page 120) and the maximal ideal $\mathfrak{m} = (x, y) \subset R = k[X, Y]/(X^3 - Y^2)$ requires *two* generators, namely x and y. On the other hand $\mathfrak{m}^2 \subset (x)$ and $\mathfrak{m}^3 \subset (y)$, since $y^2 = x^3$.

PROOF. We know that u generates

$$\frac{\mathfrak{m}}{\mathfrak{m}^2} = \frac{\mathfrak{m}_p}{\mathfrak{m}_p^2}$$

Let $N \subset \mathcal{O}_{V,p}$ be the $\mathcal{O}_{V,p}$-module generated by $\{u_1, \dots, u_r\}$. Then we have

$$\mathfrak{m} = N + \mathfrak{m} \cdot \mathfrak{m}$$

as $\mathcal{O}_{V,p}$-modules. If we form the quotient by N, we get

$$\frac{\mathfrak{m}}{N} = \mathfrak{m} \cdot \frac{\mathfrak{m}}{N}$$

as modules. Nakayama's Lemma (A.1.80 on page 378) implies that there exists an element $t \in \mathcal{O}_{V,p}$ such that

$$(3.3.9) \qquad\qquad\qquad t \cdot \frac{\mathfrak{m}}{N} = 0$$

and $t \equiv 1 \pmod{\mathfrak{m}}$. Since $\mathfrak{m}$ is maximal, this means t is invertible and equation 3.3.9 implies that $\mathfrak{m}/N = 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We can abstract out the property of $\mathcal{O}_{V,p}$ that corresponds to a simple point of an affine variety:

DEFINITION 3.3.24. Let $(R, \mathfrak{m})$ be a d-dimensional local ring with maximal ideal, $\mathfrak{m}$. The ring R is said to be *regular* if there exist elements $x_1, \ldots, x_d \in R$ such that $\mathfrak{m} = (x_1, \ldots, x_d)$. A ring, R, is *regular* if $R_\mathfrak{m}$ is regular for all maximal ideals $\mathfrak{m} \subset R$.

REMARK. This definition was first proposed by Krull in [**96**].

Suppose $p \in V$ is a simple point of an irreducible affine variety and suppose $\{u_1, \ldots, u_n\}$ is a set of local parameters. Then the subvariety $V(u_i)$ is $n-1$-dimensional by 2.8.30 on page 111 and its tangent space is also $n-1$-dimensional (because u_i is a basis element of the original n-dimensional tangent space).

Simple induction using proposition 2.8.32 on page 111 shows that:

LEMMA 3.3.25. *If $p \in V$ is a simple point of an irreducible affine variety and $\{u_1, \ldots, u_n\}$ is a set of local parameters at p then*

(1) p is a simple point in any of the subvarieties $U_i = V(u_i) \subset V$ given by $u_i = 0$
(2) the intersection

$$U(i_1, \ldots, i_t) = \bigcap_{j=1}^{t} U_{i_j}$$

is of codimension t at p.

REMARK. That proposition shows that the codimension of $U(i_1, \ldots, i_t)$ is $\leq t$ or that the dimension is $\geq \dim V - t$. But the dimension of the tangent space is precisely $\dim V - t$ because the u_i form a basis, and the dimension of a variety is never greater than that of its tangent space.

Proposition 2.8.33 on page 111 shows that an ideal of height t in an affine k-algebra is generated by $\geq t$ elements,

3.3.3. Power series expansions. We can use local parameters to give power series expansions of local rings at a point. For instance, the local ring of a simple point embeds within a ring of formal power series and inherits some of its algebraic properties. This will give interesting geometric and algebraic information.

To construct series-expansions, we need the concept of *completion* of a local ring:

DEFINITION 3.3.26. Let R be a local ring with maximal ideal $\mathfrak{m}$. The *completion of R*, denoted $\hat{R}$ is the inverse limit (see definition A.5.26)

$$\hat{R} = \varprojlim R/\mathfrak{m}^n$$

REMARK. Elements of $\hat{R}$ are infinite sequences

$$(a_1, a_2, \dots)$$

with $a_n \in R/\mathfrak{m}^n$ such that $a_n \equiv a_{n+1} \pmod{\mathfrak{m}^n}$. Since addition and multiplication is done element-wise, $\hat{R}$ is a ring.

These completions can be huge because there is a 1-1 correspondence (not a homomorphism!) between elements of $\hat{R}$ and the set

$$\prod_{n=0}^{\infty} \frac{\mathfrak{m}^n}{\mathfrak{m}^{n+1}}$$

(where $\mathfrak{m}^0 = R$). This is because, given a sequence $(a_1, \dots, a_d)$, there are $|\mathfrak{m}^d/\mathfrak{m}^{d+1}|$ possible choices for a_{d+1}.

For instance, let $R = \mathbb{Z}[1/3, 1/5, \dots] \subset \mathbb{Q}$ be the ring of all rational numbers whose denominators are odd. There is a unique maximal ideal

$$2 \cdot R \subset R$$

and the completion

$$\hat{R} = \mathbb{Z}_{(2)}$$

the ring of *2-adic integers*, which is uncountable since its cardinality is the same as that of $\mathbb{Z}_2^{\aleph_0}$.

There is an obvious ring-homomorphism

(3.3.10) $g: R \to \hat{R}$

that sends an element $r \in R$ to

$$(r \bmod \mathfrak{m}, r \bmod \mathfrak{m}^2, \dots)$$

LEMMA 3.3.27. *Let R be a local ring with maximal ideal* $\mathfrak{m}$ *and inclusion into the completion*

$$g: R \to \hat{R}$$

If R is noetherian (see definition A.1.48 on page 360) then g is an inclusion.

REMARK. Of course, coordinate rings and their localizations are always noetherian.

This gives some insight into why $\hat{R}$ is called the *completion* of R. Let $0 < \alpha < 1$ be any number and define a bizarre metric on R: if $r, s \in R$ and $r \equiv s \pmod{\mathfrak{m}^n}$ but $r \not\equiv s \pmod{\mathfrak{m}^{n+1}}$, define the *distance between r and s* to be α^n.

With this metric (it is an exercise to see that it is a metric) we can define Cauchy sequences and ask whether they converge to a limit. The answer is that they do, in $\hat{R}$ — whose relation to R is like that of $\mathbb{R}$ to $\mathbb{Q}$.

PROOF. The kernel of g is

$$N = \bigcap_{n=1}^{\infty} \mathfrak{m}^n$$

and, if R is noetherian, lemma A.1.81 on page 379 implies that this vanishes. $\square$

The reader may wonder why we are so preoccupied with the completion of a local ring (as interesting as that topic might be). The answer is:

THEOREM 3.3.28. *Let $p \in V$ be a point of an irreducible affine variety and let*

$$\{u_1, \ldots, u_r\}$$

be a set of local parameters at p. Then there exists a surjective map

$$f: k[[U_1, \ldots, U_r]] \quad \to \quad \widehat{\mathcal{O}_{V,p}}$$
$$U_i \quad \mapsto \quad u_i$$

(where the U_i are indeterminates) that is an isomorphism whenever p is a simple point.

REMARK. So, if p is simple, $\widehat{\mathcal{O}_{V,p}}$ is our ring of power-series and the image of an element under the inclusion

$$\mathcal{O}_{V,p} \to \widehat{\mathcal{O}_{V,p}}$$

is its power-series expansion.

If p is singular, $r > \dim V$, which is the transcendence degree of the field $k(V) \supset \mathcal{O}_{V,p}$. It follows that there is an algebraic dependency between the $\{u_i\}$ in $k(V)$ which (after clearing denominators) leads to one in $\mathcal{O}_{V,p}$ and $\widehat{\mathcal{O}_{V,p}}$, giving us a nontrivial "Taylor series" equal to 0!

PROOF. To see that it is surjective, note that the u_i generate $\mathfrak{m}_p$ so the set of all monomials of degree n generates $\mathfrak{m}_p^n$ and the set of all sums of the first n terms of power series in $k[[u_1, \ldots, u_r]]$ generates

$$\frac{\mathcal{O}_{V,p}}{\mathfrak{m}_p^{n+1}}$$

It follows that we get a commutative diagram

$$
\begin{array}{ccc}
k[[u_1, \ldots, u_r]] & \xdashrightarrow{\ f\ } & \widehat{\mathcal{O}_{V,p}} \\
\downarrow{\scriptstyle \pi_n} & & \downarrow \\
S_n & \twoheadrightarrow & \dfrac{\mathcal{O}_{V,p}}{\mathfrak{m}_p^{n+1}}
\end{array}
$$

for all n, where $\pi_n: k[[u_1, \ldots, u_r]] \to S_n$ projects onto the sum of the first n terms. The dotted arrow exists by the universal property of inverse limits (see statement 2 in definition A.5.26 on page 453). It is easily seen to be surjective, too.

We have to be a bit careful here: $\mathbb{Z}$ surjects to $\mathbb{Z}_{2^n}$ for all n but does *not* surject to the (uncountable) inverse limit, $\mathbb{Z}_{(2)}$. The key is that the power series ring consists of *infinite sequences* of modules (generated by all monomials of a given degree), *each* of which surjects to a corresponding $k[V]_p / \mathfrak{m}_p^{n+1}$.

If p is a simple point, we claim that f is also injective. Suppose some power series $p \in k[[U_1, \ldots, U_n]]$ is in the kernel. This means that its lowest degree term, say L of degree ℓ, has the property that

$$L(u_1, \ldots, u_n) \in \mathfrak{m}_x^{\ell+1}$$

Since L is a homogeneous polynomial in $U_1, \ldots, U_n$, we can use lemma A.3.14 on page 418 to transform it into a U_n-general polynomial (see definition A.3.10):

$$L(U_1, \ldots, U_n) =$$
$$\alpha U_n^\ell + G_1(U_1, \ldots, U_{n-1})U_n^{\ell-1} + \cdots + G_\ell(U_1, \ldots, U_{n-1})$$

Since elements of $\mathfrak{m}_x^\ell$ can be written as homogeneous polynomials in the $\{u_i\}$ of degree ℓ with coefficients in $\mathfrak{m}_x$, we get

$$L(u_1, \ldots, u_n) =$$
$$\alpha u_n^\ell + G_1(u_1, \ldots, u_{n-1})u_n^{\ell-1} + \cdots + G_\ell(u_1, \ldots, u_{n-1})$$
$$= \mu u_n^\ell + H_1 u_n^{\ell-1} + \cdots + H_\ell$$

where $\mu \in \mathfrak{m}_x$ and the H_i are homogeneous polynomials in $u_1, \ldots, u_{n-1}$. Subtraction gives us

$$(\alpha - \mu)u_n^\ell \in (u_1, \ldots u_{n-1})$$

and we claim that $\alpha - \mu$ is a unit, since $\alpha \notin \mathfrak{m}_x$. It follows that

$$u_n^\ell \in (u_1, \ldots u_{n-1})$$

which implies (by the strong Nullstellensatz) that

$$\mathcal{V}(u_n) \supset \mathcal{V}(u_1) \cap \cdots \cap \mathcal{V}(u_{n-1})$$

and

$$\mathcal{V}(u_1) \cap \cdots \cap \mathcal{V}(u_{n-1}) = \mathcal{V}(u_1) \cap \cdots \cap \mathcal{V}(u_n)$$

contradicting lemma 3.3.25 on page 134. $\square$

Actually computing these power series expansions is simple in some cases.

EXAMPLE 3.3.29. Let $V = \mathcal{V}(X^2 + Y^2 - 1)$ and let $p = (0, 1)$. Then $k[V] = k[X, Y]/(X^2 + Y^2 - 1) = k[x, y]$ where x and y are the images, respectively, of X and Y. Continuing in the vein of example 3.3.22 on page 133

$$
\begin{aligned}
(y - 1)^2 &= y^2 - 2y + 1 \\
&= 1 - x^2 - 2y + 1 \\
&= 2(1 - y) + x
\end{aligned}
$$

so

$$y - 1 \equiv \frac{x^2}{2} \in \mathfrak{m}^2 \subset k[V]$$

and we can plug this expression into itself to give

$$
\begin{aligned}
y &= 1 + \frac{1}{2}x^2 + \frac{1}{2}(y-1)^2 \\
&\equiv 1 + \frac{1}{2}x^2 \quad (\mathrm{mod}\ \mathfrak{m}^2) \\
&= 1 + \frac{1}{2}x^2 + \frac{1}{2}\left(\frac{1}{2}\left(x^2 - (y-1)^2\right)\right)^2 \\
&= 1 + \frac{1}{2}x^2 + \frac{1}{8}x^4 - \frac{1}{4}x^2(y-1)^2 + \frac{1}{8}(y-1)^4 \\
&\equiv 1 + \frac{1}{2}x^2 + \frac{1}{8}x^4 \quad (\mathrm{mod}\ \mathfrak{m}^4)
\end{aligned}
$$

The reason this is legitimate is that the new terms added in each step are contained in higher and higher powers of the maximal ideal $\mathfrak{m}$ so this process converges in $\widehat{\mathcal{O}_{V,p}} = k[[X]]$ (but not in $\mathcal{O}_{V,p}$ itself). We ultimately get the Taylor series expansion for $y = \sqrt{1 - x^2}$.

We can deduce an interesting geometric fact from the embedding $\mathcal{O}_{V,p} \subset k[[U_1, \ldots, U_n]]$:

LEMMA 3.3.30. *If $p \in V$ is a simple point in an affine variety, then p must be contained in a single irreducible component of V.*

PROOF. Consider the open set

$$
X = V \setminus \bigcup_i V_i
$$

where the V_i are the irreducible components of V that do not contain p. Then

$$
k[X] \subset \mathcal{O}_{V,p} \subset k[[U_1, \ldots, U_n]]
$$

by lemma 3.3.27 on page 135. Since $k[[U_1, \ldots, U_n]]$ is an integral domain, so is $k[[X]]$ which implies that X is irreducible (see proposition 2.4.18 on page 68). □

This means that, in varieties that are connected but not irreducible, distinct irreducible components meet at singular points:

EXAMPLE 3.3.31. Consider the variety $V = \mathcal{V}(XY) \subset \mathbb{A}^2$. This the union of the X and Y axes — see figure3.3.6 on the facing page.

Its irreducible components are the X and Y axes and they meet at the origin, which is the only singular point in V.

Lemma 3.3.30 also has interesting consequences for smooth varieties:

COROLLARY 3.3.32. *A smooth variety is a disjoint union of its irreducible components.*

REMARK. For instance, this is true of all algebraic groups, by lemma 3.3.17 on page 129.
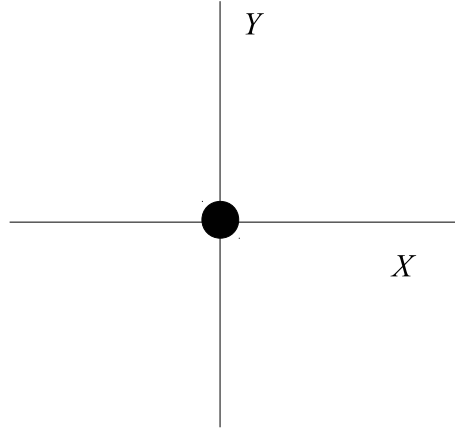
FIGURE 3.3.6.   A connected, reducible variety

EXERCISES.

7. If $p \in V$ is a simple point of a variety V, show that $\mathcal{O}_{V,x}$ (see definition 3.2.1 on page 117) is an integral domain.

8. If $p \in V$ is a simple point of a variety V, $R = \mathcal{O}_{V,x}$, and $u_1, \ldots, u_m$ are local parameters at x (so the component of V containing x is m-dimensional), show that
$$\operatorname{im} u_{i+1} \in \frac{R}{(u_1, \ldots, u_i)}$$
is a non-zero, non-zero-divisor for all $i = 1, \ldots m - 1$.

9. If G is an algebraic group and G_0 is the component of G containing the identity element, show that G_0 is a closed, normal subgroup of G and that the quotient G/G_0 is finite.

3.3.4. Intersection multiplicity of plane curves. We can give a rigorous definition of intersection multiplicity promised in the first chapter. This is a special case of a much more general construction in section 5.8.2 on page 276.

DEFINITION 3.3.33. If C_1 and C_2 are curves in $\mathbb{A}^2$ given by
$$\begin{aligned} f(X,Y) &= 0 \\ g(X,Y) &= 0 \end{aligned}$$
that intersect at a point p, we define the *intersection multiplicity* $I_p(f,g) = I_p(C_1, C_2)$ as a function with the properties
  (1) $I_p(f,g) = I_p(g,f)$
  (2) $I_p(f,g)$ is a nonnegative integer that is 0 if $C_1 \cap C_2 \cap \{p\} = \varnothing$
  (3) $I_p(f,g)$ is infinite if f and g have a common factor that vanishes at p.
  (4) $I_p(X,Y) = 1$ where $p = (0,0)$.

(5) If $\ell: \mathbb{A}^2 \to \mathbb{A}^2$ is an invertible affine linear map, then $I_{\ell(p)}(f \circ \ell^{-1}, g \circ \ell^{-1}) = I_p(f, g)$.

(6) $I_p(f, g_1 \cdot g_2) = I_p(f, g_1) + I_p(f, g_2)$

(7) $I_p(f + r \cdot g, g) = I_p(f, g)$ where $r(X, Y) \in k[X, Y]$ is any element.

REMARK. Most of these properties are intuitively clear, especially the first four. The fifth states that multiplicity does not depend on location. The sixth simply encodes the idea that $I_p(C_1, C_2 \cup C_2') = I_p(C_1, C_2) + I_p(C_1, C_2')$. The seventh is a property that our naive definition of intersection multiplicities (definition 1.4.6 on page 23) had.

Now we will give a rigorous method for computing $I_p(f, g)$. It based on the observation that, if $\mathfrak{I}_1$ and $\mathfrak{I}_2$ are ideals defining algebraic sets V_1 and V_2, respectively, then $V_1 \cap V_2$ is defined by $\sqrt{\mathfrak{I}_1 + \mathfrak{I}_2}$. In addition, example 2.3.17 on page 55 suggests that *failing* to take the radical counts intersections with the proper multiplicities.

The one remaining consideration is that we do not want to count *all* of the intersections — only those in the neighborhood of a given intersection-point. This is solved by doing computations in the local *coordinate ring at a point* rather than the complete coordinate ring.

LEMMA 3.3.34. *If C_1 and C_2 are curves in $\mathbb{A}^2$ given by*

$$
\begin{aligned}
f(X, Y) &= 0 \\
g(X, Y) &= 0
\end{aligned}
$$

then

$$
I_p(f, g) = \dim_k V_p = \mathcal{O}_{\mathbb{A}^2, p} / (f, g)
$$

where V_p is a vector-space over k.

PROOF. We go through the list of properties in definition 3.3.33 on the previous page. Properties 1, 3, and 4 are fairly clear.

Property 5 follows from the fact that such affine linear maps induce isomorphisms of $k[X, Y]$ and its localizations.

Property 2 follows from the fact that $\mathcal{O}_{\mathbb{A}^2, p}$ is a local ring: if f or g are nonzero at p they define *units* and the quotient V_p is a trivial ring, which we regard as a vector-space of dimension 0.

Property 7 follows from the fact that, as *ideals*

$$
(f, g) = (f + r \cdot g, g)
$$

The only thing that remains to be proved is property 6. Define a homomorphism of vector-spaces (and rings, for that matter)

$$
\pi: \frac{\mathcal{O}_{\mathbb{A}^2, p}}{(f, g_1 \cdot g_2)} \to \frac{\mathcal{O}_{\mathbb{A}^2, p}}{(f, g_2)}
$$

that simply maps an element to its equivalence class modulo (f, g_2). This will clearly be surjective. Now define a homomorphism of *k*-vector-spaces

$$
\iota: \frac{\mathcal{O}_{\mathbb{A}^2, p}}{(f, g_1)} \to \frac{\mathcal{O}_{\mathbb{A}^2, p}}{(f, g_1 \cdot g_2)}
$$

that sends $v \in \frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_1)}$ to $v \cdot g_2 \in \frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_1 \cdot g_2)}$. We claim that ι is injective. If $\iota(v) = 0$, then

$$v \cdot g_2 = a \cdot f + b \cdot g_1 \cdot g_2$$

with $a, b \in \mathcal{O}_{\mathbb{A}^2,p}$. Let $d \in k[X,Y]$ clear the denominators of a, b, $v \cdot g_2$ i.e. so $ad, bd, vg_2d \in k[X,Y]$. Then

$$
\begin{aligned}
v \cdot g_2 \cdot d &= a \cdot f \cdot d + b \cdot g_1 \cdot g_2 \cdot d \\
d(v - bg_1)g_2 &= a \cdot f \cdot d
\end{aligned}
$$

are equations in $k[X,Y]$. Since $f \nmid g_2$ and $k[X,Y]$ is a unique factorization domain (see lemma A.3.2 on page 412), we must have

$$f \mid d(v - bg_1)$$

so, in $\mathcal{O}_{\mathbb{A}^2,p}$

$$v = t \cdot f + bg_1$$

and $v = 0 \in \frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_1)}$.

We claim the nullspace of π is exactly equal to the image of ι. Certainly, the image of ι maps to 0 in $\frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_2)}$. Suppose $\pi(x) = 0$ for $x \in \frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_1 \cdot g_2)}$. Then $x = a \cdot f + b \cdot g_2$ so, modulo $(f, g_1 \cdot g_2)$, x is equivalent to $b \cdot g_2 = \iota(b)$.

The conclusion follows from basic properties of vector-spaces, which imply that

$$\frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_1 \cdot g_2)} \cong \frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_1)} \oplus \frac{\mathcal{O}_{\mathbb{A}^2,p}}{(f,g_2)}$$

so $I_p(f, g_1 \cdot g_2) = I_p(f, g_1) + I_p(f, g_2)$ as claimed. $\qquad\square$

Now we will look at some examples of this construction:

For instance, if

$$
\begin{aligned}
f(X,Y) &= Y - X^2 \\
g(X,Y) &= X
\end{aligned}
$$

the ideal $(X, Y - X^2) = (X, Y) \subset k[X,Y]$ and $V_p = \{1\} \cdot k$ so the intersection multiplicity is 1. This is a simple intersection.

We will apply this definition to the examples in section 1.4 on page 15:

The curves in figure 1.4.4 on page 21 are given by

$$
\begin{aligned}
X^2 + Y^2 - 4 &= 0 \\
(X-1)^2 + Y^2 - 1 &= 0
\end{aligned}
$$

and they intersect at $(2, 0)$. We transform these equations to local parameters in the neighborhood of this point by setting $X = U + 2$:

$$
\begin{aligned}
U^2 + 4U + Y^2 &= 0 \\
U^2 + 2U + Y^2 &= 0
\end{aligned}
$$

The ideal they generate is

$$\mathfrak{I} = (U^2 + 4U + Y^2, U^2 + 2U + Y^2) = (U, Y^2)$$

so

$$\mathcal{O}_{\mathbb{A}^2,(2,0)}/\mathfrak{I} = \{1, y\} \cdot k$$

so this intersection has multiplicity 2.

We return to example 2.3.17 on page 55:

EXAMPLE 3.3.35. We have the intersecting curves in figure 1.4.5 on page 23, given by equations

$$5X^2 + 6XY + Y^2 + 6Y - 5 = 0$$
$$X^2 + Y^2 - 1 = 0$$

that intersect at $(\pm 1, 0)$. We consider the point $p_- = (-1, 0)$ first and perform a change of variables, moving this point to the origin. Setting $X = U - 1$ gives

$$5U^2 - 10U + 6UY + 5Y^2 = 0$$
(3.3.11) $$U^2 - 2U + Y^2 = 0$$

In $\mathcal{O}_{\mathbb{A}^2, p_-}$, any expression of the form $a + U$ or $a + Y$ with $a \in k$ are invertible. The ideal $\mathfrak{I}_1$ generated by the two expressions in equation 3.3.11 on page 142 is

$$\begin{aligned}
\mathfrak{I}_1 &= (5U^2 - 10U + 6UY + 5Y^2, U^2 - 2U + Y^2) \\
&= (Y^3, UY, U^2 - 2UY + Y^2)
\end{aligned}$$

where the second line is a Gröbner basis (see definition 2.3.2 on page 46).

We can factor the third term as $(U - 2)U + Y^2$ where $U - 2$ is a unit of $\mathcal{O}_{\mathbb{A}^2, p}$, so we have a relation

$$U = -(U - 2)^{-1}Y^2$$

in the quotient. It follows that

$$\frac{\mathcal{O}_{\mathbb{A}^2, p_-}}{\mathfrak{I}_1} = k \cdot \{1, Y, Y^2\}$$

a three-dimensional vector-space — and p_- is an intersection of multiplicity 3.

Now we consider the point $p_+ = (1, 0)$. In this case, we set $X = U + 1$ and get

$$5U^2 + 10U + 6YU + 12Y + 5Y^2 = 0$$
(3.3.12) $$U^2 + 2U + Y^2 = 0$$

which generate an ideal $\mathfrak{I}_2 \subset k[U, Y]$ with corresponding Gröbner basis

$$\mathfrak{I}_2 = (Y^3, YU + 2Y, U^2 + 2U + Y^2)$$

Since $U + 2$ is invertible in $\mathcal{O}_{\mathbb{A}^2, p_+}$, the second term gives

$$Y \cdot (U + 2) = 0$$

in the quotient

$$Q = \frac{\mathcal{O}_{\mathbb{A}^2, p_+}}{\mathfrak{I}_2}$$

so $Y = 0$ which implies that $U = 0$ so we get

$$Q = k \cdot 1$$

a one-dimensional vector space, since p_+ is a *simple* intersection.

3.3.5. Unique factorization in $\mathcal{O}_{V,p}$. We will spend the rest of this section showing that, at a simple point, $\mathcal{O}_{V,p}$ inherits another property from power series rings — unique factorization. This results in lemma 3.3.42 on page 145 which will be applied in section 5.9 on page 284.

Most of this material appeared previously in [**121**].

ASSUMPTION 3.3.36. *Suppose that a noetherian local ring R is contained in a local ring $\hat{R}$ that is a unique factorization domain. Suppose that the maximal ideals $\mathfrak{m} \subset R$ and $\hat{\mathfrak{m}} \subset \hat{R}$ satisfy the conditions*

    (1) $\mathfrak{m} \cdot \hat{R} = \hat{\mathfrak{m}}$
    (2) $(\mathfrak{m}^n \cdot \hat{R}) \cap R = \mathfrak{m}^n$ for all $n > 0$
    (3) for any $r \in \hat{R}$ and any $n > 0$, there exists an $r_n \in R$ such that $r - r_n \in \mathfrak{m}^n \cdot \hat{R}$

PROPOSITION 3.3.37. *Under the assumption in 3.3.36, if $\mathfrak{a} \subset R$ is any ideal, then $(\mathfrak{a} \cdot \hat{R}) \cap R = \mathfrak{a}$.*

PROOF. It will suffice to prove that $(\mathfrak{a} \cdot \hat{R}) \cap R \subset \mathfrak{a}$. Let $\mathfrak{a} = (a_1, \ldots a_n)$ with the a_i in R and let $x \in \mathfrak{a} \cdot \hat{R}$. Then

$$x = \sum_{i=1}^{n} g_i \cdot a_i$$

with the $g_i \in \hat{R}$. Statement 3 implies the existence of $g_{i,j} \in R$ such that $g_i - g_{i,j} \in \mathfrak{m}^j \cdot \hat{R}$. If

$$x_j = \sum_{i=1}^{n} g_{i,j} \cdot a_i$$

then $x_j \in \mathfrak{a}$ and we have

$$x \in \mathfrak{a} + \mathfrak{m}^j \cdot \hat{R}$$

for all j. Statement 2 implies that intersecting this with R gives

$$x \in \mathfrak{a} + \mathfrak{m}^j$$

for all j or

$$x \in \bigcap_{j=1}^{\infty} \left(\mathfrak{a} + \mathfrak{m}^j \right) = \mathfrak{a}$$

by lemma A.1.81 on page 379. $\square$

This immediately implies:

COROLLARY 3.3.38. *Under the assumptions in 3.3.36, if $a, b \in R$ and $a | b \in \hat{R}$ then $a | b \in R$.*

PROOF. This is an immediate consequence of the previous statement. If $a | b \in \hat{R}$ then $b \in (a) \cdot \hat{R}$. Since $((a) \cdot \hat{R}) \cap R = (a)$, we get that $b \in (a)$ or $a | b \in R$. $\square$

The following result amplifies this a bit:

PROPOSITION 3.3.39. *If $x, y \in R$ have no common factors in R, then they have no common factors in $\hat{R}$.*

PROOF. Let $d = \gcd(x, y) \in \hat{R}$ and let $x = d \cdot u$ and $y = d \cdot v$ so that u and v have no common factors in $\hat{R}$. Then $xv - yu = 0$. By arguments like those used above, we can find $u_n, v_n \in R$ such that

$$
\begin{aligned}
u - u_n &\in \mathfrak{m} \cdot \hat{R} \\
v - v_n &\in \mathfrak{m} \cdot \hat{R}
\end{aligned}
$$

and

$$x(v - v_n) - y(u - u_n) \quad = \quad yu_n - xv_n \in (x,y)\mathfrak{m} \cdot \hat{R}$$
$$\in (x,y)\mathfrak{m}$$

This implies that

$$yu_n - xv_n = y \cdot m_n - x \cdot \bar{m}_n$$

with $m_n, \bar{m}_n \in \mathfrak{m}^n$, or

$$y(u_n + m_n) = x(v_n + \bar{m}_n)$$

Since x and y are multiples of u and v by the same factor (d) in $\hat{R}$, we get

(3.3.13) $$v(u_n + m_n) = u(v_n + \bar{m}_n)$$

Since u and v have no common factors, the fact that $\hat{R}$ is a unique factorization domain
implies that

$$v|(v_n + \bar{m}_n)$$
$$u|(u_n + m_n)$$

for all n, or

$$v_n + \bar{m}_n \quad = \quad \bar{q}_n \cdot v$$
$$u_n + m_n \quad = \quad q_n \cdot u$$

Pick a value of n such that $u, v \notin \mathfrak{m}^{n-1} \cdot \hat{R}$ — such a value exists since $\bigcap_{i=1}^{\infty} \mathfrak{m}^i \cdot \hat{R} = (0)$.
Then $v_n + \bar{m}_n, u_n + m_n \notin \mathfrak{m}^{n-1} \cdot \hat{R}$ either, which implies that $\bar{q}_n, q_n \notin \mathfrak{m} \cdot \hat{R}$. This means
they are both *units* and

(3.3.14) $$(u_n + m_n)|u$$
$$(v_n + \bar{m}_n)|v$$

and $(u_n + m_n)|x$ in $\hat{R}$, which by step 2 means that $(u_n + m_n)|x$ in R.

Let $x = h(u_n + m_n)$. Equation 3.3.13 implies that $y = h(v_n + \bar{m}_n)$. Since x and y
have no common factors in R, h must be a unit.

In light of equations 3.3.14, we have

$$h = \frac{x}{u_n + m_n} = \frac{x}{u} \cdot \frac{u}{u_n + m_n} = d \cdot q_n^{-1}$$

so $d = h \cdot q_n$ — a unit — which contradicts the assumption that d was a proper divisor
of x and y. $\square$

We can finally state a result that shows that subrings sometimes inherit the unique
factorization property from the rings that contain them:

COROLLARY 3.3.40. *Under the assumptions of 3.3.36 on the previous page, if $\hat{R}$ is a unique
factorization domain, then so is R.*

PROOF. Suppose $r, p, q \in R$ and r, p have no common divisors and $r|pq$. Then
proposition 3.3.39 on the preceding page implies that r and p have no common divi-
sors in $\hat{R}$ and $r|q$ because $\hat{R}$ is a unique factorization domain. It follows, from corol-
lary 3.3.38 on the previous page, that $r|q$ in R so R is also a unique factorization domain,
by lemma A.3.1 on page 411. $\square$

This leads to our main result:

THEOREM 3.3.41. *Let $p \in V$ be a simple point of an affine variety. Then the coordinate
ring at p, $\mathcal{O}_{V,p}$, is a unique factorization domain.*

PROOF. Because p is simple, theorem 3.3.28 on page 136 tell us that $\widehat{\mathcal{O}_{V,p}} = k[[U_1, \ldots, U_n]]$ where n is the dimension of the irreducible component of V containing p and the U_i are local parameters (see definition (3.3.21) on page 132). The power-series ring has unique factorization, by theorem A.3.16 on page 418 and we have the inclusion

$$\mathcal{O}_{V,p} \to \widehat{\mathcal{O}_{V,p}} = k[[U_1, \ldots, U_n]]$$

The maximal ideal $\mathfrak{m} \subset \mathcal{O}_{V,p}$ is $(U_1, \ldots U_n)$ and this is also the maximal ideal of $k[[U_1, \ldots, U_n]]$. Clearly, conditions 1 and 2 of3.3.36 on page 143 are satisfied. It is also not hard to see that condition 3 is met: if $x \in k[[U_1, \ldots, U_n]]$ then the set of terms of degree $\leq k$ define an element x_k of $\mathcal{O}_{V,p}$ such that $x - x_k \in \mathfrak{m}^{k+1}$. The conclusion follows from corollary 3.3.40 on the facing page. $\square$

The following application will be very important in divisor-theory in section 5.9 on page 284.

LEMMA 3.3.42. *Let $W \subset V$ be a codimension-1 irreducible subvariety of an algebraic variety and let $p \in W \subset V$ be a simple point. If U is an open set of simple points containing p, then $\mathcal{I}(W \cap U) \subset k[U]$ is a prime ideal and there exists an element $f \in \mathcal{O}_{V,p}$ such that*

$$\mathcal{I}(W \cap U) \cdot \mathcal{O}_{V,p} = (f)$$

REMARK. This result essentially says that, in the neighborhood of a simple point, a codimension-1 subvariety is defined by a *single equation*.

If U is an open set of simple points containing p, it must be irreducible since the intersection of any two distinct irreducible components is a singular point by lemma 3.3.30 on page 138. It follows that

$$k[U] \to \mathcal{O}_{V,p}$$

is an inclusion.

PROOF. As in proposition 2.8.31 on page 111, some element $H \in k[U]$ vanishes on $W \cap U$. Since it is irreducible, some irreducible factor, f, of H vanishes on it as well. Then the prime idea $\mathfrak{p} = \mathcal{I}(W \cap U)$. contains the irreducible factor, f, of H and (f) is prime (because $\mathcal{O}_{V,p}$ is a unique factorization domain). If $(f) \neq \mathfrak{p}$ then

$$W = \mathcal{V}(\mathfrak{p}) \nsubseteq \mathcal{V}((f)) \nsubseteq V$$

so

$$\dim W < \dim \mathcal{V}((f)) < \dim V$$

which contradicts the hypothesis. $\square$

3.4. Normal varieties and finite maps

3.4.1. Basic properties. We begin by defining a class of varieties that are "almost smooth":

DEFINITION 3.4.1. An affine variety, V, is said to be *normal at a point $p \in V$* if the coordinate ring at that point $\mathcal{O}_{V,p}$ is a normal ring (see definition A.4.3 on page 420), i.e., integrally closed in its field of fractions. The variety V is said to be *normal* if it is normal at all of its points.

Since we are calling normal varieties "almost smooth," it is reasonable to expect:

LEMMA 3.4.2. *Smooth varieties are normal.*

PROOF. If V is smooth, theorem 3.3.41 on page 144 proves that its coordinate ring at any point is a unique factorization domain, and proposition A.4.11 on page 422 implies that this coordinate ring is normal. $\square$

In order understand some of the properties of normal varieties, we need

LEMMA 3.4.3. *Let V be an irreducible normal variety and let $W \subset V$ be a codimension-1 subvariety. Then there exists an affine open set $U \subset V$ with $U \cap W \neq \varnothing$ such that $\bar{W} = W \cap U$ is defined in $k[U]$ by a single equation — i.e., its defining ideal is principal.*

PROOF. Suppose that $W = \bigcup_{j=1}^{m} W_i$ where the W_i are irreducible. Define $U_1 = V \setminus \bigcup_{j=2}^{m} W_j$ so that $W \cap U_1 = W_1$. Since $k[U_1]$ is a localization of $k[V]$, it is also normal by exercise 1 on page 427. Let $\mathfrak{p} \subset k[U_1]$ be the prime ideal defining W_1 and form the localization

$$k[U_1]_{\mathfrak{p}} = S^{-1}k[U_1]$$

where $S = k[U_1] \setminus \mathfrak{p}$. Since W is $n-1$-dimensional, the prime ideal $\mathfrak{p}$ is *minimal* by proposition 2.8.9 on page 101 (and the reasoning used in its proof). Then $k[U_1]_{\mathfrak{p}}$ is also normal, by exercise 1 on page 427 and it is a local domain with a *single* nontrivial prime ideal, $\mathfrak{p} \cdot k[U_1]_{\mathfrak{p}}$ — see corollary A.1.94 on page 385.

It follows that the ring $k[U_1]_{\mathfrak{p}}$ is a *discrete valuation ring,* by lemma A.4.38 on page 433, and lemma A.4.37 on page 433 implies that there exists an element, $\pi \in k[U_1]_{\mathfrak{p}}$, such that $\mathfrak{p} \cdot k[U_1]_{\mathfrak{p}} = (\pi)$.

If $\mathfrak{p} = (x_1, \ldots, x_t) \in k[U_1]$, then $x_i = \pi \cdot g_i^{-1} \in k[U_1]_{\mathfrak{p}}$ for some $g_i \notin \mathfrak{p}$. If $\pi = f/g \in k[U_1]_{\mathfrak{p}}$, where $g \notin \mathfrak{p}$, simply define

$$U = U_1 \cap D(g) \cap \bigcap_{j=1}^{t} D(g_j)$$

and then the defining ideal of $W \cap U$ will be $(\pi) \in k[U] = k[U_1]_{(g_1, \ldots, g_t, g)}$. $\square$

Now we are ready to show precisely how normal varieties are "almost smooth"

THEOREM 3.4.4. *If V is an irreducible normal variety, the set of singular points is of codimension ≥ 2.*

REMARK. In other words, the set of singular points is *small.* Since a curve is one-dimensional, a normal curve cannot have *any* singular points and must be smooth.

PROOF. Suppose V is n-dimensional and let $W \subset V$ be the set of singular points. We will assume that W has an $n-1$-dimensional component, W_1, and get a contradiction.

Let U be the open set in lemma 3.4.3 so that $\bar{W} = U \cap W$ has a defining ideal $(\pi) \in k[U]$, so that

$$k[\bar{W}] = k[U]/(\pi)$$

Let $p \in \bar{W}$ be a nonsingular point (in $\bar{W}$, not V). Then

$$\mathcal{O}_{\bar{W},p} = \mathcal{O}_{V,p}/(\pi)$$

Since p is nonsingular in $\bar{W}$, lemma 3.3.23 on page 133 implies that the maximal ideal of $\mathcal{O}_{\bar{W},p}$ is generated by uniformizing parameters $(u_1, \ldots, u_{n-1})$. Let $\bar{u}_i$ be an element of $\mathcal{O}_{V,p}$ that maps to u_i under the projection

$$\mathcal{O}_{V,p} \to \mathcal{O}_{V,p}/(\pi) = \mathcal{O}_{\bar{W},p}$$

Then $(\bar{u}_1, \ldots, \bar{u}_{n-1}, \pi)$ generates the maximal ideal of $\mathcal{O}_{V,p}$ (see lemma A.1.25 on page 352) which contradicts the fact that p is singular in V. $\qquad\square$

Although definition 3.4.1 on page 145 defines normality using local criteria, there is a simple global definition:

LEMMA 3.4.5. *An irreducible affine variety, V, is normal if and only if the coordinate ring, $k[V]$, is normal.*

PROOF. If $k[V]$ is normal, then so are all of its localizations, $\mathcal{O}_{V,p}$ for all $p \in V$ so definition 3.4.1 on page 145 is satisfied.

Suppose V is normal, and $x \in k(V)$ is integral over $k[V]$. Then it is integral over every $\mathcal{O}_{V,p}$ and, therefore, *contained* in each $\mathcal{O}_{V,p}$ since they are integrally closed. But a rational function that is regular everywhere is in the image of

$$k[V] \to k(V)$$

by proposition 2.5.21 on page 83, so $x \in k[V]$ and $k[V]$ is normal. $\qquad\square$

We are in a position to refine the results of section 2.8.3 on page 107:

LEMMA 3.4.6. *If $f: V \to W$ is a finite degree-d map of irreducible affine varieties and W is normal, then $f^{-1}(p)$ contains $\leq d$ points for any $p \in W$.*

REMARK. In general, we can only say that the fiber $f^{-1}(p)$ is *finite* for all $p \in W$ (see proposition 2.5.11 on page 77).

This can fail if W is not normal: Let W be the cubic $Y^2 = X^2 + X^3$ and let f be the map of degree 1

$$f: \mathbb{A}^1 \to W$$

defined by $f(T) = (T^2 - 1, T(T^2 - 1))$. Then $f^{-1}(0,0) = \pm 1$.

PROOF. The hypotheses imply that $k[V]$ is a finite $k[W]$-module, hence an integral extension of $k[W]$ (see proposition A.4.2 on page 419) and $[k(V):k(W)] = d$. If $p \in W$ is any point, let $g \in k[V]$ take on distinct values, $v_i \in k$, for each point of $p_i \in f^{-1}(p)$ — so the variety $f^{-1}(p)$ is be *defined* by $g(p_i) = v_i$.

Let $p_g(X)$ be the minimal polynomial of g over $k(W) \subset k(V)$. Since g is *integral* over $k[W]$, this will be a monic polynomial whose coefficients all lie in $k[W] \subset k(W)$ (by definition A.4.1 on page 419) — see the final statement of lemma A.4.14 on page 423. This polynomial has the property that $p_g(v_i) = p_g(g(p_i)) = 0$. If we map this polynomial via

$$k[W] \to k[W]/\mathfrak{m}_p = k$$

we get a polynomial over k with m distinct roots with $m \leq d$, where $\mathfrak{m}_p$ is the maximal ideal corresponding to p. The fact that it has m distinct roots, implies that the regular function, g, can take on only m *distinct* values *when restricted* to $f^{-1}(p)$. It follows that the number of points in $f^{-1}(p) = m \leq d$. $\qquad\square$

DEFINITION 3.4.7. A finite mapping $f: V \to W$ is *unramified* at a point $p \in V$ if in the induced map (see proposition 3.2.3 on page 118)

$$f_p: \mathcal{O}_{W,f(p)} \to \mathcal{O}_{V,p}$$

the maximal ideals corresponding to p and $f(p)$ satisfy $\mathfrak{m}_p = \mathfrak{m}_{f(p)} \cdot \mathcal{O}_{V,p}$. Otherwise, f is *ramified* at p and $f(p)$ is a *branch point* of W. The set of all ramification points in V is called the *ramification locus*. The set of branch points in W is called the *branch locus*.

REMARK. The map f is unramified at a point $p \in V$ if the map of maximal ideals is surjective (after localization), which happens if the map of tangent spaces is *injective* — see theorem 3.3.7 on page 123. This happens if and only if the Jacobian is nonsingular — see equation 3.3.2 on page 120.

EXAMPLE 3.4.8. Consider the projection, f, of the parabola, V, given by $Y = X^2$ to the Y-axis, W (see figure 2.2.1 on page 43). The coordinate ring is $k[V] = k[X, Y]/(Y - X^2) \cong k[X]$, and $k[W] = k[Y]$. The maximal ideal of $Y = 1$ in W is $(Y - 1)$. Its image in $k[V]$ is $(X^2 - 1) = (X + 1)(X - 1)$. In the local ring $\mathcal{O}_{V,X=1}$, $X + 1$ is invertible so $(X^2 - 1) \cdot \mathcal{O}_{V,X=1} = (X - 1) \cdot \mathcal{O}_{V,X=1}$, and f is *unramified* at the point $X = 1$. Similar reasoning applies at the point $X = -1$.

On the other hand, if p is the origin, the maximal ideal corresponding to $Y = 0$ is (Y) and its image in $k[V]$ is the ideal (X^2) and $(X^2) \cdot \mathcal{O}_{V,X=0} \neq (X) \cdot \mathcal{O}_{V,X=0}$ so f is *ramified* at $Y = 0$.

This example suggests another approach to ramification:

THEOREM 3.4.9. *Let $f: V \to W$ be a finite degree-d map of irreducible affine varieties and suppose W is smooth, and k has characteristic 0. If $y \in W$ is a point with $f^{-1}(y) = \{x_1, \dots, x_\ell\}$ then each of the x_i has a number $e_i \in \mathbb{Z}^+$ associated with it (called its ramification index) such that:*

(1) f is unramified at x_i if and only if $e_i = 1$
(2) $\sum_{i=1}^\ell e_i = d$

Consequently, $y \in W$ is a branch point if and only if $\ell < d$.

REMARK. We map V to $W \times \mathbb{A}^1$ and replace f by the projection to W:

(3.4.1)

$$V \xrightarrow{\ \iota\ } W \times I$$

with f going to W and π the projection $W \times I \to W$.

The idea is that d sheets of V map to W and ramification occurs when these sheets intersect.

PROOF. If the characteristic of k is 0, the field extension $f^*k(W) \subset k(V)$ is separable, so there exists a primitive element α such that

(3.4.2) $$k(V) = k(W)[\alpha]$$

(see theorem A.2.21 on page 394). Without loss of generality, we can assume $\alpha \in k[V]$. Since f is finite, $k[V]$ is an integral extension of $k[W]$ (see proposition A.4.2 on page 419). This implies α is integral over $k[W]$ so its *minimal*

polynomial, $p(T) \in k(W)[T]$, is a monic polynomial whose coefficients lie in $k[W]$ (see lemma A.4.14 on page 423).

We get a homomorphism

(3.4.3) $$g': k[W][T]/(p(T)) \to k[V]$$

implying the existence of a regular map

(3.4.4) $$g: V \to \bar{V} \subset W \times I$$

where $\bar{V}$ is the affine variety defined by $p(T) = 0$. Note that:

(1) $\bar{V}$ is smooth. See exercise 1 on the next page.
(2) the regular map $g: V \to \bar{V}$ is 1-1 so that the number of points in $\pi^{-1}(f(v))$ is equal to the number of points in $f^{-1}(f(v))$ for all $v \in V$. See exercise 2 on the following page.

Here ι in 3.4.1 on the preceding page is induced by g, and π is induced by $\bar{\pi}: k[W] \hookrightarrow k[W][T]$.

If $y \in W$ with maximal ideal $\mathfrak{m}_y$, then the image of $p(T)$ under the map

$$k[W][T] \to (k[W]/\mathfrak{m}_y)[T] = k[T]$$

is a polynomial, $p_y(T)$, whose roots correspond to points of $W \times \mathbb{A}^1$ in V that lie above $y \in W$, i.e., $\{x_1, \ldots, x_\ell\}$. We can factor $p_y(T)$:

$$p_y(T) = \prod_{i=1}^{\ell} (T - t_i)^{e_i}$$

Define the ramification index of x_i to be e_i. If we form the quotient by $\mathfrak{m}_y$, we get

(3.4.5) $$k[V]_{\mathfrak{m}_y}/f^*(\mathfrak{m}_y) \cong \left(k[W]_{\mathfrak{m}_y}[T]/(p(T))\right)/\mathfrak{m}_y \cong k[T]/(p_y(T))$$

Since the factors $(T - t_i)^{e_i}$ are pairwise relatively prime, we have

$$((T - t_i)^{e_i}, (T - t_j)^{e_j}) = k[T]$$

if $i \neq j$, and the Chinese Remainder Theorem (see exercise 18 on page 354) implies that

(3.4.6) $$k[T]/(p_y(T)) \cong \bigoplus_{i-1}^{\ell} k[T]/((T - t_i)^{e_i}) = \bigoplus_{i=1}^{\ell} k^{e_i}$$

(where we have "forgotten" the ring-structure of k^{e_i}). The natural map

$$\mathcal{O}_{W,y} \to \mathcal{O}_{V,x_i}$$

preserves maximal ideals if and only if

$$\mathcal{O}_{W,y}/\mathfrak{m}_y = k \xrightarrow{\cong} \mathcal{O}_{V,x_i}/\mathfrak{m}_{x_i} = k$$

which happens if and only if $e_i = 1$. Note that $\mathcal{O}_{V,x_i}$ "filters out" all factors $(T - t_j)^{e_j}$ with $j \neq i$ because they become *units*. $\square$

We can show that finite maps are "mostly unramified" — over characteristic 0, they are unramified on a dense set:

LEMMA 3.4.10. *If $f: V \to W$ is a finite map of irreducible affine varieties over k, of characteristic 0, then the ramification locus and branch locus are of codimension 1. The set of points where f is unramified is open and nonempty.*

REMARK. Over characteristic p, this can fail spectacularly:

$$f: \mathbb{A}^1 \to \mathbb{A}^1$$
$$X \mapsto X^p$$

is ramified *everywhere*.

PROOF. At a point $p \in W$, let $\{w_1, \ldots, w_n\}$ and $q \in f^{-1}(p) \{v_1, \ldots v_n\}$ be local parameters, i.e. assume $\mathfrak{m}_p = (w_1, \ldots, w_n)$ and let $\mathfrak{J}$ be the Jacobian of

$$df: T_{V,q} \to T_{W,p}$$

the induced map of tangent spaces. Then the branch locus in $B \subset W$ is given by the single equation $\det \mathfrak{J} = 0$, which means it is of codimension 1. The ramification locus, $R \subset f^{-1}(B)$ is also of codimension 1.

Since the set, S, of singular points of W is a closed subvariety (see the remark following definition 3.3.11 on page 125), it follows that $W_1 = W \setminus S$ is open and smooth, hence normal (lemma 3.4.2 on page 145). If $V_1 = f^{-1}(V)$, the finite map

$$f_1 = f|V_1: V_1 \to W_1$$

satisfies the hypotheses of lemma 3.4.9 on page 148. As in equation 3.4.3 on the preceding page, we have

$$k[V_1] = k[W_1][T]/(p(T))$$

where the discriminant, Δ, of $p(T)$ is not identically 0. Pick a point $w \in W_1$ where $\Delta \neq 0$. $f^{-1}(w)$ has d points, and f is unramified over w. $\square$

EXERCISES.

1. Why is $\bar{V}$ in the proof of theorem 3.4.9 on page 148 smooth? Hint: It "inherits" smoothness from W.

2. Why is the map $g: V \to \bar{V}$ in equation 3.4.4 on the preceding page 1-1?

3. Where do we use the fact that W is *smooth* in lemma 3.4.9 on page 148?

3.4.2. Resolving singularities. Heisuke Hironaka is a Japanese mathematician, born in 1931, who was awarded the Fields Medal in 1970 for his great contributions to algebraic geometry. In his two (long) groundbreaking papers [**78, 79**], Hironaka proved (in *characteristic zero*[1]):

---

[1] The result in characteristic $p \neq 0$ is still an open question.

If V is a variety with singularities over characteristic 0, it is possible to find a *smooth* variety, $\bar{V}$, and a regular map

$$f\colon \bar{V} \to V$$

that is a birational equivalence. This map, f, is said to *resolve the singularities* of V.

Hironaka's proof involves "blowing up" singularities — see section 5.5.3 on page 248 — using a complex strategy that is beyond the scope of this book (the survey [74] outlines the main ideas). Here, we will discuss a relatively simple procedure for resolving many singularities (including some of the "largest").

Now we can describe an algorithm for resolving many singularities of an irreducible affine variety, V. We can replace it by a normal variety V^ν that comes equipped with a regular map

$$V^\nu \to V$$

that is a birational equivalence.

The idea behind this construction is very simple:

If $k[V]$ is not normal, replace it by its *normal closure, N*, in $k(V)$. Now, build a variety, V^ν, that has N as its coordinate ring and we are done. The inclusion $k[V] \hookrightarrow N$ will induce a regular map that is a birational equivalence because $k[V]$ and N both have the same field of fractions, $k(V)$.

As always, "the devil is in the details:"

How do we know that N is an affine k-algebra?

The key to answering this question is:

LEMMA 3.4.11. *For any integer $n > 0$ let $P_n = k[X_1, \ldots, X_n]$ and let F_n be its field of fractions, $k(X_1, \ldots, X_n)$. If H is a finite extension of F_n and $\bar{P}$ is the integral closure of P_n in H, then $\bar{P}$ is a finitely generated module over P_n.*

REMARK. Corollary A.4.12 on page 423 tells us that P_n is integrally closed in F_n.

PROOF. If H is a *separable* extension of F_n, then lemma A.4.20 on page 426 implies the result (because we know that P_n is noetherian, by lemma A.1.49 on page 360). In the simple case, we are considering here (extensions of a polynomial ring), there is a trick to get around the separability requirement.

Suppose $H = F_n[\alpha_1, \ldots, \alpha_t]$ is an *inseparable* extension of P_n and α_1 has a minimal polynomial

$$f(X) = g(X^{p^r})$$

as in proposition A.2.18 on page 393, and suppose

$$g(X) = X^m + a_{m-1}X^{m-1} + \cdots + a_0$$

Now define finite algebraic extensions

$$\tilde{P} = P_n[X_1^{1/p^r}, \ldots, X_n^{1/p^r}, a_{m-1}^{1/p^r}, \ldots, a_0^{1/p^r}]$$
$$\tilde{H} = H[X_1^{1/p^r}, \ldots, X_n^{1/p^r}, a_{m-1}^{1/p^r}, \ldots, a_0^{1/p^r}]$$

and set

$$\tilde{g}(X) = X^m + a_{m-1}^{1/p^r}X^{m-1} + \cdots + a_0^{1/p^r} \in \tilde{P}[X]$$

Then
$$g(X^{p^r}) = (\tilde{g}(X))^{p^r}$$
and $\tilde{H}$ is a *separable* extension of $\tilde{P}$. If U is the integral closure of $\tilde{P}$ in $\tilde{H}$, then U is a finitely-generated module over $\tilde{P}$. Since $\tilde{P}$ is a finitely-generated module over P_n (with generators $\{X_1^{i_1/p^r}, \ldots, X_n^{i_m/p^r}, a_{m-1}^{j_{m-1}/p^r}, \ldots, a_0^{j_0/p^r}\}$ for $i_\alpha, j_\beta = 1, \ldots, p^r$), it follows that U is a finitely-generated module over P_n (see proposition A.4.5 on page 421). Since $\bar{P}$ is a submodule of U, it must be finitely-generated over P_n (see lemma A.1.70 on page 372). $\qquad\square$

THEOREM 3.4.12. *If V is an irreducible affine variety, there exists a normal variety, V^ν, called the normalization of V and a finite map*
$$n(V) \colon V^\nu \to V$$
that is a birational equivalence. This construction is functorial (see definition A.5.7 on page 442) in the sense that a dominating regular map of irreducible affine varieties
$$g \colon V_1 \to V_2$$
induces a map
$$g^\nu \colon V_1^\nu \to V_2^\nu$$
that makes the diagram

$$
\begin{array}{ccc}
V_1^\nu & \xrightarrow{\;g^\nu\;} & V_2^\nu \\
{\scriptstyle n(V_1)}\downarrow & & \downarrow{\scriptstyle n(V_2)} \\
V_1 & \xrightarrow[\;g\;]{} & V_2
\end{array}
$$

commute. If W is any normal affine variety and
$$r \colon W \to V$$
is a regular map, there exists a regular map $\bar{r} \colon W \to V^\nu$ that makes the diagram

$$
\begin{array}{ccc}
 & & V^\nu \\
 & \nearrow{\scriptstyle \bar{r}} & \downarrow{\scriptstyle n(V)} \\
W & \xrightarrow[\;r\;]{} & V
\end{array}
$$

commute.

PROOF. We cannot use lemma 3.4.11 on the previous page directly to show that the integral closure of $k[V]$ is an affine k-algebra because $k[V]$ is not a polynomial algebra. On the other hand, the Noether Normalization theorem (2.2.2 on page 40) theorem implies the existence of a polynomial ring
$$P_d = k[X_1, \ldots, X_d] \subset k[V]$$
where $d = \dim V$ such that $k[V]$ is a finitely-generated module over P_d. Using the notation of lemma 3.4.11 on page 151, let $H = k(V)$ — a finite algebraic extension of $k(X_1, \ldots, X_d)$. The normal closure, U, of P_d in $k(V)$ is a finitely-generated module over P_d.

If $x \in k(V)$ is integral over $k[V]$, it will also be integral over P_d by statement 2 of proposition A.4.5 on page 421, hence will be contained in U. This

implies that U is equal to the normal closure of $k[V]$ (compare exercise 3 on page 427) and is a finitely-generated k-algebra, i.e. an *affine k-algebra*.

Now we use theorem 2.5.5 on page 73 to define V^ν to be the irreducible affine variety whose coordinate ring is U.

The inclusion $k[V] \hookrightarrow U$ induces the regular map

$$n(V): V^\nu \to V$$

and this is a finite map and birational equivalence. The functorial properties follow from similar properties of normal closures. $\square$

Here are some examples of normalization:

EXAMPLE 3.4.13. We start with the singular curve, $W \subset \mathbb{A}^2$, defined by $Y^2 = X^3$ — see figure 3.3.1 on page 120. with coordinate ring $k[W] = k[X, Y]/(Y^2 - X^3)$. To understand this coordinate ring, we map it to the polynomial ring, $k[T]$, via

$$
\begin{aligned}
k[W] &\to k[T] \\
X &\mapsto T^2 \\
Y &\mapsto T^3
\end{aligned}
$$

The image of this map is $k[T^2, T^3]$, the set of polynomials whose *linear term* vanishes. Now the element $(Y/X) \in k(W)$ satisfies the equation

$$\left(\frac{Y}{X}\right)^2 - X = 0$$

in $k(W)$ so it is integral over $k[W]$. It is not in $k[W]$ because its image in $k[T]$ is T, which is not in the image of $k[W]$. If we adjoin this element to $k[W]$ we get

$$k[W][(Y/X)] = k[W][\sqrt{X}]$$

and, in this new ring

$$
\begin{aligned}
X &= \left(\sqrt{X}\right)^2 \\
Y &= \left(\sqrt{X}\right)^3
\end{aligned}
$$

so this new ring is a polynomial ring on $\sqrt{X}$, or $k[T]$. This is normal, by A.4.12 on page 423 and is the normal closure of $k[W]$ in $k(W)$.

Geometrically, the normalization, $\hat{W}$ is $\mathbb{A}^1$. To understand the map

$$n(W): \hat{W} \to W$$

note that the generator of $\hat{W}$ is Y/X which is the slope of a line through the origin that intersects W. In addition, note that this slope intersects the line $X = 1$ in a point whose Y-coordinate is equal to this slope. We arrive at a description of $n(W)$:

> Identify $\hat{W}$ with the line $X = 1$. If $p \in \hat{W}$ is a point, draw a line, ℓ, though the origin to p. The image of p under $n(W)$ is the intersection of this line with W — see figure 3.4.1 on the following page.
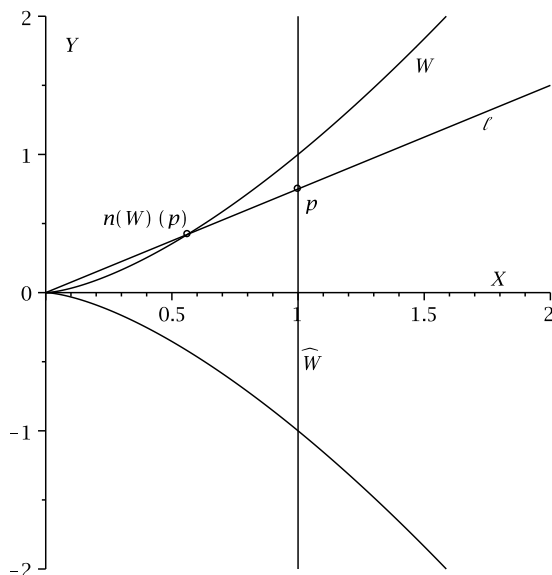
FIGURE 3.4.1.   Normalization of a curve

EXERCISES.

4. It is possible to normalize a variety in a field that is an *extension* of its field of fractions. If W is the cone $Z^2 = XY$, show that the normalization of W in the field $k(W)(\sqrt{X})$ is isomorphic to the affine plane $\mathbb{A}^2$ over coordinates U and V, and that the normalization mapping has the form $X = U^2$, $Y = V^2$, and $Z = UV$.

5. Show that the normalization of $W = \mathcal{V}\left(Y^2 - X^2(X+1)\right)$ is the affine line, with the projection defined by

$$X \mapsto T^2 - 1$$
$$Y \mapsto T(T^2 - 1)$$

6. How is the normalization of $V \times W$ connected with the normalizations of V and W?

3.5. Vector bundles on affine varieties

We conclude this chapter by using the sheaf of regular functions on an affine variety to prove a theorem of Serre (in [**150**]), classifying *algebraic vector-bundles* (see appendix C on page 507) over affine varieties. Throughout this section V will denote an irreducible affine variety and k will denote an algebraically closed field. Vector-bundles on V will be assumed to be compatible with the sheaf of regular functions, $\mathcal{O}_V$.

DEFINITION 3.5.1. If M is a module over $k[V]$, let $\underline{M}$ be the constant sheaf over V. Then we can define a presheaf $\mathcal{A}(M)$ via

$$\mathcal{A}(M)(U) = M \otimes_{k[V]} \mathcal{O}_V(U)$$

for any open subset $U \subset V$.

We should immediately note that

PROPOSITION 3.5.2. *The presheaf in definition 3.5.1 above is actually a sheaf.*

PROOF. This follows from the fact that tensor products commuted with direct limits — see exercise 19 on page 473 and its solution. The completion of $\mathcal{A}(M)$ is defined by (see lemma B.2.3 on page 500)

$$
\begin{aligned}
\overline{\mathcal{A}(M)}(U) = \varinjlim \mathcal{A}(M)(\{U_i\}) \quad &= \quad M \otimes_{k[V]} \varinjlim \mathcal{O}_V(\{U_i\}) \\
&= \quad M \otimes_{k[V]} \mathcal{O}_V(U)
\end{aligned}
$$

as $\{U_i\}$ ranges over all open covers of U. The lower equation follows from the fact that the presheaf $\mathcal{O}_V$ is *already* a sheaf. $\qquad\square$

This construction is functorial:

PROPOSITION 3.5.3. *Any homomorphism, $f\colon M_1 \to M_2$, of modules over $k[V]$ induces a homomorphism of sheaves*

$$\mathcal{A}(M_1) \to \mathcal{A}(M_2)$$

and an exact sequence

$$M_1 \to M_2 \to M_3$$

induces an exact sequence

$$\mathcal{A}(M_1) \to \mathcal{A}(M_2) \to \mathcal{A}(M_3)$$

PROOF. The first statement is clear. If $U \subset V$ is an open set, the second statement follows from the fact that $\mathcal{O}_V(U) = S^{-1}k[V]$ for some multiplicative set S (of regular functions that do not vanish on U) and the fact that $S^{-1}k[V]$ is a *flat module* over $k[V]$ (see lemma A.5.60 on page 469) so the diagram

$$
\begin{array}{ccccc}
M_1 \otimes_{k[V]} S^{-1}k[V] & \longrightarrow & M_2 \otimes_{k[V]} S^{-1}k[V] & \longrightarrow & M_3 \otimes_{k[V]} S^{-1}k[V] \\
\| & & \| & & \| \\
\mathcal{A}(M_1)(U) & \longrightarrow & \mathcal{A}(M_2)(U) & \longrightarrow & \mathcal{A}(M_3)(U)
\end{array}
$$

commutes and its bottom row is exact. $\qquad\square$

The universal property of modules of fractions (proposition A.5.24 on page 452) implies a universal property of this construction:

PROPOSITION 3.5.4. *If $V = \operatorname{Spec} R$ is an affine variety and $\mathscr{F}$ is a module over $\mathcal{O}_V$ with $\mathscr{F}(V) = M$, an R-module, then there exists a unique morphism of modules over $\mathcal{O}_V$:*

$$\mathcal{A}(M) \to \mathscr{F}$$

Now we define an important class of modules and sheaves:

DEFINITION 3.5.5. A module, M, over a ring is *coherent* if it is finitely generated and, given any surjection

$$s\colon F \to M$$

where F is a finitely generated free module, $\ker s$ is also finitely generated.

A module, M, over a *sheaf*, $\mathcal{O}_V$, is coherent if, for any open set $U \subset V$ $M(U)$ is finitely generated and the kernel of any surjection

$$s\colon \mathcal{O}_V^m(U) \to M(U)$$

for some finite $m > 0$ is also finitely generated.

LEMMA 3.5.6. *Let V be an affine variety with coordinate ring $k[V]$ and sheaf of regular functions $\mathcal{O}_V$.*

(1) If M is a finitely generated module over $k[V]$, then $\mathcal{A}(M)$ is a coherent module over $\mathcal{O}_V$, and $\mathcal{A}(M)(V) = M$.
(2) If S is a coherent module over $\mathcal{O}_V$ and $N = S(V)$, then N is finitely generated and $S = \mathcal{A}(N)$.

Consequently $\mathcal{A}(*)$ induces an equivalence of categories between the category of finitely generated modules over $k[V]$ and coherent sheaves over $\mathcal{O}_V$. This equivalence sends the subcategory of locally free modules over $\mathcal{O}_V$ to projective modules over $k[V]$.

REMARK. The significant property of coherent modules over sheaves is that their *global* behavior determines their *local* behavior — which motivates the term "coherent." It is easy to come up with examples of "incoherent" sheaves:

Let $p \in V$ be any point and let $\mathcal{O}_X$ be the sheaf defined by

(3.5.1)
$$\mathcal{O}_X(U) = \begin{cases} \mathcal{O}_V(U) & \text{if } p \notin U \\ 0 & \text{otherwise} \end{cases}$$

for every affine open set $U \subset V$. This is easily verified to be a nontrivial sheaf, but $\mathcal{O}_X(V) = 0$, so it cannot possibly be coherent.

PROOF. If M is a finitely-generated module over a noetherian ring, $R = \mathcal{O}_V(V) = k[V]$, then there exists an exact sequence

$$R^m \to R^n \to M \to 0$$

which implies that the induced sequence

$$\mathcal{O}_V^m = \mathcal{A}(R^m) \to \mathcal{O}_V^n = \mathcal{A}(R^n) \to \mathcal{A}(M) \to 0$$

is exact (by proposition 3.5.3 on the preceding page) so that $\mathcal{A}(M)$ is coherent.

If S is a coherent module over $\mathcal{O}_V$ then we have an exact sequence

$$\mathcal{O}_V^m \to \mathcal{O}_V^n \to S \to 0$$

Furthermore, restriction induces a homomorphism

$$N = S(V) \to S(U)$$

for any affine open set $U \subset V$ that extends to a homomorphism

$$N \otimes_{k[V]} \mathcal{O}_V(U) \to S(U)$$

which defines a homomorphism of sheaves

$$\mathcal{A}(N) \to S$$

This is an isomorphism if $S = \mathcal{O}_V^n$. The conclusion follows from the exactness of $\mathcal{A}(*)$ (proposition 3.5.3 on the previous page) and the commutative diagram

$$
\begin{array}{ccccc}
\mathcal{A}(R^m) & \longrightarrow & \mathcal{A}(R^n) & \longrightarrow & \mathcal{A}(N) \\
{\scriptstyle \cong}\big\downarrow & & {\scriptstyle \cong}\big\downarrow & & \big\downarrow \\
\mathcal{O}_V^m & \longrightarrow & \mathcal{O}_V^n & \longrightarrow & S
\end{array}
$$

The final statement regarding locally free sheaves follows from corollary A.5.57 on page 468 and lemma A.5.63 on page 471. □

This precise correspondence between modules and their associated coherent sheaves immediately implies a few basic properties of these sheaves:

LEMMA 3.5.7. *Let V be an affine variety and let M be a finitely generated module over $k[V]$.*

   (1) If $\mathcal{A}(M)|D(f) = 0$ (or $\mathcal{A}(M)(D(f)) = 0$), there exists an integer $n > 0$ such that $f^n \cdot M = 0$,
   (2) If $W \subset V$ is a subvariety defined by the ideal $\mathfrak{a} \subset k[V]$ and $\mathcal{A}(M)|V \setminus W = 0$, there exists an integer $n > 0$ such that $\mathfrak{a}^n \cdot M = 0$.

PROOF. The first statement implies that $M \otimes_{k[V]} k[V]_f = M_f = 0$. If $\in M$ is any element, then there exists an $n_m > 0$ such that $f^{n_m} \cdot m = 0$ (see definition A.1.89 on page 383 and the remark following it). Since M is finitely generated, it has a finite generating set $\{m_j\}$ and we can take n to be the maximum of the $\{n_{m_j}\}$.

Let $\mathfrak{a} = (f_1, \ldots, f_t)$ — a finite set because $k[V]$ is noetherian. Then $\mathcal{A}(M)|D(f_i) = 0$ for $i = 1, \ldots, t$ since $D(f_i) \subset V \setminus W$ (indeed $V \setminus W$ is the *union* of the $D(f_i)$). It follows that $f_i^{n_i} \cdot M = 0$. If we make n equal to $t \cdot \max(f_1, \ldots, f_t)$ then $\mathfrak{a}^n \cdot M = 0$. □

Even if a module is *not* projective, a coherent sheaf is "approximately" free:

LEMMA 3.5.8. *If $\mathscr{F}$ is a coherent sheaf on an irreducible affine variety, V, there exists an open dense set $W \subset V$ such that $\mathscr{F}|W$ is free.*

PROOF. Since V is irreducible, any open set is dense. We assume $R = k[V]$, an integral domain. Since $\mathscr{F}$ is coherent, it is of the form $\mathcal{A}(M)$ for some R-module M. If $\{x_i\} \in M$ is a maximal linearly independent set over R, let

$$M' = \bigoplus R \cdot x_i \subset M$$

It follows that $M' \subset M$ is free.

If $\{y_j\} \in M$ is a generating set, then the y_j are linearly dependent on the $\{x_i\}$ and there exists an $r_j \in R$ such that $r_j \cdot y_j \in M'$. If $s = \prod r_j$, then $M \otimes_R R_s = M' \otimes_R R_s$, a free R_s-module, so we can set $W = D(s) \subset V$ and $\mathscr{F}|W$ is free. □

We are finally ready to state Serre's theorem, from [**150**]:

THEOREM 3.5.9. *If V is an irreducible affine variety, the functor $\Gamma(*, V)$ (see definition C.1.10 on page 511) defines an equivalence of categories between the category of vector-bundles over V and that of projective modules over $k[V]$. The projective module associated to a vector-bundle is its module of global sections.*

REMARK. There are two mathematical theories that go by the name "K-theory": one is topological and studies vector bundles over spaces (see [**6**]), and the other is algebraic and studies (among other things) projective modules over a ring (see [**114**] or [**145**]). Serre's theorem provides the "glue" that connects these theories.

PROOF. This follows immediately from theorem C.2.7 on page 518 and lemma 3.5.6 on the preceding page. □

We immediately conclude:

COROLLARY 3.5.10. *If ξ is a vector-bundle over an irreducible affine variety, V, then there exists another vector-bundle, η, over V such that $\xi \oplus \eta$ is trivial. It follows that every vector-bundle over V is a sub-bundle of a trivial vector-bundle.*

PROOF. The space of sections, $\Gamma(\xi, V)$ is a projective module, P, over $k[V]$. This means that there exists a projective $k[V]$-module, Q, such that $P \oplus Q$ is free. Let η be the vector-bundle over V with $\Gamma(\eta, V) = Q$. The conclusion follows from exercise 274 on page 626. □

Pushing this subject further requires a bit more machinery than we have now. We revisit vector bundles in section 4.5.3 on page 204.

We close with a remark that all vector-bundles over an affine space are trivial. Although this is easy to prove topologically, Serre (in [**150**]) remarked on the difficulty in proving that they were *algebraically* trivial. The problem was eventually independently resolved by Quillen (in [**140**]) and Suslin (in [**162**]):

THEOREM 3.5.11 (Quillen-Suslin). *If $A = k[X_1, \ldots, X_n]$ then all projective modules over A are free.*

Given the correspondence between projective modules and vector bundles, it is common to regard general (i.e. non-projective) modules as *generalizations* of vector-bundles. Lemma 3.5.8 on page 157 implies that these are vector bundles on a dense subvariety that do not extend to the entire variety.

A smooth manifold, M, has a canonical vector bundle associated to it called its *tangent bundle*. This is a vector bundle whose fiber at every point of M is the tangent space (essentially defined as in section 3.3 on page 119).

Since a smooth affine variety is a smooth manifold, it also possesses a tangent bundle and we would like to get an explicit description of it. The Serre correspondence shows that this tangent bundle corresponds to a projective module over the coordinate ring.

To compute this module, we need a "global" concept of derivative on an affine variety in contrast to the "local" definition in section 3.3 on page 119. This turns out to be the module of Kähler differentials in section A.7 on page 482.

Our main result is:

THEOREM 3.5.12. *If k is a field of characteristic 0 and V is an smooth irreducible affine variety over k, then $\Omega_{k[V]/k}$ is a projective module over $k[V]$ that defines the cotangent bundle of V.*

PROOF. Corollary A.7.17 on page 491 implies that

$$\dim k(V) \otimes_{k[V]} \Omega_{k[V]/k} = n$$

— the dimension of V. If $\mathfrak{m} \subset k[V]$ is any maximal ideal then proposition A.7.9 on page 485 implies that

$$\Omega_{k[V]_{\mathfrak{m}}/k} = k[V]_{\mathfrak{m}} \otimes_{k[V]} \Omega_{k[V]/k}$$

Corollary A.7.19 on page 492 implies that

$$
\begin{aligned}
\frac{\mathfrak{m}}{\mathfrak{m}^2} &= k \otimes_{k[V]_{\mathfrak{m}}} \Omega_{k[V]_{\mathfrak{m}}/k} \\
&= k \otimes_{k[V]_{\mathfrak{m}}} \left(k[V]_{\mathfrak{m}} \otimes_{k[V]} \Omega_{k[V]/k} \right) \\
&= k \otimes_{k[V]} \Omega_{k[V]/k}
\end{aligned}
$$

The fact that V is smooth implies that

$$\dim \frac{\mathfrak{m}}{\mathfrak{m}^2} = \dim V = k(V) \otimes_{k[V]_{\mathfrak{m}}} \left(k[V]_{\mathfrak{m}} \otimes_{k[V]} \Omega_{k[V]/k} \right)$$

and lemma A.7.20 on page 493 implies that

$$k[V]_{\mathfrak{m}} \otimes_{k[V]} \Omega_{k[V]/k}$$

is a free module over $k[V]_{\mathfrak{m}}$. Since $\mathfrak{m}$ was an arbitrary maximal ideal, lemma A.5.63 on page 471 implies that $\Omega_{k[V]/k}$ is a projective module over $k[V]$. Since its localizations are free modules with the same rank as

$$\dim \frac{\mathfrak{m}}{\mathfrak{m}^2}$$

it corresponds to the cotangent bundle of V. $\square$

We conclude this section by considering the functorial properties of vector-bundles and coherent sheaves:

DEFINITION 3.5.13. If $f: V \to W$ is a morphism of affine varieties and $\mathscr{G}$ is a coherent module over $\mathcal{O}_W$, then

$$f^*\mathscr{G} = f^{-1}\mathscr{G} \otimes_{f^{-1}\mathcal{O}_W} \mathcal{O}_V$$

is a coherent module over $\mathcal{O}_V$ called the *pullback* of $\mathscr{G}$ over f.

REMARK. See exercise 8 on page 502 for the definition of $f^{-1}\mathscr{G}$.

We have an interesting way to "define" $\mathcal{A}(M)$:

PROPOSITION 3.5.14. *Let V be an affine variety with $R = k[V] = \mathcal{O}_V(V)$ and let $(\bullet, R)$ be the ringed space that consists of a single point. If M is a finitely-presented R-module, then M defines a coherent sheaf over $(\bullet, R)$.*
If $f: (V, \mathcal{O}_V) \to (\bullet, R)$ is the unique morphism of ringed spaces, then

$$\mathcal{A}(M) = f^*M$$

PROOF. If

$$R^n \to R^m \to M \to 0$$

then applying f^* to this gives

$$(f^*R)^n \to (f^*R)^m \to f^*M \to 0$$

since tensor products are right-exact and commute with direct sums — so that f^* has the same property. Since $f^*R = \mathcal{O}_V$, the conclusion follows.
This following result gives a clear idea of what pullbacks look like: □

COROLLARY 3.5.15. *Let $f: (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ be a morphism of ringed spaces (see definition B.3.1 on page 502) and let $\mathscr{G} = \mathcal{A}(M)$ where M is a module over $\mathcal{O}_Y(Y) = R$. Then*

$$f^*\mathscr{G} = \mathcal{A}(M \otimes_R S)$$

where $S = \mathcal{O}_X(X)$.

PROOF. The diagram of sheaf-morphisms

$$
\begin{array}{ccc}
(R, \mathcal{O}_Y) & \longrightarrow & (\bullet, R) \\
{\scriptstyle f^*}\downarrow & & \downarrow{\scriptstyle f^\sharp} \\
(S, \mathcal{O}_X) & \longrightarrow & (\bullet, S)
\end{array}
$$

commutes, so the conclusion follows from the fact that $(f^\sharp)(M) = M \otimes_R S$. □

EXERCISES.

1. If $V \subset \mathbb{A}^2$ is the parabola, $Y = X^2$, compute $\Omega_{k[V]/k}$.

2. If $W \subset \mathbb{A}^2$ is a smooth variety defined by a single equation $f(X, Y) = 0$, then
$$\frac{\partial f}{\partial X} \cdot dX + \frac{\partial f}{\partial Y} \cdot dY = 0$$
Show that the differential form
$$\omega = \frac{dX}{\partial f/\partial Y} = -\frac{dY}{\partial f/\partial X} \in \Omega_{k[W]/k}$$
is regular, i.e. of the form
$$u(X, Y) \cdot dX + v(X, Y) \cdot dY$$

3. If $V \subset \mathbb{A}^2$ is the affine variety defined by $X^2 + Y^2 = 1$ and k is a field of characteristic $\neq 2$, determine $\Omega_{k[V]/k}$.

CHAPTER 4

Varieties and Schemes

> "After Grothendieck's great generalization of the field ... what I myself had called abstract turned out to be a very, very concrete brand of mathematics."
> — Oscar Zariski, quoted in *The Unreal Life of Oscar Zariski,* [**135**].

4.1. Introduction

This chapter will explore two related developments in algebraic geometry in the 20th century: general algebraic varieties and the language of schemes.

Although the machinery of affine varieties is powerful, it has striking limitations. For instance, the projective spaces from section 1.2 on page 3 are not affine varieties (see example 4.4.8 on page 186). General varieties (or just varieties) are the result of gluing affine varieties together via regular maps and solve this problem. They were introduced by André Weil in 1946 in his book [**172**].

This development is related to another, the gradual and halting evolution of the subject to something much more general than algebraic varieties: *schemes.* A scheme is a kind of "variety" whose "coordinate ring" is allowed to be an *arbitrary* ring (rather than an affine *k*-algebra). This transition to the language of schemes began in the late 19th century with the work of Kronecker.

In the 1880's, Kronecker considered extending algebraic geometry to "varieties" defined over arbitrary commutative rings, envisioning a field of mathematics that combined geometry, algebra and number theory (see [**92**]). His idea was largely ignored at the time.

In the 1930's, Krull and Noether proposed a similar generalization of algebraic geometry, but their ideas were widely rejected. Critics complained that using arbitrary rings as coordinate rings was not "geometric" and did not deserve to be called algebraic geometry.

In the early 1960's a kind of critical mass was achieved when Grothendiek published a series of papers collectively called EGA — see [**63, 64, 65, 66, 67**] — which gave a very elegant and well-thought-out framework for schemes. Grothendiek and Serre subsequently published papers on the Riemann hypothesis which involved "varieties" over non-algebraically closed fields.

Alexander Grothendieck (1928 – 2014) was a mathematician who became the leading figure in the creation of modern algebraic geometry. His research extended the scope of the field and added elements of commutative algebra, homological algebra, sheaf theory and category theory to its foundations, while his so-called "relative" perspective led to revolutionary advances in many areas of pure mathematics. He is considered by many to be the greatest mathematician of the 20th century.

Born in Germany, Grothendieck was raised and lived primarily in France, and he and his family were persecuted by the Nazi regime. For much of his working life, however, he was, in effect, stateless.

EXAMPLE 4.1.1. Let $k = \bar{\mathbb{F}}_p$, the algebraic closure of $\mathbb{F}_p$, and consider $x_i \mapsto x_i^p \colon \mathbb{A}^n \to \mathbb{A}^n$. This is the Frobenius Map — see definition A.2.40 on page 404. It is a bijection but it is not an isomorphism because the corresponding map on rings

$$X_i \mapsto X_i^p \colon k[X_1, \ldots, X_n] \to k[X_1, \ldots, X_n]$$

is not surjective. Let $k = \bar{\mathbb{F}}_p$, the algebraic closure of $\mathbb{F}_p$, and let $\mathcal{F}_p$ be the Frobenius map. Theorem A.2.42 on page 405 states 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 fixed points of $\mathcal{F}_p^m$. Therefore the fixed points of $\mathcal{F}_p^m$ are precisely the points of $\mathbb{A}^m$ 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

$$f = \sum c_{i_1, \ldots, i_n} x^{i_1} \cdots x^{i_n}, \quad c_{i_1, \ldots, i_n} \in \mathbb{F}_{p^m}$$

If $f(a_1, \ldots, a_n) = 0$, then

$$0 = \mathcal{F}_p^m \left(\sum c_{i_1, \ldots, i_n} a_1^{i_1} \cdots a_n^{i_n} \right) = \sum c_{i_1, \ldots, i_n} \mathcal{F}_p^m(a_1)^{i_1} \cdots \mathcal{F}_p^m(a_n)^{i_n}$$

So $f(\mathcal{F}_p^m(a_1), \ldots, \mathcal{F}_p^m(a_n)) = 0$. Thus $\mathcal{F}_p^m$ maps $\mathcal{V}(f)$ into itself and its fixed points are the solutions of

$$f(x_1, \ldots, x_n) = 0$$

that lie in $\mathbb{F}_{p^m}$.

Examples like this were one of the motivations for defining schemes over arbitrary rings (in this case, rings of polynomials over finite fields). The Gelfand spectrum gives a another example of scheme theory's power — see section 4.2.1.

The objections to scheme-theory, then, could be phrased "where is the geometry?" If arbitrarily rings can be "coordinate rings," isn't their study a branch of algebra rather than geometry?

The answer to this is that there are natural geometric constructions involving affine varieties that produce results that are *not* affine varieties. The problem is that their "coordinate rings" are not affine k-algebras — see section 4.3.4 on page 181.

Even in classical constructions, like example 2.2.8 on page 43, there are good reasons for allowing nilpotent elements in coordinate rings: for instance, section 3.3.4 on page 139 shows that nilpotent elements in the coordinate ring at a point correctly count intersection-multiplicities. Eliminating these elements

(by taking the radical of the defining ideal) actually loses geometric information.

So, after a long and halting transition, scheme theory has finally become the main language of algebraic geometry.

4.2. Affine schemes

An affine scheme is like an affine variety whose coordinate ring can be an *arbitrary* ring. Affine varieties are, therefore, special cases of affine schemes (see corollary 4.3.20 on page 177 for the precise relationship).

In moving from affine k-algebras to arbitrary rings we encounter several problems.

(1) We seem to lose Hilbert's Nullstellensatz.
(2) Proposition 2.5.3 on page 72 is no longer true: A homomorphism $f\colon R \to S$ of rings does not induce a well-defined map

$$\operatorname{Specm} S \to \operatorname{Specm} R$$

because the inverse image of a maximal ideal in S is not necessarily maximal. In fact, the most we can say of it is that it is a *prime ideal.*

We solve the second problem by defining:

DEFINITION 4.2.1. If R is a ring, define the ringed space $V = \operatorname{Spec} R$ the *spectrum of* R to be the set of all proper prime ideals of R. In addition, define

(1) for all $r \in R$, $D(r) = \{\mathfrak{p} \in \operatorname{Spec} R | r \notin \mathfrak{p}\}$,
(2) the topology on $\operatorname{Spec} R$ for which the sets $D(r)$, for all $r \in R$, form a base.
(3) the pair $(R, \operatorname{Spec} R)$ will be called an *affine scheme*.
(4) The *dimension* of $(R, \operatorname{Spec} R)$ is defined to be the Krull dimension (see definition 2.8.7 on page 100) of R.
(5) Given two affine schemes $V_1 = (R_1, \operatorname{Spec} R_1)$ and $V_2 = (R_2, \operatorname{Spec} R_2)$ the set of morphisms from V_1 to V_2 is denoted $\hom_{\operatorname{Scheme}}(V_1, V_2)$.
(6) If $\mathbf{0}$ denotes the *trivial ring*, $\operatorname{Spec} \mathbf{0} = \varnothing$.

REMARK. Compare this to that of specm $*$ in definition 2.5.1 on page 71.

It is interesting to compare $\operatorname{Spec} R$ and $\operatorname{Specm} R$. The former has "more points" — and points with some rather odd properties.

DEFINITION 4.2.2. Let R be a ring. A point $\mathfrak{p} \in \operatorname{Spec} R$ will be said to be

(1) *closed* if it is equal to its own closure. This means that $\mathfrak{p} \subset R$ is a maximal ideal, so it also corresponds to a point in $\operatorname{Specm} R \subset \operatorname{Spec} R$,
(2) *a generic point of an irreducible subscheme,* $V \subset \operatorname{Spec} R$, if its closure is all of V.
(3) a *generic point* if its closure is $\operatorname{Spec} R$.

REMARK. For instance, if R is an integral domain, the ideal (0) is prime and is a generic point for $\operatorname{Spec} R$.

As exotic as it sounds, the notion of "generic points" in algebraic varieties goes back to the 19th century. A generic point was one whose properties reflected those of an entire subvariety — and algebraic geometers in the early 1900's identified these with prime ideals. Emmy Noether proposed making

this notion "official" by including generic points for every prime ideal of the coordinate ring, but the idea was widely rejected.

The following definition allows us to recover a version of Hilbert's Nullstellensatz (see theorem 4.3.5 on page 170):

DEFINITION 4.2.3. Let R be a ring and let $x \in R$. The *restriction of x to a point* $\mathfrak{p} \in \operatorname{Spec} R$ is defined to be the image of x under the projection

$$R \to R/\mathfrak{p} \hookrightarrow F$$

where F is the field of fractions of $R/\mathfrak{p}$.

REMARK. In this setting, the ring R is the "coordinate ring" of $\operatorname{Spec} R$, and we think of its elements as regular functions.

In affine varieties, proposition 2.4.3 on page 62 shows that the value of a function in the coordinate ring at a point with maximal ideal $\mathfrak{m}$ is its image under the projection

$$R \to R/\mathfrak{m} = k$$

One difference between the present setting and affine varieties is that a "point" in $\operatorname{Spec} R$ might be an entire irreducible subvariety. The word "restriction" is still meaningful in this case, and the restriction of a function to an actual point is the same as evaluating at that point.

Another complicating factor is that the result of restricting an element of R to a point is an element of a field that *varies,* depending on the point — see examples 4.2.6 and 4.2.8 on page 166.

Perhaps the most interesting consequence of this definition is

PROPOSITION 4.2.4. *If $r \in R$ is nilpotent, then the restriction of r to any point of* $\operatorname{Spec} R$ *is* 0.

REMARK. The classic identification of an *element* of the coordinate ring and its *values* on points breaks down in scheme theory. In a scheme, there are nonzero "functions" whose "value" on every point is 0.

PROOF. This is an immediate consequence of theorem A.1.47 on page 360 or, more simply, the fact that the only nilpotent element in a field is 0. □

Table 4.2.1 on the next page summarizes some differences between $\operatorname{Spec} R$ and $\operatorname{Specm} R$. It is not hard to see that $\operatorname{Spec} R$ and $\operatorname{Specm} R$ determine each other *uniquely.*

We can even define tangent spaces:

DEFINITION 4.2.5. Let $A = \operatorname{Spec} R$ be an affine scheme and let $\mathfrak{p} \in A$ be a point. Then $A_{\mathfrak{p}} = S^{-1}A$, where $S = A \setminus \mathfrak{p}$ is a local ring with unique maximal ideal $\mathfrak{p} \cdot A_{\mathfrak{p}}$. Let $F = A_{\mathfrak{p}}/\mathfrak{p} \cdot A_{\mathfrak{p}}$. Then

$$V = \frac{\mathfrak{p} \cdot A_{\mathfrak{p}}}{\mathfrak{p}^2 \cdot A_{\mathfrak{p}}}$$

is a vector space over the field F. Its dual is called the *tangent space* at the point $\mathfrak{p}$.

If $x \in A$ has the property that the restriction of x to $\mathfrak{p}$ vanishes — i.e. if $x \in \mathfrak{p}$ then the *differential of x at* $\mathfrak{p}$, denoted $d_{\mathfrak{p}}x$, is just the image of x in V.

Specm $*$	Spec $*$	coordinate ring
point	closed point	maximal ideal
irreducible subvariety	point	prime ideal
	generic point of an irreducible subvariety	prime ideal $\subset$ other prime ideals
evaluation at a point	restriction to a point	projection to a quotient

TABLE 4.2.1. Comparison of Spec $*$ and Specm $*$
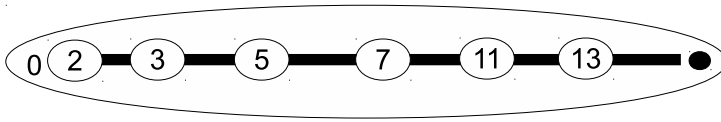
FIGURE 4.2.1.   Spec $\mathbb{Z}$ with its generic point

EXAMPLE 4.2.6. If $R = \mathbb{Z}$, the points of the space Spec $\mathbb{Z}$ correspond to the prime ideals $(p) \subset \mathbb{Z}$. All these points are closed. An integer $n \in \mathbb{Z}$ is a regular function on Spec $\mathbb{Z}$ whose value at a point (p) is its reduction $n \bmod p \in \mathbb{Z}_p$. The point defined by the ideal (0) is generic.

Since $\mathbb{Z}$ has Krull dimension 1, Spec $\mathbb{Z}$ is usually represented as a line — see figure 4.2.1.

The open sets $D(n)$ consist of Spec $\mathbb{Z} \setminus \bigcup(p_i)$ where the p_i are the primes that divide n.

The tangent space at the point $p \in \mathbb{Z}$ is just the quotient

$$\frac{(p)}{(p^2)} \cong \mathbb{Z}_p$$

Since it is a one-dimensional vector space, all points are simple points and Spec $\mathbb{Z}$ is smooth. If a number is a multiple of p its differential at p is just its value modulo p^2, so the differential of 40 at the point (5) is $40 \equiv 15 \pmod{25}$ (see lemma 3.3.9 on page 124).

We will refine the definition of an affine scheme somewhat — using the idea of *relative schemes*, proposed by Grothendiek:

DEFINITION 4.2.7. Let $V = \operatorname{Spec} R$ be an affine scheme. The *category, $\mathscr{A}_R$ of affine schemes over R or V* has

(1) objects $V = (\operatorname{Spec} S, \phi)$ where

$$\phi\colon \operatorname{Spec} S \to \operatorname{Spec} R$$

is a morphism, called the *structure morphism* of V,

(2) morphisms that preserve these maps. In other words a morphism, $f\colon (\mathrm{Spec}\, S_1, \phi_1) \to (\mathrm{Spec}\, S_2, \phi_2)$, fits into a commutative diagram

(4.2.1)
$$\begin{array}{ccc}
\mathrm{Spec}\, S_1 & \xrightarrow{\;\;f\;\;} & \mathrm{Spec}\, S_2 \\
& \phi_1 \searrow \quad \swarrow \phi_2 & \\
& \mathrm{Spec}\, R &
\end{array}$$

Sometimes we will call these *R-morphisms* or *V-morphisms* to emphasize that they are in the category of affine schemes over a fixed one.

The reader might wonder why we add this extra level of complexity.

First of all, it is not necessarily an extra level of complexity: Note that every ring, R, comes equipped with a *unique* ring-homomorphism

$$\mathbb{Z} \to R$$

so *every* affine scheme has a canonical map

$$\mathrm{Spec}\, R \to \mathrm{Spec}\, \mathbb{Z}$$

and is a scheme over $\mathbb{Z}$, and every k-algebra, A, comes equipped with a homomorphism

$$k \to A$$

inducing a map

$$\mathrm{Spec}\, A \to \mathrm{Spec}\, k$$

In the case of varieties, relative schemes can even *simplify* matters considerably by limiting the morphisms that can exist between schemes: rings have many automorphisms and any such automorphism defines a nontrivial map from a scheme to itself. We would like to rule out these maps that are really only derived from maps of coefficient-rings. A diagram like 4.2.1 guarantees that the "coefficients" always map via the *identity* map.

It follows that affine varieties are in the category $\mathscr{A}_k$ and that *all* affine schemes are in $\mathscr{A}_{\mathbb{Z}}$. So relative schemes will be useful in integrating affine *varieties* into our theory, as well as having other applications.

The following example is interesting because it interleaves geometry and algebra:

EXAMPLE 4.2.8. Let $R = \mathbb{Q}[X]$, polynomials with rational coefficients. Then $\mathrm{Spec}\, R$ is a scheme *over* $\mathbb{Q}$ (as well as one over $\mathbb{Z}$). Since this is a polynomial ring over a field, it is a principal ideal domain (see proposition A.2.4 on page 387) so all prime ideals are maximal. The points of $\mathrm{Spec}\, \mathbb{Q}[X]$ are all closed.

For each ideal of the form $(X - a)$, where $a \in \mathbb{Q}$, we get one point of $\mathrm{Spec}\, \mathbb{Q}[X]$, so $\mathrm{Spec}\, \mathbb{Q}[X]$ has a copy of $\mathbb{Q}$ embedded in it. If $p(X) \in \mathbb{Q}[X]$ is a polynomial, the result of restricting $p(X)$ to the point defined by the ideal $(X - a)$ is just $p(a) \in \mathbb{Q}$, as one might expect.

We also have irreducible polynomials of the form

$$X^2 - 2$$

and ideals like $(X^2 - 2) \subset \mathbb{Q}[X]$ define points in $\operatorname{Spec}\mathbb{Q}[X]$. Since

$$
\begin{aligned}
\mathbb{Q}[X]/(X^2 - 2) &= \mathbb{Q}[\sqrt{2}] \\
X &\mapsto \sqrt{2}
\end{aligned}
$$

the result of restricting $p(X)$ to the point defined by $(X^2 - 2)$ is $p(\sqrt{2}) \in \mathbb{Q}[\sqrt{2}]$.

So, the points of $\operatorname{Spec}\mathbb{Q}[X]$ "live in their own universes." It is not hard to see that every finite extension of $\mathbb{Q}$ occurs as the "location of a point" of $\operatorname{Spec}\mathbb{Q}[X]$.

This corresponds to Kronecker's idea that algebraic geometry is actually done over algebraic extensions of a "ground field" that is either *finite* or the *rationals.* For instance, if we have a system of algebraic equations with rational coefficients and our base field is $\mathbb{C}$, the significant points of the variety that these algebraic equations define will actually lie in a *finite extension* of $\mathbb{Q}$ embedded in $\mathbb{C}$.

Here is another example:

EXAMPLE 4.2.9. Let R be a ring and define

$$
\mathbb{A}_R^n = \operatorname{Spec} R[X_1, \ldots, X_n]
$$

the affine space with coefficients in R.

4.2.1. The Gelfand Spectrum. This is an example of the extraordinary generality of the scheme approach. Essentially every compact Hausdorff space (i.e., smooth manifold or polyhedron) can be regarded as an affine scheme — and in a way that is geometrically relevant:

> Two compact Hausdorff spaces are topologically equivalent if and only if they are isomorphic as schemes.

DEFINITION 4.2.10. Let X be a compact Hausdorff space and let $C(X)$ be the ring of complex-valued continuous functions on X. We call $\operatorname{Spec} C(X)$ the *Gelfand spectrum* of X.

REMARK. The Gelfand-Naimark Theorem implies that X is homeomorphic to $\operatorname{specm} C(X)$ — see **[54]** and **[5]**.

The original statement of the Gelfand-Naimark theorem is for locally-compact Hausdorff spaces.

If p is a point of a compact Hausdorff space and $\mathfrak{m}_p \subset C(X)$ is the ideal of functions that vanish at p, it is not hard to see that $\mathfrak{m}_p$ is often uncountably generated (i.e., if $X = S^1$, the unit circle). The structure of the quotient

$$
\frac{\mathfrak{m}_p}{\mathfrak{m}_p^2}
$$

— the cotangent space at p — is subtle and little is known about it.

EXERCISES.

1. Why isn't Proposition 2.5.3 on page 72 valid for schemes?

2. Let X be a compact Hausdorff space and let $f \in C(X)$ be a real-valued function that vanishes at $p \in X$, and is ≥ 0 everywhere on X. Show that the derivative of f at p (i.e., its image in the tangent space of $C(X)$ at p) is 0.

3. If M is the field of meromorphic functions (these are quotients of analytic functions) on $\mathbb{C}$, we can construct $\Omega_{M/\mathbb{C}}$ exactly as in section A.7 on page 482. Show that, in this vector-space, $d\left(e^x\right) \neq e^x \cdot dx$ — and $dx \nmid d\left(e^x\right)$ even! This shows that Kähler differentials should *not* be confused with calculus-differentials!

4. Consider the scheme $\operatorname{Spec}\mathbb{Z}[X]$ over $\mathbb{Z}$. Show that the function $X \in \mathbb{Z}[X]$ takes on every algebraic number at some point of $\operatorname{Spec}\mathbb{Z}[X]$.

5. Let $V = \operatorname{Spec}R$ be an affine scheme and let $x \in R$ be an element. Show that the restriction of x to every point of V (as defined in definition 4.2.3 on page 164) is *nonzero* if and only if x is a *unit* of R.

6. If $r \in R$ is an element of a ring, show that

$$D(r) = \varnothing \subset \operatorname{Spec}R$$

if and only if r is *nilpotent.*

7. If $r, s \in R$ show that $D(rs) \subseteq D(r)$.

8. If $r, s \in R$ show that $D(r) \cap D(s) = D(rs) \subset \operatorname{Spec}R$.

9. If $r \in R$ show that $D(r) \subset \operatorname{Spec}R$ is dense if and only if r is not a zero-divisor.

10. If k is an algebraically closed field, characterize the points of $\operatorname{Spec}k[X, Y]_{(X,Y)}$.

11. Describe the points of $\operatorname{Spec}\mathbb{R}[X]$. How would one draw this scheme?

12. Find a criterion for the set of closed points of an affine scheme to be dense.

13. Give a criterion for $\operatorname{Spec}R$ to be irreducible.

14. If R is a noetherian ring, the *associated points* of $\operatorname{Spec}R$ are those corresponding to associated *primes* of R, as in definition A.1.73 on page 374. Find the associated points of $\operatorname{Spec}R$ where

$$R = \frac{k[X, Y]}{(XY, Y^2)}$$

15. Show that an element of a ring R is a zero-divisor if and only if its restriction to some associated point of $\operatorname{Spec}R$ is 0.

4.3. Subschemes and ringed spaces

4.3.1. The sheaf of regular functions.

DEFINITION 4.3.1. Let R be a ring. For any ideal $\mathfrak{b} \subset R$ we define the *subscheme defined by* $\mathfrak{b}$ via

$$V(\mathfrak{b}) = \{\mathfrak{p} \in \operatorname{Spec} R | \mathfrak{b} \subset \mathfrak{p}\}$$

If R is an affine k-algebra and $\mathfrak{b}$ is a radical ideal, we get a corresponding definition of the *subvariety* defined by $\mathfrak{b}$.

Corresponding to the homomorphism $A \to A/\mathfrak{b}$ we get a regular map

$$\operatorname{Spec} R/\mathfrak{b} \to \operatorname{Spec} R$$

The image is $V(\mathfrak{b})$ and the map $\operatorname{Spec} R/\mathfrak{b} \to V(\mathfrak{b})$ is a *homeomorphism* (it induces a one-to-one correspondence between ideals of $R/\mathfrak{b}$ and ideals of R that contain $\mathfrak{b}$ — lemma A.1.25 on page 352). Thus every closed subset of $\operatorname{Spec} R$ has a natural ringed space structure making it into an affine scheme.

DEFINITION 4.3.2. Let $f\colon \operatorname{Spec} R \to \operatorname{Spec} S$ be a map of affine schemes. Then f is called a *closed immersion* if it is a homeomorphism onto a closed subscheme of $\operatorname{Spec} S$. It is an *open immersion* if it is a homeomorphism onto an open subscheme of $\operatorname{Spec} S$.

REMARK. It is important to note that this algebraic geometric term "immersion" has a *different* meaning than the term "immersion" in other areas of mathematics.

In differential geometry, an immersion is a map that is *locally* an embedding (i.e., in a neighborhood of each point) but might not *globally* be one-to-one. For instance, the *figure eight* is an immersion of a circle into a plane that intersects itself.

What is called an immersion in algebraic geometry would be called an *embedding* elsewhere.

The discussion following definition 4.3.1 essentially proved:

PROPOSITION 4.3.3. *Let $f\colon R \to S$ be a surjective homomorphism of rings. Then the induced map*

$$\operatorname{Spec} f\colon \operatorname{Spec} S \to \operatorname{Spec} R$$

is a closed immersion.

PROOF. Since f is surjective, we have

$$S \cong \frac{R}{\mathfrak{k}}$$

where $\mathfrak{k} \subset R$ is the kernel of f. The closed set that is the image of $\operatorname{Spec} S$ is the set of all prime ideals that contain $\mathfrak{k}$. $\square$

EXAMPLE 4.3.4. Let k be an algebraically closed field and consider the map

$$f\colon k[X] \to k[X]/(X^n) = R$$

This induces a map

$$f^*\colon \operatorname{Spec} R \to \operatorname{Spec} k[X] = \mathbb{A}^1$$
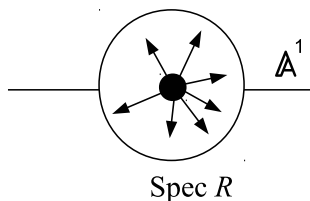
Spec R

FIGURE 4.3.1.   A nilpotent element

(where the affine line has been augmented with generic points). We claim that R has a single maximal ideal, (X). If $a \in k$ is nonzero, we have

$$X - a | X^n - a^n \in k[X]$$

so $X - a$ is a *unit* in R. It follows that Spec R consists of a single point and its image in $\mathbb{A}^1$ is the origin. This single point is "fat" (see figure 4.3.1) in the following sense:

> If we restrict a regular function on $\mathbb{A}^1$ to Spec $R \subset \mathbb{A}^1$ — mapping it to R via f — we not only get its *value* at $X = 0$, but we get its first $n-1$ *derivatives*.

Nilpotent elements of a ring serve a vital purpose: they act as "infinitesimals" or record "infinitesimal" information. Since evaluating an element (as in proposition 4.2.4 on page 164) at a point kills nilpotent elements, evaluating discards infinitesimals. This is somewhat like what happens in nonstandard analysis:

> The derivative of a function df/dx is the result of "standardizing" the quotient
> $$\frac{f(x + dx) - f(x)}{dx}$$
> — discarding infinitesimals.

Exercise 5 on page 168 is a bit like the "weak" version of Hilbert's Nullstellensatz. The following gives us the corresponding "strong" version (compare this to theorem 2.2.5 on page 41):

THEOREM 4.3.5. *Let R be a ring with an ideal $\mathfrak{b}$. If $V(\mathfrak{b}) = \emptyset$ then $\mathfrak{b} = R$.*
If the image of $x \in R$ in $R/\mathfrak{b}$ has the property that its restriction to every point of $V(\mathfrak{b}) = \text{Spec } R/\mathfrak{b}$ is 0, then

$$x^n \in \mathfrak{b}$$

for some integer n.

REMARK. It is striking that the proofs of this result and exercise 5 on page 168 are considerably shorter and simpler than that of the original Nullstellensatz. This shows some of the power of the scheme-theoretic approach.

PROOF. If $\mathfrak{b}$ has the property that $V(\mathfrak{b}) = \emptyset$, then $\mathfrak{b} \not\subset \mathfrak{p}$ for any $\mathfrak{p} \in \text{Spec } R$. But every proper ideal is contained in a maximal ideal, which is prime (see proposition A.1.20 on page 350) so $\mathfrak{b} = R$.

Suppose $\bar{x} \in R/\mathfrak{b}$ is the image of x under the projection

$$R \to R/\mathfrak{b}$$

The statement that the restriction of x to every point of $\operatorname{Spec} R/\mathfrak{b}$ vanishes is equivalent to saying that $\bar{x} \in \mathfrak{p}$ for all prime ideals $\mathfrak{p} \subset R/\mathfrak{b}$. Theorem A.1.47 on page 360 implies that $\bar{x}$ is *nilpotent,* i.e. that $\bar{x}^n = 0$ for some n. This is equivalent to $x^n \in \mathfrak{b}$. $\qquad\square$

PROPOSITION 4.3.6. *Let $V = \operatorname{Spec} R$ be an affine scheme and let $f \in R$ be nonzero.*
Then

$$D(f) = \operatorname{Spec} R_f$$

In particular, it is an affine scheme.

The map $R \to R_f$ defines a homeomorphism $\operatorname{Spec} R_{hf} \to \operatorname{Spec} A$ and the image is $D(h)$.

REMARK. This and exercise 9 on page 168 imply that $D(f) \subset \operatorname{Spec} R$ is *dense* if and only if $R \hookrightarrow R_f$ is *injective.*

PROOF. The space $D(f)$ is the set of prime ideals that do not contain f. Corollary A.1.94 on page 385 implies that this is in a one-to-one correspondence with the prime ideals of R_f. It follows that the appropriate "coordinate ring" of $D(f)$ is R_f. $\qquad\square$

As we did earlier, in section 3.2 on page 117, we can define the coordinate ring at a point:

DEFINITION 4.3.7. Let $V = \operatorname{Spec} R$ be an affine scheme and let $\mathfrak{p} \subset R$ be a prime ideal. Then the *coordinate ring at the point represented by* $\mathfrak{p}$ is

$$R_{\mathfrak{p}} = S^{-1}R$$

where $S = R \setminus \mathfrak{p}$.

REMARK. This is exactly like definition 3.2.1 on page 117, except that the point in question might be generic — i.e. an irreducible subvariety. We used this version of the "coordinate ring at a point" in the proof of lemma 3.4.3 on page 146, showing that it is useful even in classical algebraic geometry.

Definition 4.2.3 on page 164 shows how to "evaluate" functions at points of an affine scheme. In classical analysis, the set of points where a function is *nonzero* is called its *support.* This notion is much more subtle in algebraic geometry — inspired by the example of support of a section of a sheaf (see definition B.1.6 on page 497):

DEFINITION 4.3.8. Let $V = \operatorname{Spec} R$ be an affine scheme, and let $r \in R$. Then r defines a function on V and its *support* is defined to be the set of prime ideals $\mathfrak{p} \subset R$ with the property that the image of r under the localization map

$$R \to R_{\mathfrak{p}}$$

is nonzero.

REMARK. This is *not* equivalent to the definition in classical analysis: consider the function $X \in k[X]$ defined on the affine scheme $\mathbb{A}^1 = \operatorname{Spec} k[X]$. Its support in classical analysis is

$$\mathbb{A}^1 \setminus \{0\}$$

— i.e., where its *value* is nonzero — but its support in the definition above is *all* of $\mathbb{A}^1$. The difference is due to the fact that $R_{\mathfrak{p}}$ measures the behavior of a function in an "infinitesimal neighborhood" of the point defined by $\mathfrak{p}$ — called the *germ* of f at $\mathfrak{p}$. The function X vanishes *at* 0, but is nonzero in arbitrarily small *neighborhoods* of 0.

COROLLARY 4.3.9. *If R is a ring, the topological space* $\operatorname{Spec} R$ *is quasicompact.*

REMARK. "Quasicompact"="every open covering has a finite sub-cover." Although this sounds *identical* to the definition of "compact" (without the "quasi"), the Bourbaki people insist that compact spaces must be Hausdorff.

PROOF. Let

$$\operatorname{Spec} R = \bigcup_{\alpha \in S} D(f_\alpha)$$

be an open cover, where S is some set of indices. Then

$$\bigcap_{\alpha \in S} V((f_\alpha)) = \varnothing$$

so that

$$\sum_{\alpha \in S} (f_\alpha) = R$$

which means there is some equation

$$a_1 f_{\alpha_1} + \cdots + a_k f_{\alpha_k} = 1$$

and

$$\operatorname{Spec} R = \bigcup_{j=1}^{k} D(f_{\alpha_j})$$

$\square$

This is somewhat complicated by the fact that restrictions of elements of R to points live in different fields (as in example 4.2.6 on page 165).

LEMMA 4.3.10. *Let R be a ring and let $f, h \in R$ and suppose*

$$D(f) \subset D(h)$$

Then $f^n \in (h)$ for some $n > 0$.

PROOF. If $D(f) \subset D(h)$ then $f \notin \mathfrak{p} \implies h \notin \mathfrak{p}$ for all primes $\mathfrak{p} \in \operatorname{Spec} R$. This is equivalent to saying that $h \in \mathfrak{p} \implies f \in \mathfrak{p}$ for all $\mathfrak{p} \in \operatorname{Spec} R$.

Lemma A.1.25 on page 352 implies that the image, $\bar{f}$, of f under the projection

$$R \to R/(h)$$

is contained in all prime ideals of $R/(h)$. Theorem A.1.47 on page 360 implies that $\bar{f}$ is nilpotent, i.e. $\bar{f}^n = 0$ for some n, which implies that $f^n \in (h)$. $\square$

We can use this to define "restriction mappings":

DEFINITION 4.3.11. Let R be a ring and let $f, h \in R$ and suppose

$$t_{f,h}: D(f) \hookrightarrow D(h)$$

Let $f^n = a \cdot h$ for some $a \in R$ (guaranteed by lemma 4.3.10 on the preceding page). We can define a mapping

$$t^*_{f,h}: R_h \to R_f$$

that sends

$$\frac{r}{h^k} \mapsto \frac{r \cdot a^k}{f^{nk}}$$

Now we must verify the consistency condition (statement 3 in definition B.1.1 on page 495):

PROPOSITION 4.3.12. *Suppose R is a ring and $e, f, g, h \in R$ and $D(e) \subset D(f) \cap D(g)$ and $D(f) \cup D(g) \subset D(h)$. Then the diagram*

commutes.

PROOF. It suffices to show that the upper triangle commutes. The corresponding statement about the lower triangle follows by symmetry.

We have

$$\begin{align}
(4.3.1) \qquad\qquad e^{k_1} &= a \cdot f \\
(4.3.2) \qquad\qquad e^{k_3} &= c \cdot h \\
(4.3.3) \qquad\qquad f^{\ell_1} &= d \cdot h
\end{align}$$

Suppose

$$u = \frac{x}{h^k} \in A_h$$

Then $t^*_{f,h}(u)$ is

$$\frac{d^k x}{f^{\ell_1 k}}$$

and the image of that under $t^*_{e,f}$ is

$$\frac{d^k a^{\ell_1 k} x}{e^{k_1 \ell_1 k}}$$

and

$$t^*_{e,h}(u) = \frac{c^k x}{e^{kk_3}}$$

These are equal in A_e if and only if

$$z = y(d^k a^{\ell_1 k} x \cdot e^{kk_3} - c^k x \cdot e^{k_1 \ell_1 k})$$

vanishes for some $y \in R$. Equation 4.3.2 on the previous page implies that $e^{kk_3} = c^k h^k$ and equations 4.3.1 on the preceding page and 4.3.3 on the previous page implies that

$$e^{k_1 \ell_1 k} = a^{\ell_1 k} f^{\ell_1 k} = a^{\ell_1 k} d^k h^k$$

which gives

$$z = y(d^k a^{\ell_1 k} x \cdot c^k h^k - c^k x \cdot a^{\ell_1 k} d^k h^k)$$

which vanishes with $y = 1$.   □

We are now in a position to define a presheaf of regular functions on an affine scheme:

DEFINITION 4.3.13. If $V = \operatorname{Spec} R$ is an affine scheme, define a presheaf $\mathcal{O}_V$ by

$$\mathcal{O}_V(U) = \underset{D(f) \subset U}{\underleftarrow{\lim}} R_f$$

for every open set $U \subset V$.

REMARK. Proposition 4.3.12 on the preceding page implies that it is a presheaf.

PROPOSITION 4.3.14. *The presheaf in definition 4.3.13 is actually a sheaf. Furthermore*

$$\mathcal{O}_V(D(f)) = R_f$$

and $\mathcal{O}_V(V) = R$.

REMARK. The proof amounts to pointing out that the definition of an inverse limit in this case is the same as that of a sheaf.

Henceforth, we will assume that $\operatorname{Spec} R$ is equipped with the sheaf defined here and in definition 4.3.13 and that it is naturally a *locally ringed space* (see definition B.3.1 on page 502).

PROOF. The inverse limit is *all possible* sequences

$$(\ldots, r_\alpha, \ldots) \in \prod_{\alpha \in R} R_\alpha$$

where $\alpha \in R$ and $r_\alpha \in R_\alpha$, compatible with the restriction-maps defined in definition 4.3.11 on the preceding page. But this coincides with statement B.1.2 in definition B.1.1 on page 495.

The remaining statements follow from proposition A.5.32 on page 457.   □

So all closed subsets and principal open subsets of affine schemes are again affine schemes.

EXAMPLE 4.3.15. Suppose R is an integral domain and consider the "affine space," $\mathbb{A}_R^n = \operatorname{Spec} R[X_1, \ldots, X_n]$. Its sheaf of regular functions has the property that

$$\mathcal{O}_{\mathbb{A}_R^n}(D(f)) = R[X_1, \ldots, X_n]_f \subset R(X_1, \ldots, X_n)$$

Since the $D(f)$ are a basis for the topology of $\mathbb{A}_R^n$ it follows that

$$\mathcal{O}_{\mathbb{A}_R^n}(U) \subset R(X_1, \ldots, X_n)$$

for any open set $U \subset \mathbb{A}_R^n$, and

$$
\begin{aligned}
\mathcal{O}_{\mathbb{A}_R^n}(U_1 \cup U_2) &= \mathcal{O}_{\mathbb{A}_R^n}(U_1) \cap \mathcal{O}_{\mathbb{A}_R^n}(U_2) \\
\mathcal{O}_{\mathbb{A}_R^n}(U_1 \cap U_2) &= \mathcal{O}_{\mathbb{A}_R^n}(U_1) + \mathcal{O}_{\mathbb{A}_R^n}(U_2)
\end{aligned}
$$

as subrings of $R(X_1, \ldots, X_n)$. The *stalk* over a point represented by a prime ideal $\mathfrak{p} \subset R[X_1, \ldots, X_n]$ is

$$\mathcal{O}_{\mathbb{A}_R^n, \mathfrak{p}} = R[X_1, \ldots, X_n]_{\mathfrak{p}} \subset R(X_1, \ldots, X_n)$$

This the "coordinate ring at a point" as in section 3.3.2 on page 132.

If R has zero-divisors, the situation becomes considerably more complex because the localization maps fail to be injective.

Now we get a result analogous to 2.5.5 on page 73:

THEOREM 4.3.16. *If $\mathscr{R}$ is the category of commutative rings, $\mathscr{A}_{\mathbb{Z}}$ is that of affine schemes, and*

$$m\colon \mathscr{A}_{\mathbb{Z}} \to \mathscr{R}$$

is the functor that maps an affine scheme V to the ring $\mathcal{O}_V(V)$, then m and $\mathrm{Spec} *\colon \mathscr{R} \to \mathscr{A}_{\mathbb{Z}}$ *form a contravariant equivalence of categories.*

REMARK. In light of this result, the reader might wonder why we bother with the ringed space structure of $\mathrm{Spec}\, R$ — all of the algebraic and geometric information is already contained in the ring R. The point is that later (in section 4.4 on page 182), we will be concerned with general schemes where the ringed space structure is vital to their definition.

This sheaf is even necessary for defining *affine* schemes — if we want to think of them as topological spaces. Consider the affine varieties $\mathrm{Spec}\,\mathbb{R}$ and $\mathrm{Spec}\,\mathbb{C}$. As topological spaces, they both consist of single points — implying the existence of morphisms in both directions:

$$\mathrm{Spec}\,\mathbb{R} \leftrightarrow \mathrm{Spec}\,\mathbb{C}$$

Although there exists a ring-homomorphism

$$\mathbb{R} \to \mathbb{C}$$

no map going the opposite direction exists. The sheaves of regular functions allow us to recover the rings that define the schemes.

PROOF. This follows from definition 4.2.1 on page 163 and proposition 4.3.14 on the preceding page. $\qquad \square$

This gives us a criterion for ringed spaces to be affine varieties:

LEMMA 4.3.17. *Let $V \subset \mathbb{A}^n$ be a ringed space with sheaf $\mathcal{O}_V$ whose inclusion into $\mathbb{A}^n$ is a map of ringed spaces. Then $(V, \mathcal{O}_V)$ is an affine variety if and only if $\mathcal{O}_V(V)$ is an affine k-algebra and the canonical map $V \to \mathrm{specm}\,\mathcal{O}_V(V)$ is an isomorphism of ringed spaces.*

REMARK. For instance $\mathbb{A}^2 \setminus (0,0) = D(x) \cup D(y)$ has a coordinate ring of $k[X, Y]$. It gives rise to a ringed space that is not affine: the realization of $k[X, Y]$ is $\mathbb{A}^2$ and the canonical map

$$\mathbb{A}^2 \setminus (0,0) \to \mathbb{A}^2$$

is *not* an isomorphism. This example points out the limitations of affine varieties and schemes:

> *Unions* of affine varieties are not necessary affine varieties.

We will remedy this by defining *general* varieties (and schemes) in chapter 4 on page 161.

PROOF. Let $(V, \mathcal{O}_V)$ be an affine variety and let $A = \mathcal{O}_V(V)$. For any $p \in V$ define $\mathfrak{m}_p = \{f \in A | f(p) = 0\}$. Then $\mathfrak{m}_p$ is a maximal ideal in A, and it is straightforward to verify that $p \mapsto \mathfrak{m}_p$ is an isomorphism of ringed spaces.

Conversely, if $\mathcal{O}_V(V)$ is an affine k-algebra, then specm $\mathcal{O}_V(V)$ is an affine variety — see definition 4.2.1 on page 163 and there is a canonical map

$$V \to \text{specm} \, \mathcal{O}_V(V)$$

as in proposition 2.5.2 on page 72. If this map is an isomorphism of ringed spaces, $(V, \mathcal{O}_V)$ is an affine variety. □

We get a relative version of theorem 4.3.16 on the previous page if we define:

DEFINITION 4.3.18. If S is a commutative ring, the category $\mathscr{R}_S$ *of commutative rings over S* has

(1) objects that are commutative rings equipped with maps from S — i.e., an object is a pair (R, φ) where $\varphi \colon S \to R$ is a homomorphism of commutative rings.
(2) morphisms that preserve these maps, i.e. a morphism $f \colon (R_1, \varphi_1) \to (R_2, \varphi_2)$ is a homomorphism of commutative rings, $f \colon R_1 \to R_2$ that makes the diagram

$$\begin{array}{ccc} & S & \\ {\scriptstyle \varphi_1} \swarrow & & \searrow {\scriptstyle \varphi_2} \\ R_1 & \xrightarrow{\quad f \quad} & R_2 \end{array}$$

commute.

REMARK. Since every commutative ring, R, comes equipped with a *unique* map

$$\mathbb{Z} \to R$$

it follows that the category of commutative rings is identical (i.e., isomorphic) to $\mathscr{R}_{\mathbb{Z}}$.

COROLLARY 4.3.19. *If S is a commutative ring, the functors m and $\text{Spec} *$ defined in theorem 4.3.16 are a contravariant equivalence of categories*

$$\mathscr{A}_S \to \mathscr{R}_S$$

REMARK. Compare this with theorem 2.5.5 on page 73.

PROOF. This is just the relative version of theorem 4.3.16 on page 175. □

We finally conclude that Spec $*$ is just as good as specm $*$ for defining affine varieties:

COROLLARY 4.3.20. *If k is an algebraically closed field, let $\mathscr{A}\mathscr{V}_k$ denote the category of affine varieties over k and let $\mathscr{A}\mathscr{S}_k$ denote the category of affine schemes over k where $A \in \mathscr{A}\mathscr{F}_k$. Then there is an equivalence of categories*

$$\mathscr{A}\mathscr{V}_k \to \mathscr{A}\mathscr{S}_k$$

REMARK. This follows because both categories are equivalent to $\mathscr{A}\mathscr{F}_k$, the category of affine k-algebras. We can also define a direct equivalence:

- Map a scheme to the space of all of its *closed* points (corresponding to maximal ideals), with the same topology.
- Map an affine variety, V, to the scheme $\operatorname{Spec} k[V]$.

EXERCISES.

1. Do an analysis of $\mathbb{A}_R^n$ like example 4.3.15 on page 174 in the case where all of the zero-divisors of R are nilpotent elements.

2. Do an analysis of $\mathbb{A}_R^n$ like example 4.3.15 on page 174 in the case where $R = R_1 \oplus R_2$ and R_1 and R_2 are integral domains (where products of elements in R_1 and elements of R_2 vanish).

3. If R is an Artinian ring (see definition A.1.83 on page 381), show that $V = \operatorname{Spec} R$ contains a finite number of points $\{p_1, \dots, p_n\}$ and that

$$R \cong \prod_{j=1}^{n} \mathcal{O}_{V,p_i}$$

4. If $V = \operatorname{Spec} R$, show that $\mathcal{O}_V = \mathcal{A}(R)$, where $\mathcal{A}(*)$ is defined in definition 3.5.1 on page 155.

5. If $V = \operatorname{Spec} R$ and $\mathfrak{p} \subset R$ is a prime ideal representing a point $p \in V$, show that the coordinate ring of V at p is nothing but the *stalk* of $\mathcal{O}_V$ at p — see definition B.1.1 on page 495. This shows that definition 4.3.8 on page 171 is equivalent to definition B.1.6 on page 497.

6. Show that the support of a function (see definition 4.3.8 on page 171) is closed.

4.3.2. Meromorphic functions. Now that we have defined the sheaf of regular functions, we can try to define *rational* functions. Defining such objects is not as simple as "forming the field of fractions" of the ring of regular functions. For one thing, some regular functions may be divisors of zero so that defining inverses would not make sense. And in forming a sheaf, one must define *restriction-maps* — some of which might carry a potentially invertible function to a zero-divisor.

We will follow Kleiman's treatment in [**87**]:

DEFINITION 4.3.21. If $(X, \mathcal{O}_X)$ is a ringed space, the *presheaf of meromorphic functions*, $\overline{\mathcal{K}_X}$, is defined by

$$\overline{\mathcal{K}_X}(U) = \mathcal{O}_X(U)[S(U)^{-1}]$$

where:

(1) $U \subset X$ is an open set,
(2) $S(U) \subset \mathcal{O}_X(U)$ is the multiplicative set of elements whose restrictions to the stalks $\mathcal{O}_{X,x}$ are non-zero divisors for *all* $x \in X$.

The *sheaf of meromorphic functions* on X, denoted $\mathcal{K}_X$, is the completion of $\overline{\mathcal{K}_X}$ (see lemma B.2.3 on page 500).

The sheaf of *regular meromorphic functions*, $\mathcal{K}_X^* \subset \mathcal{K}_X$ is the sheaf of (abelian) groups composed of the invertible elements of $\mathcal{K}_X$.

REMARK. The expression $\mathcal{O}_X(U)[S(U)^{-1}]$ is just the ring of fractions — see definition A.1.89 on page 383.

The term "meromorphic" is taken from complex analysis — meromorphic functions are quotients of regular functions. We use it here (instead of "sheaf of rational functions") to emphasize the subtleties of this construction.

Of course, there are simple cases where no subtlety is needed:

EXAMPLE 4.3.22. If A is an integral domain with field of fractions, F, and $V = \operatorname{Spec} A$, then

$$\mathcal{K}_V = \underline{F}$$

the *constant sheaf* equal to F on every open set. In this case, we do not "need" to use a sheaf at all — we can simply use the field of rational functions.

4.3.3. Fibered Products. Since $\operatorname{Spec} *$ is a contravariant equivalence of categories (see theorem 2.5.5 on page 73 and corollary 4.3.19 on page 176), it converts *coproducts* into *products*, so

DEFINITION 4.3.23. Let S be a commutative ring and let $R_1, R_2 \in \mathscr{R}_S$. Then we define

(4.3.4) $(\operatorname{Spec} R_1) \times_S (\operatorname{Spec} R_2) = \operatorname{Spec}(R_1 \otimes_S R_2) \in \mathscr{A}_S$

— see definition A.5.44 on page 462 for this version of tensor product.

REMARK. This is the categorical product (see definition A.5.1 on page 439) in the category $\mathscr{A}_S$. This is also called the *fibered product of* $\operatorname{Spec} R_1$ *and* $\operatorname{Spec} R_2$ *over* $\operatorname{Spec} S$.

It is the object in the category of affine schemes, $\mathscr{A}$, (any object in $\mathscr{A}_S$ is also an object of $\mathscr{A}$) with the universal property that, for any morphisms

$$\begin{aligned} f \colon X &\to \operatorname{Spec} A \\ g \colon X &\to \operatorname{Spec} A \end{aligned}$$

that make the diagram of solid arrows

(4.3.5)

commute, there exists a *unique* dotted arrow making the whole diagram commute. It is probably easier to simply call it a product in the relative category $\mathscr{A}_S$ — in which case all of the maps to $\operatorname{Spec} S$ (and the commutative of the diagrams that arise) are implied.

Note that

$$(\operatorname{Spec} R_1) \times_{\mathbb{Z}} (\operatorname{Spec} R_2) = (\operatorname{Spec} R_1) \times (\operatorname{Spec} R_2)$$

is just the ordinary product.

One interesting map in algebraic geometry is the diagonal map

DEFINITION 4.3.24. Let S be a ring and let $R \in \mathscr{R}_S$. The *diagonal map*

$$\Delta \colon \operatorname{Spec} R \to \operatorname{Spec} R \times_S \operatorname{Spec} R$$

is defined to be the unique map that makes the diagram

commute, where the p_i are projection to the first and second factors.

REMARK. If V is an affine variety, the diagonal map actually maps a point $x \in V$ to $(x, x) \in V \times V$.

LEMMA 4.3.25. *If S is a ring and $R \in \mathscr{R}_S$, then the diagonal map*

$$\Delta \colon \operatorname{Spec} R \to \operatorname{Spec} R \times_S \operatorname{Spec} R$$

is a closed immersion.

REMARK. As usual, a similar result holds for affine varieties.

It is well-known that the diagonal map

$$\Delta \colon X \to X \times X$$

being closed is equivalent to X being Hausdorff — see [**175**], 13.7.

PROOF. The map of rings that induces the diagonal is the unique homomorphism

$$\delta \colon R \otimes_S R \to R$$

that makes the diagrams

commute, where the g_i are the inclusions of factors. The map defined by

$$\delta(r_1 \otimes r_2) \mapsto r_1 \cdot r_2$$

clearly fits the bill. Since this map is surjective, the conclusion follows from proposition 4.3.3 on page 169 □

Although schemes and varieties are not Hausdorff, lemma 4.3.25 on the previous page has the interesting consequence

PROPOSITION 4.3.26. *Let V and W be affine varieties and let*

$$\varphi_1, \varphi_2 \colon V \to W$$

be two regular maps. Then the set of points $x \in V$ where $\varphi_1(x) = \varphi_2(x)$ is closed.

PROOF. Consider the diagram

The set of points where the φ_i, $i = 1, 2$ agree is just $(\varphi_1 \times \varphi_2)^{-1}(\Delta(W))$, which is closed since $\varphi_1 \times \varphi_2$ is continuous. □

EXERCISES.

7. Let U, V, W be affine varieties and let $p \in W \times_V U$ be a point projecting to $p_1 \in W$, $p_2 \in U$ and $p_3 \in V$. Show that

$$\mathcal{O}_{W \times_V U, p} = \left(\mathcal{O}_{W, p_1} \otimes_{\mathcal{O}_{V, p_3}} \mathcal{O}_{U, p_2} \right)_T$$

— see definition 3.2.1 on page 117 — where T is a suitable multiplicative set in $\mathcal{O}_{W, p_1} \otimes_{\mathcal{O}_{V, p_3}} \mathcal{O}_{U, p_2}$. This is a *localized* form of equation 4.3.4 on page 178.

4.3.4. Another motivation for the use of schemes. We conclude with a motivation for allowing general rings to be "coordinate rings" of varieties — a motivation for schemes.

Let $W \subset \mathbb{A}^2$ be the parabola $Y = X^2$ with coordinate ring $k[W] = k[X,Y]/(Y - X^2)$. Projection onto the Y-axis is modeled by the ring-homomorphism

$$k[Y] \to k[X,Y]/(Y - X^2)$$

And inclusion of the origin $(0,0) \hookrightarrow Y -$ axis is modeled by

$$\begin{aligned} k[Y] &\to k \\ Y &\mapsto 0 \end{aligned}$$

It follows that the fibered product

$$W \times_{Y\text{-axis}} (0,0)$$

is modeled by

$$k[W] \otimes_{k[Y]} k = k[X]/(X^2)$$

— which is certainly *not* an affine k-algebra since it has a nilpotent element.

So this is a natural geometric construction involving affine varieties that leads one *out* of the category of affine varieties. One might argue that we should simply take the radical of the ideal (X^2) (as was done in the past) but this seems a bit artificial, and it loses actual information: in some sense the intersection of W with the X-axis is of multiplicity 2 (see definition 1.4.6 on page 23) and the nilpotent element $X \in k[X]/(X^2)$ records this fact.

Compare this with example 2.2.8 on page 43: when we took the radical, we lost the information that the intersection had multiplicity 2.

EXERCISES.

8. We know that $\operatorname{Spec} R$ is not Hausdorff, and lemma 4.3.25 on page 179 proves that its diagonal map is a closed immersion. Yet, for *any topological space, X*, the image of

$$\Delta \colon X \to X \times X$$

being closed implies that X is Hausdorff (see [**175**], 13.7). Isn't this a contradiction?

9. Suppose $j \colon X' \hookrightarrow X$ is a closed immersion of affine schemes and suppose $f \colon Y \to X$ is a morphism of affine schemes. Show that

$$Y \times_X X' \cong f^{-1}(j(X'))$$

the fiber of X'.

4.3.5. "Varieties" of affine schemes. Every one of the types of rings discussed in appendix A on page 341 gives rise to affine schemes with geometric properties derived from the ring. These properties are, of course, not mutually exclusive.

- If R is an integral domain, Spec R is called an *integral scheme*. Every principal open set $D(f) \subset$ Spec R is *dense* in Spec R (see exercise 9 on page 168) so the scheme is irreducible.
- If R is a reduced ring (i.e., it has no nonzero nilpotent elements) then Spec R is a *reduced scheme*.
- If R is a normal domain (integrally closed in its field of fractions), then Spec R is called a *normal scheme*.
- If R is a noetherian ring, Spec R is called a *noetherian scheme*. If $r \in R$ is any non zero-divisor,

$$H(r) = \{\text{prime ideals } \mathfrak{p} \subset R | r \in \mathfrak{p}\} \subset \text{Spec } R$$

  is the hypersurface generated by r — the "solution set" of the "equation" $r = 0$. Theorem 2.8.29 on page 109 implies that all of the irreducible components of $H(r)$ are of codimension 1 in Spec R.
- If R is a Jacobson ring (see definition A.4.22 on page 428), Spec R is called a *Jacobson scheme*. The fact that $\mathfrak{N}(R) = \mathfrak{J}(R)$ (see proposition A.4.23 on page 428) implies that closed points are *dense* in Spec R (also see exercise 98 on page 596). The fact that every quotient of R is also Jacobson (corollary A.4.24 on page 428) implies that closed point are dense in every *closed subscheme* of Spec R. This property is often expressed by saying closed points are *very dense* in Spec R.

4.4. Schemes

Now we are ready to define *general* schemes — *unions* of affine schemes along certain types of maps. A general algebraic *variety* is just a scheme that is locally of the form Spec A where A is a finitely generated, reduced k-algebra over an algebraically closed field, k.

DEFINITION 4.4.1. A *scheme* is a locally ringed space $(V, \mathcal{O}_V)$ such that V is quasicompact and every point of V has an open neighborhood U for which $(U, \mathcal{O}_V|U)$ is an affine scheme. A scheme, V, *over* another scheme, S, is just a scheme equipped with a morphism

$$V \to S$$

and such that morphisms are required to commute with it — as in definition 4.2.7 on page 165.

A *prevariety* over an algebraically closed field, k, is scheme over Spec k (i.e., over k) such that every point has an open neighborhood, U, with $(U, \mathcal{O}_V|U)$ finitely generated and reduced (i.e., $\mathcal{O}_V(U)$ has no nilpotent elements).

REMARK. There has been a somewhat confusing evolution of the terminology over the years. In the past, what we call a scheme here was once called a prescheme and what was once called a scheme is now called a *separated* scheme (see definition 4.6.1 on page 205).

To make matters even more confusing, the old convention was *kept* for varieties, so an object that is locally an affine variety is called a *prevariety*.

An open subset U of a scheme V such that $(U, \mathcal{O}_V|U)$ is an affine scheme is called an *open affine* or *affine chart* of V — this is just a restatement of definition 2.4.19 on page 68.

It is not hard to see that the open affines form a base for the topology of V. Indeed, if $W \subset V$ is an open affine, then *principal* open sets of W will also be open affines of V.

DEFINITION 4.4.2. Let U be an open subset of a scheme $(V, \mathcal{O}_V)$. The functions in $\mathcal{O}_V(U)$ are regular functions.

If $\{U_i\}$ is a covering by open affines, f is regular if and only if $f|U \cap U_i$ is regular for all i. The points of V are the prime ideals of $\mathcal{O}_V(U)$ for open affines $U \subset V$. Evaluating a regular function at a point is defined in definition 4.2.3 on page 164. Given a regular function, f, we can also consider its image in $\mathcal{O}_{V,x}$ — the *stalk* of $\mathcal{O}_V$ (see definition B.1.1 on page 495) at the point, $x \in V$.

REMARK. Thus understanding the regular functions on open subsets of V amounts to understanding the regular functions on the open affines of V and how they fit together to form V.

If $x \in V$ is a point of a scheme represented by a print ideal $\mathfrak{p} \subset \mathcal{O}_V(U)$, the stalk, $\mathcal{O}_{V,x}$ is a local ring $(R, \mathfrak{m})$ whose quotient $R/\mathfrak{m}$ is equal to the field of fractions of $\mathcal{O}_V(U)/\mathfrak{p}$. The stalk, $\mathcal{O}_{V,x}$, contains more information than the quotient since nilpotent elements still exist.

DEFINITION 4.4.3. Let $(V, \mathcal{O}_V)$ and $(W, \mathcal{O}_W)$ be schemes. A map $\varphi \colon V \to W$ is *regular* if it is a morphism of locally ringed spaces.

REMARK. The local nature of sheaves immediately implies that:
If $f \colon V \to W$ is a continuous map and W_i is a covering of W by open affines, such that $V_i = f^{-1}(W_i)$ is an open affine, then f is regular if and only if

$$f|V_i \colon V_i \to W_i$$

is regular.

Now we can define the relative version of definition 4.4.1 on the preceding page:

DEFINITION 4.4.4. If T is a scheme, $\mathscr{S}_T$ the category of schemes *over T* has objects $V = (R, \phi)$ where $\phi \colon R \to T$ is a regular map called the *structure map* of V and morphisms $f \colon (R_1, \phi_1) \to (R_2, \phi_2)$ that make the diagram

$$
\begin{array}{ccc}
R_1 & \xrightarrow{\ f\ } & R_2 \\
& \searrow{\scriptstyle \phi_1} \quad \swarrow{\scriptstyle \phi_2} & \\
& T &
\end{array}
$$

commute.

4.4.1. Schemes obtained by patching. We can explicitly build prevarieties and schemes by gluing together ringed spaces.

COROLLARY 4.4.5. *Let $\{U_i\}$, $i = 1, \dots, n$ be a finite set of ringed spaces with open sets $\{V_{i,j} \subset U_i\}$ where $j = 1, \dots, n$ and isomorphisms of ringed spaces*

$$\varphi_{i,j} \colon V_{i,j} \to V_{j,i}$$

such that

(1) *$\varphi_{i,i} = 1 \colon V_{i,i} \to V_{i,i}$ for all $i = 1, \dots n$,*
(2) *$\varphi_{j,i} = \varphi_{i,j}^{-1}$ for all $i, j = 1, \dots n$,*
(3) *if $\varphi'_{i,j} = \varphi_{i,j}|V_{i,j} \cap V_{i,k} \colon V_{i,j} \cap V_{i,k} \to V_{j,i}$ is an isomorphism onto $V_{j,i} \cap V_{j,k} \subset V_{j,i}$ for all $i, j, k = 1, \dots, n$ and the diagrams*

$$
\begin{array}{ccc}
V_{i,j} \cap V_{i,k} & \xrightarrow{\varphi'_{i,j}} & V_{j,i} \cap V_{j,k} \\
 & \varphi'_{j,k} \searrow & \downarrow \varphi'_{j,k} \\
 & & V_{k,i} \cap V_{k,j}
\end{array}
$$

commute, for all $i, j, k = 1, \dots n$.

Then there exists a ringed space X with maps

$$v_i \colon U_i \to X$$

that are homeomorphisms onto open sets in X. If the U_i are prevarieties, then so is X.

PROOF. Let

$$\bar{X} = \bigcup_{i=1}^{n} U_i$$

the disjoint union and define a relation on $\bar{X}$ as follows

If $p_1 \in V_{i,j} \subset U_i$ and $p_2 \in V_{j,i} \subset U_j$ then $p_1 \sim p_2$ if an only if $\varphi_{i,j}(p_1) = p_2$. The conditions on the $\varphi_{i,j}$ imply that this is an equivalence relation.

Now set $X = \bar{X}/\sim$ with canonical projection

$$p \colon \bar{X} \to X$$

Define

$$v_i \colon U_i \to X$$

to be the composites

$$U_i \hookrightarrow \bar{X} \xrightarrow{p} X$$

We define a set in $U \subset X$ to be open if $v_i^{-1}(U) \subset U_i$ is open for all i. The map p is a homeomorphism of $\psi_i \colon U_i \to V_i = p(U_i)$ and

$$X = \bigcup_i V_i$$

We define the sheaf of regular functions on X by setting, for any $W \subset V_i$,

$$\mathcal{O}_X(W) = \mathcal{O}_{U_i}(\psi_i^{-1}(W))$$

If W is also a subset of some V_j with $j \neq i$ this construction will replace $\mathcal{O}_X(W)$ with an isomorphic ring. This defines $\mathcal{O}_X$ as a presheaf on open sets of the form

$W \subset V_i$ for some i. We extend it to a sheaf on *all* open sets by defining

$$\mathcal{O}_X(U) = \underbrace{\varprojlim \mathcal{O}_X(W)}_{\substack{W \subset U, \text{ such that} \\ W \subset V_i \text{ for some } i}}$$

for any open set U.   □

EXAMPLE 4.4.6. If k is a field, consider the maps of rings

$$k[X] \to k[T, T^{-1}] \leftarrow k[Y]$$

where $X \mapsto T$ and $Y \mapsto T^{-1}$. This induces maps of affine varieties

$$U \leftarrow H \to V$$

where H is the hyperbola in $\mathbb{A}^2$ defined by $XY = 1$ and $U = V = \mathbb{A}^1$. If we glue the two copies of $\mathbb{A}^1$ together along H, we get a space containing U and V glued via the map $X \mapsto Y^{-1}$ in $U \cap V$. Figure 4.4.1 on page 185 shows a kind of crude image of what is going on. The space

$$S = U \cup_{U \cap V} V$$

is isomorphic to the one-dimensional projective space, $\mathbb{P}(\mathbb{A}^2)$, via the map

$$(x_0 : x_1) \mapsto \begin{cases} \frac{x_0}{x_1} \in U & \text{if } x_1 \neq 0 \\ \frac{x_1}{x_0} \in V & \text{if } x_0 \neq 0 \end{cases}$$

Since x_0 and x_1 are never simultaneously 0 in $k\mathbb{P}^1$, every point of $k\mathbb{P}^1$ maps to a point in S, and maps to points in $U \cap V$ in a way compatible with the gluing map.
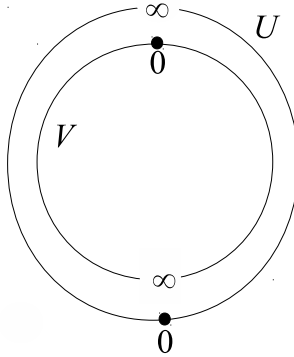

FIGURE 4.4.1.   Gluing affines, case 1

Regular functions in S are regular functions in U that extend to regular functions in V. So

$$a_n X^n + \cdots + a_0 \in k[X]$$

would map into

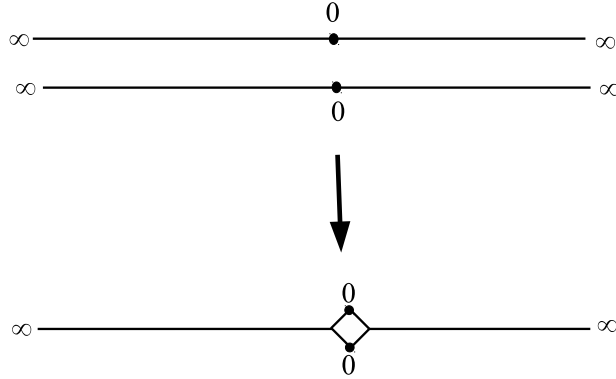$$a_n Y^{-n} + \cdots + a_0 \in k(Y)$$

FIGURE 4.4.2.   Gluing affines, case 2

under the gluing map. This is only an element of $k[Y]$ if $n = 0$, so the only regular functions on S are *constants* — as in example 4.4.8 on the following page.

It is interesting to consider impact of a slight change in the gluing maps:

EXAMPLE 4.4.7. As in example 4.4.6 on the previous page, consider the maps of rings

$$k[X] \to k[T, T^{-1}] \leftarrow k[Y]$$

where $X \mapsto T$ and $Y \mapsto T$. This also induces maps of affine varieties

$$U \leftarrow H \to V$$

where H is the hyperbola in $\mathbb{A}^2$ defined by $XY = 1$ and $U = V = \mathbb{A}^1$. If we glue the two copies of $\mathbb{A}^1$ together along H, we get a space containing U and V glued via the map $X \mapsto Y$ in $U \cap V$. Let D be the topological space that results from this gluing.

Figure 4.4.2 shows the effect of this gluing process. In this case, the affines were glued together via the identity map everywhere except at the origin. Since $H = \mathbb{A}^1 \setminus \{0\}$, V is an object that looks like $\mathbb{A}^1$ except that it has *two* copies of the origin. In this case, the ring of regular functions is simply $k[X] = k[Y]$.

Recall projective spaces as defined in chapter 1. We will devise a scheme version of them:

DEFINITION 4.4.8. Let R be a ring and define a scheme, $R\mathbb{P}^n$, over R by

$$R\mathbb{P}^n = \bigcup_{i=0}^{n} \mathbb{A}_{R,i}^n$$

where

$$\mathbb{A}_{R,i}^n = \mathbb{A}_R^n = \operatorname{Spec} R\left[X_{0,i}, \ldots, X_{i-1,i}, X_{i+1,i}, \ldots, X_{n,i}\right]$$

If $U_{i,j} = \mathbb{A}_{R,i}^n \cap \mathbb{A}_{R,j}^n$, then

(1) as a subset of $\mathbb{A}^n_{R,i}$, its coordinate ring is

$$(4.4.1) \quad R\left[X_{0,i}, \ldots, X_{i-1,i}, X_{i+1,i}, \ldots, X_{n,i}\right]_{X_{j,i}}$$
$$= R\left[X_{0,i}, \ldots, X_{i-1,i}, X_{i+1,i}, \ldots, X_{n,i}, X_{j,i}^{-1}\right]$$

(2) and, as a subset of $\mathbb{A}^n_{R,j}$,

$$(4.4.2) \quad R\left[X_{0,j}, \ldots, X_{j-1,j}, X_{j+1,j}, \ldots, X_{n,j}\right]_{X_{i,j}}$$
$$= R\left[X_{0,j}, \ldots, X_{j-1,j}, X_{j+1,j}, \ldots, X_{n,j}, X_{i,j}^{-1}\right]$$

We identify these two rings via

$$(4.4.3) \quad \begin{aligned} X_{i,j} &= X_{j,i}^{-1} \\ X_{k,j} &= X_{k,i}/X_{j,i} \text{ for } 0 \le k \le n, \, k \ne i, j \end{aligned}$$

If R is a field like $\mathbb{R}$ or $\mathbb{C}$, this corresponds to the usual definition (definition 1.2.2 on page 4) where we map a point

$$[x_0: \cdots : x_n] \in \mathbb{C}P^n = \mathbb{P}(\mathbb{C}^{n+1})$$

with $x_j \ne 0$ via

$$x_i/x_j \mapsto X_{i,j} \text{ for } 0 \le i \le n, \, i \ne j$$

Equation 4.4.3 implies that these assignments are consistent on overlaps between the $\mathbb{A}^n_{R,i}$.

The result of this gluing operation is a scheme, $R\mathbb{P}^n$, and the maps,

$$(4.4.4) \quad q_i \colon R[X_{0,i}, \ldots, X_{i-1,i}, X_{i+1,i}, \ldots, X_{n,i}]$$
$$\hookrightarrow R[X_0, \ldots, X_n][X_i^{-1}]$$

defined on open sets, $D(X_i) = \operatorname{Spec} R[X_0, \ldots, X_n][X_i^{-1}]$ by

$$(4.4.5) \quad X_{j,i} \mapsto X_j \cdot X_i^{-1}$$

induce maps, $p_i = \operatorname{Spec} q_i$, that patch together to define a regular map

$$(4.4.6) \quad p \colon \operatorname{Spec} R[X_0, \ldots, X_n] \to R\mathbb{P}^n$$

REMARK. If $R = k$, a field, then $V = \mathbb{A}^{n+1} = \operatorname{Spec} R[X_0, \ldots, X_n]$ is a vector-space (at least if you ignore the subvarieties corresponding to non-maximal prime ideals) and $k\mathbb{P}^n = \mathbb{P}(V)$ is a projective space. We will sometimes want to use the notation $\mathbb{P}(V)$ to emphasize the fact that it is a functor of V (with respect to injective morphisms of vector spaces).

Now we will examine the regular functions on $R\mathbb{P}^n$ as defined above. The image of each q_i consists of rational functions $a(X_i)/b(X_i)$ with the property that

$$\frac{a(r_0, \ldots, r_n)}{b(r_0, \ldots, r_n)} = \frac{a(tr_0, \ldots, tr_n)}{b(tr_0, \ldots, tr_n)}$$

for any $r_0, \ldots, r_n \in R$ and any $t \ne 0 \in R$. This, of course, is compatible with the classic idea (see definition 1.2.2 on page 4) that $R\mathbb{P}^n$ is the quotient of $\mathbb{A}^{n+1} = \operatorname{Spec} R[X_0, \ldots, X_n]$ by the equivalence relation

$$(r_0, \ldots, r_n) \sim (tr_0, \ldots, tr_n)$$

for all nonzero $t \in R$.

We have

PROPOSITION 4.4.9. *All regular functions on $R\mathbb{P}^n$ are elements of R (i.e., constants).*

REMARK. This, incidentally, shows that $R\mathbb{P}^n$ is not isomorphic to *any* affine scheme.

The classical proof of this result shows that regular functions on $R\mathbb{P}^n$ are polynomials $f(X_0, \ldots, X_n)$ with the property that

$$f(r_0, \ldots, r_n) = f(t \cdot r_0, \ldots, t \cdot r_n)$$

for all $r_0, \ldots, r_n, t \in R$ with $t \neq 1$ — so f must be a constant. In scheme theory, the question arises

What do we mean by "$f(r_0, \ldots, r_n)$?"

If we mean, "evaluate f at a prime ideal (as in definition 4.2.3 on page 164)" then a polynomial with *nilpotent* coefficients evaluates to 0, but *still* fails to be a regular function on $R\mathbb{P}^n$.

PROOF. Consider the commutative diagram

(4.4.7)

where $D(X_i)$ is the open set of $\mathbb{A}_R^{n+1}$ that maps to $\mathbb{A}_i^n \subset R\mathbb{P}^n$ and $p_i = \operatorname{Spec} q_i$ where q_i is defined in equation 4.4.5 on the previous page and p is defined in equation 4.4.6 on the preceding page.

Let H be the ring of regular functions on $R\mathbb{P}^n$ and let R_i be that of $\mathbb{A}_i^n$ (as given in equations 4.4.1, 4.4.2 and 4.4.3 on the previous page). If r_i is induced by the inclusion $\mathbb{A}_i^n \hookrightarrow R\mathbb{P}^n$, then

$$r_i \colon H \to R_i$$

is injective for $i = 0, \ldots, n$ because $\mathbb{A}_i^n \cap \mathbb{A}_j^n$ is *dense* in $\mathbb{A}_i^n$ and $\mathbb{A}_j^n$ for any i, j: If the restriction of a regular function to *any* of the $\mathbb{A}_i^n$ vanishes, the function must vanish.

Since the q_i are also injective, we get

$$H = \bigcap_{i=0}^{n} \operatorname{im} q_i \subset R[X_0, \ldots, X_n][X_0^{-1}, \ldots, X_n^{-1}]$$

Functions in this intersection extend to *all* open sets $D(X_i)$, hence to all of $\mathbb{A}_R^{n+1}$, which means they must be in the image of

$$R[X_0, \ldots, X_n] \hookrightarrow R[X_0, \ldots, X_n][X_0^{-1}, \ldots, X_n^{-1}]$$

Equation 4.4.5 on the preceding page shows that the image of q_i is polynomials in X_j/X_i. The only such polynomials that can be in the image of $R[X_0, \ldots, X_n]$ are *constants*. $\qquad\square$

We can also compute the sheaf of meromorphic functions:

EXAMPLE 4.4.10. If R is an integral domain, $\mathcal{K}_{R\mathbb{P}^n}$ (see definition 4.3.21 on page 178) and $\mathcal{K}^*_{R\mathbb{P}^n}$ are the constant sheaves given by

$$\mathcal{K}_{R\mathbb{P}^n} = \underline{F(X_1, \ldots, X_n)}$$
$$\mathcal{K}^*_{R\mathbb{P}^n} = \underline{F(X_1, \ldots, X_n)^\times}$$

where $F(X_1, \ldots, X_n)$ is a field and $F(X_1, \ldots, X_n)^\times = F(X_1, \ldots, X_n) \setminus \{0\}$ is an abelian group (under multiplication) — see definition A.2.1 on page 386.

The simplest approach is to note that $R\mathbb{P}^n$ is irreducible so that each open set $\mathbb{A}^n_i \subset R\mathbb{P}^n$ is dense. In the notation of definition 4.4.8 on page 186, we have

$$\mathcal{K}_{R\mathbb{P}^n}(\mathbb{A}^n_i) = \underline{F(X_{0,i}, \ldots, X_{i-1,i}, X_{i+1,i}, \ldots, X_{n,i})}$$

— since the stalks $\mathcal{O}_{R\mathbb{P}^n, x}$ are integral domains for all $x \in R\mathbb{P}^n$. It follows that the extension to all of $R\mathbb{P}^n$ is also a constant sheaf.

If we want to work this out in *minute* detail, note that the identifications in equations 4.4.3 on page 187 give isomorphisms of fields — where $X_0, \ldots, X_n$ are homogeneous coordinates of $R\mathbb{P}^n$ Now we complete this presheaf to form a sheaf. Let

$$T \subset F(X_0, \ldots, X_n)$$

be the subfield of functions.

$$\frac{f(X_0, \ldots, X_n)}{g(X_0, \ldots, X_n)} \in F(X_0, \ldots, X_n)$$

where f and g are homogeneous of the same degree. Setting $X_{i,j} = X_i/X_j$ induces isomorphisms

$$T \to F(X_{0,i}, \ldots, X_{i-1,i}, X_{i+1,i}, \ldots, X_{n,i})$$

for $i = 0, \ldots, n$. It follows that the *completion process* (see lemma B.2.3 on page 500) for the presheaf *adds* these functions to $\mathcal{K}_{R\mathbb{P}^n}(R\mathbb{P}^n)$ and the *sheaf* is given by

$$\mathcal{K}_{R\mathbb{P}^n}(R\mathbb{P}^n) = T \subset F(X_0, \ldots, X_n)$$

— the subfield generated by quotients of homogeneous polynomials of the same degree (it is a simple exercise to see that these are closed under addition and multiplication). The restriction maps

$$r_i: \mathcal{K}_{R\mathbb{P}^n}(R\mathbb{P}^n) \to \mathcal{K}_{R\mathbb{P}^n}(\mathbb{A}^n_i)$$

are *all* isomorphisms. It follows that $\mathcal{K}_{R\mathbb{P}^n}$ is isomorphic to the constant sheaf $\underline{T}$.

The reader might wonder whether there is a version of the Spec-functor that defines schemes like $R\mathbb{P}^n$ in a compact form. This is indeed the case — the corresponding functor is called Proj. To present it, we need a definition

DEFINITION 4.4.11. Given a graded ring, G, (see section A.4.4 on page 434 $n > 0$, define the nth *irrelevant ideal* to be

$$G_n^+ = \bigoplus_{j=n}^{\infty} G_j$$

Morphisms of graded rings are required to respect the grading.

REMARK. If we give the polynomial algebra $k[X_0, \dots, X_t]$ a grading by setting the grade of each of the X_i to 1, proposition A.4.45 on page 436 implies that $k_H[V]$ of any projective variety, V, is graded.

Now we are ready to define the Proj-construction, introduced by Grothendiek in section 5 of [**64**]:

DEFINITION 4.4.12. If R is a graded ring, then $\operatorname{Proj} R \subset \operatorname{Spec} R$ whose points are prime *homogeneous* ideals that do not contain R_i^+ for any $i > 0$. The topology of $\operatorname{Proj} R$ is induced by that of $\operatorname{Spec} R$. In other words, a basis for the open sets is

$$D(x) = \{\mathfrak{p} \in \operatorname{Proj} R \,|\, x \notin \mathfrak{p}\}$$

for all homogeneous $x \in R$.

REMARK. See theorem 5.1.4 on page 220 for some motivation for the condition that ideals cannot contain R_i^+ for any i.

DEFINITION 4.4.13. If R is a graded ring, $f \in R$ is a homogeneous element, and $r \in R$ is another homogeneous element, then R_f is a graded ring where the degree of

$$\frac{r}{f^t} \in R_f$$

is defined to be $\deg r - t \cdot \deg f$. Define $R_{(f)}$ to be subring consisting of the degree-0 elements of R_f.

If $\mathfrak{p} \subset R$ is a homogeneous prime ideal, we define $R_{(\mathfrak{p})}$ to be degree-0 elements of $S^{-1}R$, where $S = R \setminus \mathfrak{p}$.

With this in mind, we can prove that $\operatorname{Proj} R$ is a scheme:

PROPOSITION 4.4.14. *If R is a graded ring and $x \in R$ is a homogeneous element, then the canonical maps*

$$R \to R_x \leftarrow R_{(x)}$$

induce a homeomorphism

$$D(x) \to \operatorname{Spec} R_{(x)}$$

It follows that $\operatorname{Proj} R$ *is a scheme.*

PROOF. The points of $D(x)$ consist of graded primes that do not contain x, i.e. homogeneous primes of R_x. We define the mapping

$$D(x) \to \operatorname{Spec} R_{(x)}$$

by sending a prime ideal $\mathfrak{p}$ of $D(x)$ to $(\mathfrak{p} \cdot R_x) \cap R_{(x)}$ which is also a prime ideal representing a point of $\operatorname{Spec} R_{(x)}$. We construct an inverse: If $\mathfrak{p} \subset R_{(x)}$ is an ideal, then $\mathfrak{p} \cdot R_x$ is a homogeneous prime of R_x and $(\mathfrak{p} \cdot R_x) \cap R_{(x)} = \mathfrak{p}$.

If $u, v \in R_x$, then $u' = u^{\deg(x)} / x^{\deg(u)}$ and $v' = v^{\deg(x)} / x^{\deg(v)}$ will both be in $R_{(x)}$. If $u \cdot v \in \mathfrak{p} R_x$ then $u' \cdot v' \in \mathfrak{p}$ from which we conclude that $u' \in \mathfrak{p}$ or $v' \in \mathfrak{p}$, leading to the conclusion that $u^{\deg(x)} \in \mathfrak{p} \cdot R_x$ or $v^{\deg(x)} \in \mathfrak{p} \cdot R_x$. It follows that $\sqrt{\mathfrak{p} \cdot R_x} \subset R_x$ is prime and

$$\sqrt{\mathfrak{p} \cdot R_x} \cap R_{(x)} = \mathfrak{p}$$

So there is a 1-1 correspondence between points of $D(x)$ and those of $\operatorname{Spec} R_{(x)}$.

To see that this map is a homeomorphism, note that the image of the open set $D(g)$ is the open set $D(g^{\deg(x)}/x^{\deg(g)})$. The final statement follows from the fact that open sets $D(x_i)$ cover $\operatorname{Proj} R$ so it has a cover by open affines. □

REMARK. If $\mathfrak{a} \subset k[X_0, \ldots, X_n]$ is a homogeneous ideal with

$$R = \frac{k[X_0, \ldots, X_n]}{\mathfrak{a}}$$

the *maximal* homogeneous ideals of $\operatorname{Proj} R$ correspond to geometric points of $\mathcal{P}(\mathfrak{a})$.

Like $\operatorname{Spec} R$, $\operatorname{Proj} R$ has generic points for every irreducible subvariety.

DEFINITION 4.4.15. If R is a graded ring with homogeneous ideal $\mathfrak{a} \subset R$, $\mathcal{P}(\mathfrak{a}) = \{\mathfrak{p} \subset \operatorname{Proj} R | \mathfrak{a} \not\subset \mathfrak{p}\} \subset \operatorname{Proj} R$ is the *subscheme defined by the ideal* $\mathfrak{a}$. The projection $R \to R/\mathfrak{a}$ induces a homeomorphism

$$\mathcal{P}(\mathfrak{a}) \to \operatorname{Proj}(R/\mathfrak{a})$$

REMARK. Compare this to the similar statement for $\operatorname{Spec} R$ in definition 4.3.1 on page 169. It will turn out that *all* closed subschemes of $\operatorname{Proj} R$ arise in this way — see corollary 5.3.19 on page 241.

4.4.2. Tangent cones revisited. This allows a "coordinate free" treatment of the tangent cones defined in section 3.3.1 on page 131. Let $p \in V$ be a point in an affine variety with maximal ideal $\mathfrak{m}$ and local ring $\mathcal{O}_{V,p} = S^{-1}k[V]$ with $S = k[V] \setminus \mathfrak{m}$ (see section 3.2 on page 117). Now consider the associated graded ring (see definition A.4.40 on page 435)

$$G = \mathbf{gr}(\mathcal{O}_{V,p}, \mathfrak{m}) = \frac{\mathcal{O}_{V,p}}{\mathfrak{m}} \oplus \frac{\mathfrak{m}}{\mathfrak{m}^2} \oplus \frac{\mathfrak{m}^2}{\mathfrak{m}^3} \oplus \cdots$$

Since

$$\frac{\mathcal{O}_{V,p}}{\mathfrak{m}} = k$$

this is a *k*-algebra generated by the generators of $\mathfrak{m}$. The scheme $\operatorname{Proj} G$ has points that consist of homogeneous prime ideals of G, i.e., ideals whose generators lie in

$$G_n = \frac{\mathfrak{m}^n}{\mathfrak{m}^{n+1}}$$

for some $n > 0$. The tangent cone is defined to be the affine cone of this projective scheme.

The connection with the tangent cone's *previous* definition (in section 3.3.1 on page 131) is:

> If $k[V] = k[X_1, \ldots, X_t]/\mathfrak{I}$ and $\mathfrak{m} = (X_1, \ldots, X_t)$, then G_n "selects" the degree-*n* terms of $\mathfrak{I}$ (if they exist), or IT($\mathfrak{I}$) and *discards* higher-degree terms (because of $*/\mathfrak{m}^{n+1}$). Since $\mathfrak{m}$ has t generators, G is a finitely generated *k*-algebra with t generators. The "space" that G represents is the algebraic set defined by the kernel of the algebra-homomorphism induced by

(4.4.8) $k[X_1, \ldots, X_t] \to G$

(4.4.9) $X_i \mapsto X_i \in \mathfrak{m}/\mathfrak{m}^2$

If we look at example 3.3.20 on page 131, we have $k[V] = k[X,Y]/(Y^2 - X^2 - X^3)$, $\mathfrak{m} = (X, Y)$, and

$$G = k \oplus (k_X \oplus k_Y) \oplus \frac{\mathfrak{m}^2}{\mathfrak{m}^3} \oplus \cdots$$

where

$$Y^2 - X^2 \equiv Y^2 - X^2 - X^3 \equiv 0 \in \frac{\mathfrak{m}^2}{\mathfrak{m}^3}$$

If we consider the point $(1, \sqrt{2})$ on this curve, we get $\mathfrak{m} = (X - 1, Y - \sqrt{2})$, and the defining relation, $Y^2 - X^2 - X^3$, written in terms of these generators is

$$2\sqrt{2}(Y - \sqrt{2}) + (Y - \sqrt{2})^2 - 5(X - 1) - 4(X - 1)^2 - (X - 1)^3$$

which gives

$$2\sqrt{2}(Y - \sqrt{2}) - 5(X - 1) \equiv 0 \in \frac{\mathfrak{m}}{\mathfrak{m}^2}$$

This definition can be generalized to any noetherian affine scheme. We need the defining ring to be noetherian so that $\mathfrak{m}$ is finitely generated and

$$\bigcap_{i=1}^{\infty} \mathfrak{m}^i = (0)$$

— see lemma A.1.81 on page 379. This ensures that the kernel of the map in 4.4.8 on the previous page represents $IT(\mathfrak{I})$ and not the intersection of powers of $\mathfrak{m}$.

EXERCISES.

1. Show that, if R is a ring, then

$$R\mathbb{P}^n = \text{Proj } R[X_0, \dots, X_n]$$

as schemes, where each of the X_i are defined to have degree 1.

2. Show that $\text{Proj } R = \varnothing$ if and only if all elements of R^+ are nilpotent.

4.4.3. "Varieties" of schemes. Many of the affine schemes in section 4.4.3 can be "globalized" to define types of general schemes. This is true if the property in question is "local" — i.e. the local definition must remain valid for affine schemes.

- If the open affines of V are *reduced* affine schemes, then V is said to be *reduced*. This makes sense because the nilpotent elements of a ring, R, *inject* into all of its localizations so $\text{Spec } R$ is reduced if and only if all of its open affines are reduced.
- If the open affines of V are *normal* affine schemes, then V is said to be normal. This makes sense because normality is a local property — see A.4.13 on page 423. The usual definition states that a scheme $(V, \mathcal{O}_V)$ is normal if every stalk, $\mathcal{O}_{V,x}$, of the structure-sheaf is a normal ring. .
- If the open affines of V are noetherian, V is said to be a *noetherian scheme*.
- If the open affines of V are Jacobson (see page 182), V is called a *Jacobson scheme*. The set of closed points is dense in V and in every closed subscheme.

Varieties are examples of reduced, noetherian schemes. On an irreducible noe-
therian scheme, if f is a global regular function that is not a constant, theo-
rem 2.8.29 on page 109 implies that the solution-set of the equation $f = 0$ is
a subscheme of codimension 1 (since its intersection with all open affines is of
codimension 1).

4.4.4. Quasi-coherent sheaves. Now we will discuss subschemes of a
scheme following Grothendiek's treatment in [**63**]. We need a weaker version
of the concept of coherent module (definition 3.5.5 on page 155):

DEFINITION 4.4.16. Let $(V, \mathcal{O}_V)$ be a ringed space. A module, S, over $\mathcal{O}_V$ is
quasi-coherent if there exists an cover $V = \bigcup U_\alpha$ by open affines such that there
is a right-exact sequence

(4.4.10) $(\mathcal{O}_V|U_\alpha)^{m_\alpha} \to (\mathcal{O}_V|U_\alpha)^{n_\alpha} \to S|U_\alpha \to 0$

for each U_α, where m_α, n_α are integers that may be infinite. It is said to be
coherent if the n_α and m_α are all *finite.*

REMARK. This extends the definition of coherent sheaf (see definition 3.5.5
on page 155) to non-affine varieties and schemes. It is not hard to see that the
sheaf of sections of a vector bundle over a scheme is coherent.

Quasi-coherence is like coherence that is only required to apply over open
sets and for which the finite-generation requirement is dropped.

As remarked earlier, lemma 3.5.6 on page 156 implies that coherent mod-
ules on affine schemes are of the form $\mathcal{A}(M)$ (see definition 3.5.1 on page 155)
for $M = \mathcal{O}_V(V)$ so "global behavior determines local behavior". With coher-
ence or quasi-coherence on *general* schemes, behavior over *neighborhoods* deter-
mines behavior on *smaller* neighborhoods.

The proof of lemma 3.5.6 on page 156 can easily be adapted to show that,
for a coherent or quasi-coherent module over $\mathcal{O}_V$ and each α, $S|U_\alpha = \mathcal{A}(M_\alpha)$
where $M_\alpha = S(U_\alpha)$ is a module over the ring $\mathcal{O}_V(U_\alpha)$. The main distinction
between coherent and quasi-coherent modules is that coherent ones are finitely
generated.

The definition of pullbacks of coherent sheaves (see definition 3.5.13 on
page 159) easily generalizes to quasi-coherent sheaves on quasi-compact
schemes: simply pull back the restrictions to open affines.

The incoherent module in equation 3.5.1 on page 156 *also* fails to be *quasi-
coherent.*

LEMMA 4.4.17. *Let $V = \operatorname{Spec} R$ be an affine scheme and let $\mathscr{M}$ be a quasi-
coherent module over $\mathcal{O}_V$. If $s \in R$, let $D(s)$ be the open set (as in definition 4.2.1 on
page 163).*

  (1) *If $x \in \mathscr{M}(V)$ has the property that $x|D(s) = p^V_{D(r)}(x) = 0$, then $s^n \cdot x = 0$
      for some n (compare with lemma 3.5.7 on page 157).*
  (2) *If $x \in \mathscr{M}(D(s))$, then for some $n \geq 0$, $s^n \cdot x = p^V_{D(r)}(y) = y|D(s)$ for
      some $y \in \mathscr{M}(V)$.*

PROOF. Since V is affine, the $D(s)$ form a basis for its topology and $D(s) =$
$\operatorname{Spec} R_s$ (see proposition 4.3.6 on page 171). It follows that there exists a finite

open cover of V by open affines $D(s_i)$ and that

$$\mathscr{M}|D(s_i) = \mathcal{A}(M_i)$$

for some module M_i over R_{s_i}. If $x \in \mathscr{M}(V)$ then x gives rise to elements $x_i \in M_i$ over $D(s_i)$. If $x|D(s) = 0$, then the image of x_i in $D(s) \cap D(s_i) = D(s \cdot s_i)$ is 0. But

$$M_i = \mathscr{M}(D(s_i)) \to \mathscr{M}(D(s \cdot s_i)) = (M_i)_s$$

is just the localization map. If the image of x_i in $(M_i)_s$ vanishes, then $s^{n_i} \cdot x = 0$ for some s_i (see the remark following definition A.1.89 on page 383). If we set $n = \max(\{n_i\})$. Then $s^n \cdot x_i = 0$ for all i and $s^n \cdot x$ restricts to 0 in all of the $D(s_i)$ — which makes it 0, by the separation-property of a sheaf.

If $x \in \mathscr{M}(D(s))$, then x defines images in each $\mathscr{M}(D(s \cdot s_i)) = (M_i)_s$. It follows that there exists elements $x_i \in M_i$ whose image in $(M_i)_s$ coincides with $s^{n_i} \cdot x$ — set $n = \max(\{n_i\})$. On every overlap $D(s_i) \cap D(s_j) = D(s_i \cdot s_j)$ the elements $x_i = s^n \cdot x$ must agree on $D(s)$ so the differences $x_i - x_i$ define an element of $\mathscr{M}(D(s_i \cdot s_i))$ that vanishes on $D(s) \cap D(s_i) \cap D(s_j)$. The first statement above (applied to the affine sheaf $D(s_i) = \operatorname{Spec} R_{s_i}$) implies that there exists an m_i such that

$$s^{m_i}(x_i - x_j) = 0$$

If we set $m = \max(\{m_i\})$, we get that the $s^m \cdot x_i$ patch together to form a global section — i.e. there exists a $y \in \mathscr{M}(V)$ such that $y|D(s_i) = s^m \cdot x_i = s^{m+n} \cdot s$.   □

This immediately implies that, over affine schemes, quasi-coherent sheaves are essentially coherent:

COROLLARY 4.4.18. *Let $\mathscr{F}$ be a quasi-coherent sheaf over a scheme $V = \operatorname{Spec} R$. Then there exists an R-module M such that*

$$\mathscr{F} = \mathcal{A}(M)$$

— see definition 3.5.1 on page 155. If R is noetherian and $\mathscr{F}(V)$ is finitely generated, $\mathscr{F}$ is coherent.

PROOF. Set $M = \mathscr{F}(V)$ — then the universal property of the $\mathcal{A}(*)$ construction (proposition 3.5.4 on page 155, which also applies to affine schemes) there exists a homomorphism

(4.4.11) $u\colon \mathcal{A}(M) \to \mathscr{F}$

As before, there exists a finite open cover $\{D(s_i)\}$ of V with the property that

$$\mathscr{F}|D(s_i) = \mathcal{A}(M_i)$$

The second statement of lemma 4.4.17 on the previous page implies that $M_i = M_{s_i}$, i.e. that any element $x \in M_i$ has the property that $s_i^n \cdot x \in M$ — so $x = m/s_i^n$ and $u|D(s_i)$ in 4.4.11 is an isomorphism. Since the $D(s_i)$ cover V it follows that u is an isomorphism.   □

We also get a version of lemma 3.5.8 on page 157, regarding the behavior of coherent sheaves:

LEMMA 4.4.19. *If $\mathscr{F}$ is a coherent sheaf on an irreducible noetherian, reduced scheme, V, there exists an open dense set $W \subset V$ such that $\mathscr{F}|W$ is free.*

PROOF. Since V is irreducible, every open set is dense. It suffices to prove the statement for one of the open affines, so we may assume $V = \operatorname{Spec} R$ for some reduced noetherian ring, R. At this point, the proof of lemma 3.5.8 on page 157 goes through unchanged. $\qquad \square$

Indeed, coherent sheaves are free except for proper subsets of the space where they are defined:

PROPOSITION 4.4.20. *Let V be an irreducible noetherian reduced scheme. If $\mathscr{F}$ is a coherent sheaf over V then there exists a coherent sheaf, $\mathscr{G}$ such that $\mathcal{O}_V^r \subset \mathscr{G}$ for some r and a homomorphism*

$$g: \mathscr{F} \to \mathscr{G}$$

such that the supports of the coherent sheaves, $\ker g$ and $\mathscr{G}/\mathcal{O}_V^r$ lie on closed proper subschemes of V.

PROOF. Let W be as in lemma 4.4.19 on the preceding page and let

$$f: \mathscr{F}|W \to \mathcal{O}_V^r|W$$

be the isomorphism constructed there. Now we define $\mathscr{G}$ by

$$\mathscr{G}(U) = (f|U \cap W) \circ p_{U \cap W}^U \mathscr{F}(U) + p_{U \cap W}^U \mathcal{O}_V^r(U)$$

where $p_{U \cap W}^U$ are the restriction maps of their respective sheaves (see definition B.1.1 on page 495). Both summands are subgroups of $\mathcal{O}_V^r(U \cap W)$ and we can define an action of $r \in \mathcal{O}_V(U)$ on $x \in \mathscr{G}(U)$ by $p_{U \cap W}^U(r) \cdot x$ Since both summands are finitely generated modules over $\mathcal{O}_V(U)$, the same is true of the sum. We define the map $g: \mathscr{F} \to \mathscr{G}$ by

$$g = (f|U \cap W) \circ p_{U \cap W}^U: \mathscr{F}(U) \to \mathscr{G}(U)$$

We claim that $\mathcal{O}_V(U)^r$ is a subsheaf of $\mathscr{G}(U)$. This follows from the fact that the maps $p_{U \cap W}^U$ are injective — which follows from the fact that $\mathcal{O}_V(U)$ is an integral domain (because V is irreducible). Furthermore $g|W$ is an isomorphism to $\mathcal{O}_V^r|W$ so that $\ker g|W = 0$ and $\mathscr{G}/\mathcal{O}_V^r|W = 0$ so their supports are on proper subschemes of $V \setminus W$. $\quad \square$

4.4.5. Subschemes. Now we can define a closed subscheme of a scheme — this is a direct generalization of definition 4.3.1 on page 169 with some added complexity due to a scheme being patched together from affine schemes.

DEFINITION 4.4.21. Let $(V, \mathcal{O}_V)$ be a scheme and let $\mathfrak{I} \subset \mathcal{O}_V$ be a quasi-coherent subsheaf of ideals. Suppose

$$V = \bigcup U_\alpha$$

be an open cover such that each U_α is an open affine of V, and $\mathfrak{I}|U_\alpha$ satisfies an equation like 4.4.10 on page 193. Then the *closed subscheme*, S, *defined by* $\mathfrak{I}$ is the union of the closed subschemes

$$Z_\alpha \subset U_\alpha$$

defined by the ideals $\mathfrak{I}(U_\alpha) \subset \mathcal{O}_V(U_\alpha)$ (see definition 4.3.1 on page 169) with structure sheaf $\mathcal{O}_S = \mathcal{O}_V/\mathfrak{I}$.

REMARK. We have merely provided a global construct (ideal *sheaf*) that "localizes" to ideals over local coordinate rings. The sheaf conditions (see definition B.1.1 on page 495) ensure that the $\{Z_\alpha\}$ all patch together properly.

It is reassuring that this is consistent with definition 4.3.1 on page 169:

PROPOSITION 4.4.22. *If $V \subset W = \operatorname{Spec} R$ is a closed subscheme of an affine scheme, then $V = \operatorname{Spec} S$ is affine and the canonical map*

$$f: R \to S$$

is surjective.

PROOF. If $\mathscr{I} \subset \mathcal{O}_W$ is a quasi-coherent ideal defining V, then corollary 4.4.18 on page 194 implies that

$$\mathscr{I} = \mathcal{A}(\mathfrak{J})$$

(see definition 3.5.1 on page 155) for an ideal $\mathfrak{J} \subset R$, $V = W(\mathfrak{J})$, and $S = R/\mathfrak{J}$. $\square$

Open subschemes are easier to define because open sets are "large":

> If $(V, \mathcal{O}_V)$ is a scheme and $U \subset V$ is an open set, then $(U, \mathcal{O}_V|U)$ is an open subscheme of $(V, \mathcal{O}_V)$.

We come to general subschemes of a scheme:

DEFINITION. A sub-ringed space $(Y, \mathcal{O}_Y)$ of a scheme $(X, \mathcal{O}_X)$ is a *subscheme* if:

(1) It is locally closed in X,
(2) If U is the largest open set in X such that $Y \subset U$ is closed (for instance, $U = X \setminus (\bar{Y} \setminus Y)$, where $\bar{Y}$ is the closure of Y in X), then $(Y, \mathcal{O}_Y)$ is the closed subscheme of $(U, \mathcal{O}_X|U)$ defined by a quasi-coherent sheaf of ideals $\mathfrak{J} \subset \mathcal{O}_U = \mathcal{O}_X|U$.

REMARK. Note that if Y is closed, $U = X$. Given a subscheme $(Y, \mathcal{O}_Y)$ of a scheme $(X, \mathcal{O}_X)$, we have a canonical inclusion $j: Y \hookrightarrow X$ and $j^*(\mathcal{O}_X) = \mathcal{O}_X|Y$. This comes with a canonical map

$$j^*(\mathcal{O}_X) \to \mathcal{O}_X/\mathfrak{J}$$

which makes j a morphism of ringed spaces (see definition B.3.1 on page 502).

Note that our definition of subscheme is transitive: subschemes of subschemes are subschemes.

Now that we know what a subscheme is, we can define immersions

DEFINITION 4.4.23. A morphism of schemes $f: (V, \mathcal{O}_V) \to (W, \mathcal{O}_W)$ is an *immersion* if there exists a subscheme $(X, \mathcal{O}_X) \subset (W, \mathcal{O}_W)$ and an isomorphism $g: (V, \mathcal{O}_V) \to (X, \mathcal{O}_X)$ that makes the diagram

$$(V, \mathcal{O}_V) \xrightarrow{f} (W, \mathcal{O}_W)$$
$$\searrow_{g} \qquad \uparrow_{j}$$
$$(X, \mathcal{O}_X)$$

commute, where $j: (X, \mathcal{O}_X) \to (W, \mathcal{O}_W)$ is the canonical inclusion.

REMARK. Again, we point out that this use of the word "immersion" is peculiar to algebraic geometry. In differential geometry, this would be called an embedding and the term "immersion" (which also *exists* in differential geometry) means something completely different.

EXERCISES.

3. If $(Y, \mathcal{O}_Y)$ and $(X, \mathcal{O}_X)$ are ringed spaces with $Y \subset X$, and there exists an open cover

$$X = \bigcup U_\alpha$$

such that $(Y \cap U_\alpha, \mathcal{O}_Y | Y \cap U_\alpha)$ is a closed subscheme of $(U, \mathcal{O}_X | U)$, show that $(Y, \mathcal{O}_Y)$ is a (closed) subscheme of $(X, \mathcal{O}_X)$.

4. If R is an integral domain, compute the sheaf of regular functions on $R\mathbb{P}^n$ in a manner similar to that of example 4.3.15 on page 174.

5. If $\mathscr{F}$ is a quasi-coherent sheaf on a scheme, V, and $U \subset V$ is an affine open set, show that

$$i_*(\mathscr{F}|U)$$

is also quasi-coherent, where i_* is the direct image sheaf (see exercise 3 on page 498).

6. Show that, if $f: X \to Y$ is a closed immersions of schemes and $\mathscr{F}$ is a coherent sheaf on X, then $f_*\mathscr{F}$ is quasi-coherent on Y.

7. if X is a noetherian scheme, its *Grothendiek group* — denoted $K(X)$ — is defined to be the free abelian group generated by coherent sheaves $\mathscr{F}$ defined on X subject to the relations

$$[\mathscr{F}_2] = [\mathscr{F}_1] + [\mathscr{F}_3]$$

whenever there exists a short exact sequence

$$0 \to \mathscr{F}_1 \to \mathscr{F}_2 \to \mathscr{F}_3 \to 0$$

Here $[\mathscr{F}]$ denotes the element of $K(X)$ defined by a coherent sheaf $\mathscr{F}$.
    Show that $K(\mathbb{A}^1) = \mathbb{Z}$.

8. Let $V = \operatorname{Spec} R$ be an affine scheme and let $W \subset V$ be a closed subscheme defined by $s = 0$ for some $s \in R$. if $\mathscr{M}$ is a coherent sheaf on V whose support lies in W show that there exists a filtration

$$0 \subset \mathscr{M}_1 \subset \cdots \subset \mathscr{M}_n = \mathscr{M}$$

such that

$$\frac{\mathscr{M}_{i+1}}{\mathscr{M}_i}$$

is an $\mathcal{O}_W$-module for all i.

4.5. Products

4.5.1. Construction by patching. Now we define products of general varieties and schemes. Products of affine schemes were defined in section 4.3.3 on page 178 and were expressed in terms of tensor products, in definition 4.3.23 on page 178. We would like to have a category-theoretic product like that for affine schemes, but it is not obvious that such objects exist. Unfortunately, we have no algebraic functor like the coordinate ring that faithfully models all of a scheme's behavior.

We will work with schemes *over* a fixed scheme, T — see definition 4.4.4 on page 183 — and follow Grothendiek's original proof of the existence of products of schemes in [**63**]. This involves a somewhat cumbersome process of patching together products of *open affines.*

We first consider what properties a product would have *if* it existed.

LEMMA 4.5.1. *Suppose* $T' \subset T$ *is a open affine subscheme and X and Y are schemes over T'. Then*

$$X \times_{T'} Y = X \times_T Y$$

if they exist.

REMARK. Note that X and Y are also schemes over T. This clarifies the role of T in the product: it "selects" acceptable morphisms to X and Y.

PROOF. There exist canonical maps

$$X \xrightarrow{\alpha} T' \hookrightarrow T$$
$$Y \xrightarrow{\beta} T' \hookrightarrow T$$

Suppose

$$f\colon A \to X$$
$$g\colon A \to Y$$

are T-morphisms, i.e., they make the diagram

$$
\begin{array}{ccc}
A & \xrightarrow{f} & X \\
\downarrow{\scriptstyle g} & & \downarrow \\
Y & \longrightarrow & T
\end{array}
$$

commute. Since the images of X and Y actually lie within T' and T' injects into T, the diagram,

$$
\begin{array}{ccc}
A & \xrightarrow{f} & X \\
\downarrow{\scriptstyle g} & & \downarrow \\
Y & \longrightarrow & T'
\end{array}
$$

also commutes and f and g automatically become T'-morphisms as well. It follows that both products will have the same universal category-theoretic properties, so they will be isomorphic (if either of them exists). □

Next, we show that the property of being a product is inherited by open sets:

LEMMA 4.5.2. *Suppose X and Y are schemes over T and suppose Z is a fibered product. In other words suppose there exist T-morphisms*

$$Z \xrightarrow{p_1} X$$
$$Z \xrightarrow{p_2} Y$$

and whenever there are T-morphisms

$$f\colon A \to X$$
$$g\colon A \to Y$$

that make the solid arrows in the diagram

commute, there exists a unique T-morphism $A \to Z$ represented by the dotted arrow, that makes the whole diagram commute.

If $U \subset A$, and $V \subset B$ are open sets, then

$$p_1^{-1}(U) \cap p_2^{-1}(V) = U \times_T V$$

PROOF. This is just a matter of verifying that $p_1^{-1}(U) \cap p_2^{-1}(V)$ has the same universal property as $U \times_T V$: Any object A with T-morphisms

$$u\colon A \to U$$
$$v\colon A \to V$$

will also have T-morphisms to X and Y by composing with the inclusions. Consequently, A there will exist a *unique* T-morphism to Z and its image will lie in $p_1^{-1}(U) \cap p_2^{-1}(V)$. $\qquad\square$

Next, we prove a kind of converse: if a scheme "locally behaves" like a product, then it *is* a product.

LEMMA 4.5.3. *Let X, Y, Z be schemes over T, with T-morphisms*

$$Z \xrightarrow{p_1} X$$
$$Z \xrightarrow{p_2} Y$$

and suppose

$$X = \bigcup_i U_i$$
$$Y = \bigcup_j V_j$$

are open covers. If

$$p_1^{-1}(U_i) \cap p_2^{-1}(V_j) = U_i \times_T V_j \subset Z$$

for all i, j, then

$$Z = X \times_T Y$$

PROOF. Suppose we have T-morphisms

$$\begin{aligned} f \colon A &\rightarrow X \\ g \colon A &\rightarrow Y \end{aligned}$$

that make the diagram

and set

$$Z_{i,j} = p_1^{-1}(U_i) \cap p_2^{-1}(V_j) \subset Z$$

$$W_{i,j} = f^{-1}(U_i) \cap g^{-1}(V_j) \subset A$$

for all i, j. Then

$$A = \bigcup_{i,j} W_{i,j}$$

is an open covering with $f(W_{i,j}) \subset U_i$ and $g(W_{i,j}) \subset V_j$. It follows that there exists a *unique T-morphism*

$$\varphi_{i,j} \colon W_{i,j} \rightarrow U_i \times_T V_j = Z_{i,j} = p_1^{-1}(U_i) \cap p_2^{-1}(V_j)$$

that makes the diagram

commute.

We claim that the $\varphi_{i,j}$ are compatible where the $W_{i,j}$ overlap: If $W \subset W_{i,j} \cap W_{i',j'}$ is an open set, then lemma 4.5.2 on the preceding page implies that

$$Z_{i,j} \cap Z_{i',j'} = (U_i \cap U_{i'}) \times_T (V_j \cap V_{j'})$$

so that there exists a *unique T-morphism*

$$W \rightarrow Z_{i,j} \cap Z_{i',j'} \subset Z$$

Its uniqueness implies that it must coincide with the morphisms that result from including $W \subset W_{i,j}$ or $W \subset W_{i',j'}$ and then mapping via $\varphi_{i,j}$ or $\varphi_{i',j'}$, respectively.

It follows that the $\{\varphi_{i,j}\}$ patch together to define a T-morphism

$$\Phi \colon A \rightarrow Z$$

that makes the diagram

commute. It is unique because its restriction to each of the $W_{i,j}$ is unique.   □

Now we are ready to actually construct products. We do this in two steps:

LEMMA 4.5.4. *If T is an affine scheme and X, Y are schemes over T, then the product*

$$X \times_T Y$$

exists.

PROOF. Suppose

$$X = \bigcup_i U_i$$

$$Y = \bigcup_j V_j$$

are decompositions into unions of open affines. There are gluing maps connecting the U_i to each other and the V_j to each other in such a way that the application of corollary 4.4.5 on page 184 would reconstruct X and Y. Furthermore, the U_i and V_j are affine schemes over T — they inherit this from X and Y.

So we construct all of the products

$$U_i \times_T V_j$$

and patch them together via corollary 4.4.5 on page 184. It is a tedious exercise that all of the induced patching maps of products are compatible with each other. The result is a scheme, Z, that satisfies the hypotheses of lemma 4.5.3 on page 199 so that

$$Z = X \times_T Y$$

□

We are finally ready to present Grothendiek's construction:

THEOREM 4.5.5. *Let T be a scheme and let X and Y be schemes over T. Then the (fibered) product*

$$X \times_T Y$$

exists and has the universal property:
    Given any morphisms of schemes over T

$$f: U \quad \to \quad X$$
$$g: U \quad \to \quad Y$$

making the solid arrows in the diagram

(4.5.1)

commute, there exists a unique dotted arrow making the whole diagram commute.

PROOF. If T is an affine scheme, we are done by the previous result. Consequently, suppose that T is not affine.

Since X and Y are schemes over T, there exist canonical maps

$$X \xrightarrow{\alpha} T$$
$$Y \xrightarrow{\beta} T$$

If

$$T = \bigcup_i T_i$$

is a decomposition into open affines, set

$$X_i = \alpha^{-1}(T_i)$$
$$Y_i = \beta^{-1}(T_i)$$

and form the products

$$X_i \times_{T_i} Y_i$$

for all i, using lemma 4.5.4 on the preceding page. Lemma 4.5.1 on page 198 implies that

$$X_i \times_{T_i} Y_i = X_i \times_T Y_i$$

We finish the construction by gluing together all of the $X_i \times_T Y_i$ via corollary 4.4.5 on page 184.   □

REMARK. The fiber over a point of a morphism of *varieties* over an algebraically closed field, k, is

$$V \times_W \operatorname{Spec} k$$

4.5.2. Fibers of maps revisited. Section 2.5.1 on page 75 studied fibers of maps in a very limited setting. For instance, it did not consider fibers over open sets or general subvarieties.

PROPOSITION 4.5.6. *If $j: U \hookrightarrow V$ is the inclusion of any subscheme, and $f: W \to V$ is a regular map, the fiber of f over U is the fibered product*

$$f^{-1}(U) = U \times_V W$$

PROOF. This is exactly like the solution to exercise 9 on page 181: We consider the universal properties that $f^{-1}(U)$ and the fibered product share. If X is *any* scheme and we have maps

$$\begin{aligned} a\colon X &\to U \\ b\colon X &\to W \end{aligned}$$

that make a diagram like equation 4.5.1 on the preceding page commute, then the image of b must map to U under f and $\operatorname{im} b \subset f^{-1}(U)$. It follows that the regular maps a and b induce a *unique* map $X \to f^{-1}(U)$. This is the universal property of the fibered product. □

Definition 4.3.23 on page 178 and the discussion following it implies that

COROLLARY 4.5.7. *If* $\operatorname{Spec} T \hookrightarrow \operatorname{Spec} S$ *is the inclusion of a subscheme and* $f\colon \operatorname{Spec} R \to \operatorname{Spec} S$ *is a regular map then*

$$f^{-1}(\operatorname{Spec} T) = \operatorname{Spec}(T \otimes_S R)$$

In analogy with lemma 2.5.8 on page 75, we have

LEMMA 4.5.8. *Let* $f\colon \operatorname{Spec} S \to \operatorname{Spec} R$ *is a morphism of affine schemes and let* $X \subset \operatorname{Spec} R$ *be a closed subscheme corresponding to the ideal* $\mathfrak{r}$. *Then* $f^{-1}(X)$ *is the closed subscheme of* $\operatorname{Spec} S$ *corresponding to the ideal*

$$f^*(\mathfrak{r}) \cdot S \subset S$$

and is the affine scheme

$$Z = \operatorname{Spec}\left(\frac{S}{f^*(\mathfrak{r}) \cdot S}\right)$$

REMARK. In many cases the *geometric* fiber, defined in section 4.6.4 on page 215 is more geometrically relevant than this.

PROOF. Simply note that X is defined by ideal $\mathfrak{r}$, and its coordinate ring is $R/\mathfrak{r}$, and

$$S \otimes_R \left(\frac{R}{\mathfrak{r}}\right) = \frac{S}{f^*(\mathfrak{r}) \cdot S}$$

□

EXERCISES.

1. Compute the fibers of the maps:
   a. $\operatorname{Spec} \mathbb{C} \to \operatorname{Spec} \mathbb{Z}$
   b. $\operatorname{Spec} \mathbb{Z}[X] \to \operatorname{Spec} \mathbb{Z}$

4.5.3. Vector-bundles revisited. We begin this section by asking the question:

Is the total space of a vector-bundle over an affine scheme *also* an affine scheme? And, if so, what is its coordinate ring?

The following result answers both questions:

THEOREM 4.5.9. *Let ξ be a vector-bundle over an affine scheme $V = \operatorname{Spec} R$ corresponding to the projective module, M, over R and let $M^* = \hom_R(M, R)$ be the dual (see definition A.5.64 on page 472). If the fibers of ξ are vector-spaces over the field, k, assume that V is a scheme over k as well.*

Then the total space of ξ is the affine scheme $\operatorname{Spec} \mathcal{S}(M^*)$ — defined by the symmetric algebra, $\mathcal{S}(M^*)$ (see definition A.6.4 on page 475). The morphism

$$f \colon \operatorname{Spec} \mathcal{S}(M^*) \to \operatorname{Spec} R$$

corresponding to the inclusion

$$R = \mathcal{S}(M^*)_0 \hookrightarrow \mathcal{S}M^*$$

is the projection from the total space to the base.

PROOF. Since V is a scheme over k, we have a homomorphism of rings $k \to R$ making R an algebra over k. Assume the rank of ξ is t. A global section of $f \colon \operatorname{Spec} \mathcal{S}(M^*) \to \operatorname{Spec} R$ corresponds (uniquely!) to an algebra homomorphism

$$\mathcal{S}(M^*) \to R$$

that splits the inclusion $R \to \mathcal{S}(M^*)$. Proposition A.6.6 on page 476 shows that such homomorphisms are in a 1-1 correspondence with module-homomorphisms

$$M^* \to R$$

— i.e., elements of $\hom_R(M^*, R) = M^{**} = M$. So — theorem 3.5.9 on page 157 implies that, *if* f is a vector-bundle it is isomorphic to ξ. So see that it *is* a vector-bundle, we must show that its sheaf of sections is *coherent* — i.e., is of the form $\mathcal{A}(M)$ (see proposition C.2.5 on page 516 and lemma 3.5.6 on page 156).

If $\{s_1, \ldots, s_t\} \in R$ are a set elements, $S \subset R$ is the multiplicative set generated by the s_i and

$$U = \bigcap_{i=1}^{t} D(s_i) \subset \operatorname{Spec} R$$

is an open set (see definition 4.2.1 on page 163 and proposition 4.3.6 on page 171), then $U = \operatorname{Spec} S^{-1}R$ and the fiber

$$f^{-1}(U) = \operatorname{Spec}(\mathcal{S}M^* \otimes_R S^{-1}R) = \operatorname{Spec}(\mathcal{S}(M^* \otimes_R S^{-1}R))$$

(by proposition A.6.10 on page 478). As before, the module of sections of this is in a 1-1 correspondence with homomorphisms

$$\mathcal{S}(M^* \otimes_R S^{-1}R) \to S^{-1}R$$

i.e., elements of

$$\hom_{S^{-1}R}(M^* \otimes_R S^{-1}R, S^{-1}R)$$

Consider the natural (see definition A.5.10 on page 444) map

$$M \otimes_R S^{-1}R \quad \to \quad \hom_{S^{-1}R}(M^* \otimes_R S^{-1}R, S^{-1}R)$$
$$m \otimes s^{-1} \quad \mapsto \quad (\mu \otimes t^{-1} \mapsto \mu(m) \cdot s^{-1}t^{-1})$$

for $\mu \in M^*$ and $s, t \in S$. We claim that this is an isomorphism. To see that note that:

(1) it is (trivially) true if M is *free*
(2) all of the constructions in this expression (i.e. $\hom_R$ and $\otimes_R$) preserve finite direct sums,
(3) projective modules are direct summands of free modules.

The conclusion follows. $\qquad\square$

4.6. Varieties and separated schemes

One might think that, since schemes are the result of gluing together affine schemes, they would automatically satisfy a separation condition like that in lemma 4.3.25 on page 179. Unfortunately, there are clever ways (see example 4.4.7 on page 186) to glues affines together to violate this condition, so we must impose it independently.

General varieties and schemes are *not* determined by their coordinate rings, as example 4.4.8 on page 186 shows. One needs the full structure of a ringed space to define them.

DEFINITION 4.6.1. Let X be a scheme over another scheme, T. Then T is a *separated scheme* if the diagonal map

$$\Delta \colon T \to T \times_{\mathbb{Z}} T$$

is a closed immersion. If this is true, then X is a *separated scheme over T* if the diagonal map

$$\Delta_T \colon X \to X \times_T X$$

is a closed immersion. If X is a prevariety and

$$\Delta_k \colon X \to X \times_k X$$

is a closed immersion, then X is called a *variety*.

EXAMPLE 4.6.2. Compare exercise 7 on page 90. The open subspace $U = \mathbb{A}^2 \setminus \{(0,0)\} \subset \mathbb{A}^2$ becomes a variety with the sheaf $\mathcal{O}_{\mathbb{A}^2}|U$. It *cannot* be affine since its coordinate ring is $B = k[X, Y]$ and $\operatorname{Specm} B = \mathbb{A}^2 \neq U$.

The reader may wonder whether there exist algebraic ways of distinguishing U in the example above and $\mathbb{A}^2$ using the sheaf of regular functions. The answer is a resounding yes, but involves more advanced methods — see example D.3.19 on page 555

EXAMPLE 4.6.3. In example 4.4.7 on page 186, define a function on V 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

$$f_i \colon \mathbb{A}^1 = U, V \hookrightarrow V$$

agree exactly on $\mathbb{A}^1 \setminus \{0\}$ which is *not* closed in $\mathbb{A}^1$.

The following example shows what can go wrong on a non-separated scheme:

EXAMPLE 4.6.4. This is a two-dimensional version of example 4.4.7 on page 186. Simply take two copies of $\mathbb{A}^2$ and identify them everywhere except at the origin:

$$V = \mathbb{A}_1^2 \cup \mathbb{A}_2^2 / \sim$$

where $\sim$ identifies $(x,y) \in \mathbb{A}_1^2$ with $(x,y) \in \mathbb{A}_2^2$ unless $(x,y) = (0,0)$. This is just $\mathbb{A}^2$ with a doubled origin: $(0,0)_1, (0,0)_2$. Then $\mathbb{A}_1^2, \mathbb{A}_2^2 \subset V$ are affine open sets since

$$\mathbb{A}_i^2 = V \setminus (0,0)_{3-i}$$

but

$$\mathbb{A}_1^2 \cap \mathbb{A}_2^2 = \mathbb{A}^2 \setminus (0,0)$$

which is open but *not* affine — see exercise 7 on page 90. Exercise 1 on page 210 shows that this never happens on a separated scheme.

One important consequence of separation properties is that the graph of a regular map is closed:

DEFINITION 4.6.5. Let $f\colon V \to W$ be a morphism of schemes. Define the *graph* of f, denoted Γ_f to be the fiber (see section 4.5.2 on page 202) of $\Delta(W)$ under the morphism

$$(f \times \mathrm{id})\colon V \times W \to W \times W$$

REMARK. The graph, Γ_f, is also the image of $(1, f)\colon V \to V \times W$.

The continuity (in the Zarisky sense) of a regular map immediately implies:

PROPOSITION 4.6.6. *If W is separated or a variety, then the graph of any morphism $f\colon V \to W$ is a closed subset of $V \times W$.*

Corollary 4.3.20 on page 177 characterizes affine varieties as types of affine schemes. We can generalize this to general varieties. We need the definition:

DEFINITION 4.6.7. If $(S, \mathcal{O}_S)$ is a scheme over an algebraically closed field k, S is said to be locally

- of *finite type*, for each open affine $(U, \mathcal{O}_S|U)$, the ring $\mathcal{O}_S(U)$ is a finitely generated k-algebra.
- *reduced*, for each open affine $(U, \mathcal{O}_S|U)$, the ring $\mathcal{O}_S(U)$ is a reduced k-algebra.

The following result shows that varieties are simply a type of scheme:

THEOREM 4.6.8. *Let $\mathscr{V}_k$ denote the category of varieties over an algebraically closed field, k, and let $\mathscr{RS}_k$ denote the category of locally reduced, separated schemes over k that are locally of finite type. Then there is an equivalence of categories*

$$\mathscr{V}_k \leftrightarrow \mathscr{RS}_k$$

PROOF. Every variety is a scheme and every morphism of varieties defines a unique morphism of schemes. To go in the other direction, simply map a scheme in $\mathscr{RS}_k$ to its subset of closed points: the result is a separated union of affine varieties, hence a variety. Every morphism of schemes induces morphisms of open affines which induce morphisms of open affine varieties, hence it induces a morphism of varieties. $\square$

The following result addresses an issue raised earlier (in theorem 2.5.27 on page 86):

COROLLARY 4.6.9. *If $U \subset V$ is an open set in an algebraic variety, then U is also an algebraic variety.*

PROOF. The open set, U, is a subscheme of V (see definition 4.4.5 on page 196) that also satisfies all of the requirements in theorem 4.6.8 on the facing page for being a variety. $\square$

Theorem 2.5.27 on page 86, applied to open affines, implies that:

PROPOSITION 4.6.10. *Two varieties X and Y are birationally equivalent if and only if there exist dense open subsets $U \subset X$ and $V \subset Y$ that are isomorphic.*

We can define familiar concepts for general varieties:

DEFINITION 4.6.11. Given an irreducible variety, V, define
 (1) The *ring of regular functions, $k[V]$* is defined to be functions $f : V \to k$ whose restriction to any open affine set $U \subset V$ is an element of $k[U]$.
 (2) The *field of rational functions, $k(V)$*, is defined to be equivalence classes of pairs (f, U), where $U \subset V$ is an open affine and $f \in k[U]$, where (f_1, U_1) is equivalent to (f_2, U_2) if $f_1|U_1 \cap U_2 = f_2|U_1 \cap U_2$.

REMARK. Since V is irreducible, all pairs of open affines have nonempty intersections.

Note that $U_1 \cap U_2$ is also an affine variety (see exercise 1 on page 210).

Also note that $k(V)$ is a field: one can add, multiply elements, and if $(f, U) \in k(V)$, the set of points where f is nonzero is an open affine $U' \subset U$, so (f^{-1}, U') is well-defined.

For *non-affine* varieties, $k(V)$ is *not* the field of fractions of $k[V]$. For instance, $k[k\mathbb{P}^n] = k$ and $k(k\mathbb{P}^n) = k(X_1, \dots, X_n)$.

The following result is key to understanding the properties of rational function fields:

PROPOSITION 4.6.12. *If V is an irreducible variety with an open affine $U \subset V$, then the inclusion*

$$U \hookrightarrow V$$

induces an isomorphism $k(V) \to k(U)$.

REMARK. Recall that $k(U)$ is the field of fractions of $k[U]$.

PROOF. Every element (f, U') restricts to an element $(f, U' \cap U)$ of $k(U)$. Conversely, every element $f/g \in k(U)$ gives rise to a regular function $(f/g|U', U')$, where $U' \subset U$ is the open set where $g \neq 0$, and so, induces an element of $k(V)$. $\square$

The local definition of normal varieties, 3.4.1 on page 145, immediately extends to general varieties, in which case the normalization of affine varieties in section 3.4.2 on page 150 also extends to general varieties:

THEOREM 4.6.13. *If V is a variety, there exists a normal variety, V^ν, called the normalization of V and a finite map*

$$n(V): V^\nu \to V$$

that is a birational equivalence.

PROOF. Simply apply theorem 3.4.12 on page 152 to open affines. The functoriality of the construction implies that the normalizations patch together to form a normal general variety.   □

This normalization also has the universal property listed in theorem 3.4.12 on page 152:

THEOREM 4.6.14. *If W is any normal variety and*

$$r: W \to V$$

is a regular map, there exists a unique regular map $\bar{r}: W \to V^\nu$ that makes the diagram

commute.

PROOF. Let V be covered by open affines $\{U_\alpha\}$, such that $f^{-1}(U_\alpha)$ is an open affine of W (see definition 4.4.3 on page 183 and the remark following it). The conclusion follows from the commutativity of the diagrams

Exercise 1 on page 210 implies that $U_\alpha \cap U_\beta$ and $f^{-1}(U_\alpha) \cap f^{-1}(U_\beta)$ are also affine and the properties of normalization in theorem 3.4.12 on page 152 imply that $(U_\alpha \cap U_\beta)^\nu = U_\alpha^\nu \cap U_\beta^\nu$. The uniqueness of the maps $\bar{r}_\alpha$ implies that they patch together on the intersections $f^{-1}(U_\alpha) \cap f^{-1}(U_\beta)$ to define a global map.
                                                                       □

4.6.1. Maps to affine schemes. We defined morphisms of schemes in definition 4.4.3 on page 183, and can simplify this in the special case where the target of the map is affine. Although the coordinate ring of a general variety doesn't determine its geometric properties, it does determine how it can map to an affine variety:

PROPOSITION 4.6.15. *For a scheme V and a ring, R, there is a canonical one-to-one correspondence*

$$\mathrm{hom}_{\mathrm{Scheme}}(V, \mathrm{Spec}\, R) \longleftrightarrow \mathrm{hom}(R, \mathcal{O}_V(V))$$

REMARK. A corresponding result is true for varieties.

PROOF. Let $(V, \mathcal{O}_V)$ be a scheme and let $\beta: R \to \mathcal{O}_V(V)$ be a homomorphism. If $p \in V$ is a point, let $U \subset V$ be an open affine containing p. Then restriction defines a homomorphism

$$\mathcal{O}_V(V) \to \mathcal{O}_V(U)$$

and composing this with β gives a homomorphism

$$R \to \mathcal{O}_V(U)$$

which defines a map

$$U \to \operatorname{Spec} R$$

The compatibility conditions of a sheaf imply that these maps patch together to give a map

$$V \to \operatorname{Spec} R$$

Conversely, from a regular map $\varphi: V \to \operatorname{Spec} R$ we get a homomorphism $f_\alpha \mapsto f \circ \varphi: R \to \mathcal{O}_V(U_\alpha)$ for any open affine U_α. The compatibility conditions on overlaps of open affines imply that

$$\begin{array}{ccc}
 & \mathcal{O}_V(U_\alpha) & \\
\overset{f_\alpha}{\nearrow} & & \searrow \\
R & & \mathcal{O}_V(U_\alpha \cap U_\beta) \\
\underset{f_\beta}{\searrow} & & \nearrow \\
 & \mathcal{O}_V(U_\beta) &
\end{array}$$

commutes for all α, β. The sheaf condition 5 of definition B.1.1 on page 495 implies that every such family of maps $\{f_\alpha\}$ are the restrictions of a map

$$R \to \mathcal{O}_V(V)$$

$\square$

This allows us to define a universal affine version of a scheme:

DEFINITION 4.6.16. Let $(V, \mathcal{O}_V)$ be a scheme. The identity map $1: \mathcal{O}_V(V) \to \mathcal{O}_V(V)$ corresponds to a morphism

$$a_V: V \to \operatorname{Spec} \mathcal{O}_V(V)$$

where we call $\operatorname{Spec} \mathcal{O}_V(V)$ the *affine image* of V.

Since every homomorphism of rings factors through the identity map, we get

PROPOSITION 4.6.17. *If $f: (V, \mathcal{O}_V) \to A$ is a morphism from a scheme to an affine scheme, then there exists a map $f': \operatorname{Spec} \mathcal{O}_V(V) \to A$ that makes the diagram*

$$\begin{array}{ccc}
 & \operatorname{Spec} \mathcal{O}_V(V) & \\
\overset{a_V}{\nearrow} & & \downarrow f' \\
V & \xrightarrow{\ f\ } & A
\end{array}$$

commute.

For instance:

EXAMPLE. The coordinate ring of the projective space, $k\mathbb{P}^n$, consists of constants (see example 4.4.8 on page 186), so its affine image is $\operatorname{Spec} k$, which is a single (closed) point. This implies that morphisms $f\colon k\mathbb{P}^n \to V$, where V is affine, send $k\mathbb{P}^n$ to a *single point*.

EXERCISES.

1. If S is a separated scheme with two affine open sets A_1 and A_2, show that $A_1 \cap A_2$ is also affine. This suggests what can go "wrong" in non-separated schemes.

2. Give an example of a non-separated scheme N with two affine open sets A_1 and A_2 such that $A_1 \cap A_2$ is not affine.

4.6.2. Differential forms. In this section, we will explore a subject touched upon in section A.7 on page 482, namely that of *differentials* on a variety. Proposition 3.3.4 on page 121 shows that, on an affine variety, the set of differentials is the dual to the tangent space at a point. We will assume that all varieties are smooth. For a glimpse of what happens in the singular case, see section 4.4.2 on page 191.

DEFINITION 4.6.18. If $f\colon V = \operatorname{Spec} R \to W = \operatorname{Spec} S$ is a morphism of irreducible affine schemes, define the *sheaf of relative differentials* via:

$$\Omega_{V/W} = \mathcal{A}(\Omega_{R/S})$$

where $\Omega_{R/S}$ is the R-module defined in section A.7 on page 482, and given by

$$\Omega_{R/S} = \mathfrak{K}/\mathfrak{K}^2$$

where $\mathfrak{K} = \ker *\colon R \otimes_S R \to R$ (see theorem A.7.10 on page 486).

If $f\colon V \to W$ is a morphism of schemes, cover V with open affines $\{A_\alpha = \operatorname{Spec} R_\alpha\}$ and W with $\{B_\alpha = \operatorname{Spec} S_\alpha\}$ such that $f(A_\alpha) \subset B_\alpha$ and define

$$\Omega_{V/W}|A_\alpha = \mathcal{A}(\Omega_{R_\alpha/S_\alpha})$$

These sheaves patch together to form a sheaf over V. The universal derivations $d_\alpha\colon R_\alpha \to \Omega_{R_\alpha/S_\alpha}$ patch together to give a universal map

$$d\colon \mathcal{O}_V \to \Omega_{V/W}$$

of sheaves.

REMARK. If V is a *variety* and $S = k$, theorem 3.5.12 on page 158 implies that this is a locally free sheaf defining the cotangent bundle of V. In this case, we suppress the subscript

$$\Omega_{V/\operatorname{Spec} k} = \Omega_V$$

PROPOSITION 4.6.19. *If V is an irreducible n-dimensional variety over a field of characteristic 0, and $p \in V$ is a simple point (see definition 3.3.11 on page 125), then there exists an open affine $U \subset V$ containing p such that*

(1) $\Gamma(U, \Omega_V^1) = \Omega_V^1(U)$ *is a free module over* $k[U] = \mathcal{O}_V(U)$ *of rank equal to* n, *and equal to* $\Omega_{k[U]/k}$,

(2) $\Omega_V^t(U) = \Lambda^t \, \Omega_V^1(U)$ *is a free module over* $k[U]$ *of rank equal to* $\binom{n}{t}$.

REMARK. If $\{u_1, \ldots, u_n\}$ is a system of parameters at p then

$$\{du_{i_1} \wedge \cdots \wedge du_{i_t}\}$$

with $1 \le i_1 < \cdots < i_t \le n$, will be a free basis for Ω_U^t over $k[U]$.

There is some variation in notation: some authors use $\Omega^i[U]$ for what we would call $\Omega_V^i(U)$ or use the same notion for the sheaf and global sections interchangeably — leading to some ambiguity. We will always use sheaf-notation.

PROOF. Since p is simple, it is contained in $V \setminus S$, where S is the closed subvariety of singular points. Without loss of generality, we can assume V is smooth and $\Omega_V^1(V) = \Omega_{k[V]/k}$ represents its cotangent bundle, ξ (see theorem 3.5.12 on page 158). Let $U \subset V$ be an open affine with the property that $\xi|U$ is *trivial.* Then $\Omega_{k[U]/k} = \Omega_V^1(U)$ is free on a basis of differentials of local parameters. It is not hard to see that

$$\Omega_V^t = \Lambda^t \, \Omega_V^1$$

so proposition A.6.9 on page 477 implies the second statement. $\qquad \square$

We extend this to *general* varieties via

DEFINITION 4.6.20. If V is a smooth n-dimensional variety over a field of characteristic 0, the sheaves Ω_V^i can be defined, on open affines, $\{U_\alpha\}$, by

$$\begin{aligned} \Omega_V^1|U_\alpha &= \mathcal{A}(\Omega_{k[U_\alpha]/k}) \\ \Omega_V^i|U_\alpha &= \Lambda^i \, \Omega_V^1|U_\alpha \end{aligned}$$

where $\Omega_{k[U_\alpha]/k}$ is the Kähler module (see definition A.7.5 on page 483) over $k[U_\alpha] = \mathcal{O}_V(U_\alpha)$.

REMARK. The functorial properties of Kähler modules implies that regular maps induce homomorphisms of the sheaves Ω_V^1. It is a combination of pullback and "differentiation."

PROPOSITION 4.6.21. *Let* $F : X \to Y$ *be a morphism, let* $g : Y' \to Y$ *be another morphism, and let* $f' : X' = X \times_Y Y'$. *Then*

$$\Omega_{X'/Y'} = g'^*(\Omega_{X/Y})$$

where $g' : X' \to X$ *is the first projection.*

PROOF. This is proposition A.7.9 on page 485, applied to open affines. $\qquad \square$

The first exact sequence of Kähler modules (see theorem A.7.11 on page 487)

PROPOSITION 4.6.22. *Given morphisms of schemes* $f : X \to Y$, *and* $g : Y \to Z$ *the following sequence of sheaves is exact*

$$f^* \Omega_{Y/Z} \xrightarrow{v} \Omega_{X/Z} \xrightarrow{u} \Omega_{X/Y} \to 0$$

PROOF. This follows from A.7.11 on page 487 applied to open affines. $\qquad \square$

By restricting to open affines, we can get a conormal (see lemma A.7.12 on page 488) sequence of sheaves:

LEMMA 4.6.23. *If V is a smooth scheme and $j\colon W \hookrightarrow V$ is a closed subscheme defined by a quasi-coherent sheaf of ideals $\mathscr{I} \subset \mathcal{O}_V$, then the sequence*

$$\mathscr{I}/\mathscr{I}^2 \to j^*\Omega_V \xrightarrow{\varphi} \Omega_W \to 0$$

is exact. If W is also smooth, we get an exact sequence

(4.6.1) $$0 \to \mathscr{I}/\mathscr{I}^2 \to j^*\Omega_V \xrightarrow{\varphi} \Omega_W \to 0$$

REMARK. Since $\mathscr{I}$ is a module over $\mathcal{O}_V$, it follows that $\mathscr{I}/\mathscr{I}^2$ is naturally a module over $\mathcal{O}_V/\mathscr{I} = \mathcal{O}_W$.

PROOF. The top exact sequence is just the conormal exact sequence on open affines — noting that $j^*\Omega_V = j^{-1}\Omega_V \otimes_{\mathcal{O}_V} \mathcal{O}_W$ (see definition 3.5.13 on page 159). If W is smooth and $p \in W$ is a point, then proposition 4.6.19 on page 210 implies that, in a neighborhood, U, of p, $\ker\varphi$ locally free of rank $\dim V - \dim W = r$. It follows that there exist elements $x_1, \ldots, x_r \in \mathscr{I}(U)$ such that $dx_1, \ldots, dx_r$ generate $\ker\varphi$. If $\mathscr{I}' \subseteq \mathscr{I}$ is the ideal sheaf on U generated by $x_1, \ldots, x_r$ and $W' \supseteq W$ is the subvariety defined by $\mathscr{I}'$, then

$$0 \to \mathscr{I}'/(\mathscr{I}')^2 \to \Omega_V \otimes_{\mathcal{O}_V} \mathcal{O}_{W'} \xrightarrow{\varphi} \Omega_{W'} \to 0$$

is exact by construction (at least on U). Since W' is defined by the vanishing of r equations, $\dim W' \geq \dim V - r = \dim W$ (by theorem 2.8.29 on page 109). Considering the stalks at any point q of $W' \cap U$, we get an exact sequence

$$0 \to \mathscr{I}'/(\mathscr{I}')^2 = k^r \to \Omega_V \otimes_{\mathcal{O}_V} \mathcal{O}_{W'} = k^{\dim V} \xrightarrow{\varphi} \Omega_{W'} = k^{\dim W'} \to 0$$

which implies that $\dim W' = \dim W$, so that $W' = W$ and $\mathscr{I}' = \mathscr{I}$.   □

DEFINITION 4.6.24. *If V is a smooth n-dimensional variety, Ω_V is locally free of rank n. This means that $\Lambda^n \Omega_V = \Omega_V^n$ is locally free of rank 1 — i.e., an invertible sheaf called the* canonical sheaf, *ω_V, of V.*

REMARK. The canonical sheaf will play an important part in the Riemann-Roch theorem and its many generalizations.
We have the Adjunction Formula for smooth varieties:

THEOREM 4.6.25 (Adjunction Formula). *If $j\colon W \hookrightarrow V$ is a closed codimension-r immersion of smooth varieties and W is defined by the quasi-coherent ideal-sheaf, $\mathscr{I} \subset \mathcal{O}_V$, then*

$$\omega_W = j^*\omega_V \otimes_{\mathcal{O}_W} \left(\Lambda^r\left(\mathscr{I}/\mathscr{I}^2\right)\right)^\vee$$

where $\mathscr{G}^\vee = \mathscr{H}om(\mathscr{G}, \mathcal{O}_W)$ is the dual sheaf — as defined in equation D.4.5 on page 566.

PROOF. The exact sequence in 4.6.1, coupled with exercise 2 on page 518 implies that

$$\omega_W \otimes_{\mathcal{O}_W} \Lambda^r\left(\mathscr{I}/\mathscr{I}^2\right) = j^*\omega_V$$

where $\Lambda^r\left(\mathscr{I}/\mathscr{I}^2\right)$ is an invertible sheaf, or line-bundle (see definition C.1.3 on page 507 and theorem C.2.7 on page 518). The conclusion follows from the fact that

$$\Lambda^r\left(\mathscr{I}/\mathscr{I}^2\right) \otimes_{\mathcal{O}_W} \left(\Lambda^r\left(\mathscr{I}/\mathscr{I}^2\right)\right)^\vee = \mathcal{O}_W$$

so that $\left(\Lambda^r\left(\mathscr{I}/\mathscr{I}^2\right)\right)^\vee$ *cancels out* the factor of $\Lambda^r\left(\mathscr{I}/\mathscr{I}^2\right)$ — see exercise 4 on page 505.
                                                                                              □

4.6.3. Finite Maps. The concept of finite maps, defined for affine varieties in definition 2.5.10 on page 76, can be extended to general varieties. To do this, we must first show that finiteness of a morphism is a *local property:*

PROPOSITION 4.6.26. *A morphism $f: V \to W$ of affine varieties is finite if and only if every point $p \in W$ has an affine open set U containing it, such that $U' = f^{-1}(U)$ is affine and $f|U': U' \to U$ is finite in the sense of 2.5.10 on page 76.*

PROOF. Since sets of the form $D(g)$ (see definition 2.4.4 on page 63) for $g \in k[W]$ form a base for its topology, it will suffice to prove the result for them.

Let $g_i \in k[W], i = 1, \dots, t$, be a set of elements such that $(g_1, \dots, g_t) = k[W]$ so that

$$W = \bigcup_{i=1}^{t} D(g_i)$$

— see corollary 2.4.7 on page 64. Then $f^{-1}(D(g_i)) = D(f^*g_i))$ and, by hypothesis,

$$k[D(g_i)] = k[W]_{g_i} = \bigoplus_{j=1}^{n_i} \omega_{i,j} \cdot k[D(f^*g_i))] = \bigoplus_{j=1}^{n_i} \omega_{i,j} \cdot k[V]_{f^*g_i}$$

(see proposition 2.4.5 on page 63) where the $\omega_{i,j}$ are generators of the finitely generated module, $k[D(g_i)]$, over $k[D(f^*g_i))]$.

We claim that the set $\{\omega_{i,j}\}$ for $i = 1, \dots, t$ and $j = 1, \dots, n_i$ generates $k[V]$ as a module over $k[W]$. If $v \in k[V]$, then for each $i = 1, \dots, t$,

$$(4.6.2) \qquad v = \sum_{j=1}^{n_i} \omega_{i,j} \cdot \frac{h_{i,j}}{g_i^{\alpha_i}}$$

Claim: The $g_j^{\alpha_i}$ generate the whole ring, $k[W]$. That is because $\sqrt{(g_1^{\alpha_1}, \dots, g_t^{\alpha_t})} = (g_1, \dots, g_t) = k[W]$ and $1 \in \sqrt{\mathfrak{a}}$ if and only if $1 \in \mathfrak{a}$ for any ideal $\mathfrak{a} \subset k[W]$.

It follows that we can find $z_{i,j}$ such that

$$1 = \sum_{i=1}^{t} z_{i,j} \cdot g_i^{\alpha_i}$$

Since

$$v = v \cdot 1 = v \cdot \sum_{i=1}^{t} z_{i,j} \cdot g_i^{\alpha_i}$$

we plug this into equation 4.6.2 to get

$$\begin{aligned} v &= \sum_{i=1}^{t} z_{i,j} \cdot g_i^{\alpha_i} \cdot \left(\sum_{j=1}^{n_i} \omega_{i,j} \cdot \frac{h_{i,j}}{g_i^{\alpha_i}} \right) \\ &= \sum_{i=1}^{t} \sum_{j=1}^{n_i} \omega_{i,j} \cdot h_{i,j} z_{i,j} \end{aligned}$$

$\square$

In light of this result, we can simply *define* a finite map of schemes to satisfy the hypotheses of proposition 4.6.26:

DEFINITION 4.6.27. A morphism $f\colon V \to W$ of schemes is *finite* if and only if every point $p \in W$ has an affine open set U containing it, such that $U' = f^{-1}(U)$ is affine and $f|U'\colon U' \to U$ is finite in the sense of 2.5.10 on page 76.

The local nature of finite maps and proposition 2.5.11 on page 77 imply:

PROPOSITION 4.6.28. *Let $f\colon V \to W$ be a finite morphism of locally noetherian schemes. Then*

   (1) f is surjective
   (2) if $p \in W$, then $f^{-1}(p)$ consists of a finite number of points.

We must revise the definition of "unramified" a bit:

DEFINITION 4.6.29. A finite mapping $f\colon V \to W$ of schemes is *unramified* at a point $p \in V$ if in the induced map (see proposition 3.2.3 on page 118)

$$f_p\colon \mathcal{O}_{W,f(p)} \to \mathcal{O}_{V,p}$$

the maximal ideals of the local rings corresponding to p and $f(p)$ satisfy $\mathfrak{m}_p = \mathfrak{m}_{f(p)} \cdot \mathcal{O}_{V,p}$, and $k(p) = \mathcal{O}_{V,p}/\mathfrak{m}_p$ is a separable algebraic extension of $k(f(p)) = \mathcal{O}_{W,f(p)}/\mathfrak{m}_{f(p)}$.

LEMMA 4.6.30. *If a finite mapping $f\colon V \to W$ of schemes is unramified at a point $p \in V$ then $(\Omega_{V/W})_p = 0$.*

REMARK. It is interesting that the converse of this result is also true. Many authors *define* unramified maps in terms of this.

PROOF. Since this is a local condition, assume $V = \operatorname{Spec} A$, $W = \operatorname{Spec} B$, $p = \mathfrak{p}$ and $f(p) = \mathfrak{q}$. Since f is unramified at p, we have $\mathfrak{q}A_\mathfrak{p} = \mathfrak{p}A_\mathfrak{p}$ and

$$(4.6.3) \qquad\qquad A_\mathfrak{p} \otimes_{B_\mathfrak{q}} k(\mathfrak{q}) = A_\mathfrak{p}/\mathfrak{q}A_\mathfrak{p} = A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p} = k(\mathfrak{p})$$

so we have a commutative diagram of rings

$$
\begin{array}{ccc}
A_\mathfrak{p} & \longrightarrow & k(\mathfrak{p}) \\
\uparrow & & \uparrow \\
B_\mathfrak{q} & \longrightarrow & k(\mathfrak{q})
\end{array}
$$

which, coupled with equation 4.6.3 and proposition A.7.9 on page 485 implies that

$$\Omega_{A_\mathfrak{p}/B_\mathfrak{q}} \otimes_{A_\mathfrak{p}} k(\mathfrak{p}) = \Omega_{k(\mathfrak{p})/k(\mathfrak{q})} = 0$$

since $k(\mathfrak{p})/k(\mathfrak{q})$ is a finite separable extension (see corollary A.7.16 on page 490).

Since $\Omega_{A_\mathfrak{p}/B_\mathfrak{q}}$ is a finitely generated $A_\mathfrak{p}$-module (because we assumed f is finite) we can conclude from Nakayama's lemma (corollary A.1.82 on page 379) that $\Omega_{A_\mathfrak{p}/B_\mathfrak{q}} = 0$. Now, note that $\Omega_{A_\mathfrak{p}/B_\mathfrak{q}} = \Omega_{A_\mathfrak{p}/B}$. This is because the morphisms of rings

$$B \to B_\mathfrak{q} \to A_\mathfrak{p}$$

gives a short exact sequence (see theorem A.7.11 on page 487)

$$\Omega_{B_\mathfrak{q}/B} \otimes_{B_\mathfrak{q}} A_\mathfrak{p} \to \Omega_{A_\mathfrak{p}/B} \to \Omega_{A_\mathfrak{p}/B_\mathfrak{q}} \to 0$$

and equation A.7.5 on page 485 implies that

$$\Omega_{B_{\mathfrak{q}}/B} = (\Omega_{B/B})_{\mathfrak{q}} = 0$$

and $\Omega_{A_{\mathfrak{p}}/B} = (\Omega_{A/B})_{\mathfrak{p}} = 0$, which is what we wanted to prove. $\square$

EXERCISES.

3. Prove the converse of lemma 4.6.30 on the preceding page, i.e., show that if $f\colon \operatorname{Spec} A \to \operatorname{Spec} B$ is a finite map of schemes and $(\Omega_{A/B})_p = 0$ for a point $p \in \operatorname{Spec} A$, then f is unramified at p in the sense of definition 4.6.29 on the facing page. Hint: reverse the steps in the proof of lemma 4.6.30 on the preceding page.

4.6.4. S-valued points of a scheme. We will fix a scheme, S, throughout this section.

DEFINITION 4.6.31. If V is a scheme *over* a scheme T, and S is *another* scheme over T, then an *S-valued point* or *S-rational* point of V is a morphism

$$S \to V$$

of schemes over T. The set of S-valued points of V is denoted $V(S)$.

REMARK. A point of an algebraic variety is a $\operatorname{Spec} k$-valued point.
Suppose $S = \operatorname{Spec} R$ and

$$V = \operatorname{Spec}\left(\frac{\mathbb{Q}[X_1, \ldots, X_n]}{(f_1, \ldots, f_t)}\right)$$

Then an S-valued point of V is a morphism

$$\frac{\mathbb{Q}[X_1, \ldots, X_n]}{(f_1, \ldots, f_t)} \to R$$

that extends to a morphism

$$h\colon \mathbb{Q} \to \mathbb{Q}[X_1, \ldots, X_n] \to \frac{\mathbb{Q}[X_1, \ldots, X_n]}{(f_1, \ldots, f_t)} \to R$$

and represents a solution to the equations

$$\bar{f}_1(X_1, \ldots, X_n) \;=\; 0$$
$$\vdots$$
$$\bar{f}_t(X_1, \ldots, X_n) \;=\; 0$$

in R. Here, $\bar{f}_i$ is the result of mapping the coefficients of f_i via the map h.
By abuse of notation, when $S = \operatorname{Spec} R$, we often write $V(\operatorname{Spec} R) = V(R)$.

EXAMPLE 4.6.32. Consider the affine scheme over $\mathbb{R}$, $\operatorname{Spec}\mathbb{R}[X]$ — as in exercise 11 on page 168. Its $\operatorname{Spec}\mathbb{R}$-valued (or just $\mathbb{R}$-valued) points correspond to points of $\mathbb{R}$. Its complex points (i.e., like the one corresponding to the principle ideal $(X^2 + 1) \subset \mathbb{R}[X]$ do not induce any $\mathbb{R}$-valued points because there does not exist a ring-homomorphism[1]

$$\mathbb{R}[X]/(X^2 + 1) = \mathbb{C} \to \mathbb{R}$$

Similar reasoning shows that, if V is any variety *over* $\operatorname{Spec}\mathbb{C}$ (i.e., over $\mathbb{C}$), then $V(\mathbb{R}) = \varnothing$: If $\operatorname{Spec}S$ is an open affine of V, an $\mathbb{R}$-valued point would be a ring-homomorphism

$$\mathbb{C} \to S \to \mathbb{R}$$

Here is another interesting example:

EXAMPLE 4.6.33. There is a field called *real algebraic geometry* that studies varieties over the real numbers (see [**18**]). It fits into the current framework: a real algebraic variety is the set of $\mathbb{R}$-valued points of a scheme over $\mathbb{R}$. The results of Nash in [**125**] and Tognoli, in [**166**], imply that every compact smooth manifold is the set of $\operatorname{Spec}\mathbb{R}$-valued points of a scheme over $\mathbb{R}$.

DEFINITION 4.6.34. If V is a scheme over a field, F, and $\bar{F}$ is the algebraic closure of F, then the $\operatorname{Spec}\bar{F}$-valued points are called the *geometric points of V*.

REMARK. Suppose V is a scheme over $\mathbb{Q}$ in the sense that each of its open affines, $\{A_j\}$, is an affine scheme over $\mathbb{Q}$ of finite type. Then the set of geometric points of V is just the algebraic variety (over the algebraic closure, $\bar{\mathbb{Q}}$) defined using the same equations as the $\{A_j\}$ and glued together. Any real variety R has, as its set of geometric points, a complex variety $R(\operatorname{Spec}\mathbb{C}) = R(\mathbb{C})$ obtained by "complexifying" R.

We can extend the concept of geometric points:

DEFINITION 4.6.35. If $g: V \to W$ is a morphism of schemes over F, then the *geometric fiber* of g is defined to be

$$V \times_W \operatorname{Spec}\bar{F}$$

where $\bar{F}$ is the algebraic closure of F.

REMARK. The notion of geometric fiber first appeared in [**64**]. In some respects, it reflects the geometry better than the fiber as defined in section 4.5.2 on page 202. For instance, $\operatorname{Spec}\mathbb{R}$ and $\operatorname{Spec}\mathbb{C}$ are both single points so the fiber over a point of the map

$$\operatorname{Spec}\mathbb{C} \to \operatorname{Spec}\mathbb{R}$$

is a single point, $\operatorname{Spec}\mathbb{C}$, by lemma 4.5.8 on page 203 or definition 4.5.6 on page 202. The geometric fiber is

$$F = \operatorname{Spec}(\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C})$$

We determine the maximal ideals, $\mathfrak{m}$, of $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}$ by realizing that the quotient $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}/\mathfrak{m}$ is a field. Consequently, we map $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C}$ to a field — essentially $\mathbb{C}$ — and consider the kernels. Suppose we consider maps of the form

$$f: \mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \to \mathbb{C}$$

---

[1]Homomorphisms of fields must be *injective* — the kernel is the *only* ideal in a field: (0).

with $f|1 \otimes \mathbb{C} = 1\colon \mathbb{C} \to \mathbb{C}$. All such maps will be of the form

$$\begin{aligned} f\colon \mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} &\to \mathbb{C} \\ x \otimes y &\mapsto h(x) \cdot y \end{aligned}$$

where $h\colon \mathbb{C} \to \mathbb{C}$ is a morphism that *fixes* $\mathbb{R} \subset \mathbb{C}$, so $h|\mathbb{R} = f|1 \otimes \mathbb{R}$. It is easily verified that there are *two* such homomorphisms: the *identity* map and *complex-conjugation.* The kernels of these two versions of f will be two distinct maximal ideals.

It follows that the geometric fiber has *two* closed points. This reflects the notion that $\operatorname{Spec}\mathbb{C} \to \operatorname{Spec}\mathbb{R}$ is somehow collapsing two dimensions down to one.

Another interesting angle on S-valued points of a scheme is that they allow us to regard schemes as functors:

DEFINITION 4.6.36. Let V be a scheme and define the *functor*

$$h_V\colon \mathscr{R} \to \mathscr{S}$$

from the category of rings to that of sets via

$$h_V(R) = \operatorname{hom}_{\operatorname{Scheme}}(\operatorname{Spec} R, V) = V(R)$$

In greater generality, we can define a functor from the category of *schemes* to that of sets in the same way:

$$h_V(S) = \operatorname{hom}_{\operatorname{Scheme}}(S, V) = V(S)$$

These h-functors represent schemes in a certain sense:

PROPOSITION 4.6.37. *Let V and W be schemes. Then any morphism $f\colon V \to W$ induces a natural transformation (see definition A.5.10 on page 444)*

$$h_f\colon h_V \to h_W$$

and any natural transformation

$$t\colon h_V \to h_W$$

is induced by a morphism $f_t\colon V \to W$.

REMARK. This is a special case of a general category-theoretic result called the Yoneda Lemma — see [**106**].

PROOF. The first statement is clear: given a morphism $f\colon V \to W$, any morphism $\operatorname{Spec} R \to V$ can be composed with f to give a morphism $\operatorname{Spec} R \to W$. It is easy to verify that this defines a natural transformation.

To prove the second statement, we first consider the case where V is affine, i.e. $V = \operatorname{Spec} S$ for some ring S. Then the identity map

$$1\colon V \to V$$

defines an S-valued point of V and its image under the natural transformation, t, defines an S-valued point of W, i.e., a morphism

$$g\colon \operatorname{Spec} S = V \to W$$

That $t = h_g$, as defined above, follows from the properties of a natural transformation and is left to the reader as an exercise.

In the general case, we have a decomposition into open affines

$$V = \bigcup_\alpha \operatorname{Spec} S_\alpha$$

and the affine case, above, provides morphisms

$$\operatorname{Spec} S_\alpha \to W$$

that are compatible with each other (because t is a natural transformation — see definition A.5.10 on page 444) so they induce a morphism

$$V \to W$$

$\square$

One motivation for considering the functors h_* is how these sets of R-valued points behave under products

LEMMA 4.6.38. *If V, W are schemes over T and*

$$Z = V \times_T W$$

then

$$h_Z(S) = h_V(S) \times h_W(S)$$

(Cartesian product) for all schemes, S, over T.

PROOF. This follows from the diagram in lemma 4.5.2 on page 199 that defines the category-theoretic properties of a fibered product. Every pair of morphisms over T

$$S \xrightarrow{f} V$$
$$S \xrightarrow[g]{} W$$

induces a *unique* morphism $f \times g \colon S \to Z$. Furthermore, given a morphism $F \colon S \to Z$, the projections

$$Z \to V$$
$$Z \to W$$

define morphisms from S to V and W. $\square$

EXERCISES.

4. Show that

$$\operatorname{Spec} (\mathbb{Q}(X) \otimes_{\mathbb{Q}} \mathbb{C})$$

has a closed point for every transcendental number over $\mathbb{Q}$.

CHAPTER 5

Projective varieties

> "...to look upon affine geometry as well as metric geometry as special
> cases of projective geometry..."
> —Felix Klein, *Elementary Mathematics from an Advanced Standpoint*
> (see [**88**]).

5.1. Introduction

Projective varieties are solution-sets of algebraic equations in projective space. We have already seen how big a part they played in classical algebraic geometry with Bézout's Theorem. In general, projective varieties are useful in studying objects that are "compact in the usual sense". Although all affine varieties are compact in the Zariski topology, they are seldom compact in any more conventional topology.

This is why Nash could only show that manifolds occurred as compact *components* of a real affine variety in [**125**]. Tognoli's trick moved the problem into a real projective space and modeled a manifold with a real projective variety[1] in [**166**].

Every projective space is a general variety (see example 4.4.8 on page 186), so projective varieties are also general varieties or schemes. We could, consequently, use the machinery developed in the previous chapters to study these objects. Although we will do this to a large extent, it is often possible to give "intrinsic" descriptions of projective varieties that are more succinct — i.e., via homogeneous coordinates.

Suppose we have an algebraically closed field k and an n-dimensional projective space $k\mathbb{P}^n = \mathbb{P}(\mathbb{A}^{n+1})$ — see definition 4.4.8 on page 186. If $f(X_0,\ldots,X_n) \in k[X_0,\ldots,X_n]$ is a polynomial of degree m in homogeneous coordinates then

$$f(X_0,\ldots,X_n) = 0$$

defines an algebraic set, V, in $\mathbb{P}(\mathbb{A}^{n+1})$. The properties of homogeneous coordinates imply that we must have

$$f(t \cdot x_0,\ldots,t \cdot x_n) = 0$$

for $[x_0{:}\cdots{:}x_n] \in V$ and any nonzero $t \in k$. If we split f into its homogeneous components we get

$$f(t \cdot x_0,\ldots,t \cdot x_n) = f_0 + t \cdot f_1 + t^2 \cdot f_2 + \cdots + t^m \cdot f_m$$

---

[1]Strictly speaking, a real projective variety is the set of $\mathbb{R}$-valued points of a projective scheme over $\mathbb{R}$.

and, since k is infinite, each of these must vanish on V. Consequently, we can restrict our attention to homogeneous polynomials.

As before, we define the Zariski topology on $k\mathbb{P}^n$:

DEFINITION 5.1.1. A set in $k\mathbb{P}^n$ is said to be *closed* if it is the zero-set of finitely many polynomials.

REMARK. This definition is compatible with the topology gotten by decomposing $k\mathbb{P}^n$ into a union of open affines equipped with the conventional Zariski topology. For instance

$$f(x_0, \ldots, x_n) = 0$$

implies that on $\mathbb{A}_i^n \subset k\mathbb{P}^n$ (in the notation of definition 4.4.8 on page 186), we get

$$f\left(\frac{x_0}{x_i}, \ldots, 1, \ldots, \frac{x_n}{x_i}\right) = 0$$

The definition above is "more compact" than the one involving open affines.

As before, zero-sets of polynomials are really zero-sets of ideals within $k[X_0, \ldots, X_n]$:

DEFINITION 5.1.2. An ideal $\mathfrak{a} \in k[X_0, \ldots, X_n]$ will be said to be *homogeneous* if whenever a polynomial, $f \in \mathfrak{a}$, then every homogeneous component of f is in $\mathfrak{a}$.

Given a homogeneous ideal $\mathfrak{a}$, we define $\mathcal{P}(\mathfrak{a}) \subset k\mathbb{P}^n$ to be the set of points

$$(a_0 : \cdots : a_n) \in k\mathbb{P}^n$$

such that $f(a_0, \ldots, a_n) = 0$ for all $f \in \mathfrak{a}$.

REMARK. It is not hard to see that ideals are homogeneous if and only they are generated by homogeneous polynomials.

Since at least one homogeneous coordinate must be nonzero, we have:

PROPOSITION 5.1.3. *If the ideal $\mathfrak{i} \in k[X_0, \ldots, X_n]$ is given by $\mathfrak{i} = (X_0, \ldots, X_n)$, then $P(\mathfrak{i}) = \varnothing$.*

REMARK. This ideal is called the *irrelevant ideal* since it generates the empty set.

Now we can state the projective form of the Nullstellensatz:

THEOREM 5.1.4. *Let $\mathfrak{h} \in k[X_0, \ldots, X_n]$ be a homogeneous ideal. Then $\mathfrak{h}$ defines the empty set in $k\mathbb{P}^n$ if and only if $\sqrt{\mathfrak{h}}$ contains the irrelevant ideal.*

REMARK. The correspondence between ideals and projective varieties is somewhat more complex than in the affine case. Exercises 8 on page 226, 9, and 10 clarify this relationship.

PROOF. Certainly, if $\sqrt{h} = (X_0, \ldots, X_n)$, then $X_i^{k_i} \in \mathfrak{h}$ for all i and suitable $k_i > 0$, so $\mathfrak{h}$ defines the origin in $\mathbb{A}^{n+1}$ and the empty set in $k\mathbb{P}^n$.

Conversely, suppose $\mathfrak{h}$ defines the empty set in $k\mathbb{P}^n$. Then its intersection with each of the open affines $\mathbb{A}_i^n$, $i = 0, \ldots, n$ is also empty (see definition 4.4.8 on page 186). If

$$\mathfrak{h} = (F_1(X_0, \ldots, X_n), \ldots, F_t(X_0, \ldots, X_n))$$

then, in $\mathbb{A}_i^n$, this is the ideal

$$\mathfrak{h}_i = \left(F_1 \left(\frac{X_0}{X_i}, \ldots, 1, \ldots, \frac{X_n}{X_i} \right), \ldots, F_t \left(\frac{X_0}{X_i}, \ldots, 1, \ldots, \frac{X_n}{X_i} \right) \right)$$

Since this defines the empty set, the regular Nullstellensatz (theorem 2.2.3 on page 40) implies that

$$1 \in \mathfrak{h}_i$$

or that there exist polynomials $p_j \left(\frac{X_0}{X_i}, \ldots, \frac{X_n}{X_i} \right)$ such that

$$1 = \sum_{j=1}^t p_j \left(\frac{X_0}{X_i}, \ldots, \frac{X_n}{X_i} \right) \cdot F_j \left(\frac{X_0}{X_i}, \ldots, 1, \ldots, \frac{X_n}{X_i} \right)$$

and after multiplying by a suitable power of X_i to clear denominators, we get

$$X_i^{N_i} = \sum_{j=1}^t p_i \cdot F_j$$

Since this is true for any i, $\sqrt{\mathfrak{h}}$ contains the irrelevant ideal. $\square$

Since the tangent space is a local property of a variety,

DEFINITION 5.1.5. Let $W \subset k\mathbb{P}^n$ be a projective variety and let $p \in W$ be contained in an open affine $\mathbb{A}_i^n \subset k\mathbb{P}^n$. The *tangent space* of W at p is just the tangent space of the affine variety $W \cap \mathbb{A}_i^n \subset \mathbb{A}^n$ at p.

If V is a vector-space over k, every point, p, in $\mathbb{P}(V)$ represents a one-dimensional subspace, ℓ_p, of V. Since the point p is gotten by collapsing this subspace to a point, intuition suggests

LEMMA 5.1.6. *If V is a vector-space and $p \in \mathbb{P}(V)$ corresponds to the line $\ell_p \subset V$, then the tangent space of $\mathbb{P}(V)$ at p is V / ℓ_p. In particular, $\mathbb{P}(V)$ is smooth.*

PROOF. Suppose $p \in \mathbb{A}_i^n$, the coordinates of a point in V are given by $(x_0, \ldots, x_n)$. Then the map

$$V \setminus \{0\} \to \mathbb{P}(V)$$

defines a map

$$V \setminus \{x_i = 0\} \xrightarrow{f} \mathbb{A}_i^n$$
$$(x_0, \ldots, x_n) \mapsto \left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i} \right)$$

Let $q \in V \setminus \{0\}$ map to $p \in \mathbb{P}(V)$. At q, the induced map of tangent spaces is surjective and given by

$$d_q f =$$
$$\left(\frac{(dx_0) \cdot x_i(q) - x_0(q) \cdot (dx_i)}{x_i^2(q)}, \ldots, \frac{(dx_n) \cdot x_i(q) - x_n(q) \cdot (dx_i)}{x_i^2(q)} \right)$$

where $x_i(q)$ is just the coordinates of $q \in V$.

The kernel of this map is given by

$$(dx_j) \cdot x_i(q) - x_j(q) \cdot (dx_i) = 0$$

Under the identification of the tangent space of V with V, $dx_j = x_j$ for all j so we get

$$x_j \cdot x_i(q) - x_i \cdot x_j(q) = 0$$

for all i, j. This implies that $(x_0, \ldots x_n)$ lies on the same line through the origin as $(x_0(q), \ldots, x_n(q))$ or $(x_0, \ldots x_n) \in \ell_p$. $\square$

As we have seen, projective varieties cannot be studied in isolation: one must have knowledge of associated affine varieties to understand them.

DEFINITION 5.1.7. Let $\mathfrak{a} \in k[X_0, \ldots, X_n]$ be a homogeneous ideal and let $\mathcal{P}(\mathfrak{a}) \subset k\mathbb{P}^n = \mathbb{P}(\mathbb{A}^{n+1})$ be the associated projective variety. Then the affine variety, $\mathcal{V}(\mathfrak{a}) \subset \mathbb{A}^{n+1}$ is called the *affine cone* of $P(\mathfrak{a})$. The quotient ring

$$k[\mathcal{V}(\mathfrak{a})] = \frac{k[X_0, \ldots, X_n]}{\mathfrak{a}}$$

— the coordinate ring of $\mathcal{V}(\mathfrak{a})$ — is called the *homogeneous coordinate ring* of $\mathcal{P}(\mathfrak{a})$, and denoted $k_H[\mathcal{P}(\mathfrak{a})]$.

REMARK. The homogeneous coordinate ring is also frequently called the *projective coordinate ring*.

Although the homogeneous coordinate ring of $P(\mathfrak{a})$ defines the regular functions on the *cone,* these generally do *not* induce regular functions on $P(\mathfrak{a})$. In fact, for many projective varieties, the only regular functions are *constants* — see 5.5.17 on page 257. Nevertheless, the homogeneous coordinate ring of $P(\mathfrak{a})$ contains a great deal of useful geometric information — see section 5.7 on page 266.

It is possible for *distinct* homogeneous ideals $\mathfrak{a}_1, \mathfrak{a}_2 \subset k[X_0, \ldots, X_n]$ to determine the *same* projective variety since they determine distinct affine varieties $\mathcal{V}(\mathfrak{a}_i) \subset \mathbb{A}^{n+1}$ with the same image under the projection

$$\mathbb{A}^{n+1} \to \mathbb{P}(\mathbb{A}^{n+1}) = k\mathbb{P}^n$$

— see exercises 7 on page 226 and 8 on page 226.

EXAMPLE 5.1.8. The affine cone has a line through the origin for each point of $\mathcal{P}(\mathfrak{a})$. For instance

$$(X^2 + 2Y^2 - Z^2) \subset k[X, Y, Z]$$

defines an ellipse in $k\mathbb{P}^2 = \mathbb{P}(\mathbb{A}^3)$ and its affine cone is depicted in figure 5.1.1

The affine cone of a projective variety is important because it determines many of the variety's properties.

PROPOSITION 5.1.9. *Let $\mathfrak{a} \subset k[X_0, \ldots, X_n]$ be a homogeneous ideal. The projective variety $\mathcal{P}(\mathfrak{a}) \subset k\mathbb{P}^n$ is irreducible if and only if its affine cone $\mathcal{V}(\mathfrak{a}) \subset \mathbb{A}^{n+1}$ is irreducible, which happens if and only if $\mathfrak{a}$ is prime.*

PROOF. Any decomposition of $\mathcal{P}(\mathfrak{a}) = P_1 \cup P_2$ induces a decomposition of the affine cone and vice-versa. $\square$

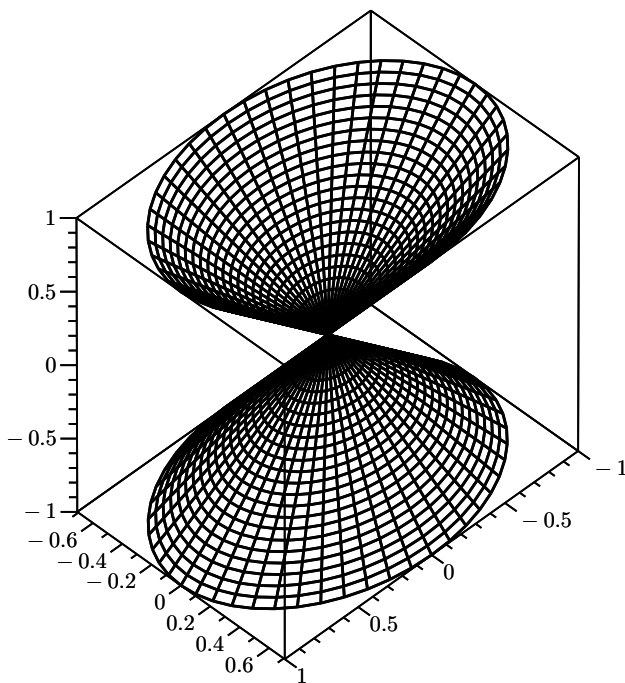The affine cone also determines the tangent space:

FIGURE 5.1.1.   Affine cone of an ellipse

LEMMA 5.1.10. *Let $\mathfrak{a} \in k[X_0, \ldots, X_n]$ determine a projective variety $\mathcal{P}(\mathfrak{a}) \subset k\mathbb{P}^n = \mathbb{P}(\mathbb{A}^{n+1})$ with affine cone $\mathcal{V}(\mathfrak{a}) \subset \mathbb{A}^{n+1}$. If $p \in P(\mathfrak{a})$ is in the image of $q \in \mathcal{V}(\mathfrak{a})$ then*

(1) $\ell_p \subset T_{\mathcal{V}(\mathfrak{a}),p}$, where $\ell_p \subset \mathbb{A}^{n+1}$ is the one-dimensional subspace corresponding to $p \in \mathbb{P}(\mathbb{A}^{n+1})$ and

(2) $T_{\mathcal{P}(\mathfrak{a}),p} = T_{\mathcal{V}(\mathfrak{a}),q}/\ell_p$

REMARK. It follows that singular points in $\mathcal{P}(\mathfrak{a})$ correspond to singular lines in $\mathcal{V}(\mathfrak{a})$. If $\mathcal{V}(\mathfrak{a})$ has no such singular lines, then $\mathcal{P}(\mathfrak{a})$ is smooth.

Note that $\mathcal{V}(\mathfrak{a})$ will always have a singular point at the origin because the ideal $\mathfrak{a}$ is homogeneous. This point-singularity does not contribute to possible singularities in $\mathcal{P}(\mathfrak{a})$.

PROOF. The fact that $p \in \mathcal{P}(\mathfrak{a})$ implies that $\ell_p \subset \mathcal{V}(\mathfrak{a})$, which implies that it is in the tangent space. The second statement follows from the fact that the equations defining the tangent space in $\mathcal{P}(\mathfrak{a})$ are essentially the same as those defining it in $\mathcal{V}(\mathfrak{a})$, except for the projection

$$V = \bigcup_{j=0}^{n} V_j T_{\mathbb{A}^{n+1},p} \to T_{\mathbb{P}(\mathbb{A}^{n+1})} = T_{\mathbb{A}^{n+1},p}/\ell_p$$

$\square$

We can easily decompose projective varieties into a union of affines:

PROPOSITION 5.1.11. *Let $V \subset k\mathbb{P}^n$ be a projective variety. Then*

$$V = \bigcup_{j=0}^{n} V_j$$

with $V_j = V \cap \mathbb{A}_j^n$ with $j = 0, \ldots, n$ is a decomposition into affines (see definition 4.4.8 on page 186).

In addition,

$$k[V_j] = k_H[V]_{(X_j)}$$

where $k_H[V]_{(X_j)}$ is defined in 4.4.13 on page 190 and

$$k(V) = k_H[V]_{((0))}$$

PROOF. For every defining equation

$$f(X_0, \ldots, X_n) = 0$$

for V, V_j has a defining equation

$$\bar{f}(x_1, \ldots, x_n) = f\left(\frac{X_0}{X_j}, \ldots, \frac{X_{j-1}}{X_j}, 1, \frac{X_{j+1}}{X_j}, \ldots, \frac{X_n}{X_j}\right) = 0$$

$$= \frac{f(X_0, \ldots, X_n)}{X_j^{\deg f}}$$

which makes it an affine variety in $\mathbb{A}_j^n$. Here

$$x_i = \begin{cases} \frac{X_{i-1}}{X_j} & \text{if } i \leq j \\ \frac{X_i}{X_j} & \text{if } i > j \end{cases}$$

are the coordinates of $\mathbb{A}_j^n \subset \mathbb{P}(\mathbb{A}^{n+1})$. This clearly coincides with the definition of $k_H[V]_{(X_j)}$ in 4.4.13 on page 190.

The final statement follows from the fact that $k(V) \cong k(V_j)$ (see proposition 4.6.12 on page 207) which is the field of fractions of $k[V_j]$, i.e. a field of fractions of the form

$$\frac{\frac{r(X_0, \ldots, X_n)}{X_j^{\deg r}}}{\frac{g(X_0, \ldots, X_n)}{X_j^{\deg g}}} = \frac{r(X_0, \ldots, X_n)}{g(X_0, \ldots, X_n)} X_j^{\deg g - \deg r}$$

— in other words, we invert *all* nonzero graded elements of $k_H[V]$ and select quotients of grade 0.   □

There is a standard way of converting an affine variety into a projective one:

DEFINITION 5.1.12. Let $V \subset \mathbb{A}^n$ be an affine variety and embed $\mathbb{A}^n$ into $\mathbb{A}^{n+1}$, adding an additional coordinate X_0. If

$$f_i(X_1, \ldots, X_n) = 0$$

are the defining equations of V, for $i = 1, \ldots m$, then

$$X_0^{\deg f_i} \cdot f_i \left(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0} \right) = 0$$

are homogeneous equations defining a projective variety $\bar{V} \subset k\mathbb{P}^n$ called the *projective closure* of V.

REMARK. We have already seen this in chapter 1 in proposition 1.2.6 on page 5. Clearly $\bar{V} \cap \mathbb{A}_0^n = V$.

Recall the Proj defined in definition 4.4.12 on page 190. If $R = k[X_0, \ldots, X_n]$, the inclusion

$$\mathrm{Proj}\, R \hookrightarrow \mathrm{Spec}\, R$$

is induced by the projection of the affine cone to a projective variety.

We can generalize our notion of projective variety:

DEFINITION 5.1.13. If Y is any scheme and $\mathbf{P} = \mathrm{Proj}\, \mathbb{Z}[X_0, \ldots, X_n]$, the *projective n-space over Y*, denoted $Y\mathbb{P}^n$, is defined to be the fibered product

$$\mathbf{P} \times_{\mathrm{Spec}\,\mathbb{Z}} Y$$

A morphism $f\colon X \to Y$ of schemes is said to be *projective* if it factors as

$$X \xrightarrow{\iota} Y\mathbb{P}^n \to Y$$

where ι is a closed immersion and $Y\mathbb{P}^n \to Y$ is the canonical projection.

EXERCISES.

1. Each point $p \in k\mathbb{P}^n$ represents a line in $\mathbb{A}^{n+1}$. Consider the subset of $k\mathbb{P}^n \times \mathbb{A}^{n+1}$ composed of pairs

$$(p, \ell_p)$$

of points paired with the lines that they represent. This is called the *tautological line bundle*, η, on $k\mathbb{P}^n$. Show that

$$\eta \oplus \tau = k\mathbb{P}^n \times \mathbb{A}^{n+1}$$

a trivial bundle, where τ is the tangent bundle of $k\mathbb{P}^n$.

2. If V is a vector space over a field k, show that (as schemes)

$$\mathbb{P}(V) = \mathrm{Proj}\, \mathcal{S}V$$

where the symmetric algebra, $\mathcal{S}V$, is defined in A.6.4 on page 475 and elements of V are defined to have degree 1.

3. If R is the graded ring $k[X]$, where X has degree 1, list all of the points of $\mathrm{Proj}\, R$.

4. Show that the set of all cubic hypersurfaces in $k\mathbb{P}^3$ are parametrized by a projective space. What is its dimension?

5. Show that the set of all smooth cubic hypersurfaces in $k\mathbb{P}^3$ are parametrized by an open set in projective space.

6. Suppose $V \subset \mathbb{C}^{n+1}$ is a complex analytic cone — i.e.
- V is defined by the vanishing of a finite number of complex analytic functions (functions equal to their Taylor series expansions about 0 in a small neighborhood of 0), and
- if $(x_0, \ldots, x_n) \in V$ then $(tx_0, \ldots, tx_n) \in V$ for all $t \in \mathbb{C}$

Show that V is defined by the vanishing of a *finite* number of *homogeneous polynomials.* This is one step in the proof of Chow's Theorem that a complex *analytic* projective variety is actually *algebraic.*

7. Let $R = k[X_0, X_1, X_2]$ and let

$$\mathfrak{a} = \left(X_0 \cdot (X_1^2 - X_2^2), X_1 \cdot (X_1^2 - X_2^2), X_2 \cdot (X_1^2 - X_2^2) \right)$$
$$\mathfrak{b} = (X_1^2 - X_2^2)$$

Show that $\mathcal{P}(\mathfrak{a}) = \mathcal{P}(\mathfrak{b})$.

8. If $\mathfrak{a} \subset k[X_0, \ldots, X_n]$ is a radical homogeneous ideal and $f \in k[X_0, \ldots, X_n]$ is an element such that $f \cdot \mathfrak{i}^t \subset \mathfrak{a}$, where $\mathfrak{i} = (X_0, \ldots, X_n)$ is the irrelevant ideal and t is some integer, show that

$$\mathcal{P}(\mathfrak{a}) = \mathcal{P}(\mathfrak{a} + (f))$$

which means that more than one radical ideal can define the *same* projective variety.

9. If $\mathfrak{a} \subset k[X_0, \ldots, X_n]$ is a radical homogeneous ideal and $f \in k[X_0, \ldots, X_n]$ is a homogeneous element that vanishes on $\mathcal{P}(\mathfrak{a})$, show that $X_i^{n_i} \cdot f \in \mathfrak{a}$ for all i and suitable values of n_i.

10. Recall the definition of *saturation of an ideal* with respect to another in exercise 9 on page 71. If $\mathfrak{i} = (X_0, \ldots, X_n)$ is the irrelevant ideal and $\mathfrak{a}, \mathfrak{b} \subset k[X_0, \ldots, X_n]$, are two ideals of $R = k[X_0, \ldots, X_n]$, show that $\mathcal{P}(\mathfrak{a}) = \mathcal{P}(\mathfrak{b})$ if and only if $(\mathfrak{a} : \mathfrak{i}^\infty) = (\mathfrak{b} : \mathfrak{i}^\infty)$.

5.2. Grassmannians

5.2.1. Introduction. Grassmannians are an important class of projective varieties that have many applications in algebraic geometry, topology (see [**115**], [**6**], and [**72**]), and even genetics (see the survey [**133**]). Just as projective spaces are sets of lines through the origin within a vector space (see definition 1.2.2 on page 4), Grassmannians are spaces whose points are d-dimensional subspaces of a vector space.

Hermann Grassmann (1809 – 1877) was a German mathematician, physicist and linguist whose mathematical contributions were not recognized until he was in his sixties. He is responsible for many modern concepts in linear algebra including linear independence and exterior algebras (in his great work, [**60**]). He generalized work of Julius Plücker and defined general Grassmann Manifolds or Grassmannians.

5.2.2. Plücker coordinates. Throughout this section, we will use the $\mathbb{P}(V)$-notation for projective spaces, since the functorial properties of this construction will be important.

We will make extensive use of *exterior algebras* (see section A.6 on page 474, which defined them and develops their basic properties).

DEFINITION 5.2.1. If V is an n-dimensional vector space with a basis

$$\{e_1, \ldots, e_n\}$$

and W is a subspace with basis

$$\{b_1, \ldots, b_k\}$$

then

$$b_1 \wedge \cdots \wedge b_k = \sum_{i_1 < \cdots < i_k} a_{i_1,\ldots,i_k} e_{i_1} \wedge \cdots \wedge e_{i_k}$$

with $a_{i_1,\ldots,i_k} \in F$ (by proposition A.6.9 on page 477). The coefficients, $\{a_{i_1,\ldots,i_k}\}$, are called the *Plücker coordinates* of W.

REMARK. Corollary A.6.15 on page 480 implies that changing the basis of W multiplies all of the Plücker coordinates of W by the same nonzero scalar. Corollary A.6.13 on page 479 implies that the Plücker coordinates of W uniquely determine it.

> Julius Plücker (1801 – 1868) was a German mathematician and physicist who made many contributions to analytic geometry. His research on cathode rays paved the way for the discovery of the electron and X-rays (see [**109**]).

It follows that

PROPOSITION 5.2.2. *Plücker coordinates define a one-to-one correspondence between the k-dimensional subspaces of V and points of the projective space $\mathbb{P}(\Lambda^k V)$ defined by decomposable elements of $\Lambda^k V$.*

REMARK. Decomposable elements are defined in definition A.6.11 on page 478.

PROOF. Corollary A.6.15 on page 480 implies that the Plücker coordinates of a subspace determine a unique point of $\mathbb{P}(\Lambda^k V)$ corresponding to a decomposable element of $\Lambda^k V$. Corollary A.6.13 on page 479 implies that every such decomposable element determines a unique subspace.   □

Our next challenge will be to develop criteria for an element of $\Lambda^k V$ to be decomposable.

5.2.3. Interior products. Grassmann's original treatment of decomposable elements of an exterior algebra involved many tedious algebraic computations. We will adopt a simpler approach, using a construction that goes by various names including *interior products*, *convolutions,* and *contractions*.

Interior products were invented by Élie Cartan in his applications of exterior calculus to differential geometry.

DEFINITION 5.2.3. Let V be a vector space and let V^* be the dual vector space of linear functions on V. Define the *interior product*

$$\llcorner: V^* \otimes \Lambda^k V \to \Lambda^{k-1} V$$

for all $k \geq 1$, inductively by

(1) If $v \in F = \Lambda^0 V$, then $u \llcorner v = 0$.
(2) If $k = 1$, $u \llcorner v = u(v) \in F = \Lambda^0 V$
(3) $(u_1 + u_2) \llcorner v = u_1 \llcorner v + u_2 \llcorner v$ for all $u_1, u_2 \in V^*$ and $v \in \Lambda^k V$
(4) $u \llcorner (v_1 + v_2) = u \llcorner v_1 + u \llcorner v_2$ for all $u \in V^*$ and $v_1, v_2 \in \Lambda^k V$
(5) If $a \in \Lambda^i V$, $b \in \Lambda^j V$, then $u \llcorner (a \wedge b) = (u \llcorner (a) \wedge b + (-1)^i a \wedge (u \llcorner b)$

We can extend this definition:

PROPOSITION 5.2.4. *Let $u_1, u_2 \in V^*$. Then, for any k and any $v \in \Lambda^k V$ we get*

$$u_1 \llcorner (u_2 \llcorner v) = -u_2 \llcorner (u_1 \llcorner v)$$

Consequently we can define

$$\llcorner: \Lambda^j V^* \otimes \Lambda^k V \to \Lambda^{k-j} V$$

by setting

$$(u_1 \wedge \cdots u_j) \llcorner v = u_1 \llcorner (\cdots (u_j \llcorner v) \cdots)$$

PROOF. Suppose

$$v = a \wedge b \wedge c$$

where $a, b \in V$ and $c \in \Lambda^{k-2} V$. Then

$$
\begin{aligned}
u_1 \llcorner (u_2 \llcorner (a \wedge b \wedge c) &= u_1 \llcorner (u_2(a) b \wedge c - u_2(b) \cdot a \wedge c) \\
&= u_1(b) u_2(a) \cdot c - u_2(b) u_1(a) \cdot c \\
&= \big(u_1(b) u_2(a) - u_2(b) u_1(a)\big) \cdot c
\end{aligned}
$$

and

$$
\begin{aligned}
u_2 \llcorner (u_1 \llcorner (a \wedge b \wedge c) &= u_2 \llcorner (u_1(a) b \wedge c - u_1(b) a \wedge c) \\
&= u_2(b) u_1(a) \cdot c - u_2(a) u_1(b) \cdot c \\
&= \big(u_2(b) u_1(a) - u_2(a) u_1(b)\big) \cdot c
\end{aligned}
$$

$\square$

The following result (and the proof of proposition 5.2.4) illustrate the purpose of interior products — they "rip apart" exterior products:

LEMMA 5.2.5. *Let $\{e_1, \ldots, e_n\}$ be a basis for V and let $\{e^1, \ldots, e^n\}$ be a dual basis for V^* so*

$$(5.2.1) \qquad\qquad e^i(e_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Then

$$e^i \llcorner (e_{j_1} \wedge \cdots \wedge e_{j_k}) =$$

$$\begin{cases} 0 & \text{if } i \notin \{j_1, \ldots, j_k\} \\ (-1)^{t-1} e_{j_1} \wedge \cdots e_{j_{t-1}} \wedge e_{j_{t+1}} \wedge \cdots \wedge e_{j_k} & \text{if } i = j_t \end{cases}$$

PROOF. This is a straightforward inductive application of rule 5 of definition 5.2.3 on the facing page and equation 5.2.1 on the preceding page. □

DEFINITION 5.2.6. If V is a vector space with subspace W, the inclusion $W \hookrightarrow V$ induces an inclusion

$$\Lambda^k W \hookrightarrow \Lambda^k V$$

for any k. If $x \in \Lambda^k V$ then we say that W *envelopes* x if $x \in \Lambda^k W$.

LEMMA 5.2.7. *If V is a vector space with subspace W, a nonzero element $x \in \Lambda^k V$ is enveloped by W if and only if*

$$y \llcorner x \in W$$

for all $y \in \Lambda^{k-1} V^*$.

PROOF. Let $\{e_1, \ldots, e_t\}$ be a basis for W and extend it to a basis $\{e_1, \ldots, e_n\}$ for all of V with a corresponding dual basis $\{e^1, \ldots, e^n\}$ of V^*.
Let

$$x = \sum_{i_1 < \cdots < i_k} a_{i_1, \ldots, i_k} e_{i_1} \wedge \cdots \wedge e_{i_k}$$

Then $x \in \Lambda^k W$ if and only if $a_{i_1, \ldots, i_k} = 0$ whenever $i_k > t$. An inductive application of lemma 5.2.5 on the facing page gives

$$e^{i_1} \wedge \cdots \wedge e^{i_{k-1}} \llcorner x \quad = \quad \sum_{j > i_{k-1}} \pm a_{i_1, \ldots, j} e_j$$
$$+ \text{ other terms}$$

which is contained in W if and only if $a_{i_1, \ldots, i_k} = 0$ whenever $i_k > t$. □

We can use this to determine the minimal subspace that envelopes an element $x \in \Lambda^k V$:

LEMMA 5.2.8. *If V is a vector space with dual V^* and $x \in \Lambda^k V$ is a nonzero element, then the smallest subspace, W_x, of V that envelopes x is the set of*

$$y \llcorner x \in V$$

as y runs over all elements of $\Lambda^{k-1} V^*$.

This will be instrumental in determining when an element is reducible.

THEOREM 5.2.9. *Let V be a vector space with dual V^* and let $x \in \Lambda^k V$ be a nonzero element. Then x is reducible if and only if the minimal subspace of V that envelopes x is k-dimensional. This happens if and only if*

(5.2.2) $$(y \llcorner x) \wedge x = 0$$

for all $y \in \Lambda^{k-1} V^*$.

REMARK. If x is not reducible, $\dim W_x > k$.
In applying this test, it suffices to check it on basis elements of $\Lambda^{k-1} V^*$ since equation 5.2.2 is *linear* in y,

PROOF. Let W_x denote the minimal subspace of V that envelopes x.

Suppose $x = v_1 \wedge \cdots \wedge v_k$. Since $x \neq 0$ the set of vectors $\{v_1, \ldots, v_k\}$ must be linearly independent by lemma A.6.12 on page 479. It follows that the W_x is the span of $v_1, \ldots, v_k$ so it is k-dimensional.

We claim that W_x must have more than k dimensions if x is not decomposable. Suppose

$$x = v_1 \wedge \cdots \wedge v_k + w_1 \wedge \cdots \wedge w_k + \text{ other terms}$$

At least one of the w_i is not in the span of $v_1, \ldots, v_k$, since otherwise

$$w_1 \wedge \cdots \wedge w_k = c \cdot v_1 \wedge \cdots \wedge v_k$$

for some $c \in F$ by corollary A.6.15 on page 480, so $\dim W_x \geq k + 1$.

If $x = v_1 \wedge \cdots \wedge v_k$, then W_x is the subspace annihilated by x so equation 5.2.2 on the preceding page holds.

Suppose $\dim W_x = m > k$ and let $\{v_1, \ldots, v_m\}$ be a basis. There exists a set $\{y_1, \ldots, y_m\} \in \Lambda^{k-1} V$ such that $v_i = y_i \lrcorner x$.

Since W_x envelopes x, we have

$$x = \sum_{j_1 < \cdots < j_k} a_{j_1, \ldots, j_k} v_{j_1} \wedge \cdots \wedge v_{j_k}$$

where all of the $\{a_{j_1, \ldots, j_k}\}$ are nonzero. Since $m > k$, at least one of these terms must omit some subscript, say $1 \leq \ell \leq m$. Then

$$v_\ell \wedge x = (y_\ell \lrcorner x) \wedge x =$$
$$\sum_{j_1 < \cdots < j_k} (-1)^{r-1} a_{j_1, \ldots, j_k} v_{j_1} \wedge \cdots \wedge v_{j_{r-1}} \wedge v_\ell \wedge v_{j_r} \wedge \cdots \wedge v_{j_k} \wedge v_m$$

which is a linear combination of distinct basis elements of $\Lambda^{k+1} W_x$ (see proposition A.6.9 on page 477) and, therefore, nonzero. $\square$

Now we are ready to define Grassmann varieties:

DEFINITION 5.2.10. If V is an n-dimensional vector space with dual V^* and $1 \leq r \leq n$, the *Grassmann variety*

$$\mathbb{G}_r(V) \subset \mathbb{P}(\Lambda^r V)$$

is a projective variety defined by the $\binom{n}{r-1}$ homogeneous quadratic equations

$$(c^i \lrcorner x) \wedge x = 0$$

where the $\{c^i\}$ are a basis for $\Lambda^{r-1} V^*$. The points of $\mathbb{G}_r(V)$ are in a 1-1 correspondence with the r-dimensional subspaces of V. If we are not interested in emphasizing the vector space, V, we use the notation $\mathbb{G}_{r,n}$ for $\mathbb{G}_r(\mathbb{A}^n)$.

REMARK. Grassmannians are a simple example of a *moduli space* — a topological space (or algebraic variety) whose *points* correspond to geometric objects (subspaces of an affine space, in this case). By introducing *parameters* on a moduli space (Plücker coordinates, in this case) we can study properties of geometric objects. Exercises 4 and 5 on page 225 and exercise 3 on page 245 give other examples of moduli spaces.

The $\binom{n}{r-1}$ equations are not necessarily all distinct. For instance, if $r = 2$, and the characteristic of F is not 2, we get

$$c^i \lrcorner (x \wedge x) = 2(c^i \lrcorner x) \wedge x$$

so that

$$(c^i \lrcorner x) \wedge x = 0$$

for all i, implies that

$$c^i \lrcorner (x \wedge x) = 0$$

for all i, and lemma 5.2.5 on page 228 implies that

$$x \wedge x = 0$$

so we only get *one* equation. It follows that $G_2(V)$ is a *hypersurface* in $\mathbb{P}(\Lambda^2 V)$.

Now we will look at some of these varieties. Suppose V is four-dimensional with basis $\{e_1, e_2, e_3, e_4\}$. Then a typical element of $\Lambda^2 V$ is a linear combination

$$(5.2.3) \qquad \begin{aligned} x \;=\; & a_{1,2} e_1 \wedge e_2 + a_{1,3} e_1 \wedge e_3 + a_{1,4} e_1 \wedge e_4 \\ & + a_{2,3} e_2 \wedge e_3 + a_{2,4} e_2 \wedge e_4 + a_{3,4} e_3 \wedge e_4 \end{aligned}$$

and

$$x \wedge x = 2(a_{1,2} a_{3,4} - a_{1,3} a_{2,4} + a_{1,4} a_{2,3}) e_1 \wedge e_2 \wedge e_3 \wedge e_4$$

so the equation defining $G_2(V)$ is

$$(5.2.4) \qquad a_{1,2} a_{3,4} - a_{1,3} a_{2,4} + a_{1,4} a_{2,3} = 0$$

in Plücker coordinates. The vector space $\Lambda^2 V$ is $\binom{4}{2} = 6$ dimensional so $\mathbb{P}(\Lambda^2 V)$ is 5 dimensional and the hypersurface defined by equation 5.2.4 is 4-dimensional, by corollary 2.8.30 on page 111.

If

$$v = b_1 e_1 + b_2 e_2 + b_3 e_3 + b_4 e_4$$

is a vector in V, we get

$$(5.2.5) \qquad x \wedge v \;=\; (b_1 a_{2,3} - b_2 a_{1,3} + b_3 a_{1,2}) e_1 \wedge e_2 \wedge e_3$$
$$(5.2.6) \qquad \qquad + (b_1 a_{2,4} - b_2 a_{1,4} + b_4 a_{1,2}) e_1 \wedge e_2 \wedge e_4$$
$$\qquad \qquad + (b_1 a_{3,4} - b_3 a_{1,4} + b_4 a_{1,3}) e_1 \wedge e_3 \wedge e_4$$
$$\qquad \qquad + (b_2 a_{3,4} - b_3 a_{2,4} + b_4 a_{2,3}) e_2 \wedge e_3 \wedge e_4$$

Given a set of Plücker coordinates, $S = \{a_{1,2}, a_{1,3}, a_{1,4}, a_{2,3}, a_{2,4}, a_{3,4}\}$, satisfying equation 5.2.4 on page 231, we can recover the subspace they represent by setting

$$x \wedge v = 0$$

In light of equation 5.2.5, we get

$$\begin{bmatrix} a_{2,3} & -a_{1,3} & a_{1,2} & 0 \\ a_{2,4} & -a_{1,4} & 0 & a_{1,2} \\ a_{3,4} & 0 & -a_{1,4} & a_{1,3} \\ 0 & a_{3,4} & -a_{2,4} & a_{2,3} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = 0$$

for the subspace that S represents.

Assuming none of the $a_{i,j}$ are zero, reducing this matrix to echelon form (without doing divisions), and applying equation 5.2.4 on the preceding page five times gives:

$$\begin{bmatrix} a_{2,3} & -a_{1,3} & a_{1,2} & 0 \\ 0 & a_{1,2}\,a_{3,4} & -a_{1,2}\,a_{2,4} & a_{1,2}\,a_{2,3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = 0$$

which has a 2-dimensional solution.

5.2.4. Real Grassmannians as affine varieties. I am indebted to Thomas Yu for this example. Consider the space $\mathbb{R}^{n^2}$ of $n \times n$ real matrices, A, satisfying:

(1) $A^2 = A$
(2) $A^T = A$

The first condition implies that A is a projection to a subspace and the second condition implies that this subspace is uniquely determined. If we impose a third condition

$$\text{Trace}(A) = t$$

for an integer $1 \le t < n$, A becomes a projection to a t-dimensional subspace of $\mathbb{R}^n$. Since there is clearly a 1-1 correspondence between such subspaces and matrices satisfying these conditions, we get a representation of the Grassmannian $\mathbb{G}_t(\mathbb{R}^n)$ as a (quadratic) affine variety in $\mathbb{R}^{n^2}$. In fact, the second condition implies that this is an embedding in $\mathbb{R}^{n(n+1)/2}$.

EXAMPLE 5.2.11. If $n > 1$, we can consider the projective space $\mathbb{P}(\mathbb{R}^{n+1}) = \mathbb{R}P^n = \mathbb{G}_1(\mathbb{R}^{n+1})$ in this context. This is the set of all $(n+1) \times (n+1)$ real matrices, A with

(1) $A^2 = A$
(2) $A^T = A$
(3) $\text{Trace}(A) = 1$

Since the image of the projection is one-dimensional, the row or column space is one-dimensional and every column is a multiple of every other.

For instance, if $n = 2$, our A-matrices look like

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{1,2} & A_{2,2} & A_{2,3} \\ A_{1,3} & A_{2,3} & A_{3,3} \end{bmatrix}$$

— we have already applied rule 2. Rules 1 and 3 give algebraic conditions on the elements of A that can be simplified by finding a Gröbner basis:

$$\mathfrak{P} = (-A_{2,2} + A_{2,3}{}^2 + A_{2,2}A_{1,1} + A_{2,2}{}^2,$$
$$-A_{1,1} + A_{1,3}{}^2 + A_{1,1}{}^2 + A_{2,2}A_{1,1},$$
$$A_{1,1}A_{1,2} + A_{1,2}A_{2,2} + A_{1,3}A_{2,3} - A_{1,2},$$
$$-A_{2,2}A_{1,3} + A_{1,2}A_{2,3}, -A_{1,1}A_{2,3} + A_{1,2}A_{1,3},$$
$$-A_{2,2}A_{1,1} + A_{1,2}{}^2, A_{1,1} + A_{2,2} + A_{3,3} - 1)$$

We can see what this matrix is like if, say $A_{1,1} = 1/3$ and $A_{2,2} = 1/3$ by adding terms $A_{1,1} - 1/2$ and $A_{2,2} - 1/3$ to $\mathfrak{P}$ and taking another Gröbner basis

$$
\begin{aligned}
\mathfrak{P} + (A_{1,1}, -1/2, A_{2,2} - 1/3) \quad = \quad & (3\,A_{2,2} - 1, 2\,A_{1,1} - 1, \\
& 18\,A_{2,3}{}^2 - 1, 12\,A_{1,3}{}^2 - 1, \\
& -6\,A_{1,3}A_{2,3} + A_{1,2}, 6\,A_{3,3} - 1)
\end{aligned}
$$

from which we get the matrix

$$
A = \begin{bmatrix}
\frac{1}{2} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} \\
\frac{1}{\sqrt{6}} & \frac{1}{3} & \frac{1}{\sqrt{18}} \\
\frac{1}{\sqrt{12}} & \frac{1}{\sqrt{18}} & \frac{1}{6}
\end{bmatrix}
$$

The affine subspaces $\mathbb{R}_i^2$, $i = 1, 2, 3$, are the sets of matrices for which row and column i are nonzero.

This discussion raises an interesting question:

> Proposition 4.4.9 on page 187 show that the only regular functions on projective space (which is a Grassmannian) are *constants,* yet affine varieties have *plenty* of nonconstant regular functions. Isn't this a contradiction?

The answer comes from remembering how functions on a scheme are evaluated. Our real projective space is actually the set of *real points* of $\mathbb{RP}^n$ or $\mathbb{RP}^n(\mathbb{R}) \subset \mathbb{RP}^n$ (as in definition 4.4.8 on page 186)

Recall that regular functions on *subspaces* of $\mathbb{RP}^n$ are defined as rational functions, f, whose denominators are nonvanishing on the subspace. The set of real points, $\mathbb{RP}^n(\mathbb{R})$, has plenty of these — for instance

$$
\frac{X^2 - X + 5}{X^2 + 1}
$$

This is not a regular function on $\mathbb{RP}^n$ because the denominator gets evaluated in $\mathbb{C}$ (see definition 4.2.3 on page 164) where it *does* vanish.

Consider the case where $n = 1$. The space $\mathbb{RP}^1(\mathbb{R})$ is well known to be homeomorphic to S^1, the unit circle. The map is given by

$$
(5.2.7) \qquad\qquad (X_0 : X_1) \mapsto \left(\frac{X_0^2 - X_1^2}{X_0^2 + X_1^2}, \frac{2X_0 X_1}{X_0^2 + X_1^2} \right)
$$

so regular functions on S^1 pull back to rational functions that are *homogeneous* of degree 0 on $\mathbb{RP}^1(\mathbb{R})$. This is a birational equivalence that is nowhere singular (over *real* points).

Another way to think of this: over an algebraically closed field, a rational function is everywhere nonsingular if and only if it is a polynomial. Over other non-closed fields, this correspondence breaks down.

In real algebraic geometry, all projective varieties are also affine — see [18] for more on this fascinating field.

EXERCISES.

1. If V is n-dimensional, show that the interior product
$$\llcorner : \Lambda^k V^* \otimes_F \Lambda^n V \to \Lambda^{n-k} V$$
is an isomorphism for all $0 \leq k \leq n$. This induces an isomorphism
$$\mathbf{G}_k(V^*) \to \mathbf{G}_{n-k}(V)$$

2. Find an inverse to the map in equation 5.2.7 on the preceding page.

3. Show the set of hyperplanes in $k\mathbb{P}^n$ is parametrized by a projective space.

4. Find a Grassmannian that parametrizes lines in $k\mathbb{P}^3$.

5.3. Invertible sheaves on projective varieties

This presents Serre's (largely successful) efforts to extend his theory of vector-bundles (or locally free sheaves) onto projective varieties. Of course every projective variety is a union of affine ones, and vector bundles could simply be described locally. The interesting aspect of this subject is that it is possible to get a *global* description of locally free sheaves on projective varieties, somewhat like that in section 3.5 on page 154. As one might expect, the projective case is much more complicated.

We begin this section with *Serre twists* — line-bundles (see definition C.1.3 on page 507) defined by Serre in [**150**]:

DEFINITION 5.3.1. Let k be a field and let
$$k\mathbb{P}^n = \bigcup_{i=0}^{n} \mathbb{A}_i^n$$
be the standard decomposition into open affines — where $[X_0 : \cdots : X_n]$ are the homogeneous coordinates of $k\mathbb{P}^n$ and $\mathbb{A}_i^n$ has $X_i \neq 0$. If $m \in \mathbb{Z}$, define the *Serre twist*, $\mathcal{O}_{k\mathbb{P}^m}(m)$, to be the line bundle, ξ, defined by gluing together the $\{\mathbb{A}_i^n \times k\}$ via the transition functions
$$1 \times \varphi_{j,i} = 1 \times \left(\frac{X_i}{X_j} \right)^m : \mathbb{A}_i^n \cap \mathbb{A}_j^n \times k \to \mathbb{A}_j^n \cap \mathbb{A}_i^n \times k$$

REMARK. The term Serre twist is usually applied to the *invertible sheaf* associated to this vector bundle — which is the reason for the notation, $\mathcal{O}_V(m)$.

DEFINITION 5.3.2. The *Serre-twist sheaf* is defined over $R\mathbb{P}^n$ by
$$\mathcal{O}_V(m)(\mathbb{A}_i^n) = \mathcal{O}_{R\mathbb{P}^n}(\mathbb{A}_i^n)$$
and these are glued together via homomorphisms
$$\begin{aligned}
\mathcal{O}_V(m)(\mathbb{A}_i^n)|\mathbb{A}_i^n \cap \mathbb{A}_j^n &\to \mathcal{O}_V(m)(\mathbb{A}_j^n)|\mathbb{A}_j^n \cap \mathbb{A}_i^n \\
f &\mapsto f \cdot (X_i/X_j)^m
\end{aligned}$$
We will use $\mathcal{O}_{R\mathbb{P}^n}(m)$ for both the line bundle and invertible sheaf when R is a field.

REMARK. Again, the notation $\mathbb{A}_j^n \cap \mathbb{A}_i^n$ represents the intersection *regarded as part of the* $\mathbb{A}_j^n$ *chart.*

Serre defined these sheaves in his groundbreaking paper, [**150**]. Most of our results regarding these sheaves also came from that paper.

Although the definition above describes the sheaf, it doesn't quite compute it:

PROPOSITION 5.3.3. *If* $\mathscr{F} = \mathcal{O}_{R\mathbb{P}^n}(d)$ *for* $d \in \mathbb{Z}$, *then*

(1) the global evaluation is

$$\mathcal{O}_{R\mathbb{P}^n}(d)(R\mathbb{P}^n) = R[X_0, \ldots, X_n]_d$$

— the free R-module generated by homogeneous monomials of degree d,

(2) over the open sets $\mathbb{A}_{i_1}^n \cap \cdots \cap \mathbb{A}_{i_t}^n$, *we have*

$$\mathcal{O}_{R\mathbb{P}^n}(d)(\mathbb{A}_{i_1}^n \cap \cdots \cap \mathbb{A}_{i_t}^n) = R[X_0, \ldots, X_n][X_{i_1}^{-1}, \ldots, X_{i_t}^{-1}]_d$$

— the free R-module generated by homogeneous monomials of degree d, (where some of the X's may have negative degree).

The restrictions $\mathcal{O}_{R\mathbb{P}^n}(d)|\mathbb{A}_i^n$ *are coherent, so* $\mathcal{O}_{R\mathbb{P}^n}(d)$ *is quasi-coherent. If* $U_1, U_2 \subset R\mathbb{P}^n$ *are open sets, then*

$$\mathcal{O}_{R\mathbb{P}^n}(d)(U_1) \hookrightarrow \mathcal{O}_{R\mathbb{P}^n}(d)(U_1 \cap U_2 \cap \mathbb{A}_0^n \cap \cdots \cap \mathbb{A}_n^n)$$
$$\mathcal{O}_{R\mathbb{P}^n}(d)(U_2) \hookrightarrow \mathcal{O}_{R\mathbb{P}^n}(d)(U_1 \cap U_2 \cap \mathbb{A}_0^n \cap \cdots \cap \mathbb{A}_n^n)$$

and, as submodules of $\mathcal{O}_{R\mathbb{P}^n}(d)(U_1 \cap U_2 \cap \mathbb{A}_0^n \cap \cdots \cap \mathbb{A}_n^n)$

$$\mathcal{O}_{R\mathbb{P}^n}(d)(U_1 \cup U_2) = \mathcal{O}_{R\mathbb{P}^n}(d)(U_1) \cap \mathcal{O}_{R\mathbb{P}^n}(d)(U_2)$$

REMARK. This implies that the vector-bundle $\mathcal{O}_{k\mathbb{P}^n}(d)$ has no nonzero global sections if $d < 0$.

It is not hard to see (by multiplication of gluing maps) that

(5.3.1) $$\mathcal{O}_{R\mathbb{P}^n}(r) \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(s) = \mathcal{O}_{R\mathbb{P}^n}(r+s)$$

with $\mathcal{O}_{R\mathbb{P}^n}(0) = \mathcal{O}_{R\mathbb{P}^n}$ as the unit so that the sheaves $\mathcal{O}_{R\mathbb{P}^n}(d)$ form a group under tensor-product that is isomorphic to $\mathbb{Z}$.

PROOF. To prove the first statement, note that an element of $\mathcal{O}_{R\mathbb{P}^n}(d)(R\mathbb{P}^n)$ is given by a set of functions g_j on $\mathbb{A}_j^n$ that are compatible with the gluing map. Define a homomorphism

$$R[X_0, \ldots, X_n]_d \rightarrow \mathcal{O}_{R\mathbb{P}^n}(d)(\mathbb{A}_i^n)$$
$$f(X_0, \ldots, X_n) \mapsto f/X_i^d$$
$$= f\left(\frac{X_0}{X_i}, \ldots, \frac{X_{i-1}}{X_i}, 1, \frac{X_{i+1}}{X_i}, \ldots, \frac{X_n}{X_i}\right)$$

This is clearly injective (if applied to *all* of the $\mathbb{A}_i^n$) and compatible with the gluing functions in the remarks following definition 5.3.1 on the preceding page. $\square$

On the other hand, given polynomials $\{g_i\}$, the condition that

$$g_i\left(\frac{X_0}{X_i}, \ldots, \frac{X_{i-1}}{X_i}, \frac{X_{i+1}}{X_i}, \ldots, \frac{X_n}{X_i}\right) \cdot \left(\frac{X_i}{X_j}\right)^d = g_j\left(\frac{X_0}{X_j}, \ldots, \frac{X_{j-1}}{X_j}, \frac{X_{j+1}}{X_j}, \ldots, \frac{X_n}{X_j}\right)$$

implies that all of the $\{g_i\}$ are of degree d. If they *are* of degree d and satisfy this gluing condition, they are in the image of some homogeneous f of degree d.

The second statement follows from the fact that $\mathcal{O}_{R\mathbb{P}^n}(d)|\mathbb{A}_i^n$ is isomorphic to a *free sheaf*

$$\mathcal{O}_{R\mathbb{P}^n}(d)(\mathbb{A}_i^n) = R[x_1, \ldots, x_n]$$

— i.e., no gluing is done. To get a form of this compatible with gluing maps, define an isomorphism of free R-modules

$$
\begin{aligned}
R[x_1, \ldots, x_n] &\rightarrow R[X_0, \ldots, X_n][X_i^{-1}]_d \\
g(x_1, \ldots, x_n) &\mapsto X_i^d \cdot g\left(\frac{X_0}{X_i}, \ldots, \frac{X_{i-1}}{X_i}, \frac{X_{i+1}}{X_i}, \ldots, \frac{X_n}{X_i}\right)
\end{aligned}
$$

We conclude that

$$
\begin{aligned}
\mathcal{O}_{R\mathbb{P}^n}(d)(\mathbb{A}_{i_1}^n \cap \cdots \cap \mathbb{A}_{i_t}^n) &= R[x_1, \ldots, x_n][x_{i_2}^{-1} \cdots x_{i_t}^{-1}] \\
&= R[X_0, \ldots, X_n][X_{i_1}^{-1}]_d[X_{i_2}^{-1}X_{i_1}, \ldots, X_{i_n}^{-1}X_{i_1}] \\
&= R[X_0, \ldots, X_n][X_{i_1}^{-1}, \ldots, X_{i_t}^{-1}]_d
\end{aligned}
$$

The remaining statements follow from the fact that every open set is dense in $R\mathbb{P}^n$.

EXAMPLE 5.3.4. Proposition 5.3.3 on the preceding page implies that the sheaf-completion operation (lemma B.2.3 on page 500) is actually *necessary* in computing a tensor product (in definition B.3.5 on page 504):

The sheaf $\mathcal{O}_{R\mathbb{P}^1}(1)$ has $\mathcal{O}_{R\mathbb{P}^1}(1)(R\mathbb{P}^1) = R \cdot X_0 \oplus R \cdot X_1$. If we compute the tensor-product

$$\mathcal{O}_{R\mathbb{P}^1}(1) \otimes_{\mathcal{O}_{R\mathbb{P}^1}} \mathcal{O}_{R\mathbb{P}^1}(1)$$

treating both factors as *presheaves*, we get a presheaf P with

$$P(R\mathbb{P}^1) = R \cdot X_0 \otimes X_0 \oplus R \cdot X_0 \otimes X_1 \oplus R \cdot X_1 \otimes X_0 \oplus R \cdot X_1 \otimes X_1 = R^4$$

On the other hand — as *sheaves*

$$\mathcal{O}_{R\mathbb{P}^1}(1) \otimes_{\mathcal{O}_{R\mathbb{P}^1}} \mathcal{O}_{R\mathbb{P}^1}(1) = \mathcal{O}_{R\mathbb{P}^1}(2)$$

and

$$\mathcal{O}_{R\mathbb{P}^1}(2)(R\mathbb{P}^1) = R \cdot X_0^2 \oplus R \cdot X_0 X_1 \oplus R \cdot X_1^2 = R^3$$

This example shows that the presheaf, P, is not even *separated* — $P(R\mathbb{P}^1)$ contains *distinct* elements whose restrictions to *every* open affine are the *same* (namely $X_0 \otimes X_1$ and $X_1 \otimes X_0$).

DEFINITION 5.3.5. If $f: V \rightarrow R\mathbb{P}^n$ is a closed immersion of a projective scheme, the pullback

$$\mathscr{L}_V = f^*(\mathcal{O}_{R\mathbb{P}^n}(1))$$

is called a *very ample sheaf* on V. It is often written $\mathcal{O}_V(1)$.

REMARK. When a result mentions that a variety or scheme has a very ample sheaf, this often merely says that there *exists* a closed immersion into a projective space.

The Serre twists play an important part in the theory of invertible sheaves on a $R\mathbb{P}^n$:

DEFINITION 5.3.6. If $(X, \mathcal{O}_X)$ is a ringed space, a sheaf $\mathcal{F}$ of $\mathcal{O}_X$ modules is said to be *generated by global sections* if there exist a set of indices, I $f_i \in \mathcal{F}(X)$ for $i \in I$, and a surjection

$$\bigoplus_{i \in I} \mathcal{O}_X \;\; \rightarrow \;\; \mathcal{F}$$
$$\oplus s_i \;\; \mapsto \;\; \sum s_i \cdot f_i$$

REMARK. For instance, any sheaf of the form $\mathcal{A}(M)$ (see definition 3.5.1 on page 155) is generated by global sections — just take a set of generators of M.

One way to characterize such sheaves is:

PROPOSITION 5.3.7. *If $(X, \mathcal{O}_X)$ is a ringed space, a sheaf $\mathcal{F}$ of $\mathcal{O}_X$ modules is generated by global sections if and only if there exists a set $f_i \in \mathcal{F}(X)$ such that $\mathcal{F}_x$ is generated by $f_i|x$ (as a module over $\mathcal{O}_{X,x}$) for all $x \in X$.*

PROOF. This follows from the fact that a map of sheaves is surjective if and only if it is surjective on *stalks* (see exercise 6 on page 502). $\square$

We have an analogue of lemma 4.4.17 on page 193:

LEMMA 5.3.8. *If $\mathcal{F}$ is a quasi-coherent sheaf on a quasi-compact scheme, X, $\mathcal{L}$ is an invertible sheaf and $f \in \mathcal{L}(X)$, let X_f be the open set where f is nonvanishing.*

(1) *if $s \in \mathcal{F}(X)$ has the property that $s|X_f = 0$, then there exists an integer n such that*

$$s \cdot f^n = 0 \in (\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{L}^n)(X)$$

(2) *if $\{U_i\}$ is an open affine cover of X such that each intersection $U_i \cap U_j$ is quasi-compact and $\mathcal{L}|U_i$ is free, and $s \in \mathcal{F}(X_f)$, then there exists an integer n such that*

$$f^n \cdot s \in (\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{L}^n)(X_f)$$

is the restriction of an element $s' \in (\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{L}^n)(X)$.

REMARK. Recall that "quasi-compact"="every open covering has a finite sub-cover." In other word, when one must cover a portion of X, the number of open sets required is always finite. Affine schemes are *automatically* quasi-compact (see corollary 4.3.9 on page 172).

PROOF. The restriction of $\mathcal{F}$ to each of the U_i is coherent — see 4.4.18 on page 194. The proof, here, follows an argument virtually identical to that used in the proof of lemma 4.4.18 on page 194 with a little care to account for f being in a sheaf. $\square$

We have an immediate consequence

COROLLARY 5.3.9. *If $\mathcal{F}$ is a finitely-generated coherent module over $\mathcal{O}_V$ on a scheme V and $\mathcal{L} = \mathcal{O}_V(1)$ is a very ample sheaf, then there exists an integer $n > 0$ such that*

$$\mathcal{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(1)^d$$

is generated by global sections for all $d \geq n$.

PROOF. The *existence* of $\mathscr{L}$ implies the existence of a closed immersion

$$f : V \hookrightarrow R\mathbb{P}^t$$

for some t. The open affines $\{\mathbb{A}_i^t\}$, $i = 0, \dots, t$ defines open affines of V, namely

$$X_i = f^{-1}\left(f(V) \cap \mathbb{A}_i^t\right)$$

Proposition 5.3.3 on page 235 implies that the twist, $\mathcal{O}_{R\mathbb{P}^t}(1)$, has global evaluations $X_i \in \mathcal{O}_{R\mathbb{P}^t}(1)(R\mathbb{P}^t)$ and $\mathbb{A}_i^t$ is defined by the nonvanishing of X_i. It follows that the open affine V_i is defined by the nonvanishing of a function $f_i \in \mathscr{L}(V)$ that maps to X_i. We have

$$\mathscr{F}|V_i = \mathcal{A}(M_i)$$

for a finitely generated module M_i over $\mathcal{O}_V(V_i)$ (see definition 3.5.1 on page 155 and lemma 3.5.6 on page 156). Suppose $\{m_{i,j}\}$ is a finite set of generators of M_i.

Lemma 5.3.8 on the preceding page implies that there exists $n_{i,j}$ such that $f_i^{n_{i,j}} \cdot m_{i,j}$ is the restriction of something in $(\mathscr{F} \otimes_{\mathcal{O}_V} \mathscr{L}^{n_{i,j}})(V)$. If n is the maximum of all of the $\{n_{i,j}\}$, it has the required properties. $\square$

In projective spaces, Serre twists play an important part:

LEMMA 5.3.10. *Let $\mathscr{F}$ be a coherent sheaf on $R\mathbb{P}^n$. Then there exists a surjective map*

$$\bigoplus_{j=1}^{t} \mathcal{O}_{R\mathbb{P}^n}(d) \to \mathscr{F}$$

for some value of t, and d. Moreover, d can be chosen arbitrarily small.

REMARK. Why is this interesting? Over an affine variety, A, we already know that coherent sheaves are surjective images of $\mathcal{O}_A^t$ for some t (see definition 3.5.5 on page 155). Over a general variety or scheme, we can only say something like this in *open affines*.

The present result is interesting because it says something like this is true for *projective spaces* (even though they are not affine). Over a projective space, one cannot simply use $\mathcal{O}_{R\mathbb{P}^t}$ — one must be willing to use Serre twists.

PROOF. Corollary 5.3.9 on page 237 implies that

$$\mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(m)$$

is generated by global sections, say $s_1, \dots, s_t$. Then there exists a surjective map

$$\bigoplus_{j=1}^{t} \mathcal{O}_{R\mathbb{P}^n} \quad \to \quad \mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(m)$$
$$\oplus r_i \quad \mapsto \quad \sum r_i \cdot s_i$$

of sheaves. If we take the tensor product with $\mathcal{O}_{R\mathbb{P}^n}(-m)$ we get a surjection

$$\bigoplus_{j=1}^{t} \mathcal{O}_{R\mathbb{P}^n} \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(-m) \to \mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(m) \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(-m)$$

since $\otimes$ is right-exact (see proposition A.5.55 on page 467). It is not hard to see that

$$\bigoplus_{j=1}^{t} \mathcal{O}_{R\mathbb{P}^n} \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(-m) = \bigoplus_{j=1}^{t} \mathcal{O}_{R\mathbb{P}^n}(-m)$$

and

$$\mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(m) \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(-m) = \mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^n}} \mathcal{O}_{R\mathbb{P}^n}(0) = \mathscr{F}$$

(see equation 5.3.1 on page 235) so the conclusion follows. Since m can be arbitrarily large in corollary 5.3.9 on page 237, $d = -m$ can be arbitrarily small. $\square$

We consider a more abstract version of Serre-twists on schemes of the form $\operatorname{Proj} R$ (see definition 4.4.12 on page 190) for a graded ring, R.

We begin with Serre's generalization of his $A(*)$-construction:

DEFINITION 5.3.11. Let M be a graded module over a graded ring, R, and let $x \in R$ be a homogeneous element. Then $M_{(x)}$ is defined to be the degree-0 elements of the localization M_x.

REMARK. This is nothing but a *module*-version of definition 4.4.13 on page 190.

DEFINITION 5.3.12. If M is a graded module over a graded ring, R, define $\mathcal{A}(M)$ to be the sheaf over $\operatorname{Proj} R$ defined by

$$\mathcal{A}(M)|D(x) = \mathcal{A}(M_{(x)})$$

where $D(x)$ is the *affine* scheme, $\operatorname{Spec} R_{(x)}$, in proposition 4.4.14 on page 190.

Now we can generalize the definition of $\mathcal{O}_X$ given above:

DEFINITION 5.3.13. If R is a graded module and $V = \operatorname{Proj} R$, then define

$$\mathcal{O}_V(n) = \mathcal{A}(R(n))$$

for n an integer ≥ 0, where $R(n)$ is the graded R-module

$$R(n) = \bigoplus_{j=n}^{\infty} R_j$$

REMARK. These first appeared in [150]. It is not hard to see that $\mathcal{O}_V = \mathcal{O}_V(0)$: on the open set, $\operatorname{Spec} R_{(x)}$, for any homogeneous $x \in R$ we get $\mathcal{O}_V(0)|\operatorname{Spec} R_{(x)} = \mathcal{A}(R_{(x)})$, which is the definition of $\mathcal{O}_{\operatorname{Spec} R_{(x)}}$ (see exercise 4 on page 177).

PROPOSITION 5.3.14. *If R is a graded ring that is finitely-generated by R_1 as an R_0-algebra and $V = \operatorname{Proj} R$, then:*

(1) $\mathcal{O}_V(n)$ is an invertible sheaf (i.e., locally-free of rank 1).
(2) $\mathcal{O}_V(n) \otimes_{\mathcal{O}_V} \mathcal{O}_V(m) = \mathcal{O}_V(n+m)$

REMARK. The proof shows that for different values of n, the $\mathcal{O}_V(n)$ mainly differ from one another in their gluing maps — or how they behave on overlaps of open-sets.

Even if the hypotheses are not satisfied, there exists a homomorphism

$$\mathcal{O}_V(n) \otimes_{\mathcal{O}_V} \mathcal{O}_V(m) \to \mathcal{O}_V(n+m)$$

that may fail to be an isomorphism. It is induced by the homomorphism of modules

$$R(n) \otimes R(m) \to R(n+m)$$

given by multiplication.

PROOF. If $x \in R_1$, then on the open affine

$$\mathcal{O}_V(n)|D(x) = \mathcal{A}(R(n)_{(x)})$$

Since x is of degree 1, multiplication by x^n defines a homomorphism of R-modules

$$x^n \cdot : R \to R(n)$$

Since x is invertible in R_x, we get an *isomorphism* of R_x-modules

$$x^n \cdot : R_x \to R(n)_x$$

which preserves elements of degree 0. It follows that $\mathcal{O}_V(n)|D(x)$ is free of rank 1. Since R is generated by elements of degree 1, these open sets cover V.

The remaining statement is also verified on each open set $D(x)$, where $x \in R_1$:

$$\mathcal{A}(R(n)_{(x)}) \otimes_{\mathcal{O}_V} \mathcal{A}(R(m)_{(x)}) = \mathcal{A}(R(n)_{(x)} \otimes R_{(x)} R(m)_{(x)}) = \mathcal{A}(R(n+m)_{(x)})$$

by the argument used above. □

The comparison of gluing maps implies the comforting conclusion that:

PROPOSITION 5.3.15. *If $S = R[X_0, \ldots, X_n]$ and $V = \operatorname{Proj} R$, then the Serre twists, $\mathcal{O}_V(n)$ in definition 5.3.13 on the preceding page are isomorphic to those in definition 5.3.2 on page 234.*

PROOF. The proof that $\mathcal{O}_V(n)$ is locally free of rank 1 show that, on the overlap $\mathbb{A}_i^n \cap \mathbb{A}_j^n$, the gluing map is given by

$$S_{(X_i)} \xrightarrow{X_i^n \cdot} S(n)_{(X_i \cdot X_j)} \xrightarrow{X_j^{-n}} S_{(X_j)}$$

or $\left(X_i / X_j\right)^n$. □

DEFINITION 5.3.16. If $V = \operatorname{Proj} R$ and $\mathscr{F}$ is a module over $\mathcal{O}_V$, then define the *graded module associated* to $\mathscr{F}$ to be

$$\Gamma_*(\mathscr{F}) = \bigoplus_{n=0}^{\infty} \left(\mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)\right)(V)$$

REMARK. If $u \in \left(\mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)\right)(V)$ and $x \in R_d$ determines an element of $\mathcal{O}_V(d)(V)$ then $x \cdot u$ is the image of $u \otimes x$ under the mapping

$$\mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n) \otimes_{\mathcal{O}_V} \mathcal{O}_V(d) \to \mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n+d)$$

In some cases, the graded module recovers important information about the underlying ring:

PROPOSITION 5.3.17. *If $R = S[X_0, \ldots, X_n]$ and $V = \operatorname{Proj} R$, then*

$$\Gamma_*(\mathcal{O}_V) = R$$

PROOF. This immediately follows from propositions 5.3.15 and 5.3.3 on page 235, which show that

$$\mathcal{O}_V(d)(V) = S[X_0, \ldots, X_n]_d$$

the subgroup of degree-d monomials. The direct sum is all of R. □

PROPOSITION 5.3.18. *If S is a graded ring with is finitely generated by S_1 as an algebra over S_0, $V = \operatorname{Proj} S$, and $\mathscr{F}$ is a quasi-coherent sheaf on V, then there is a natural isomorphism*

$$\beta : \mathcal{A}(\Gamma_*(\mathscr{F})) \to \mathscr{F}$$

PROOF. We begin by *defining* β for a module over $\mathcal{O}_V$. If $x \in S_1$ we define

$$\mathcal{A}(\Gamma_*(\mathscr{F}))|D(x) \to \mathscr{F}|D(x)$$

by specifying an $S_{(x)}$-module-homomorphism

$$\bar{\beta}: \Gamma_*(\mathscr{F})_{(x)} \to \mathscr{F}(D(x))$$

Elements of $\Gamma_*(\mathscr{F})_{(x)}$ are of the form f/x^n where $f \in \left(\mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n) \right)(V)$ and $x^{-n} \in \mathcal{O}_V(-n)(D(x))$, so

$$f/x^n \in \left(\mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n) \otimes_{\mathcal{O}_V} \mathcal{O}_V(-n) \right)(D(x)) \to \mathscr{F}(D(x))$$

and this induces $\beta|D(x)$.

Having *defined* β, we now need to show that it is an *isomorphism* for a quasi-coherent sheaf. If $\mathscr{F}$ is quasi-coherent, this involves showing that

$$\bar{\beta}: \Gamma_*(\mathscr{F})_{(x)} \to \mathscr{F}(D(x))$$

is an isomorphism for $x \in S_1$. Set $\mathscr{L} = \mathcal{O}_V(1)$ so $\mathcal{O}_V(n) = \mathscr{L}^n$ — a sheaf that is invertible and free on each open set $D(x_i)$ where the $x_i \in S_1$ generate S.

Statement 1 of lemma 5.3.8 implies that $\bar{\beta}$ is *injective,* since any $s \in \mathscr{F}(V)$ with $s|D(x_i) = 0$ must have the property that $s \cdot x_i^n = 0 \in (\mathscr{F} \otimes_{\mathcal{O}_V} \mathscr{L}^n)(V)$ for some n so s is in the kernel of $\Gamma_*(\mathscr{F}) \to \Gamma_*(\mathscr{F})_{(x_i)}$.

Statement 2 of lemma 5.3.8 shows that every $f \in \mathscr{F}(D(x_i))$ has the property that $x_i^k \cdot f \in (\mathscr{F} \otimes_{\mathcal{O}_V} \mathscr{L}^k)(D(x_i))$ is the restriction of $s \in (\mathscr{F} \otimes_{\mathcal{O}_V} \mathscr{L}^k)(V)$ for some k. It follows that $f = s/x_i^k \in (\mathscr{F} \otimes_{\mathcal{O}_V} \mathscr{L}^k)(V) \cdot x_i^{-k} \subset \Gamma_*(\mathscr{F})_{(x)}$, making $\bar{\beta}$ *surjective.* $\square$

The Proj-construction allows us to describe *all* projective varieties and schemes in a compact way:

COROLLARY 5.3.19. *If $W \subset V = R\mathbb{P}^n$ is a closed subscheme, then there exists a graded ring S with $S_0 = R$ such that*

$$W = \mathrm{Proj}\, S$$

PROOF. According to definition 4.4.21 on page 195, there exists an quasi-coherent ideal $\mathscr{I} \subset \mathcal{O}_{R\mathbb{P}^n}$ that defines W. Since $\Gamma_*(\mathcal{O}_V) = R[X_0, \ldots, X_n]$ and $\mathscr{I} \subset \mathcal{O}_V$, we get

$$\Gamma_*(\mathscr{I}) \subset \Gamma_*(\mathcal{O}_V) = R[X_0, \ldots, X_n]$$

is a graded ideal. The ideal, $\Gamma_*(\mathscr{I})$, determines a closed subscheme,

$$\mathcal{P}(\Gamma_*(\mathscr{I})) = \mathrm{Proj}(S) \subset V$$

(see definition 4.4.15 on page 191) with $S = R[X_0, \ldots, X_n]/\Gamma_*(\mathscr{I})$ whose sheaf of ideals is $\mathcal{A}(\Gamma_*(\mathscr{I}))$ (see definition 5.3.13 on page 239). Proposition 5.3.18 on the preceding page implies that $\mathcal{A}(\Gamma_*(\mathscr{I})) = \mathscr{I}$ so that $W = \mathrm{Proj}\, S$. $\square$

EXAMPLE 5.3.20. Given a ring R and integers n and d, we can define the Grassmannian over R as

$$R\mathbb{G}_{n,d} = \mathrm{Proj}\, R[X_1, \ldots, X_{\binom{n}{d}}]/P$$

where P is the homogeneous ideal of Plücker relations (see definition 5.2.10 on page 230).

EXERCISES.

1. Show that a separated presheaf, $\mathscr{F}$, generated by global sections is actually a sheaf — so sheaves generated by global sections have this property in common with flasque sheaves.

2. If $\mathscr{L}$ is a very ample sheaf on a variety V, show that there exists $f_0, \ldots, f_t \in \mathscr{L}(V)$, for some value of t, that define a closed immersion

$$
\begin{aligned}
V &\to k\mathbb{P}^t \\
x &\mapsto (f_0(x) \colon \cdots \colon f_t(x))
\end{aligned}
$$

This is often used as the definition of a very ample sheaf.

5.4. Regular and rational maps

5.4.1. Definitions. We could define them as we would for general varieties:

DEFINITION 5.4.1. Let $V \subset k\mathbb{P}^n$ be a projective variety with standard decomposition (as in proposition 5.1.11 on page 224) $V = \bigcup_{j=0}^n V_j$. Then we define a function

$$ f \colon V \to k $$

to be *regular* or *rational* if the $f|V_i$ are, respectively, regular or rational for all $i = 0, \ldots, n$.

REMARK. This is just the definition of a regular function on a general variety — see 4.4.2 on page 183.

When we are mapping a projective variety to a *projective space,* we can get a somewhat intrinsic definition:

PROPOSITION 5.4.2. *Let $V \subset k\mathbb{P}^n$ be a projective variety with standard decomposition (as in proposition 5.1.11 on page 224) $V = \bigcup_{j=0}^n V_j$. Then a regular map*

$$ f \colon V \to k\mathbb{P}^m $$

is given by a sequence of regular functions

$$ f = (f_0 \colon \cdots \colon f_m) $$

that never simultaneously vanish on V. These, in turn, can be defined by a sequence of homogeneous polynomials of the same degree (that never simultaneously vanish on V)

$$ f = (F_0(X_0, \ldots, X_n) \colon \cdots \colon F_m(X_0, \ldots, X_n)) $$

where $X_0, \ldots, X_n$ are homogeneous coordinates in $k\mathbb{P}^n$. Two such sequences of homogeneous polynomials

$$ (F_0(X_0, \ldots, X_n) \colon \cdots \colon F_m(X_0, \ldots, X_n)) $$
$$ (G_0(X_0, \ldots, X_n) \colon \cdots \colon G_m(X_0, \ldots, X_n)) $$

define the same regular map if and only if

$$ F_i G_j = F_j G_i $$

for all $i, j = 0, \ldots, m$.

Rational mappings

$$f\colon V \to k\mathbb{P}^m$$

are defined by a sequence

$$f = (f_0\colon \cdots \colon f_m)$$

of rational functions, at least one of which is not identically zero on V. These, in turn, can be defined by a sequence of homogeneous polynomials of the same degree such that at least one polynomial does not identically vanish on V

$$f = (F_0(X_0, \ldots, X_n)\colon \cdots \colon F_m(X_0, \ldots, X_n))$$

REMARK. Note that the definitions of regular and rational functions in terms of homogeneous coordinates are very similar. Regular maps are essentially rational maps that are regular at every point of V.

The homogeneous polynomials in a *rational* map may simultaneously vanish at some points of V (where the map becomes undefined). This means that a rational map is *not necessarily well-defined* at every point of V — it is well defined (and regular) on a dense open subset (in analogy with rational maps of affine varieties).

PROOF. The first statement is clear. The second follows by reasoning exactly like that used in proposition 1.2.6 on page 5, applied to each of the $V_i = V \cap \mathbb{A}_i^n$. The third statement follows from the definition of homogeneous coordinates. $\square$

The following example shows that we can regard regular and rational functions as corresponding maps:

EXAMPLE 5.4.3. We can regard a regular function on $k\mathbb{P}^n$ as a regular map

$$k\mathbb{P}^n \to k\mathbb{P}^1$$

whose image lies in $\mathbb{A}_0^1 \subset k\mathbb{P}^1$. This is given by a sequence of homogeneous polynomials of the same degree

$$(F_0(X_0, \ldots, X_n)\colon F_1(X_0, \ldots, X_n))$$

such that $F_1(X_0, \ldots, X_n)$ *never vanishes* on $\mathbb{A}^{n+1}$. The Nullstellensatz tells us that such a function must be a nonzero *constant.* Since both homogeneous polynomials must be of the same degree, we conclude that $F_0(X_0, \ldots, X_n)$ is also a constant, and the only regular functions on $k\mathbb{P}^n$ are constants.

Any sequence

$$(F_0(X_0, \ldots, X_n)\colon F_1(X_0, \ldots, X_n))$$

of homogeneous polynomials of the same degree (where at least one of them is not identically 0) defines a rational map

$$\mathbb{P}(\mathbb{A}^{n+1}) \to \mathbb{P}(\mathbb{A}^2)$$

This can be regarded as a rational map to $\mathbb{A}_0^1 \subset \mathbb{P}(\mathbb{A}^2)$ if $F_1 \neq 0$, in which case the rational function is

$$\frac{F_0(X_0, \ldots, X_n)}{F_1(X_0, \ldots, X_n)}$$

We conclude that rational functions on $\mathbb{P}(\mathbb{A}^{n+1})$ are quotients of homogeneous polynomials of the *same degree.*

DEFINITION 5.4.4. If $V \subset k\mathbb{P}^n$ and $W \subset k\mathbb{P}^m$, then a *regular* (or *rational*) map
$$f: V \to W$$
is a regular (respectively, rational) map
$$f: V \to k\mathbb{P}^m$$
whose image lies in W. The varieties V and W are *isomorphic* (respectively, *birationally equivalent*) if there exist regular (rational) maps $f: V \to W$ $g: W \to V$ such that $f \circ g = 1: W \to W$ and $g \circ f = 1: V \to V$.

The following example will be very useful in the future:

EXAMPLE 5.4.5. Let $V \subset k\mathbb{P}^n$ be the d-dimensional linear subspace defined by linearly independent linear forms
$$L_1 = \cdots = L_{n-d} = 0$$
Then the *projection*
$$\pi_V: k\mathbb{P}^n \to k\mathbb{P}^{n-d-1} = \mathbb{P}(\mathbb{A}^{n-d})$$
with center V is defined by
$$\pi_V(x) = (L_1: \cdots : L_{n-d})$$
is a rational map that is regular on $k\mathbb{P}^n \setminus V$.

EXERCISES.

1. If $L \subset \mathbb{P}(\mathbb{A}^{n+1})$ is a hyperplane (i.e., a variety defined by a single linear equation), show that $\mathbb{P}(\mathbb{A}^{n+1}) \setminus L$ is affine.

5.4.2. The Veronese embedding. In this section, we discuss a very important class of regular maps named after the mathematician who discovered them:

Giuseppe Veronese (1854–1917) was an Italian mathematician responsible for much of what we know about projective and non-Archimidean geometry (see [**24**] and [**31**]). He also did research on transfinite numbers and was a contemporary (rival?) of Georg Cantor.

We begin with an elementary fact:

PROPOSITION 5.4.6. *If S is the set of degree-m monomials $X_0^{d_0} \cdots X_n^{d_n}$, with all of the $d_i \geq 0$ and*
$$d_1 + \cdots + d_n = m$$
then the number of elements in S is
$$\binom{n+m}{n}$$

PROOF. Consider sequences of m 1's and n dividers, $|$,

$$(1,1,1,|,1,1,\ldots,|,1,1)$$

The exponent of X_0 is the number of 1's before the first divider, and that of X_i the number between the ith divider and the next. There is clearly a one-to-one correspondence between the monomials and these sequences. $\square$

Now let $\mathbb{P}(\mathbb{A}^N)$ be a projective space whose homogeneous coordinates correspond to every possible monomial as defined above. We can denote these coordinates by $Y_{d_0,\ldots,d_n}$. The mth *Veronese map* is a map

$$v_{n,m} \colon k\mathbb{P}^n \to \mathbb{P}(\mathbb{A}^N) = k\mathbb{P}^{N-1}$$

defined by

$$Y_{d_0,\ldots,d_n} = X_0^{d_0} \cdots X_n^{d_n}$$

This is regular because, among the monomials in S in proposition 5.4.6 on the preceding page, we have monomials X_i^m which do not all vanish because at least one of the X_i must be nonzero.

We can also define an inverse map. On the open affine $\mathbb{A}^{N-1} \subset k\mathbb{P}^{N-1}$ with $Y_{m,\ldots,0} = X_0^m \neq 0$, we map

$$(\cdots \colon Y_{m,\ldots,0} \colon \cdots \colon Y_{m-1,\underbrace{\ldots 0,1,0,\ldots 0}_{i^{\text{th}} \text{ position}}} \colon \cdots)$$

$$\mapsto (X_0^m \colon \cdots \colon X_0^{m-1} X_i \colon \cdots) \sim (X_0 \colon \cdots \colon X_i \colon \cdots)$$

This is clearly a regular map and an inverse to $v_{n,m}$ so the Veronese map is an *embedding* (or closed immersion) of $k\mathbb{P}^n$ in $k\mathbb{P}^{N-1}$.

The Veronese embedding has a number of interesting properties, including its ability to map an arbitrary degree-m hypersurface in $k\mathbb{P}^n$ like

$$\sum_{d_0+\cdots+d_n=m} a_{d_0,\ldots,d_n} X_0^{d_0} \cdots X_n^{d_n} = 0$$

into a *hyperplane*

$$\sum_{d_0+\cdots+d_n=m} a_{d_0,\ldots,d_n} Y_{d_0,\ldots,d_n} = 0$$

in $k\mathbb{P}^{N-1}$. In other words, every hypersurface in $k\mathbb{P}^n$ is isomorphic to the intersection of $v_{n,m}(k\mathbb{P}^n)$ with a suitable hyperplane in $k\mathbb{P}^{N-1}$.

EXERCISES.

2. If $V = \mathbb{P}(\mathbb{A}^{n+1})$ and W is the hypersurface $\mathcal{P}(F)$, where $F(X_0,\ldots,X_n)$ is homogeneous polynomial of degree d, show that

$$V \setminus W$$

is an open affine.

3. Show that the set of degree-d hypersurfaces in $k\mathbb{P}^n$ is parametrized by a projective space.

4. Show that the set of degree-d hypersurfaces in $k\mathbb{P}^n$ that contain a given point is parametrized by a hyperplane in a projective space.

5.5. Products

5.5.1. Introduction. Although we know how to compute products of general varieties (see section 4.5 on page 198) and could use those techniques to compute products of projective varieties, it is natural to ask

(1) Is there a more intrinsic definition of products (i.e., one using homogeneous coordinates)?
(2) Is the product of projective varieties a projective variety?

Key to answering both questions lies in:

5.5.2. The Segre Embedding. Since projective varieties are schemes, their products are well-defined and can be explicitly constructed (see section 4.5 on page 198). It is not obvious but true that a product of projective varieties is *also* a projective variety. The *Segre Embedding* is the key to proving this.

> Corrado Segre (1863 – 1924) was an Italian mathematician who was responsible for many early developments in algebraic geometry — see [**8**].

If U and V are vector spaces over the field, k, consider the natural map

$$t\colon U \times V \quad \to \quad U \otimes_k V$$
$$u \times v \quad \mapsto \quad u \otimes v$$

where $U \times V$ is the Cartesian product. The map t is not injective since, for any $c \in k^\times$,

$$c \cdot u \otimes c^{-1} \cdot v$$

defines the same element as $u \otimes v$. If we pass to projective spaces, we get

$$s(U,V)\colon \mathbb{P}(U) \times \mathbb{P}(V) \to \mathbb{P}(U \otimes V)$$

called the *Segre map.* The main result is:

THEOREM 5.5.1. *The Segre map, defined above, is a closed immersion of projective varieties, embedding the product $\mathbb{P}(U) \times \mathbb{P}(V)$ into $\mathbb{P}(U \otimes V)$.*

REMARK. The word "product" here is in the categorical sense — i.e. a product of varieties as defined in section 4.5 on page 198.

This result gives a functorial closed immersion

$$k\mathbb{P}^n \times k\mathbb{P}^m \to k\mathbb{P}^{(n+1)(m+1)-1}$$

PROOF. Let $U = \mathbb{A}^{n+1}$ and $V = \mathbb{A}^{m+1}$. If $u = (x_0\colon\cdots\colon x_n) \in \mathbb{P}(U)$ and $v = (y_0\colon\cdots\colon y_m) \in \mathbb{P}(V)$ then the image of $u \times v$ is $(w_{i,j}) \in \mathbb{P}(U \otimes V)$, with $i = 0,\ldots,n$ and $j = 0,\ldots,m$ and where

$$(5.5.1) \qquad\qquad\qquad w_{i,j} = x_i \cdot y_j$$

These quantities satisfy the identities

(5.5.2) $$w_{i,j}w_{s,t} = w_{i,s}w_{j,t}$$

for all $i, s = 0, \ldots, n$ and $j, t = 0, \ldots, m$. If $s = t = 0$, and $w_{0,0} \neq 0$, equation 5.5.2 implies

$$w_{i,j}w_{0,0} = w_{i,0}w_{0,j}$$

or

$$\frac{w_{i,j}}{w_{0,0}} = \frac{w_{i,0}}{w_{0,0}} \cdot \frac{w_{0,j}}{w_{0,0}}$$

so that, on $\mathbb{A}_{0,0}^{(n+1)(m+1)}$ (i.e., where $w_{0,0} \neq 0$) a set of quantities $(w_{i,j})$ satisfying equation 5.5.2 is of the form given in equation 5.5.1 on the preceding page if we define $x_i = w_{i,0}$ and $y_j = w_{0,j}/w_{0,0}$. This implies that the image of the Segre map is a *projective variety* (i.e., defined by the equations 5.5.2).

If $x_0 \neq 0$ and $y_0 \neq 0$, we get $w_{0,0} \neq 0$ so

$$\frac{x_i}{x_0} = \frac{w_{i,0}}{w_{0,0}} \text{ and } \frac{y_i}{y_0} = \frac{w_{0,j}}{w_{0,0}}$$

which implies that

$$s \colon \mathbb{A}_0^{n+1} \times \mathbb{A}_0^{m+1} \to \operatorname{im} s \cap \mathbb{A}_{0,0}^{(n+1)(m+1)}$$

is 1-1. This shows that the restriction to open affines of the Segre map is a map to the product and an isomorphism. We conclude that the Segre map as a whole is an isomorphism and a map to the product (see lemma 4.5.3 on page 199). □

This implies that:

COROLLARY 5.5.2. *Subvarieties of $k\mathbb{P}^n \times k\mathbb{P}^m$ are defined by sets of equations*

$$f_j(X_0, \ldots, X_n, Y_0, \ldots, Y_m) = 0$$

that are homogeneous in the X_i and the Y_j.

REMARK. The degrees of homogeneity of the X_i's can be different from that of the Y_j's.

Since we can regard an affine variety as a subset of a projective variety (i.e., its projective closure — see definition 5.1.12 on page 224), we get

COROLLARY 5.5.3. *Subvarieties of $k\mathbb{P}^n \times \mathbb{A}^m$ are defined by sets of equations*

(5.5.3) $$f_j(X_0, \ldots, X_n, Y_1, \ldots, Y_m) = 0$$

for $j = 1, \ldots, t$, that are homogeneous in the X_i.

REMARK. If an f_i is homogeneous in the X_j of degree d, it can be written

$$f_i = \sum_{|\alpha|=d} X^\alpha \cdot g_\alpha(Y)$$

where $\alpha = (\alpha_0, \ldots, \alpha_n)$, $|\alpha| = \sum_{t=0}^n \alpha_t$, and $X^\alpha = X_0^{\alpha_0} \cdots X_n^{\alpha_n}$. The *scheme* they define can be regarded as Proj R (see definition 4.4.12 on page 190), where

$$R = \frac{k[X_0, \ldots, X_n, Y_1, \ldots, Y_m]}{(f_1, \ldots, f_t)}$$

and the Y_i have degree 0 and the X_j have degree 1. The quotient is a graded ring computed as in equation A.4.8 on page 436.

Projective varieties have a number of rigidity properties:

LEMMA 5.5.4. *Let X and Y be projective varieties and let $\varphi, \psi \colon X \to Y$ be morphisms. If there is a nonempty open set $U \subset X$ such that $\varphi|U = \psi|U$, then $\varphi = \psi$.*

PROOF. Since $Y \subset k\mathbb{P}^n$ for some n, it will suffice to prove it for the case where $Y = k\mathbb{P}^n$. Consider the diagonal map

$$\Delta \colon k\mathbb{P}^n \to k\mathbb{P}^n \times k\mathbb{P}^n \subset k\mathbb{P}^{2n+1}$$

where the image of Δ is given by $\{x_i y_j = x_j y_i\}$ where the $\{x_i\}, \{y_i\}$ are the homogeneous coordinates of the two copies of $k\mathbb{P}^n$. It follows that $\operatorname{im}\Delta$ is closed. Now consider

$$\varphi \times \psi \colon X \to k\mathbb{P}^n \times k\mathbb{P}^n$$

The hypotheses imply that $\operatorname{im}\varphi \times \psi(U) \subset \operatorname{im}\Delta$. Since U is an open dense subset and $\operatorname{im}\Delta$ is closed, we conclude that $\operatorname{im}\varphi \times \psi(X) \subset \operatorname{im}\Delta$. $\qquad\square$

EXERCISES.

1. If $V = \mathcal{P}(X_0 X_3 - X_1 X_2) \subset k\mathbb{P}^3$, (a quadric surface), show that
   a. $V = k\mathbb{P}^1 \times k\mathbb{P}^1$
   b. V is birationally equivalent to $k\mathbb{P}^2$.

5.5.3. "Blowing up". We will discuss a construction that is widely used in algebraic geometry (and is a key element of Hironaka's remarkable papers [**78, 79**]). The basic idea is that "singular varieties are the image of smooth varieties under a suitable projection" or what Hauser ([**74**]) called "an almost philosophical speculation:"

> "Singular curves are the shadow of smooth curves in higher dimensional space."

For instance the singular curve $V \subset \mathbb{A}^2$ given by $Y^2 = X^3$ is the image of the smooth curve $W \subset \mathbb{A}^3$ defined by $X = Z^3, Y = Z^2$ as in figure 5.5.1 on the next page.

The new parameter, Z, that we introduced is equal to Y/X — defining a line through the origin — i.e., a point of $k\mathbb{P}^1$. If we try to extend this "blow up" to *all* of $\mathbb{A}^2$ (so $\mathbb{A}^2$ get blown up and V is "carried along"), the result is the *graph,* Γ_F, (see definition 4.6.5 on page 206) of a map

$$\begin{aligned} F \colon \mathbb{A}^2 \setminus (0,0) &\to k\mathbb{P}^1 \\ (X, Y) &\mapsto [X \colon Y] \end{aligned}$$
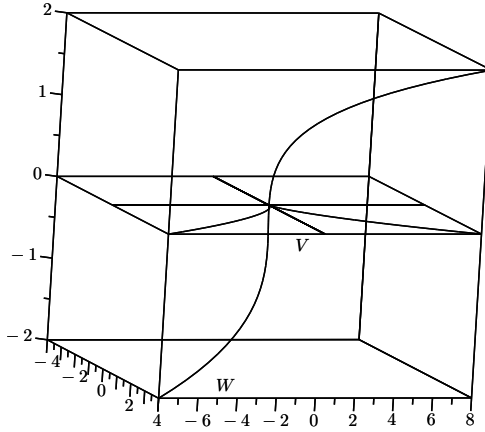
FIGURE 5.5.1.   Blowing up

— a closed subset of $(\mathbb{A}^2 \setminus (0,0)) \times k\mathbb{P}^1$ (see proposition 4.6.6 on page 206). If we regard it as a subset of $\mathbb{A}^2 \times k\mathbb{P}^1$, the Zariski closure is a subvariety $B \subset \mathbb{A}^2 \times k\mathbb{P}^1$ and projection onto the first factor is a map

$$f : B \to \mathbb{A}^2$$

since $F(X,Y) = [X{:}Y]$, B is the subset of $\mathbb{A}^2 \times k\mathbb{P}^1$ consisting of points $(X, Y, R{:}S)$ such that

$$\frac{S}{R} = \frac{Y}{X}$$

for $X \neq 0$ or

$$S \cdot X = R \cdot Y$$

Since *this* equation is well-defined for *all* X, Y, we conclude that it defines B completely. If $p \in \mathbb{A}^2$ is any point except the origin, $f^{-1}(p)$ is a single point. In fact, $f^{-1}(\mathbb{A}^2 \setminus (0,0) \subset B$ is *isomorphic* to $\mathbb{A}^2 \setminus (0,0)$ — which implies that f is a *birational equivalence* (see theorem 2.8.23 on page 107). On the other hand $f^{-1}((0,0)) = k\mathbb{P}^1$, so the point $(0,0)$ has been "blown up" to a *whole copy* of $k\mathbb{P}^1$ — see figure 5.5.2 on the next page. Since the top of the figure is identified with the bottom, B looks like a Möbius strip.

On the chart (or open affine $\mathbb{A}_0^1 \subset k\mathbb{P}^1$), we set $R = 1$ and $B \cap \mathbb{A}^2 \times \mathbb{A}_0^1$ is defined by the equation
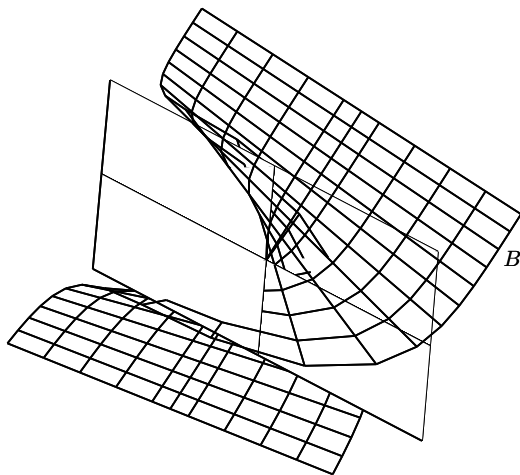
$$S \cdot X = Y$$

which defines a copy of $\mathbb{A}^2$ embedded in $\mathbb{A}^2 \times \mathbb{A}_0^1$. The subvariety $f^{-1}((0,0))$ is defined by the single equation $X = 0$. On the chart $\mathbb{A}_1^1 \subset k\mathbb{P}^1$, we have $S = 1$ and $B \cap \mathbb{A}^2 \times \mathbb{A}_1^1$ is defined by

$$X = R \cdot Y$$

and $f^{-1}((0,0))$ is defined by the single equation $Y = 0$.

We can easily generalize this:

FIGURE 5.5.2.   Blowing up $\mathbb{A}^2$

DEFINITION 5.5.5. If $S = \{f_1, \ldots, f_t\} \subset k[X_1, \ldots, X_n]$ is a set of polynomials and $t \geq 2$, the *blowup of $\mathbb{A}^n$ induced by S* is the subvariety $W \subset \mathbb{A}^n \times k\mathbb{P}^{t-1}$ that is the Zariski-closure of the graph of

$$f: \mathbb{A}^n \setminus \mathcal{V}\left((f_1, \ldots, f_t)\right) \quad \to \quad k\mathbb{P}^{t-1}$$
$$(X_1, \ldots, X_n) \quad \mapsto \quad [f_1(X){:}\cdots{:}f_t(X)]$$

This comes with a canonical projection

$$\sigma: W \to \mathbb{A}^n$$

REMARK. If the f_i are algebraically independent, W is the variety defined by the equations

(5.5.4) $f_i(X) \cdot Y_j = f_j(X) \cdot Y_i$

for $i, j = 1, \ldots, t$, and where $\{Y_i\}$ are the homogeneous coordinates of $k\mathbb{P}^{t-1}$ — or (as a *scheme*) Proj R with

(5.5.5) $$R = \frac{k[X_1, \ldots, X_n, Y_1, \ldots, Y_t]}{(\{f_i(X) \cdot Y_j - f_j(X) \cdot Y_i\})}$$

where the Y_i have degree 1 and the X_j have degree 0. Since $k\mathbb{P}^0$ is a *point,* the blowup of $\mathbb{A}^n$ by a *principal* ideal (f) is just $\mathbb{A}^n$ and the canonical projection is the identity map[2].

We will ultimately show that a blowup only depends on the *ideal* $(f_1, \ldots, f_t)$. The first step is

PROPOSITION 5.5.6. *Let $\{f_1, \ldots, f_t\} \subset k[X_1, \ldots, X_n]$ be a set of polynomials as in definition 5.5.5 inducing the blowup*

$$\sigma: W \to \mathbb{A}^n$$

---

[2]We could also say that the blowup by (f) is Proj $k[X_1, \ldots, X_n, Y]/(f \cdot Y - f \cdot Y)$.

If $g \in (f_1, \ldots, f_t) = \mathfrak{a}$ and $W' \subset \mathbb{A}^n \times k\mathbb{P}^{t-1}$ with projection

$$\sigma': W' \to \mathbb{A}^n$$

is the blow up induced by the expanded set $\{f_1, \ldots, f_t, g\}$, there exist isomorphisms $\varphi: W \to W'$ and $\psi: W' \to W$ with $\psi \circ \varphi = 1: W \to W$ and $\varphi \circ \psi = 1: W' \to W'$ that makes the diagram

(5.5.6)

$$
\begin{array}{ccc}
W & \underset{\psi}{\overset{\varphi}{\rightleftarrows}} & W' \\
& \underset{\sigma}{\searrow} \quad \underset{\sigma'}{\swarrow} & \\
& \mathbb{A}^n &
\end{array}
$$

commute.

REMARK. When two blowups fit into a commutative diagram like 5.5.6 we will say that they are *equivalent*.

PROOF. Since $g \in (f_1, \ldots, f_t)$, we have an expression

$$g = \sum_{i=1}^{t} r_i f_i$$

We define φ by

$$(X_1, \ldots, X_n) \times [Y_1 : \cdots : Y_t] \mapsto (X_1, \ldots, X_n) \times \left[Y_1 : \cdots : Y_t : \sum_{i=1}^{t} r_i(X) \cdot Y_i \right]$$

This is a regular map that sends $\mathbb{A}^n \times k\mathbb{P}^{t-1} \setminus \sigma^{-1}(\mathcal{V}(\mathfrak{a}))$ to $\mathbb{A}^n \times k\mathbb{P}^t \setminus \sigma'^{-1}(\mathcal{V}(\mathfrak{a}))$ so it extends to a regular map of Zariski closures. The inverse, ψ, is defined by

$$(X_1, \ldots, X_n) \times [Y_1 : \cdots : Y_t : Y_{t+1}] \mapsto (X_1, \ldots, X_n) \times [Y_1 : \cdots : Y_t]$$

and the composite $\psi \circ \varphi$ carries W to W and is the identity map. It is also not hard to verify that $\varphi \circ \psi: W' \to W'$ is the identity map. $\square$

This immediately implies

COROLLARY 5.5.7. *In the context of definition 5.5.5 on the facing page, suppose $(f_1, \ldots, f_t) = (g_1, \ldots, g_s) \subset k[X_1, \ldots, X_n]$ are sets of polynomials that generate the same ideal. If $\sigma: W \to \mathbb{A}^n$ is the blowup of $\mathbb{A}^n$ with respect to $\{f_1, \ldots, f_t\}$ and $\sigma': W' \to \mathbb{A}^n$ is that with respect to $\{g_1, \ldots, g_s\}$, there exist isomorphisms $\varphi: W \to W'$ and $\psi: W' \to W$ with $\psi \circ \varphi = 1: W \to W$ and $\varphi \circ \psi = 1: W' \to W'$ that makes diagram 5.5.6 commute.*

PROOF. Inductively apply proposition 5.5.6 on the facing page to show that W and W' are both equivalent to the blowup with respect to

$$\{f_1, \ldots, f_t, g_1, \ldots, g_s\}$$

$\square$

With this in mind, we can define

DEFINITION 5.5.8. If $\mathfrak{a} \subset k[X_1, \ldots, X_n] = k[\mathbb{A}^n]$ is an ideal, the *blowup* of $\mathbb{A}^n$ *induced by* $\mathfrak{a}$, denoted $\mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n)$, is the blowup with respect to the ideal $\mathfrak{a}$. This comes with a canonical map $\sigma \colon \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n) \to \mathbb{A}^n$ with the property that

$$\sigma | \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n) \setminus \sigma^{-1}(V) \colon \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n) \setminus \sigma^{-1}(V) \to \mathbb{A}^n \setminus V$$

with $V = \mathcal{V}(\mathfrak{a})$ is an isomorphism. The subvariety $\sigma^{-1}(V) \subset \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n)$ is called the *exceptional fiber* of σ. The variety $\mathcal{V}(\mathfrak{a})$ is called the *center* of the blowup.

REMARK. The exceptional fiber is *exceptional* because it is the only one that is not a point. If $\mathfrak{a}$ is a principal ideal, there is no exceptional fiber.

Blowups are often defined with respect to a *subvariety* $V = \mathcal{V}(\mathfrak{a})$. It is possible for distinct ideals with the same zero-set to give different blowups — see exercise 3 on page 254, which shows that $\mathbb{B}_{(X_1, X_2)}(\mathbb{A}^2) \neq \mathbb{B}_{(X_1^2, X_2)}(\mathbb{A}^2)$. Since *subschemes* of $\mathbb{A}^n$ are in a 1-1 correspondence with their defining ideals, it makes sense to write $\mathbb{B}_V(\mathbb{A}^n)$ if V is a subscheme of $\mathbb{A}^n$.

The following result gives an interesting property of blowups[3]:

PROPOSITION 5.5.9. *Suppose* $\mathfrak{a} = (f_1, \ldots, f_t) \subset k[X_1, \ldots, X_n]$ *is not a principal ideal and* E *is the exceptional fiber of* $\sigma \colon \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n) \to \mathbb{A}^n$. *Then*

$$E \subset \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n) \subset \mathbb{A}^n \times k\mathbb{P}^{t-1} = \bigcup_{j=1}^{t} \mathbb{A}^n \times \mathbb{A}_j^{t-1}$$

On the ith *open affine,* $E \cap (\mathbb{A}^n \times \mathbb{A}_i^{t-1})$, *is defined by the single equation:*

$$f_i = 0$$

It follows that E *is* $n - 1$-*dimensional.*

REMARK. Strictly speaking, the statement is true even if $\mathfrak{a}$ is principal — but is less interesting. In effect, the ideal $\mathfrak{a}$ has been "blown up" into a set of principal ideals — each "valid" on a different subspace of $\mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n)$.

This turns out to show that the exceptional fiber is something called a *Cartier divisor* — see example 5.9.14 on page 290.

Since blow-ups are isomorphisms outside the exceptional fiber, they are *always* birational equivalences.

PROOF. Equations 5.5.4 on page 250 are *valid* on $\mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n)$ even if they do not completely define it (see exercise 2 on page 254). On the chart $\mathbb{A}^n \times \mathbb{A}_i^t \subset \mathbb{A}^n \times k\mathbb{P}^{t-1}$, we have $Y_i = 1$, which transforms equations 5.5.4 on page 250 into

$$f_i(X) \cdot Y_j = f_j(X)$$

Consequently, $f_i(X) = 0 \implies f_j(X) = 0$ for *all* j on this chart and $E \cap (\mathbb{A}^n \times \mathbb{A}_i^t)$ is defined by the *single* equation $f_i = 0$. It follows that E is $n - 1$-dimensional (see corollary 2.8.30 on page 111). $\qquad\square$

We can also blow up affine varieties (and schemes):

---

[3]In a certain sense, it can be regarded as a *defining* property — see exercises 5 and 6 on page 254.

PROPOSITION 5.5.10. *Suppose* $\mathfrak{a} \subset k[X_1, \ldots, X_n]$ *and suppose* $\mathcal{V}(\mathfrak{a}) \subseteq W \subset \mathbb{A}^n$, *and* $\sigma: \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n) \to \mathbb{A}^n$ *is the blowup. Then* $\mathbb{B}_{\mathfrak{a}}(W) \subset \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n)$ *is the Zariski-closure of* $\sigma^{-1}(W \setminus \mathcal{V}(\mathfrak{a}) \subset \mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n)$.

REMARK. Since the exceptional fiber is defined by a single equation on each chart (proposition 5.5.9 on page 5.5.9), corollary 2.8.30 on page 111 implies that its intersection with $\mathbb{B}_{\mathfrak{a}}(W)$ has codimension 1.

If W is an affine variety (or scheme) and $\mathfrak{a} \subset k[W]$, we can embed W in an affine space $\mathbb{A}^n$ so that $\mathfrak{a}$ lifts to an ideal $\mathfrak{a}' \subset k[X_1, \ldots, X_n]$ and we can compute $\mathbb{B}_{\mathfrak{a}}(W) = \mathbb{B}_{\mathfrak{a}'}(W) \subset \mathbb{B}_{\mathfrak{a}'}(\mathbb{A}^n)$.

It's a small step from this to blowing up portions of *general* varieties and schemes: If W is an open affine of a general scheme, S, and $\mathfrak{a} \subset \mathcal{O}_S(W)$, then the fact that $\sigma|\mathbb{B}_{\mathfrak{a}}(W) \setminus \operatorname{Spec}(\mathcal{O}_S(W)/\mathfrak{a})$ is an *isomorphism* implies that we can "glue" $\mathbb{B}_{\mathfrak{a}}(W)$ to the rest of S to produce $\mathbb{B}_{\mathfrak{a}}(S)$.

EXAMPLE 5.5.11. The Whitney Umbrella, $W \subset \mathbb{A}^3$, given by the equation $X_1^2 + X_2 X_3^2 = 0$. Computing the Jacobian shows that the its singular set is the X_2-axis. Since the X_2 axis is given by the equations $X_1 = 0$ and $X_3 = 0$, so we can blow up $\mathbb{A}^3$ to get $\mathbb{B}_{(X_1, X_3)}(\mathbb{A}^3)$

$$\begin{aligned} g: \mathbb{A}^3 &\to k\mathbb{P}^1 \\ (X_1, X_3) &\mapsto [Y_1 : Y_2] \end{aligned}$$

As before, we get equations for $\mathbb{B}_{(X_1, X_3)}(\mathbb{A}^3)$

$$X_1 \cdot Y_2 = X_3 \cdot Y_1 \subset \mathbb{A}^3 \times k\mathbb{P}^1$$

and $\mathbb{B}_{(X_1, X_3)}(W) \subset \mathbb{B}_{(X_1, X_3)}(\mathbb{A}^3) \subset \mathbb{A}^3 \times k\mathbb{P}^1$ is the Zariski closure of $g^{-1}(W \setminus X_2\text{-axis})$. Since $g|g^{-1}(\mathbb{A}^3 \setminus X_2\text{-axis})$ is an isomorphism and W is irreducible, it follows that $g^{-1}(W \setminus X_2\text{-axis})$ is irreducible in $g^{-1}(\mathbb{A}^3 \setminus X_2\text{-axis})$. On the chart with $Y_2 = 1$, plug $X_1 = Y_1 X_3$ into the defining equation to get

(5.5.7) $$Y_1^2 X_3^2 + X_2 X_3^2 = 0$$

or $X_2 + Y_1^2 = 0$, when $X_3 \neq 0$, so the Zariski-closure, $\mathbb{B}_{(X_1, X_3)}(W) \cap (\mathbb{A}^2 \times \mathbb{A}_2^1)$, is given by

$$\begin{aligned} X_2 + Y_1^2 &= 0 \\ X_1 - Y_1 X_3 &= 0 \end{aligned}$$

It is easily verified that the Jacobian is of rank 2 everywhere so it is smooth. On the chart with $Y_1 = 1$, $X_3 = Y_2 X_1$ and we get

$$X_1^2 + X_2 Y_2^2 X_1^2 = 0$$

or $Y_2^2 X_2 + 1 = 0$, when $X_1 \neq 0$, and $\mathbb{B}_{(X_1, X_3)}(W) \cap (\mathbb{A}^2 \times \mathbb{A}_1^1)$ is given by

$$\begin{aligned} Y_2 X_1 - X_3 &= 0 \\ Y_2^2 X_2 + 1 &= 0 \end{aligned}$$

and computation of the Jacobian shows that this has a singularity defined by $X_1 = Y_2 = 0$. If E is the exceptional fiber, $E \cap \mathbb{B}_{(X_1, X_3)}(W)$ is equal the parabola $X_2 + Y_1^2 = 0$ on the first chart and the curve $Y_2^2 X_2 + 1 = 0$ on the second. These curves intersect at the single point $X_1 = X_3 = 0$, $Y_1 = Y_2 = 1$, $X_2 = -1$.

EXERCISES.

2. Why can't we *always* use equations 5.5.4 on page 250 to compute the result of blowing up?

3. Show that the blowup $\mathbb{B}_{(X_1^2, X_2)}(\mathbb{A}^2)$ has a singularity. This shows that blowing-up can create singularities where none existed before.

4. Suppose $\mathfrak{a} = (f_1, \ldots, f_t) \subset k[X_1, \ldots, X_n] = R$ and suppose that the f_i are algebraically independent. Show that, as a scheme,

$$\mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n) = \text{Proj}\left(R \oplus \mathfrak{a} \oplus \mathfrak{a}^2 \oplus \cdots\right)$$

where the summand R has degree 0, $\mathfrak{a}$ has degree 1, and so on — i.e., we take Proj of the Rees algebra of $\mathfrak{a} \subset R$ described in example A.4.43 on page 436. With a little more work (and keeping in mind the solution to exercise 2) one can show that this formula for $\mathbb{B}_{\mathfrak{a}}(\mathbb{A}^n)$ is *always* true.

5. Suppose $f: V \to W$ is a regular map of affine varieties with induced homomorphism of coordinate rings $f^*: k[W] \to k[V]$ and suppose $\mathfrak{b} \subset k[W]$ is an ideal. Show that there exists a *unique* map

$$g: \mathbb{B}_{f^*(\mathfrak{b})}(V) \to \mathbb{B}_{\mathfrak{b}}(W)$$

that makes the diagram

$$
\begin{array}{ccc}
\mathbb{B}_{f^*(\mathfrak{b})}(V) & \xrightarrow{\ g\ } & \mathbb{B}_{\mathfrak{b}}(W) \\
{\scriptstyle \sigma_1}\downarrow & & \downarrow{\scriptstyle \sigma_2} \\
V & \xrightarrow{\ f\ } & W
\end{array}
$$

commute, where σ_1 and σ_2 are the standard maps from blowups to bases.

6. Suppose $f: V \to W$ is a regular map of affine varieties with induced homomorphism of coordinate rings $f^*: k[W] \to k[V]$ and suppose $\mathfrak{b} \subset k[W]$ is an ideal such that $f^*(\mathfrak{b}) \subset k[V]$ is principal. Show that there exists a *unique* morphism

$$g: V \to \mathbb{B}_{\mathfrak{b}}(W)$$

that makes the diagram

$$
\begin{array}{ccc}
 & & \mathbb{B}_{\mathfrak{b}}(W) \\
 & {\scriptstyle g}\nearrow & \downarrow{\scriptstyle \sigma} \\
V & \xrightarrow{\ f\ } & W
\end{array}
$$

commute.

5.5.4. Projective elimination theory. This section is a projective version of section 2.5.2 on page 80. In the projective case, we can accomplish considerably more than in the affine case. For instance, we can compute the image of a variety under the projection $k\mathbb{P}^n \times \mathbb{A}^m \to \mathbb{A}^m$ — rather than its Zariski closure. This leads to several interesting and useful results that emphasize some unique features of projective varieties.

DEFINITION 5.5.12. Let $\mathfrak{a} \subset k[X_0, \ldots, X_n, Y_1, \ldots, Y_m]$ be an ideal whose generators are homogeneous in the X_i. Then the *projective elimination ideal*, $\hat{\mathfrak{a}} \subset k[Y_1, \ldots, Y_m]$ is defined by

$$\hat{\mathfrak{a}} = \left\{ f \in k[Y_1, \ldots, Y_m] \big| f \cdot X_j^{e_j} \in \mathfrak{a} \right\}$$

for *each* $j = 0, \ldots, n$ and some exponent $e_j \geq 0$.

REMARK. Compare this to the *saturation-ideals* in exercises 9 on page 71 and 10 on page 226.

That this actually defines an ideal is left to the reader as an exercise. Compare this to the affine case, definition 2.5.15 on page 80. The exponents e_j only have to be "sufficiently large" — if $f \cdot X_j^{e_j} \in \mathfrak{a}$ then $f \cdot X_j^r \in \mathfrak{a}$ for any $r \geq e_j$.

This is a projective version of proposition 2.5.16 on page 80:

PROPOSITION 5.5.13. *Let $\mathcal{V}(\mathfrak{a}) \subset k\mathbb{P}^n \times \mathbb{A}^m$ be a variety defined by an ideal $\mathfrak{a} \subset k[X_0, \ldots, X_n, Y_1, \ldots, Y_m]$ that is homogeneous in the X_i and arbitrary in the Y_j. If W is the image of $\mathcal{V}(\mathfrak{a})$ under the projection $k\mathbb{P}^n \times \mathbb{A}^m \to \mathbb{A}^m$, then*

$$W \subset \mathcal{V}(\hat{\mathfrak{a}})$$

where $\mathcal{V}(\hat{\mathfrak{a}}) \subset \mathbb{A}^m$ is the set of points where the polynomials in $\hat{\mathfrak{a}}$ vanish and $\hat{\mathfrak{a}}$ is defined in 5.5.12.

REMARK. By abuse of notation, we use the notation $\mathcal{V}(\mathfrak{a})$ to denote a variety in $k\mathbb{P}^n \times \mathbb{A}^m$.

PROOF. In other words, we are claiming that the polynomials in $\hat{\mathfrak{a}}$ all vanish on W. If Let $(u_1, \ldots, u_m) \in W$ be the image of a point

$$(t_0, \ldots t_n, u_1, \ldots, u_m) \in \mathcal{V}(\mathfrak{a})$$

and suppose some $t_i \neq 0$ (there must be at least one). If $g \in \hat{\mathfrak{a}}$, then $X_i^{e_i} \cdot g \in \mathfrak{a}$ and $t_i^{e_i} \cdot g(u_1, \ldots u_m) = 0$. Since $t_i \neq 0$, we conclude that $g(u_1, \ldots u_m) = 0$. $\square$

The interesting thing about the projective case is that we can prove a converse that completely characterizes the image of a variety:

THEOREM 5.5.14 (Projective elimination). *Let $\mathcal{V}(\mathfrak{a}) \subset k\mathbb{P}^n \times \mathbb{A}^m$ be a variety defined by an ideal $\mathfrak{a} \subset k[X_0, \ldots, X_n, Y_1, \ldots, Y_n]$ that is homogeneous in the X_i and arbitrary in the Y_j — and assume k is algebraically closed. If W is the image of $\mathcal{V}(\mathfrak{a})$ under the projection $k\mathbb{P}^n \times \mathbb{A}^m \to \mathbb{A}^m$, then*

$$W = \mathcal{V}(\hat{\mathfrak{a}})$$

where $\hat{\mathfrak{a}}$ is defined in 5.5.12.

PROOF. It suffices to show that $V(\hat{\mathfrak{a}}) \subset W$. Suppose

$$\mathfrak{a} = (f_1, \ldots, f_t) \subset k[X_0, \ldots, X_n, Y_1, \ldots, Y_m]$$

where the f_i are homogeneous in the x_i. Now suppose that $u = (u_1, \ldots, u_m) \in \mathcal{V}(\mathfrak{a})$ but $u \notin W$. Then the equations

$$f_i(X_0, \ldots, X_n, u) = 0$$

have no solution in $k\mathbb{P}^n$. The Projective Nullstellensatz (theorem 5.1.4 on page 220) implies that

$$(f_1(X, u), \ldots, f_t(X, u)) \supset (X_0^s, \ldots, X_n^s) \subset k[X_0, \ldots, X_n]$$

for some value of $s > 0$. This means that all monomials in the X_i^s can be expressed in terms of the $f_i(X, u)$, and all monomials of total degree $N = n \cdot s$ can be expressed in terms of the $f_i(X, u)$:

$$X^\alpha = \sum_{i=1}^{t} f_i(X, u) \cdot h_{i,\alpha}(X)$$

where $\alpha = (\alpha_0, \ldots, \alpha_n)$ is a multi-index with $|\alpha| = \sum_{i=0}^{n} \alpha_i = N$ and the $h_{i,\alpha}(X)$ are homogeneous.

We can choose

$$g_j = X^{\beta_j} \cdot f_{i_j}$$

for $j = 1, \ldots, \binom{n+N}{N}$ so the $g_j(X, u)$ are a basis for all of the monomials of total degree N. Since the $g_j(X, Y)$ are homogeneous in the X_i's, we can write

$$g_j = \sum_{|\alpha|=N} X^\alpha p_{j,\alpha}(Y_1, \ldots, Y_m)$$

where $p_{j,\alpha}(Y_1, \ldots, Y_m)$ is an $\binom{n+N}{N} \times \binom{n+N}{N}$ matrix of polynomials in the Y_j. Let

$$A(Y_1, \ldots, Y_m) = \det\left(p_{j,\alpha}(Y_1, \ldots, Y_m)\right)$$

Since when $Y_i = u_i$ for $i = 1, \ldots, m$, the g_j are a basis for the monomials X^α with $|\alpha| = N$, we have

(5.5.8) $A(u_1, \ldots, u_m) \neq 0$

On the other hand, Cramer's Rule implies that, for some matrix $d_{j,\alpha}$

$$A(Y_1, \ldots, Y_m) \cdot X^\alpha = \sum_{j=1}^{N} g_j(X, Y) \cdot d_{j,\alpha}$$

for every multi-index $\alpha = (\alpha_0, \ldots, \alpha_n)$ with $|\alpha| = N$. This implies that

$$A(Y_1, \ldots, Y_m) \in \hat{\mathfrak{a}}$$

so $A(u_1, \ldots, u_m) = 0$ (by proposition 5.5.13 on the previous page), which contradicts equation 5.5.8. $\square$

COROLLARY 5.5.15. *If V is a projective variety and W is a general variety, then the projection*

$$V \times W \to W$$

maps closed sets in $V \times W$ to closed sets in W.

REMARK. The remarkable property described here is called *completeness*.

PROOF. Since V is a projective variety, it suffices to prove this for $V = k\mathbb{P}^n$ — i.e. a closed set in $V \times W$ will still be closed in $k\mathbb{P}^n \times W \supset V \times W$. Since closed sets in W are characterized by their intersections with open affines, it suffices to prove the result for $W = \mathbb{A}^m$.

It follows that we only need to prove the conclusion for

$$k\mathbb{P}^n \times \mathbb{A}^m \to \mathbb{A}^m$$

This follows immediately from theorem 5.5.14 on page 255. $\square$

This leads to the remarkable result:

COROLLARY 5.5.16. *The image of a projective variety under a regular map is closed.*

REMARK. This is very different from what happens with affine varieties. For instance the image of

$$XY = 1 \subset \mathbb{A}^2$$

under the projection to the X-axis is $\mathbb{A}^1 \setminus \{0\}$, which is open.

This also implies that the general variety, $\mathbb{A}^2 \setminus \{(0,0)\}$, is not projective since its inclusion into $\mathbb{A}^2$ is open. We have already seen that it is not affine.

PROOF. Let

$$f\colon V \to W$$

be a regular map of a projective variety to a general variety. Then the graph of this map, Γ_f (see definition 4.6.5 on page 206) is a closed subset of $V \times W$, by proposition 4.6.6 on page 206. The image of Γ_f under the projection

$$V \times W \to W$$

is equal to $f(V)$ and this is closed, by corollary 5.5.15 on the facing page. $\square$

COROLLARY 5.5.17. *Let V be an irreducible projective variety. Then the only regular functions on V are constants.*

PROOF. A regular function on V is a regular map

$$f\colon V \to k = \mathbb{A}^1$$

and its image must be closed, by corollary 5.5.16. A closed subset of $\mathbb{A}^1$ consists of a finite set of points $\{k_1, \ldots, k_s\}$. If $s > 0$, $f^{-1}(k_i)$ is a component of V. Since V is irreducible, we must have $s = 1$. $\square$

A map from a product can be regarded as a *set* of maps indexed by points of a factor:

DEFINITION 5.5.18. A regular map

$$f\colon X \times Y \to Z$$

defines a *family of maps $f_y\colon X \to Z$* for each $y \in Y$.

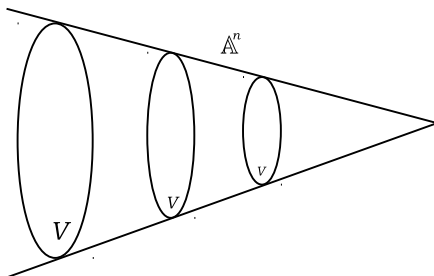With this definition in place, we get a rather surprising result:

FIGURE 5.5.3.   Shrinking an affine variety

LEMMA 5.5.19. *Suppose X is an irreducible projective variety and*

$$f: X \times Y \to Z$$

is a family of maps for some varieties Y and Z, where Y is irreducible. If there exists a point $y_0 \in Y$ with $f_{y_0}(X)$ equal to a single point in Z, then $f_y(X)$ is a single point of Z for all $y \in Y$.

REMARK. This is a kind of "rigidity" property of projective varieties: It is impossible to "smoothly deform" a projective variety to a point. This is most definitely *not* true of *affine* varieties. For instance, if $V \subset \mathbb{A}^n$ is *any* affine variety, the family of maps

$$\mathbb{A}^n \times \mathbb{A}^1 \quad \to \quad \mathbb{A}^n$$
$$\{X_1, \ldots, X_n\} \times Y \quad \to \quad \{X_1 \cdot Y, \ldots, X_n \cdot Y\}$$

smoothly shrinks V to a point as $Y \to 0$ — see figure 5.5.3.

PROOF. Let $\Gamma_f \subset X \times Y \times Z$ be the graph of f (see definition 4.6.5 on page 206), and let $\bar{\Gamma}$ be its image in $Y \times Z$ under the projection

$$X \times Y \times Z \to Y \times Z$$

Corollary 5.5.15 on page 256 implies that $\bar{\Gamma} \subset Y \times Z$ is closed, hence a variety. Consider the projection

$$g: \bar{\Gamma} \to Y$$

to the first factor. The fiber of this map over a point $y \in Y$ is $g^{-1}(y) = y \times f(X \times y)$ so the map is surjective. Since $f(X \times y_0)$ is a single point, proposition 5.6.4 on page 263 implies that $\dim \bar{\Gamma} = \dim Y$.

If $x_0 \in X$ is any point, the subvariety $S = \{(y \times f(x_0, y)|y \in Y\} \subset \bar{\Gamma}$ is isomorphic to Y (via g). Since they are of the same dimension and irreducible, we conclude that $S = \bar{\Gamma} = Y$ and that g is an isomorphism. This means that $g^{-1}(y) = y \times f(X \times y)$ is a single point for every $y \in Y$.   □

It is possible to *compute* the projective elimination ideal using *Gröbner bases:*

DEFINITION 5.5.20. If R is a commutative ring and $\mathfrak{a}, \mathfrak{b} \subset R$ are ideals, define the *ideal quotient*, $(\mathfrak{a}: \mathfrak{b})$ by

$$(\mathfrak{a}: \mathfrak{b}) = \{r \in R | r \cdot \mathfrak{b} \subset \mathfrak{a}\}$$

PROPOSITION 5.5.21. *If R is an integral domain, $r \neq 0 \in R$ and $\mathfrak{a} \subset R$, and $\mathfrak{a} \cap (r) = (r \cdot g_1, \dots, r \cdot g_n)$ then*

$$(\mathfrak{a} : (r)) = (g_1, \dots, g_n)$$

REMARK. This is why $(\mathfrak{a} : (r))$ is sometimes written

$$\frac{1}{r} \cdot \mathfrak{a} \cap (r)$$

PROOF. If s is a linear combination of the $\{g_j\}$, then $r \cdot s \in \mathfrak{a} \cap (r) \subset \mathfrak{a}$ so that $s \in (\mathfrak{a} : (r))$.

Conversely, if $s \in (\mathfrak{a} : (r))$, then $sr \in \mathfrak{a}$. Since $sr \in (r)$, it follows that $sr \in \mathfrak{a} \cap (r)$. If

$$sr = \sum_{j=1}^{n} a_j \cdot r \cdot g_j$$

then

$$s = \sum_{j=1}^{n} a_j \cdot g_j$$

since R is an integral domain. □

It follows that

PROPOSITION 5.5.22. *Let $\mathfrak{a} \subset k[X_0, \dots, X_n, Y_1, \dots, Y_m]$ be an ideal whose generators are homogeneous in the X_i. Then the projective elimination ideal, $\hat{\mathfrak{a}} \subset k[Y_1, \dots, Y_m]$, defined in 5.5.12 on page 255 is given by*

$$\hat{\mathfrak{a}} = \left(\bigcap_{j=1}^{n} (\mathfrak{a} : (X_j^{e_j})) \right) \cap k[Y_1, \dots, Y_m]$$

for each $j = 0, \dots, n$ and "sufficiently large" exponent $e_j \geq 0$.

REMARK. This allows for a purely algorithmic computation of $\hat{\mathfrak{a}}$ using Gröbner bases. "Sufficiently large" means large enough that $(\mathfrak{a} : (X_j^{e_j})) = (\mathfrak{a} : (X_j^{e_j+1}))$.

For instance, let

(5.5.9) $\qquad \mathfrak{a} = (X^2Y - XY^2, X^2Z^2 - 2ZXY - Y^2) \subset k[X, Y, Z]$

Then

$$(\mathfrak{a} : (Y)) = (\mathfrak{a} : (Y^2)) = (Z^2Y^3 - 2ZY^3 - Y^3,$$
$$Z^2XY - 2ZXY - Y^2, XY^2 - Y^3, X^2 - XY)$$

while

$$(\mathfrak{a} : (Y^3)) = (Z^2Y - 2ZY - Y, X - Y)$$

and we "stabilize" at Y^4:

(5.5.10) $\qquad (\mathfrak{a} : (Y^4)) = (\mathfrak{a} : (Y^5)) = \cdots = (Z^2 - 2Z - 1, X - Y)$

Compare this to the concept of *saturation* of an ideal in exercise 10 on page 226. The ideal $\hat{\mathfrak{a}}$ is a kind of saturation with respect to the X-variables *intersected* with the subring of Y-variables.

PROOF. This is a straightforward consequence of definitions 5.5.12 on page 255 and 5.5.20 on page 258. □

We conclude this section with an example:

EXAMPLE 5.5.23. Let $W = \mathcal{V}(\mathfrak{a}) \subset k\mathbb{P}^1 \times \mathbb{A}^1$ be defined by the ideal in equation 5.5.9 on the preceding page. We have already computed $(\mathfrak{a} \colon (Y^t))$ for "sufficiently large" t — it is given by equation 5.5.10 on page 259. We compute (using proposition 5.5.21 on page 259, Gröbner bases, and proposition 2.3.14 on page 53)

$$
\begin{aligned}
(\mathfrak{a} \colon (X)) &= \frac{1}{X} \cdot \mathfrak{a} \cap (X) \\
&= (-Y^3 - 2ZY^3 + Y^3 Z, \\
&\quad XY - Y^2, Y^2 - 2Y^2 Z + X^2 Z^2)
\end{aligned}
$$

If we cut to the chase and compute

$$
\begin{aligned}
(\mathfrak{a} \colon (X^{20})) &= \frac{1}{X^{20}} \cdot \mathfrak{a} \cap (X^{20}) \\
&= (-Z^2 - 2Z^3 + Z^4, \\
&\quad YZ^2 - 2YZ - Y, \\
&\quad XZ^2 - 2YZ - Y, XY - Y^2)
\end{aligned}
$$

Now we compute the intersection, to get

$$
\begin{aligned}
(\mathfrak{a} \colon (X^{20})) \cap (\mathfrak{a} \colon (Y^4)) &= (-Z^2 - 2Z^3 + Z^4, \\
&\quad Z^2 Y - 2ZY - Y, Z^2 X - 2ZY - Y, \\
&\quad XY - Y^2)
\end{aligned}
$$

Since this is a Gröbner basis for the intersection with lexicographic ordering with

$$
X \succ Y \succ Z
$$

we have

$$
\hat{\mathfrak{a}} = (\mathfrak{a} \colon (X^{20})) \cap (\mathfrak{a} \colon (Y^4)) \cap k[Z] = (-Z^2 - 2Z^3 + Z^4)
$$

So the image of W in $\mathbb{A}^1$ under the projection $k\mathbb{P}^1 \times \mathbb{A}^1 \to \mathbb{A}^1$ is the set of points $\{0, 1 \pm \sqrt{2}\}$.

Now we are ready for the main result — the final "rigidity" property of projective varieties:

THEOREM 5.5.24. *If $\mathscr{F}$ is a coherent sheaf over a projective variety $V \subset k\mathbb{P}^n$, then $\mathscr{F}(V)$ is a finite-dimensional vector space over k.*

PROOF. This follows immediately from the fact that $\mathscr{F}(V) = H^0(V, \mathscr{F})$ (corollary D.3.16 on page 553) and the cohomology of $\mathscr{F}$ is finitely generated, by theorem D.3.22 on page 558. □

5.6. Noether Normalization

It is possible to prove a version of theorem 2.5.12 on page 77 for projective varieties. Definition 4.6.27 on page 214 shows what a finite map *is* in the projective context. Corollary 5.5.16 on page 257 allows us to use a geometric argument to prove the result.

We begin with:

PROPOSITION 5.6.1. *Let $V \subset k\mathbb{P}^n$ be a projective variety and let $L \subset k\mathbb{P}^n$ be the d-dimensional linear subspace defined by linearly independent linear forms*

$$L_0 = \cdots = L_{n-d-1} = 0$$

such that $V \cap L = \varnothing$. Then the projection *(see example 5.4.5 on page 244)*

$$\pi_L : k\mathbb{P}^n \to k\mathbb{P}^{n-d-1} = \mathbb{P}(\mathbb{A}^{n-d})$$

defined by

$$\pi_L(x) = (L_0 : \cdots : L_{n-d-1})$$

is a finite map

$$\pi_L : V \to \pi_L(V)$$

REMARK. The fiber over every point of $k\mathbb{P}^{n-d-1}$ is a $d+1$-dimensional linear subspace of $k\mathbb{P}^n$. This can be seen by going to open affines or by simple computation: if $p \in k\mathbb{P}^{n-d-1}$ is in the image of π_L with homogeneous coordinates (x_i) then set

(5.6.1) $L_i = x_i$

where not all $x_i = 0$. We can use the linear independence of the L_i to get a set of $n - d - 1$ linear forms equivalent to this.

PROOF. Let $X = (X_0 : \cdots : X_n)$ be homogeneous coordinates of $k\mathbb{P}^n$. Since the linear forms, L_i, are linearly independent, there exists an $n \times n$ matrix, A, that transforms L_i into X_i — in other words, there exists an *isomorphism*

$$\mathbb{P}(A) : k\mathbb{P}^n \to k\mathbb{P}^n$$

that transforms π_L to projection onto the first $n - d$ coordinates. The open affine

$$U_i = \pi_L^{-1}(\mathbb{A}_i^{n-d-1}) \cap V$$

is given by $X_i \neq 0$, i.e. it is $V \cap \mathbb{A}_i^n$. We claim that

$$\pi_L : U_i \to \mathbb{A}_i^{n-d-1} \cap \pi_L(V)$$

is a finite map.

CLAIM 5.6.2. *An arbitrary element, $g \in k[U_i]$, satisfies a monic polynomial over $k[\mathbb{A}_i^{n-d-1} \cap \pi_L(V)]$.*

Proof of claim: We will begin by using $g \in k[U_i]$ to construct a regular map

$$\Pi_g : X \to k\mathbb{P}^{n-d}$$

and will use that to prove the claim.

Since $k[U_i]$ is a quotient of $k[\mathbb{A}_i^n]$, g is of the form

$$g\left(\frac{X_0}{X_i}, \ldots, \frac{X_{i-1}}{X_i}, \frac{X_{i+1}}{X_i}, \ldots, \frac{X_n}{X-i}\right) = \frac{G_i(X_0, \ldots, X_n)}{X_i^m}$$

where G_i is homogeneous of degree m. Define

$$\Pi_g \colon X \to k\mathbb{P}^{n-d}$$

by

(5.6.2)
$$Z_j = \begin{cases} X_j^m & \text{for } j = 0, \ldots, n-d-1 \\ G_i & \text{for } j = n-d \end{cases}$$

where $(Z_0 \colon \cdots \colon Z_{n-d})$ are the homogeneous coordinates of $k\mathbb{P}^{n-d}$. This is a regular map and its image in $k\mathbb{P}^{n-d}$ is closed by corollary 5.5.16 on page 257. Suppose $\Pi_g(X)$ is given by homogeneous equations

$$F_0 = \cdots = F_t = 0$$

Since V was disjoint from L it follows, in the present context (i.e., after transforming everything with $\mathbb{P}(A)$) that

$$Z_0 = \cdots = Z_{n-d-1} = F_0 = \cdots = F_t = 0$$

have no common solution in $k\mathbb{P}^{n-d}$. The Projective Nullstellensatz (theorem 5.1.4 on page 220) implies that

$$Z_j^r \in (Z_0, \ldots, Z_{n-d-1}, F_0, \ldots, F_t)$$

for all $j = 0, \ldots, n-d-1$ and some $r > 0$. In particular

$$Z_{n-d}^r \in (Z_0, \ldots, Z_{n-d-1}, F_0, \ldots, F_t)$$

which gives an equation

(5.6.3)
$$Z_{n-d}^r = \sum_{j=0}^{n-d-1} Z_j^{n_j} \cdot h_j(Z) + \sum_{\ell=1}^{t} F_\ell \cdot w_\ell(Z)$$

for some polynomials h_j, w_ℓ. Break the polynomial

$$Q(Z) = Z_{n-d}^r - \sum_{j=0}^{n-d-1} Z_j^{n_j} \cdot h_j(Z) - \sum_{\ell=1}^{t} F_\ell \cdot w_\ell(Z)$$

into its homogeneous components

$$Q(Z) = Q_0 + Q_1 + \cdots$$

The rth component will look like

(5.6.4)
$$Q_r(Z) = Z_{n-d}^r - \sum_{j=0}^{n-d-1} Z_j^{n_j} \cdot h_j'(Z) - \sum_{\ell=1}^{t} F_\ell \cdot w_\ell'(Z)$$

Since $Q(Z)$ vanishes for all Z, by equation 5.6.3, the same must be true of its homogeneous components, so equation 5.6.4 gives

$$Z_{n-d}^r - \sum_{j=0}^{n-d-1} Z_j^{n_j} \cdot h_j'(Z) - \sum_{\ell=1}^{t} F_\ell \cdot w_\ell'(Z) = 0$$

Now we substitute equation 5.6.2 into this to get

$$G_i^r - \sum_{j=0}^{n-d-1} X_j^{m \cdot n_j} \cdot h_j'(X_*^m, G_i) - \sum_{\ell=1}^{t} F_\ell(X_*^m, G_i) \cdot w_\ell'(X_*^m, G_i) = 0$$

Dividing by $X_i^{r \cdot m}$ gives

$$g^r - \sum_{j=0}^{n-d-1} \left(\frac{X_j}{X_i}\right)^{m \cdot n_j} \cdot h_j' \left(\left(\frac{X_*}{X_i}\right)^m, g\right)$$

$$- \sum_{\ell=1}^{t} F_\ell \left(\left(\frac{X_*}{X_i}\right)^m, g\right) \cdot w_\ell' \left(\left(\frac{X_*}{X_i}\right)^m, g\right) = 0$$

or

$$g^r - \sum_{j=0}^{r-1} g^j a_j(x_0, \dots, x_{i-1}, 1, x_{i+1}, \dots x_{n-d-1}) = 0$$

where $x_j = X_j/X_i$ are the coordinates in $\mathbb{A}_i^{n-d-1}$, and the a_j are polynomials. Claim 5.6.2 on page 261 and the result are proved. $\qquad\square$

Now we can prove a *projective* version of Noether Normalization:

COROLLARY 5.6.3. *Let $V \subset k\mathbb{P}^n$ be an irreducible projective variety. Then there exists a dominating finite map*

$$V \to k\mathbb{P}^d$$

for some d.

REMARK. Corollary 5.5.16 on page 257 allows us to give a proof that is considerably more geometric than that of the affine case (theorem 2.5.12 on page 77).

Since finite maps preserve dimension, the number d in the statement is equal to the *dimension* of V.

PROOF. If $V = k\mathbb{P}^n$, we are done. Otherwise, there exists a point $p \in k\mathbb{P}^n \setminus V$. A linear projection centered at this point

$$k\mathbb{P}^n \to k\mathbb{P}^{n-1}$$

defines a finite map

$$\pi_p \colon V \to \pi_p(V)$$

and $\pi_p(V)$ is a projective variety, by corollary 5.5.16 on page 257. Continue this process until an image of V is equal to all of its containing projective space. $\qquad\square$

Because dimension is a local property, most of the dimension-theoretic results for affine varieties carry over to projective ones:

PROPOSITION 5.6.4. *Let $V, W \subset k\mathbb{P}^n$ be projective varieties*

(1) *If $F \in k[X_0, \dots, X_n]$ is a form that does not vanish identically on V, then the dimension of each component of $V \cap \mathcal{P}((F))$ is $\dim V - 1$ (see definition 5.1.2 for the terminology $\mathcal{P}((F))$).*
(2) *The intersection satisfies*

$$\dim(V \cap W) \geq \dim V + \dim W - n$$

(3) *If $f \colon V \to W$ is a surjective morphism and if $p \in W$ is a point,*

$$\dim F \geq \dim V - \dim W$$

for any component, F, of $f^{-1}(p)$.

REMARK. Just apply corollary 2.8.30 on page 111, corollary 2.8.34 on page 112 and corollary 2.8.36 on page 113, respectively, to open affines of V and W.

Another variation of this is:

LEMMA 5.6.5. *Let $V \subset k\mathbb{P}^N$ be an m-dimensional variety. Then there exists $m + 1$ homogeneous polynomials $f_1 \ldots, f_{m+1} \in k[X_0, \ldots, X_n]$ that do not simultaneously vanish at any point of V.*

PROOF. We prove this by induction on m. If $m = -1$ (the empty set), there is nothing to prove. Now choose a point p_i in each irreducible component of V and a homogeneous polynomial f_1, that does not vanish at any of these points. The subvariety
$$f_1 = 0$$
has $m - 1$ dimensional irreducible components (by proposition 5.6.4 on the previous page) and the inductive hypothesis implies the conclusion. $\qquad \square$

Statement 2 above leads to an interesting way of characterizing the dimension of a variety:

COROLLARY 5.6.6. *Let $V \subset k\mathbb{P}^n$ be an irreducible projective variety and suppose j is the maximum dimension of a linear subspace, L, that is disjoint from V. Then $\dim V = n - j - 1$.*

PROOF. A linear subspace of dimension $j + 1$ is defined by $\leq n - j - 1$ linear forms. Proposition 5.6.4 on the preceding page implies that intersection of V with the linear subspace of $k\mathbb{P}^n$ defined by these forms is of dimension ≥ 0, i.e. is nonempty. $\qquad \square$

Now we will consider the question:

> If V is a projective variety of dimension d, what is the smallest value of n such that there exists an isomorphic embedding of V into $k\mathbb{P}^n$.

We begin with

LEMMA 5.6.7. *If V is a smooth variety and $f \colon V \to W$ is a finite map, then f is an isomorphic embedding if*

(1) *$f^{-1}(w)$ is a single point, for any $w \in W$*
(2) *$df \colon T_{V,v} \to T_{W,f(v)}$ is a isomorphic embedding of tangent spaces for every $v \in V$.*

REMARK. This intuitively clear result is surprisingly difficult to prove.

PROOF. Condition 1 above implies the *existence* of a map $f^{-1} \colon \bar{V} = f(V) \to V$. We need to verify that it is algebraic and regular. We will prove this in a neighborhood of every point. Let $v \in V$ with $f(v) = w$ and let A be an open affine of V containing v with affine image $B = f(A) \subset \bar{V}$.

Since the tangent space is dual to $\mathfrak{m}_v / \mathfrak{m}_v^2$ (where $\mathfrak{m}_v \subset \mathcal{O}_{V,v}$ is the maximal ideal of functions that vanish at v) the second condition is equivalent to saying that
$$f^* \colon \frac{\mathfrak{m}_{f(v)}}{\mathfrak{m}_w^2} \to \frac{\mathfrak{m}_v}{\mathfrak{m}_v^2}$$

is surjective for all $v \in V$. If $\mathfrak{m}_{f(v)} = (x_1, \ldots, x_n) \subset \mathcal{O}_{V,w}$, then $f^*(x_i) + \mathfrak{m}_v^2$ generate $\mathfrak{m}_v / \mathfrak{m}_v^2$. If we define the module over $\mathcal{O}_{V,v}$

$$M = \frac{\mathfrak{m}_v}{f^*(\mathfrak{m}_w) \cdot \mathcal{O}_{V,v}}$$

then $M/\mathfrak{m}_v \cdot M = 0$ or $M = \mathfrak{m}_v \cdot M$, and Nakayama's Lemma (A.1.80 on page 378) implies that $M = 0$ and

(5.6.5) $$\mathfrak{m}_v = f^*(\mathfrak{m}_w) \cdot \mathcal{O}_{V,v}$$

We claim that $\mathcal{O}_{V,v}$ is a finitely-generated module over $f^*(\mathcal{O}_{W,w})$. Since we already know that $k[A]$ is a finitely-generated module over $f^*k[B]$ and $\mathcal{O}_{W,w}$ is a localization of $k[B]$, it suffices to show that $\mathcal{O}_{V,v} \subset (f^*(S))^{-1} k[A]$ where S is a set of invertible elements of $\mathcal{O}_{W,w}$.

If $x \in k[A] \setminus \mathfrak{m}_v$, we claim that there exists an element $b \in k[B]$ with $b \notin \mathfrak{m}_w$ such that $f^*(b) = xy$ with $y \in k[A]$. The set of points Z where x vanishes is a closed subset of V, hence its image under f is closed by corollary 5.5.16 on page 257. Since f is 1-1, $w \notin f(Z)$, and it follows that there is a function $c \in k[B]$ with $c|f(Z) = 0$ and $c(w) \neq 0$. So $f^*(c)|Z = 0$ and $f^*(c)(x) \neq 0$. The Strong Nullstellensatz (2.2.5 on page 41) implies that $f^*(c)^n \in (x)$ for some value of n. If we set $b = c^n$, $f^*(b) = xy$ as claimed.

We claim that $\mathcal{O}_{V,v} = f^*(\mathcal{O}_{W,w})$. Consider the $\mathcal{O}_{W,w}$-module $M = \mathcal{O}_{V,v}/f^*(\mathcal{O}_{W,w})$. Equation 5.6.5 implies that $M/\mathfrak{m}_w \cdot M = 0$ so $M = \mathfrak{m}_W \cdot M$ and $M = 0$.

If $u_1, \ldots, u_t \in k[A]$ are a basis for it as a $k[B]$-module, we can conclude that $u_i \in \mathcal{O}_{V,v} = f^*(\mathcal{O}_{W,w})$. Let $v_i = (f^*)^{-1}(u_i) = r_i/s_i$, let $h = \prod s_i$, and let

$$\begin{aligned} B' &= B \setminus B(h) \\ A' &= A \setminus A(f^*(h)) \end{aligned}$$

where $A(f^*(h)$ and $B(h)$ are the (closed) sets of points where $f^*(h)$ and h vanish, respectively. Then $k[A'] = \sum f^*(k[B'] \cdot u_i$ and f induces an isomorphism between A' and B'.   □

This and proposition 5.6.1 on page 261 immediately imply that:

COROLLARY 5.6.8. *Let $V \subset k\mathbb{P}^n$ be a variety and let $p \in k\mathbb{P}^n \setminus V$. If every line through p intersects V in at most one point and is not contained in the tangent space of V at any point, then projection from p is an isomorphic embedding $V \to k\mathbb{P}^{n-1}$.*

We come to the main result:

THEOREM 5.6.9. *An n-dimensional smooth projective variety is isomorphic to a subvariety of $k\mathbb{P}^{2n+1}$.*

REMARK. This is an algebraic analogue of the Whitney Embedding Theorem which states that a compact n-manifold can be embedded into $\mathbb{R}^{2n}$.

PROOF. We will show that $V \subset k\mathbb{P}^M$ is a smooth n-dimensional subvariety and $M > 2n + 1$ then there exists a point $p \in k\mathbb{P}^M$ satisfying the hypotheses of corollary 5.6.8. Let $U_1 \subset k\mathbb{P}^M$ denote the set of points that violate its first hypothesis — there are lines through points in U_1 that intersect V in more than

one point. Let $U_2 \subset k\mathbb{P}^M$ denote the set of points that violate its second hypothesis: lines through points in U_2 lie in tangent spaces of V. These are both subvarieties of $k\mathbb{P}^M$ and we will estimate their dimensions.

Let $\Gamma \subset k\mathbb{P}^M \times V \times V$ be the set of triples (a, b, c) that are collinear. The set Γ is closed and projections to the factors give regular maps

$$f:\Gamma \rightarrow k\mathbb{P}^m$$
$$g:\Gamma \rightarrow V \times V$$

If $(b, c) = g((a, b, c)) \in V \times V$ with $b \neq c$, then $g^{-1}(b, c)$ is the set of points on the line through b and c. It follows that $\dim g^{-1}(b, c) = 1$ and proposition 5.6.4 on page 263 implies that

$$\dim \Gamma \leq 2n + 1$$

Since $f(\Gamma) = U_1$, we conclude that $\dim U_1 \leq 2n + 1$.

To analyze U_2 consider the set $\Gamma' \subset k\mathbb{P}^M \times V$ of point (a, b) where $a \in T_{V,b}$ — the tangent space. We get projections

$$f:\Gamma' \rightarrow k\mathbb{P}^M$$
$$g:\Gamma' \rightarrow V$$

with $\dim g^{-1}(b) = n$, for any $b \in V$. Proposition 5.6.4 on page 263 implies that $\dim \Gamma' \leq 2n$ and $\dim U_2 = \dim f(\Gamma') \leq 2n$.

It follows that, if $M > 2n + 1$ then $U_1 \cup U_2 \neq k\mathbb{P}^M$ and we can apply corollary 5.6.8 on the previous page. $\square$

EXERCISES.

1. Suppose $V = \mathcal{P}((F_1, \ldots, F_n)) \in k\mathbb{P}^t$ is a one-dimensional smooth variety. Show that there exists homogeneous forms $G_1, G_2 \in k[X, Y, Z, W]$ such that

$$\frac{k[X, Y, Z, W]}{(G_1, G_2)} = k_H[V]$$

2. Let $f_i \in k[X_0, \ldots, X_n]$, for $i = 1, \ldots, r$ be homogeneous polynomials of the same degree that never simultaneously vanish when at least one of the X_i is nonzero. Show that they define a finite map

$$F:k\mathbb{P}^n \rightarrow \operatorname{im} F \subset k\mathbb{P}^{r-1}$$

5.7. Graded ideals and modules

5.7.1. Introduction. In this section, we analyze the homogeneous coordinate ring of a projective variety. As has been remarked before, the coordinate ring of a projective variety does not determine its regular functions. Nevertheless its structure as a *graded* ring (see definition A.4.39 on page 434) contains a great deal of geometric information about the variety.

5.7.2. Hilbert functions and polynomials. It is interesting that the number of monomials in a graded ring as a function of their degree contains useful geometric information.

We begin by defining:

DEFINITION 5.7.1. Let M be a graded module over a graded k-algebra. Then the *Hilbert function* of M, $h_M(s)$ is a function

$$h_M(s) = \dim M_s$$

for all integral $s \geq 0$, as a vector space over k.

REMARK. Naturally, this function can be defined for graded k-algebras too.

The following is clear:

PROPOSITION 5.7.2. *Given a short exact sequence of graded modules*

$$0 \to M_1 \to M_2 \to M_3 \to 0$$

over a graded ring, R, we get

$$h_{M_2}(s) = h_{M_1}(s) + h_{M_3}(s)$$

This and equation A.4.9 on page 436 immediately implies:

COROLLARY 5.7.3. *Let M be a graded module over a graded ring with prime filtration*

$$0 = M_0 \subset M_1 \subset \cdots \subset M_t = M$$

such that for each i

$$\frac{M_{i+1}}{M_i} \cong \frac{R}{\mathfrak{p}_i}(\ell_i)$$

Then

(5.7.1)
$$h_M(s) = \sum_{j=1}^{t} h_{R/\mathfrak{p}_i(\ell_i)}(s)$$

for all $s \geq 0$ and

(5.7.2)
$$\mathcal{P}(\text{Ann}(M)) = \bigcup_{j=1}^{t} \mathcal{P}(\text{Ann}(R/\mathfrak{p}_i(\ell_i)))$$

PROOF. If

$$0 \to M_1 \to M_2 \to M_3 \to 0$$

is a short exact sequence of modules, it is not hard to see that

$$h_{M_2}(s) = h_{M_1}(s) + h_{M_3}(s)$$

and exercise 3 on page 38 shows that

$$\mathcal{V}(\text{Ann}(M_2)) = \mathcal{V}(\text{Ann}(M_1)) \cup \mathcal{V}(\text{Ann}(M_3))$$

which implies that

$$\mathcal{P}(\text{Ann}(M_2)) = \mathcal{P}(\text{Ann}(M_1)) \cup \mathcal{P}(\text{Ann}(M_3))$$

and a simple induction proves the result. $\square$

In [**77**], Hilbert proved the remarkable result that the Hilbert function of a finitely generated graded module over a noetherian graded ring becomes a *polynomial*, for s sufficiently large. This polynomial is called the *Hilbert Polynomial* of M, denoted $\mathcal{H}_M(s)$.

EXAMPLE 5.7.4. If $G = k[X_0, \ldots, X_n] = k_H[k\mathbb{P}^n]$, the number of monomials of degree s is

$$h_G(s) = \binom{s+n}{n} = \frac{(s+1)\cdots(s+n)}{n!}$$

a polynomial of degree n (see proposition 5.4.6 on page 244). In this case, the Hilbert function is a polynomial for *all* $s \geq 0$.

Even a trivial example can be instructional:

EXAMPLE 5.7.5. If G is a finite-dimensional graded vector space over k, then

$$h_G(s) = 0$$

for s sufficiently large. For small values of s ($< \dim G$), $\dim G_i$ can be assigned at random. Above this range, $\dim G_i = 0$, consistently. In this case, the Hilbert polynomial is zero,.

Here's an example with a quotient ring

EXAMPLE 5.7.6. Let $G = k[X_0, \ldots, X_n]$ and let $\mathfrak{I} = (f)$ where f is a homogeneous polynomial of degree d and suppose $R = G/\mathfrak{I}$. Then multiplication by f defines an isomorphism

$$\times f \colon G_t \to \mathfrak{I}_{t+d}$$

or $\mathfrak{I}_r = G_{r-d}$, for $r \geq d$. Corollary 5.7.3 on the preceding page (applied to a short exact sequence) gives

$$h_R(s) = \begin{cases} \binom{s+n}{n} & \text{if } s < d \\ h_G(s) - h_G(s-d) = \binom{s+n}{n} - \binom{s+n-d}{n} & \text{if } s \geq d \end{cases}$$

In this case, we conclude that the Hilbert polynomial is

$$\mathcal{H}_R(s) = \binom{s+n}{n} - \binom{s+n-d}{n}$$

for all s. Since $\binom{s+n}{n}$ and $\binom{s+n-d}{n}$ both have $s^n/n!$ as their leading term, $\mathcal{H}_R(s)$ is a degree-$n-1$ polynomial. If $n = 2$, we get

$$\mathcal{H}_R(s) = ds + \frac{3d - d^2}{2}$$

The last two examples clarify the *sense* in which $h_R(s)$ is a polynomial for "sufficiently large values of s." For small values of s, the dimension of G_s may be irregular but it "stabilizes" for s sufficiently large. The Hilbert polynomial is the polynomial that $h_R(s)$ eventually becomes.

Corollary 5.7.3 on the previous page shows that, in order to compute Hilbert polynomials of *any* module, it suffices to compute them for modules of the form $R/\mathfrak{p}$, where $\mathfrak{p}$ is a prime ideal. Example 5.7.6 easily generalizes to a method for computing many Hilbert polynomials:

EXAMPLE 5.7.7. Suppose $\mathfrak{I} = (g_1, \ldots, g_m) \subset R = k[X_0, \ldots, X_n]$ is an ideal, where the g_i are monomials. If $\succ$ is an ordering on monomials (see section 2.3 on page 45), we can compute

$$\mathcal{H}_{R/\mathfrak{I}}(s)$$

by induction on m. If $m = 1$, use the computation in example 5.7.6 on the facing page. Otherwise, let g_1 be the highest ranked monomial in $\mathfrak{I}$ and write

$$\mathfrak{I} = (g_1) + \mathfrak{I}'$$

where $\mathfrak{I}'$ is generated by $< m$ monomials. We have an exact sequence

$$R(-d) \xrightarrow{\times g_1} \frac{R}{\mathfrak{I}'} \to \frac{R}{\mathfrak{I}} \to 0$$

where d is the degree of g_1 and the kernel of $\times g_1$ is

$$
\begin{aligned}
(\mathfrak{I}' : g_1) &= \{ g \in R | g \cdot g_1 \in \mathfrak{I}' \} \\
&= \left(\frac{g_2}{\gcd(g_1, g_2)}, \ldots, \frac{g_m}{\gcd(g_m, g_1)} \right)
\end{aligned}
$$

which is also generated by $< m$ monomials[4]. We get an exact sequence

$$0 \to \frac{R}{(\mathfrak{I}' : g_1)}(-d) \to \frac{R}{\mathfrak{I}'} \to \frac{R}{\mathfrak{I}} \to 0$$

and

$$\mathcal{H}_{R/\mathfrak{I}}(s) = \mathcal{H}_{R/\mathfrak{I}'}(s) - \mathcal{H}_{R/(\mathfrak{I}' : g_1)}(s - d)$$

It follows that Hilbert polynomials are *always* alternating sums of suitable binomial coefficients.

We give one last example:

EXAMPLE 5.7.8. Let $p = (a_0, \ldots, a_n) \in k\mathbb{P}^n$ be a point with $a_0 \neq 0$. This lies in $\mathbb{A}_0^n$ and is given by

$$\frac{X_j}{X_0} = \frac{a_j}{a_0}$$

so that the homogeneous coordinate ring is $G = k[X_0, \ldots, X_n]/\mathfrak{P}$, where $\mathfrak{P} = (a_0 \cdot X_i - a_j \cdot X_0)$ for $j = 1, \ldots .n$. It follows that a monomial

$$X_0^{j_0} \cdots X_n^{j_n}$$

is equivalent, modulo $\mathfrak{P}$, to

$$\frac{a_1 \cdots a_n}{a_0^n} \cdot X_0^{\sum j_i}$$

so $h_G(s) = \mathcal{H}_G(s) = 1$, for all $s \geq 0$.

In order to prove our main result, we will need:

LEMMA 5.7.9. *Let $f : \mathbb{Z} \to \mathbb{Z}$ be a function with the property that $\Delta f(n) = f(n) - f(n-1) \in \mathbb{Q}[n]$ is a degree-d polynomial for $n > n_0$ and leading coefficient $a \in \mathbb{Q}$. Then, for $n > n_0 + 1$, $f(n) \in \mathbb{Q}[n]$ is a degree-$(d+1)$ polynomial function with leading coefficient $a/(d+1)$.*

---

[4]Note that the greatest common divisor of two *monomials* is easy to compute.

PROOF. This follows quickly from calculus of finite differences — see [**59**].
If S is the $\mathbb{Q}$-module of all functions $f: \mathbb{Q} \to \mathbb{Q}$, define two module homomorphisms:

(1) $\Delta: S \to S$, $(\Delta f)(n) = f(n) - f(n-1)$ and
(2) $\Sigma: S \to S$, $(\Sigma f)(n) = \sum_{j=1}^{n} f(j)$

It is not hard to see that

(5.7.3) $$\Delta \circ \Sigma = 1: S \to S$$

and $(\Sigma \circ \Delta f)(n) = f(n) - f(0)$ so that they are (almost) inverses. Now define

$$X^{\underline{d}} = \begin{cases} 1 & \text{if } d = 0 \\ X(X-1) \cdots (X-d+1) & \text{if } d > 0 \end{cases}$$

for $d \geq 0$ — a kind of "power" of X called the *falling factorial* or Pochhammer symbol[5]. A simple calculation shows that

$$\Delta X^{\underline{d}} = d X^{\underline{d-1}}$$

for $d > 0$, and equation 5.7.3 implies that

$$\Sigma X^{\underline{d}} = \frac{X^{\underline{d+1}}}{d+1}$$

for $d \geq 0$. Now, given a degree-d polynomial $p(X)$, we can write it as a linear combination of the $\{X^{\underline{j}}\}$ with the highest value of j being d (for instance, by plugging it into the "finite difference Taylor series"). Then apply the Σ operator to get a linear combination, $q(X)$, of the $\{X^{\underline{j}}\}$ with the highest value of j equal to $d+1$. This is a degree-$d+1$ polynomial of X such that $\Delta q(X) = p(X)$. If $p(X) \in \mathbb{Z}$ whenever $X \in \mathbb{Z}$, the same is true of $q(X)$.

If $\Delta f(n) = p(n)$ for $n > n_0$, then $q(n) = f(n) + c$ for $n > n_0 + 1$ for some constant c. $\qquad \square$

Note that the degree of the Hilbert polynomial is equal to the *dimension* of $k\mathbb{P}^n$.

This is no coincidence and leads to our main result:

THEOREM 5.7.10. *Let G be a finitely-generated graded module over the graded k-algebra $R = k[X_0, \ldots, X_n]$. Then, for s sufficiently large, the Hilbert function $h_G(s)$ is a polynomial, $\mathcal{H}_G(s)$, and*

$$\deg \mathcal{H}_G(s) = \dim \mathcal{P}(\mathrm{Ann}(G))$$

If $d = \deg \mathcal{H}_G(s)$ and c_d is the leading coefficient of $\mathcal{H}_G(s)$, then $c_d \cdot d! > 0$ is an integer.

REMARK. If $V \subset k\mathbb{P}^n$ is a projective variety and $G = k_H[V]$, then the degree of the Hilbert polynomial equals the dimension of V. Dimension can even be *defined* this way. Exercise 3 on page 273 shows that the Hilbert polynomial depends on V and not necessarily the ideal used to define it.

PROOF. Corollary 5.7.3 on page 267 implies that it suffices to prove the result for $G = R/\mathfrak{p}(\ell)$:

---

[5]The standard notation for falling factorial is $X_{(d)}$ but we use Knuth's notation to suggest "raising X to a power."

In equation 5.7.2 on page 267, the dimension is the *maximum* of the dimensions of the right side and in equation 5.7.1 on page 267 the degree is the maximum of the *degrees* that occur on the right side (where leading terms of polynomials of the same degree *never cancel*). So if the result is true for $G = R/\mathfrak{p}(\ell)$, it is true in general.

Since

$$\mathcal{H}_{R/\mathfrak{p}(\ell)}(s) = \mathcal{H}_{R/\mathfrak{p}}(s + \ell)$$

it suffices to prove the result for $G = R/\mathfrak{p}$, representing an irreducible projective variety. We do this by induction on the dimension of $\mathcal{P}(\mathrm{Ann}(G))$. When $\mathcal{P}(\mathrm{Ann}(G))$ is zero-dimensional, it is a single point. Example 5.7.8 on page 269 shows that the Hilbert polynomial is equal to 1, establishing the ground case.

Now suppose $\mathfrak{p}$ is a homogeneous prime ideal and suppose $X_i \notin \mathfrak{p}$, and $G = R/\mathfrak{p}$. If $\dim \mathcal{P}(G) = d$, then

$$\dim \left(\mathcal{P}(G) \cap \mathcal{P}((X_i)) \right) = d - 1$$

by proposition 5.6.4 on page 263 and we get an exact sequence

$$0 \to X_i \cdot G \to G \to G/(X_i) \to 0$$

Since multiplication by X_i raises degree by 1, this is not a sequence of *graded* modules. To get that, we need to shift grading

$$0 \to X_i \cdot G(-1) \to G \to G/(X_i) = Q \to 0$$

Since $X_i \cdot G \cong G$, proposition 5.7.2 on page 267 gives

$$\mathcal{H}_G(s) = \mathcal{H}_Q(s) + \mathcal{H}_G(s - 1)$$

or

$$\mathcal{H}_G(s) - \mathcal{H}_G(s - 1) = \mathcal{H}_Q(s)$$

where $\mathcal{P}(\mathrm{Ann}(Q)) = \mathcal{P}(G) \cap \mathcal{P}((X_i))$. By induction, $\mathcal{H}_Q(s)$ is a polynomial of degree $d - 1$ with leading coefficient $a/(d-1)!$ and lemma 5.7.9 on page 269 implies that $\mathcal{H}_G(s)$ is a polynomial of degree d with leading coefficient $a/d!$. This proves the result. $\square$

This gives a geometric proof of the following:

COROLLARY 5.7.11. *Let $\mathfrak{I} \subset R = k[X_0, \dots, X_n]$ be an ideal. Then*

$$\deg \mathcal{H}_{R/\mathfrak{I}}(s) = \deg \mathcal{H}_{R/\sqrt{\mathfrak{I}}}(s)$$

PROOF. This follows immediately from theorem 5.7.10 on the preceding page and the fact that

$$\mathcal{P}(\mathrm{Ann}(R/\mathfrak{I})) = \mathcal{P}(\mathrm{Ann}(R/\sqrt{\mathfrak{I}}))$$

(see corollary 2.2.6 on page 42). $\square$

There is a great deal of other information we can extract from the Hilbert polynomial. The fact that we kept track of its leading coefficient hints that this may be significant.

DEFINITION 5.7.12. Let $V \subset k\mathbb{P}^n$ be a projective variety with homogeneous coordinate ring R. Then we define the Hilbert polynomial of V via

$$\mathcal{H}_V(s) = \mathcal{H}_R(s)$$

and the *degree of* V, denoted $\deg V$ as

$$(\deg \mathcal{H}_V(s))! \cdot c$$

where c is the leading coefficient of $\mathcal{H}_V(s)$.

REMARK. Although this use of the term "degree" is rather confusing (especially since it is computed using a Hilbert polynomial whose *degree* is completely different), it is standard. It is always a positive integer.

Roughly speaking, the degree of a variety is the number of times it intersects "most" hyperplanes of complementary dimension. See proposition 5.8.8 on page 280 for a precise statement.

We need a little lemma:

LEMMA 5.7.13. *Let $p(X) \in k[X]$ be a polynomial with leading coefficient, c, of degree n. Then the leading coefficient of*

$$p(X) - p(X - d)$$

is $n \cdot d \cdot c$ in degree $n - 1$.

PROOF. Just use the binomial theorem. $\square$

Here are some examples of degrees of varieties:

(1) Example 5.7.4 on page 268 implies that $\deg k\mathbb{P}^n = 1$ for any n.
(2) Example 5.7.6 on page 268 implies if $H \subset k\mathbb{P}^n$ is a hypersurface defined by a degree-d homogeneous polynomial, its Hilbert polynomial is

$$\binom{s+n}{n} - \binom{s+n-d}{n} = \frac{d}{(n-1)!}s^{n-1} + \cdots$$

(by lemma 5.7.13) so the degree of H is d. This justifies the use of the word "degree" in definition 5.7.12 on page 272: the degree of a hypersurface is equal to the degree of its defining polynomial.
(3) Example 5.7.8 on page 269 shows that the degree of a point in $k\mathbb{P}^n$ is always 1.

One important property of degrees

LEMMA 5.7.14. *Let $V, W \subset k\mathbb{P}^n$ be projective varieties of the same dimension. If $\dim(V \cap W) < \dim V = \dim W$, then*

$$\deg(V \cup W) = \deg V + \deg W$$

PROOF. Let V and W be defined by homogeneous ideals $\mathfrak{v}, \mathfrak{w} \subset R = k[X_0, \ldots, X_n]$. Then $\mathcal{I}(V \cup W) = \mathfrak{v} \cap \mathfrak{w}$ and we claim that there exists an exact sequence:

$$0 \to \frac{R}{\mathfrak{v} \cap \mathfrak{w}} \to \frac{R}{\mathfrak{v}} \oplus \frac{R}{\mathfrak{w}} \to \frac{R}{\mathfrak{v} + \mathfrak{w}} \to 0$$

The leftmost map is defined by

$$\frac{R}{\mathfrak{v} \cap \mathfrak{w}} \quad \rightarrow \quad \frac{R}{\mathfrak{v}} \oplus \frac{R}{\mathfrak{w}}$$
$$r \quad \mapsto \quad (r, -r)$$

If $r \in R/(\mathfrak{v} \cap \mathfrak{w})$ maps to 0 it must be contained in $\mathfrak{v}$ and $\mathfrak{w}$, hence in $\mathfrak{v} \cap \mathfrak{w}$, so it is 0 in $R/(\mathfrak{v} \cap \mathfrak{w})$, and this map is well-defined.

The maps

$$\frac{R}{\mathfrak{v}} \quad \rightarrow \quad \frac{R}{\mathfrak{v} + \mathfrak{w}}$$
$$\frac{R}{\mathfrak{w}} \quad \rightarrow \quad \frac{R}{\mathfrak{v} + \mathfrak{w}}$$

are induced by inclusions of the ideals. If $(r_1, r_2) \in R/\mathfrak{v} \oplus R/\mathfrak{w}$ maps to 0 in $R/(\mathfrak{v} + \mathfrak{w})$ then $r_1 + r_2 \in \mathfrak{v} + \mathfrak{w}$, so $r_2 \equiv -r_1 \pmod{\mathfrak{v} + \mathfrak{w}}$. Since we may vary r_1 within $\mathfrak{v}$ and r_2 within $\mathfrak{w}$, we get $r_1' \equiv r_1 \pmod{\mathfrak{v}}$ and $r_2' \equiv r_2 \pmod{\mathfrak{w}}$ such that $r_2' = -r_1'$, which is in the image of $r_1' \in R/\mathfrak{v} \cap \mathfrak{w}$.

We get

(5.7.4) $$\mathcal{H}_{R/\mathfrak{v} \cap \mathfrak{w}}(s) + \mathcal{H}_{R/(\mathfrak{v} + \mathfrak{w})}(s) = \mathcal{H}_{R/\mathfrak{v}}(s) + \mathcal{H}_{R/\mathfrak{w}}(s)$$

Corollary 5.7.11 on page 271 implies that

$$\deg \mathcal{H}_{R/(\mathfrak{v} + \mathfrak{w})}(s) = \deg \mathcal{H}_{R/\sqrt{(\mathfrak{v} + \mathfrak{w})}}(s)$$
$$= \dim V \cap W < \dim V = \dim W$$

so $\mathcal{H}_{R/(\mathfrak{v} + \mathfrak{w})}(s)$ in equation 5.7.4 contributes *nothing* to the *leading* terms of the other polynomials. It follows that the leading term of $\mathcal{H}_{R/\mathfrak{v} \cap \mathfrak{w}}(s)$ is the *sum* of the leading terms of $\mathcal{H}_{R/\mathfrak{v}}(s)$ and $\mathcal{H}_{R/\mathfrak{w}}(s)$, proving the result. $\qquad\square$

EXERCISES.

1. Use example 5.7.7 on page 269 to compute $\mathcal{H}_{R/\mathfrak{J}}(s)$ where $\mathfrak{J} = (f, g) \subset k[X_0, X_1, X_2, X_3]$, where f and g are monomials of degrees d_1 and d_2, respectively and $\gcd(f, g) = 1$.

2. If $\mathfrak{a} \subset k[X_0, \ldots, X_n] = R$ is a radical homogeneous ideal and $f \in k[X_0, \ldots, X_n]$ is a homogeneous element such that $f \cdot \mathfrak{i}^t \subset \mathfrak{a}$ for some t (where $\mathfrak{i} = (X_0, \ldots, X_n)$ is the irrelevant ideal) show that

$$\mathcal{H}_{R/\mathfrak{a}}(s) = \mathcal{H}_{R/(\mathfrak{a} + (f))}(s)$$

3. If $\mathfrak{a}, \mathfrak{b} \subset k[X_0, \ldots, X_n] = R$ and $\mathcal{P}(\mathfrak{a}) = \mathcal{P}(\mathfrak{b}) \subset k\mathbb{P}^n$, show that

$$\mathcal{H}_{R/\mathfrak{a}}(s) = \mathcal{H}_{R/\mathfrak{b}}(s)$$

5.8. Bézout's Theorem revisited

5.8.1. Preliminaries. Suppose $V \subset k\mathbb{P}^n$ is a projective variety of dimension r and $W \subset k\mathbb{P}^n$ is a hypersurface. If the associated homogeneous ideals are $\mathcal{I}(V) = \mathfrak{v}$ and $\mathcal{I}(W) = \mathfrak{w} = (F)$ for some $F \in k[X_0, \ldots, X_n]$, respectively, then the homogeneous coordinate ring of $V \cap W$ is given by

$$(5.8.1) \qquad R = \frac{k[X_0, \ldots, X_n]}{\sqrt{\mathfrak{v} + \mathfrak{w}}}$$

Furthermore, the irreducible components of this intersection have coordinate rings $k[X_0, \ldots, X_n]/\mathfrak{p}_i$ where the $\mathfrak{p}_i$ are the primes that occur in the *prime filtration* (see theorem A.1.77 on page 375) of R — see exercise 4 on page 79.

Unfortunately, it is not enough to determine the intersections as varieties — chapter 1 shows that these intersections must be counted with *multiplicities* in order to get a correct theorem.

Example 2.2.8 on page 43, lemma 3.3.34 on page 140, and example 4.3.4 on page 169 suggest the correct course of action: in equation 5.8.1, *refrain* from taking the radical. Instead, define

$$(5.8.2) \qquad \hat{R} = \frac{k[X_0, \ldots, X_n]}{\mathfrak{v} + \mathfrak{w}}$$

This will generally not be a reduced ring but there will exist a canonical map

$$\hat{R} \to R$$

induced by the inclusion $\mathfrak{v} + \mathfrak{w} \hookrightarrow \sqrt{\mathfrak{v} + \mathfrak{w}}$.

We will have to study how the rings $\hat{R}$ and R are related. At first glance, we conclude:

LEMMA 5.8.1. *Let $\mathfrak{I} \subset A = k[X_0, \ldots, X_n]$ be an ideal. Then the canonical map*

$$f \colon \hat{R} = \frac{A}{\mathfrak{I}} \to \frac{A}{\sqrt{\mathfrak{I}}} = R$$

induces a one-to-one correspondence between prime ideals

$$\hat{\mathfrak{p}} \leftrightarrow \mathfrak{p}$$

where $\hat{\mathfrak{p}} \subset \hat{R}$, and $\mathfrak{p} \subset R$ and $f(\hat{\mathfrak{p}}) = \mathfrak{p}$. Furthermore, for every prime ideal $\hat{\mathfrak{p}} \subset \hat{R}$, there is a canonical isomorphism

$$\frac{\hat{R}}{\hat{\mathfrak{p}}} \cong \frac{R}{\mathfrak{p}}$$

PROOF. First note that the kernel of f is the ideal, $\mathfrak{N}$, of all *nilpotent* elements of $\hat{R}$. Theorem A.1.47 on page 360 shows that $\mathfrak{N}$ is equal to the intersection of all of the prime ideals of $\hat{R}$, hence contained in each of them. Lemma A.1.25 on page 352 then implies the one-to-one correspondence stated above. The final statement follows from lemma A.1.26

$$\frac{\hat{R}}{\hat{\mathfrak{p}}} \cong \frac{\hat{R}/\mathfrak{N}}{\hat{\mathfrak{p}}/\mathfrak{N}} \cong \frac{R}{\mathfrak{p}}$$

$\square$

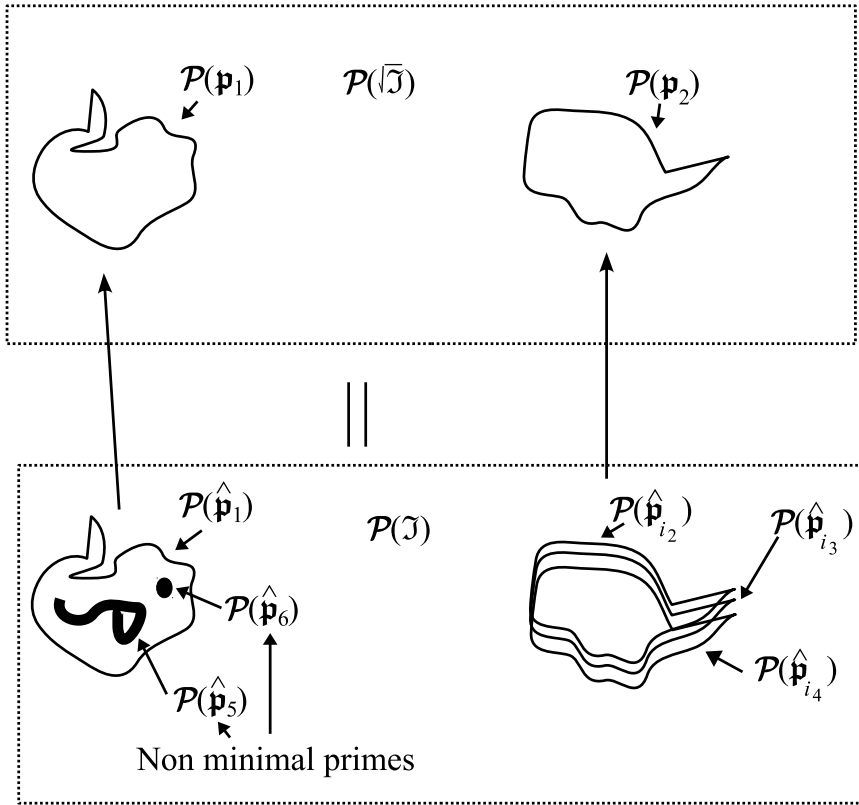This allows us to give a geometric proof of the well-known result:

FIGURE 5.8.1.   Minimal prime decomposition

THEOREM 5.8.2. *Let $\mathfrak{I} \subset A = k[X_0, \ldots, X_n]$ be an ideal, let $\{\hat{\mathfrak{p}}_{i_1}, \ldots, \hat{\mathfrak{p}}_{i_r}\}$ denote the minimal primes that occur in the prime filtration of $\hat{R} = A/\mathfrak{R}$ and let $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_s\}$ be the primes that occur in the prime filtration of $R = A/\sqrt{\mathfrak{I}}$. Then there exists a surjective function $g: \{1, \ldots, r\} \to \{1, \ldots, s\}$ such that*

$$f(\hat{\mathfrak{p}}_{i_j}) = \mathfrak{p}_{g(j)}$$

and

$$\frac{\hat{R}}{\hat{\mathfrak{p}}_{i_j}} \cong \frac{R}{\mathfrak{p}_{g(j)}}$$

REMARK. We know that the primes in a reduced ring's prime filtration are uniquely determined by the ring — see exercise 4 on page 79. When the ring is *not* reduced, this result shows that at least the *minimal* primes are still uniquely determined.

PROOF. Suppose $\{\hat{\mathfrak{p}}_1, \ldots, \hat{\mathfrak{p}}_t\}$ are *all* of the primes that occur in the prime filtration of $\hat{R}$, with $t \geq r$. In addition, suppose that $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_s\}$ are all the

primes that occur in the prime filtration of R. Corollary 5.7.3 on page 267 implies that

$$\mathcal{P}(\mathfrak{I}) = \bigcup_{i=1}^{t} \mathcal{P}(\hat{\mathfrak{p}}_i)$$

$$\mathcal{P}(\sqrt{\mathfrak{I}}) = \bigcup_{j=1}^{s} \mathcal{P}(\mathfrak{p}_j)$$

Since $\mathcal{P}(\mathfrak{I}) = \mathcal{P}(\sqrt{\mathfrak{I}})$, these two unions of irreducible varieties must be the *same* — where the second (of $\mathcal{P}(\sqrt{\mathfrak{I}})$) is irredundant (so $\mathfrak{p}_i \not\subset \mathfrak{p}_i$ for all $i \neq j$ and *all* are minimal).

We conclude that every irreducible component, $\mathcal{P}(\mathfrak{p}_j)$, must occur in the decomposition of $\mathcal{P}(\mathfrak{I})$, above. Since the two have the same union, the first may have

(1) multiple copies of the *same* irreducible component.
(2) irreducible components embedded in *other* irreducible components.

See figure 5.8.1 on the previous page for examples of these cases. Since the components of *maximal* dimension contain all other components, they must be the same in both cases so we must have a correspondence

$$\hat{\mathfrak{p}}_{i_j} \leftrightarrow \mathfrak{p}_{g(j)}$$

and we define $g(j)$ to be the value that makes $f(\hat{\mathfrak{p}}_{i_j}) = \mathfrak{p}_{g(j)}$. Since $\hat{R}/\hat{\mathfrak{p}}_{i_j} = R/\mathfrak{p}_{g(j)}$ (see lemma 5.8.1 on page 274), we have $\mathcal{P}(\hat{\mathfrak{p}}_{i_j}) = \mathcal{P}(\mathfrak{p}_{g(j)})$. In diagram 5.8.1 on the previous page, $g(1) = 1$, $g(i_2) = g(i_3) = g(i_4) = 2$ and $g(5), g(6)$ are undefined since $\hat{\mathfrak{p}}_5$ and $\hat{\mathfrak{p}}_6$ are not minimal. Components of *maximal* dimension correspond to *minimal* prime ideals (see the proof of proposition 2.8.9 on page 101). □

5.8.2. Bézout's Theorem. Now we are ready to prove a generalization of Bézout's Theorem as introduced in chapter 1. We will calculate the leading term of a Hilbert polynomial of $\hat{R}$ in equation 5.8.2 on page 274 in two different ways — a "global" way and a "local" way — and Bézout's Theorem will assert that these two results are equal. The glue that connects these approaches is corollary 5.7.3 on page 267.

In what follows, $V \subset k\mathbb{P}^n$ is a projective variety whose components are all of dimension t and $W \subset k\mathbb{P}^n$ is a hypersurface of degree d not contained in V. Then $V \cap W$ is a projective variety whose components are of dimension $t - 1$ (see proposition 5.6.4 on page 263).

We begin with the "global" computation:

LEMMA 5.8.3. *Let $V, W \subset k\mathbb{P}^n$ be projective varieties, where V is of dimension t, W is a hypersurface not contained in V and defined by a homogeneous polynomial $F \in k[X_0, \ldots, X_n]$, and $k_H[V] = k[X_0, \ldots, X_n]/\mathfrak{I}$. If*

$$\hat{R} = \frac{k[X_0, \ldots, X_n]}{\mathfrak{I} + (F)}$$

then the leading term of $\mathcal{H}_{\hat{R}}(s)$ is

$$E = \frac{(\deg V) \cdot (\deg W)}{(t-1)!}$$

REMARK. Degrees, here, are in the sense of definition 5.7.12 on page 272.

Note that $E \cdot (t-1)!$ is "like" the degree of $V \cap W$. It would be *exactly* $\deg(V \cap W)$ if we computed it using the ring

$$\frac{k[X_0, \ldots, X_n]}{\sqrt{\mathfrak{J} + (F)}}$$

rather than $\hat{R}$.

PROOF. First note that

$$\hat{R} = \frac{k_H[V]}{(F)}$$

Suppose F is homogeneous of degree d. Then we have an exact sequence

$$0 \to k_H[V] \xrightarrow{\times F} k_H[V] \to \hat{R} \to 0$$

Since F raises degree by d, we must rewrite this as

$$0 \to k_H[V](-d) \xrightarrow{\times F} k_H[V] \to \hat{R} \to 0$$

to get a sequence of graded rings. Proposition 5.7.2 on page 267 implies that

$$\mathcal{H}_{\hat{R}}(s) = \mathcal{H}_{k_H[V]}(s) - \mathcal{H}_{k_H[V]}(s-d)$$

Definition 5.7.12 on page 272 implies that the leading term of $\mathcal{H}_{k_H[V]}(s)$ is $(\deg V)/t!$. If the leading term of $\mathcal{H}_{\hat{R}}(s)$ is E, lemma 5.7.13 on page 272 implies that

$$\begin{aligned} E = d \cdot t \cdot \text{leading term of } \mathcal{H}_{k_H[V]}(s) &= d \cdot t \cdot (\deg V)/t! \\ &= d \cdot (\deg V)/(t-1)! \end{aligned}$$

Example 2 on page 272 shows that $d = \deg W$, and the conclusion follows. $\square$

Now we do the "local" computation: we add up the contributions of all of the components of $V \cap W$. Since some components occur more than once (as in the bottom half of diagram 5.8.1 on page 275), we have to multiply their contribution by the number of times they occur:

DEFINITION 5.8.4. Let $V, W \subset k\mathbb{P}^n$ be projective varieties, where W is a hypersurface not contained in V and let $\{Z_1, \ldots, Z_r\}$ be the irreducible components of $V \cap W$— defined by primes $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_r\}$, respectively. Suppose $k_H[V] = k[X_0, \ldots, X_n]/\mathfrak{J}$ and W is a hypersurface not contained in V and defined by a homogeneous polynomial $F \in k[X_0, \ldots, X_n]$ and

$$\begin{aligned} R &= \frac{k[X_0, \ldots, X_n]}{\sqrt{\mathfrak{J} + (F)}} \\ \hat{R} &= \frac{k[X_0, \ldots, X_n]}{\mathfrak{J} + (F)} \end{aligned}$$

Recall that there is a standard map $f: \hat{R} \to R$. If $\{\hat{\mathfrak{p}}_1, \ldots, \hat{\mathfrak{p}}_t\}$ is the set of primes that occur in the prime filtration of $\hat{R}$, then, the *intersection multiplicity*

$$\mu(V, W, \mathfrak{p}_j) = \mu(V, W, Z_j)$$

is defined as the *number* of such $\hat{\mathfrak{p}}_i$ with the property that $f(\hat{\mathfrak{p}}_i) = \mathfrak{p}_j$.

REMARK. Diagram 5.8.1 on page 275 shows why these quantities can be called *multiplicities* and

$$
\begin{aligned}
\mu(V, W, \mathfrak{p}_1) &= 1 \\
\mu(V, W, \mathfrak{p}_2) &= 3
\end{aligned}
$$

The primes that occur this way will all be of height 1.

Now we do the "local" computation of the leading term of $\mathcal{H}_{\hat{R}}(s)$:

LEMMA 5.8.5. *Let $V, W \subset k\mathbb{P}^n$ be projective varieties, where W is a hypersurface not contained in V. If*

$$V \cap W = \{Z_1, \ldots, Z_r\}$$

where the Z_i are irreducible, then the leading term of $\mathcal{H}_{\hat{R}}(s)$ is

(5.8.3) $$E = \sum_{i=1}^{r} \frac{\deg Z_i}{(t-1)!} \cdot \mu(V, W, Z_i)$$

PROOF. Let

$$0 = R_0 \subsetneq R_1 \subsetneq \cdots \subsetneq R_n = \hat{R}$$

be the prime filtration of $\hat{R}$ and suppose $\{\hat{\mathfrak{p}}_1, \ldots, \hat{\mathfrak{p}}_t\}$ are the primes that occur in it. Then corollary 5.7.3 on page 267 implies that

$$\mathcal{H}_{\hat{R}}(s) = \sum_{j=1}^{t} \mathcal{H}_{\hat{R}/\hat{\mathfrak{p}}_j(\ell_j)}(s) = \sum_{j=1}^{t} \mathcal{H}_{\hat{R}/\hat{\mathfrak{p}}_j}(s + \ell_j)$$

If E is the leading term of $\mathcal{H}_{\hat{R}}(s)$ and e_j is the leading term of $\mathcal{H}_{\hat{R}/\hat{\mathfrak{p}}_j}(s + \ell_j)$ which is equal to the leading term of $\mathcal{H}_{\hat{R}/\hat{\mathfrak{p}}_j}(s)$, then

(5.8.4) $$E = \sum_{j=1}^{r} e_{i_j}$$

where $\{\hat{\mathfrak{p}}_{i_1}, \ldots, \hat{\mathfrak{p}}_{i_r}\}$ is the subset of *minimal* primes: They represent components of maximal dimension and have Hilbert polynomials of maximal *degree* (in the usual sense!). Lemma 5.8.1 on page 274 implies that $\hat{R}/\hat{\mathfrak{p}}_{i_j} = R/f(\hat{\mathfrak{p}}_{i_j}) = R/\mathfrak{p}_{g(j)} = k_H[Z_{g(j)}]$, where g is the function defined in theorem 5.8.2 on page 275.

Definition 5.7.12 on page 272 shows that $e_{i_j} = (\deg Z_{g(j)})/(t-1)!$ and the number of times a term equal to e_{i_j} occurs in the sum in equation 5.8.4 on page 278 is $\mu(V, W, Z_{g(j)})$, so we get equation 5.8.3 on page 278. Combining this with lemma 5.8.3 on page 276 gives

$$E = \sum_{i=1}^{r} \frac{\deg Z_i}{(t-1)!} \cdot \mu(V, W, Z_i)$$

$\square$

THEOREM 5.8.6 (Bézout'sTheorem). *Let $V, W \subset k\mathbb{P}^n$ be projective varieties, where W is a hypersurface not contained in V. If*

$$V \cap W = \{Z_1, \ldots, Z_r\}$$

where the Z_i are irreducible, then

$$(5.8.5) \qquad (\deg V) \cdot (\deg W) = \sum_{i=1}^{r} (\deg Z_i) \cdot \mu(V, W, Z_i)$$

REMARK. In the case where $V, W \subset k\mathbb{P}^2$ are curves, the Z_j are points and of degree 1 (see statement 3 on page 272), so we get

$$(\deg V) \cdot (\deg W) = \sum_{i=1}^{r} \mu(V, W, Z_i)$$

Since V and W are both "hypersurfaces," their degrees in the sense of definition 5.7.12 on page 272 are equal to the degrees of their defining polynomials — so we get a rigorous proof of the classical Bézout theorem.

We have an enormous generalization of that classical result though: a theorem that is valid even when the intersections are not points.

It is interesting to consider what happens when W is *not* a hypersurface. In [**168, 169**], Van der Waerden showed that the theorem fails in that case — requiring a more complex definition of intersection multiplicity. See [**172**] for a general statement.

PROOF. Just combine equation 5.8.3 on the facing page with the conclusion of lemma 5.8.3 on page 276 and multiply by $(t-1)!$. □

This proof shows what we would have to do if we wanted to dispense with multiplicities and simply count intersections: we would have to compute the leading term of the Hilbert polynomial of

$$\frac{k[X_0, \ldots, X_n]}{\sqrt{\mathfrak{I} + (F)}}$$

rather than

$$\frac{k[X_0, \ldots, X_n]}{\mathfrak{I} + (F)}$$

Unfortunately, there is no obvious way to do this — certainly nothing as elementary as lemma 5.8.3 on page 276.

EXAMPLE 5.8.7. We will consider an example from chapter 1, depicted in figure 1.4.5 on page 23:

$$\mathfrak{a} = (X^2 + Y^2 - 1, 5\,X^2 + 6\,XY + 5\,Y^2 + 6\,Y - 5) \subset k[X, Y]$$

This has a Gröbner basis of

$$\mathfrak{a} = (Y^3, XY + Y, X^2 + Y^2 - 1)$$

which makes it clear that $Y \in \sqrt{\mathfrak{a}}$ so

$$\begin{aligned} \sqrt{\mathfrak{a}} \quad &\supset \quad (Y, XY + Y, X^2 + Y^2 - 1) \\ &= \quad (Y, X^2 + Y^2 - 1) \text{ since } XY + Y \text{ is a multiple of } Y \\ &= \quad (Y, X^2 - 1) \quad \text{since we can subtract multiples} \\ &\qquad\qquad\qquad \text{of generators from other generators} \end{aligned}$$

so $\sqrt{\mathfrak{a}} = \sqrt{(Y, X^2 - 1)}$. Using Gröbner bases, we conclude that

$$
\begin{aligned}
X &\notin \sqrt{(Y, X^2 - 1)} \\
X^2 &\notin \sqrt{(Y, X^2 - 1)} \\
XY &\in \sqrt{(Y, X^2 - 1)}
\end{aligned}
$$

so $\sqrt{\mathfrak{a}} = (Y, XY, X^2 - 1)$. A Gröbner basis for $(Y, XY, X^2 - 1)$ is $(Y, X^2 - 1)$ and the term XY is not needed.

Let

$$
\begin{aligned}
R &= \frac{k[X, Y]}{(Y, X^2 - 1)} = k \oplus k \\
\hat{R} &= \frac{k[X, Y]}{(Y^3, XY + Y, X^2 + Y^2 - 1)}
\end{aligned}
$$

If $x, y \in \hat{R}$ are the images of $X, Y \in k[X, Y]$ under the projection, example A.1.79 on page 376 works out the prime filtration of $\hat{R}$ in great detail. It is

$$
0 \subsetneq y^2 \cdot \hat{R} \subsetneq y \cdot \hat{R} \subsetneq (x - 1) \cdot \hat{R} \subsetneq \hat{R}
$$

where the prime $\hat{\mathfrak{p}}_2 = (x + 1, y)$ occurs three times (representing the threefold multiplicity of the intersection $(-1, 0)$ in figure 1.4.5 on page 23) and the prime $\hat{\mathfrak{p}}_1 = (x - 1, y)$ occurs once.

We can get a geometric interpretation of *degree of a variety* (as defined in 5.7.12 on page 272):

PROPOSITION 5.8.8. *If $V \subset k\mathbb{P}^n$ is a d-dimensional projective variety and H is any $n - d$ dimensional hyperplane not contained in V then*

$$
V \cap H
$$

consists of $\deg V$ points (counted with multiplicities).

PROOF. The degree of H is 1 so theorem 5.8.6 on the previous page implies that

$$
\deg V = \sum_{j=1}^{r} \mu(V, H, Z_j) \cdot \deg Z_j
$$

where $V \cap H = Z_1 \cup \cdots \cup Z_r$. If the Z_j are not points, simply intersect their union with a *second* hyperplane not contained in them. This lowers the dimensions of the intersections by 1 and we get a sum

$$
\deg V = \sum_{j=1}^{r} \mu(V, H, Z_j') \cdot \deg Z_j'
$$

We continue this process until the intersections are points.   □

The alert reader will have noticed a discrepancy between definition 5.8.4 on page 277 and lemma 3.3.34 on page 140: the *first* defines multiplicity as the number of times a prime appears in a prime filtration and the *second* shows that it is equal to the dimension of a vector space.

PROPOSITION 5.8.9. *Under the assumptions and notation of definition 5.8.4 on page 277, let $\hat{\mathfrak{p}}$ be a minimal prime in a prime filtration of $\hat{R}$. Then*

$$\hat{R}_{\hat{\mathfrak{p}}}$$

vector space of dimension $\mu(V, W, \mathfrak{p})$ over the field

$$\frac{\hat{R}}{\hat{\mathfrak{p}} \cdot \hat{R}} = \frac{R}{\mathfrak{p} \cdot R}$$

where $\mathfrak{p} = f(\hat{\mathfrak{p}})$ under the map $f: \hat{R} \to R$.

REMARK. This is frequently stated as:

"The length of the module, $\hat{R}_{\hat{\mathfrak{p}}}$, is $\mu(V, W, \mathfrak{p})$."

where the *length* of a module, M, is the largest value of n such that there exists a filtration

$$0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_n = M$$

If a module is a vector space, its length is equal to its dimension.

PROOF. Since $\hat{\mathfrak{p}} \cdot \hat{R} \subset \hat{R}$ is a maximal ideal, the quotient is a field. Now suppose

$$0 = R_0 \subsetneq R_1 \subsetneq \cdots \subsetneq R_n = \hat{R}_{\hat{\mathfrak{p}}}$$

is a prime filtration of $\hat{R}_{\hat{\mathfrak{p}}}$. Corollary A.5.62 on page 470 implies that the *only* primes that will occur in this filtration will be ones equal to $\hat{\mathfrak{p}}$. Each short exact sequence

$$0 \to R_i \to R_{i+1} \to \frac{\hat{R}}{\hat{\mathfrak{p}} \cdot \hat{R}} \to 0$$

can be *split* (see exercise 41 on page 380) by mapping $1 \in \hat{R}_{\hat{\mathfrak{p}}} / \hat{\mathfrak{p}} \cdot \hat{R}_{\hat{\mathfrak{p}}}$ to any element of R_{i+1} that maps to it. It follows that

$$R_{i+1} = R_i \oplus \frac{\hat{R}}{\hat{\mathfrak{p}} \cdot \hat{R}}$$

for all i, and a simple induction proves the result.   □

EXAMPLE 5.8.10. In example 5.8.7 on page 279, if we localize

$$\hat{R} = \frac{k[X, Y]}{(Y^3, XY + Y, X^2 + Y^2 - 1)}$$

by $\hat{\mathfrak{p}}_2 = (x + 1, y)$, we invert $x - 1$ and $y + 1$. This is accomplished by defining auxiliary variables Z, W and adding $(X - 1)Z - 1$ and $(Y + 1)W - 1$ to the ideal's generators. Now we take a Gröbner basis in lexicographic ordering that eliminates Z and W as much as possible (i.e., it orders them *above* the other variables — see proposition 2.3.13 on page 52). We get

$$(Y^3, -Y^2 + 2 + 2X, -1 + Y - Y^2 + W, 4 + 8Z + Y^2)$$

or, if $S = \langle X - 1, Y + 1 \rangle$, then (assuming that the characteristic of k is $\neq 2$)

$$S^{-1}\hat{R} =$$

$$\frac{k[X, Y, Z, W]}{(Y^3, -Y^2 + 2 + 2X, 4 + 8Z + Y^2, -1 + Y - Y^2 + W)}$$

$$= \frac{k[X, Y, Z]}{(Y^3, -Y^2 + 2 + 2X, 4 + 8Z + Y^2)}$$

$$= \frac{k[X, Y]}{(Y^3, -Y^2 + 2 + 2X)}$$

$$= \frac{k[Y]}{(Y^3)} = k^3$$

so the multiplicity of the intersection-point, $(-1, 0)$, is 3 (as we knew).

If we localize at $\hat{\mathfrak{p}}_1 = (x - 1, y)$. we invert $X + 1$ and $Y + 1$ in the same way as before: add elements $(X + 1)Z - 1$ and $(Y + 1)W - 1$ to the ideal and form a lexicographic Gröbner basis with the ordering $\{Z, W, X, Y\}$ to get

$$(Y, X - 1, W - 1, 2Z - 1)$$

so the quotient is

$$\hat{R}_{\hat{\mathfrak{p}}_1} = \frac{k[X, Y, Z, W]}{(Y, X - 1, W - 1, 2Z - 1)} = \frac{k[Y]}{(Y)} = k$$

if the characteristic of k is $\neq 2$.

Here is another example:

EXAMPLE 5.8.11. Let $V = \mathcal{P}((Y^2Z - X^3 + X^2Z)) \subset k\mathbb{P}^2$ and $W = \mathcal{P}((X - Z))$. Then (assuming the characteristic of k is $\neq 3$)

$$\hat{R} = \frac{k[X, Y, Z]}{(Y^2Z - X^3 + X^2Z, X - Z)}$$

The lexicographic Gröbner basis for the denominator is $(YZ^2, -Z + X)$ so we get

$$\hat{R} = \frac{k[X, Y, Z]}{(YZ^2, -Z + X)} = \frac{k[Y, Z]}{(YZ^2)}$$

The minimal prime ideals that contain (YZ^2) are $\mathfrak{p}_1 = (Y)$ and $\mathfrak{p}_2 = (Z)$ and we get

$$\hat{R}_{\mathfrak{p}_1} = \frac{S_1^{-1}k(Z)[Y]}{(Y)} = k(Z)$$

where S_1 consists of all polynomials in Y and Z that are not multiples of Y. The quotient is a one-dimensional vector space over itself, so the multiplicity of this intersection is 1.

Inverting $\mathfrak{p}_2$ gives

$$\hat{R}_{\mathfrak{p}_2} = \frac{S_2^{-1}k(Y)[Z]}{(Z^2)} = k(Y) \cdot 1 \oplus k(Y) \cdot Z$$

where S_2 is all polynomials in Y and Z that are *not* multiples of Z — giving a two-dimensional vector space over $k(Y)$. It follows that there are *two* intersections between V and W:

- $(1\!:\!0\!:\!1)$, of multiplicity 1, and
- $(0\!:\!1\!:\!0)$, of multiplicity 2.

5.8.3. Eigenvalues of tensors. In this section, we discuss an application of Bézout's theorem to computing the number of eigenvalues of a tensor following the treatment in [**26**]. Tensors may be regarded as "higher dimensional matrices" — i.e., matrices with *more* than two subscripts. They are very extensively used in differential geometry (see [**158**]), the theory of relativity (see [**116**]), and many other fields. For instance, the elasticity tensor in continuum mechanics is a *fourth* order tensor — see [**163**].

DEFINITION 5.8.12. If A is an order-m tensor over $\mathbb{C}^n$ — i.e. $A = (A_{i_1,\dots,i_m})$ — and $\mathbf{x} \in \mathbb{C}^n$ is a vector, define

$$(A\mathbf{x}^{m-1})_j = \sum_{i_2=1}^{n} \cdots \sum_{i_m=1}^{n} A_{j,i_2,\dots,i_m} x_{i_2} \cdots x_{i_m} \in \mathbb{C}^n$$

Then $\lambda \in \mathbb{C}$ is an *eigenvalue* of A with *eigenvector* $\mathbf{x}$ with $\mathbf{x} \cdot \mathbf{x} = 1$, if

$$A\mathbf{x}^{m-1} = \lambda \cdot \mathbf{x}$$

Two such pairs $(\mathbf{x}_1, \lambda_1)$ and $(\mathbf{x}_2, \lambda_2)$ will be considered *equivalent* if their exists a $c \in \mathbb{C}$ such that $\lambda_2 = c^{m-2}\lambda_1$ and $\mathbf{x}_2 = c\mathbf{x}_1$.

REMARK. Eigenvalues of tensors have applications to a number of areas — see [**105**].

THEOREM 5.8.13. *If A is an order-m tensor over $\mathbb{C}^n$, then the number of equivalence classes of eigenvalues (counted with multiplicities) is*

$$\frac{(m-1)^n - 1}{m-2} = \sum_{i=0}^{n-1} (m-1)^i$$

REMARK. When $m = 2$, A is an ordinary matrix and this count of eigenvalues is n. See [**25**] for many more details and complements to this.

PROOF. If $\mu^{m-2} = \lambda$, our equation for eigenvalues becomes

$$A\mathbf{x}^{m-1} = \mu^{m-2}\mathbf{x}$$

which is a system of n homogeneous polynomials of degree $m-1$ in the variables $x_1, \dots, x_n, \mu$, i.e., n degree-$m-1$ hypersurfaces in $\mathbb{C}P^n$. Bézout's theorem implies that the number of intersections is $(m-1)^n$. If we eliminate the trivial solution $\mathbf{x} = 0$ (it violates our requirement that $\mathbf{x} \cdot \mathbf{x} = 1$) we get $(m-1)^n - 1$ solutions.

Now we distinguish two cases:

(1) *None of the eigenvalues is* 0. Multiplying μ by $e^{2\pi i \cdot k/m-2}$ for $k = 0, \dots, m-3$ gives rise to the same value of μ^{m-2}, hence the same solution. Consequently, the number of solutions is

$$\frac{(m-1)^n - 1}{m-2}$$

(2) *We have an eigenvalue equal to* 0. If the solution has $x_i = d_i$ for $i = 1, \ldots, n$ (some set of constants) and $\mu = 0$, we get

$$\frac{\mathbb{C}[x_1, \ldots, x_n, \mu]}{(\mu^{m-2} \cdot \mathbf{x}, \{x_i - d_i\})} = \mathbb{C}^{m-2}$$

so proposition 5.8.9 on page 281 implies that the *multiplicity* of this solution is $m - 2$. It follows that this *single* solution contributes $m - 2$ to the total $(m - 1)^n - 1$.

$\square$

5.9. Divisors

5.9.1. Weil and Cartier divisors. We have seen how uninteresting the ring of regular functions on a projective variety can be (corollary 5.5.17 on page 257). The same is *not* true of rational functions, and we can deduce interesting relations between the rational functions on a variety and the geometry.

We start with particularly simple case. Suppose $f(z)$ is a rational function on $\mathbb{C}$:

$$f(z) = A \frac{(z - \alpha_1)^{n_1} \cdots (z - \alpha_s)^{n_s}}{(z - \beta_1)^{m_1} \cdots (z - \beta_t)^{m_t}}$$

where $A \in \mathbb{C}$. This function is completely characterized by:

(1) the constant $A \in \mathbb{C}$,
(2) the $\alpha_i \in \mathbb{C}$ and the degrees, n_i,
(3) the $\beta_i \in \mathbb{C}$ and the degrees m_j.

After a bit of thought, it seems reasonable to *combine* this data and say that the β_j occur with *negative* degrees, $-m_j$, so that zeros and singularities are treated the same way.

Now we take a step that seems rather bizarre at first: define a group, G, whose generators are the *points* of $\mathbb{C}$ — its elements are finite "linear combinations"

$$n_1 \cdot \langle \alpha_1 \rangle + \cdots + n_t \cdot \langle \alpha_t \rangle$$

where $\alpha_1, \ldots, \alpha_t \in \mathbb{C}$ are points, the $\langle \alpha_i \rangle$ are *symbols* representing those points, and the n_i are integers. Having taken this odd step, we can represent the function f above, by an element

(5.9.1)
$$(f) = \sum_{i=1}^{s} n_i \langle \alpha_i \rangle - \sum_{j=1}^{t} m_j \langle \beta_j \rangle \in G$$

We would like to do something like this with rational functions on an algebraic variety:

DEFINITION 5.9.1. If V is an irreducible, normal, affine variety, a *prime divisor* on V is an irreducible codimension-1 subvariety. A *divisor* is a formal linear combination of prime divisors with integer coefficients. The (free abelian) group of divisors on V is denoted $\mathrm{Div}(V)$. An *effective divisor*, E, is an integer linear combination of prime divisors with *nonnegative* coefficients, written

$$E \geq 0$$

If $W \subset V$ is a codimension-1 irreducible subvariety, the divisor it represents is denoted $\langle W \rangle$.

If $f \in k(V)$ is a meromorphic function, as in equation 5.9.1 on the preceding page, we have

$$
(f)_0 \;=\; \sum_{i=1}^{s} n_i \langle \alpha_i \rangle
$$

$$
(f)_\infty \;=\; \sum_{j=1}^{t} m_j \langle \beta_j \rangle
$$

$$
(f) \;=\; (f)_0 - (f)_\infty
$$

where $(f)_0$ is the *zero-divisor* of f and $(f)_\infty$ is called its *polar divisor*.

REMARK. The term "divisor" for a *subvariety* seems odd but is justified (to some extent) by the remarks following definition 5.9.3 on the following page.

Definition 3.4.1 on page 145 defines normal varieties, and lemma 3.4.2 on page 145 shows that all smooth varieties are normal.

What we have defined here is technically called a *Weil divisor*.

We would like to associate a divisor with a rational function on an affine variety. On an irreducible affine variety, V, a rational function is a quotient of regular functions

$$
f = \frac{u}{v},\ u, v \in k[V]
$$

and corollary 2.8.30 on page 111 shows that:

(1) the zero set of u is a union of irreducible codimension-1 subvarieties

$$
\bigcup_{i=1}^{n} W_i
$$

and

(2) the zero-set of v is a union of irreducible codimension-1 subvarieties

$$
\bigcup_{j=1}^{m} Z_j
$$

Consequently, we could represent f by an expression

$$
(5.9.2) \qquad \sum_{i=1}^{n} \alpha_i \cdot \langle W_i \rangle - \sum_{j=1}^{m} \beta_j \cdot \langle Z_j \rangle \in \mathrm{Div}(V)
$$

We still need to define the coefficients, $\alpha_i, \beta_j \in \mathbb{Z}$ in some reasonable way. The following result shows how to do this:

LEMMA 5.9.2. *Let $W \subset V$ be a codimension-1 irreducible subvariety of an n-dimensional normal, irreducible, affine variety V. Then $\mathcal{O}_{V,W}$ — the coordinate ring at W— is a discrete valuation ring.*

REMARK. Lemma A.4.37 on page 433 implies that there exists an element, $\pi \in \mathcal{O}_{V,W}$, such that $\mathfrak{p} \cdot \mathcal{O}_{V,W} = (\pi)$ and all ideals are of the form (π^t) for some integer, t. Furthermore, $\mathcal{O}_{V,W}$ has a discrete valuation

$$
v \colon \mathcal{O}_{V,W} \to \mathbb{Z}
$$

where $v(x)$ is the highest power of π divides x.

PROOF. We have already essentially proved this in section 3.4.2 on page 150. If $\mathfrak{p} \subset k[V]$ is the prime ideal representing W, we have

$$\mathcal{O}_{V,W} = k[V]_{\mathfrak{p}}$$

so that $\mathcal{O}_{V,W}$ is also normal by exercise 1 on page 427. Since W is $n - 1$-dimensional, the prime ideal $\mathfrak{p}$ is *minimal* by proposition 2.8.9 on page 101 (and the reasoning used in its proof). It follows that $\mathcal{O}_{V,W}$ is a normal domain with a single prime ideal, $\mathfrak{p} \cdot \mathcal{O}_{V,W}$, and is a *discrete valuation ring,* by lemma A.4.38 on page 433. $\square$

This motivates the definition:

DEFINITION 5.9.3. If p is a regular function on V that vanishes on a codimension-1 subvariety, W, then the *order of p at W*, denoted $\mathrm{ord}_W(p)$, is defined to be the valuation of p in $\mathcal{O}_{V,W}$.

In equation 5.9.2 on the previous page, set

$$
\begin{aligned}
\alpha_i &= \mathrm{ord}_{W_i}(u) \\
\beta_j &= \mathrm{ord}_{Z_j}(v)
\end{aligned}
$$

The *divisor* of a rational function f, computed this way, is denoted

$$(f) \in \mathrm{Div}(V)$$

Elements of $\mathrm{Div}(V)$ of the form (f) for $f \in \mathcal{K}_V$ are called *principal divisors*. If

$$D = \sum_{i=1}^{n} \alpha_i \cdot \langle W_i \rangle - \sum_{j=1}^{m} \beta_j \cdot \langle Z_j \rangle \in \mathrm{Div}(V)$$

is a divisor, the *degree of D* is

$$\deg D = \sum_{i=1}^{n} \alpha_i - \sum_{j=1}^{m} \beta_j \in \mathbb{Z}$$

REMARK. If $\pi_i \in \mathcal{O}_{V,W_i}$ is the element that generates the only prime ideal, then $\pi_i^{\alpha_i} | u$ and $\pi_i^{\alpha_i+1} \nmid u$. This justifies the terminology "divisor" for elements of $\mathrm{Div}(V)$ — they actually *divide* the functions u and v.

One important question in the theory of divisors is:

> How much freedom do we have in constructing rational functions? Does every possible divisor correspond to a rational function?

In other words, are all divisors principal? To study this question, we define:

DEFINITION 5.9.4. Let V be an irreducible normal variety. The set of principal divisors, $P(V) \subset \mathrm{Div}(V)$, is a subgroup and the quotient

$$\mathrm{Cl}(V) = \frac{\mathrm{Div}(V)}{P(V)}$$

is called the *divisor class group*. Two divisors $d_1, d_2 \in \mathrm{Div}(V)$ will be said to be *linearly equivalent*, denoted $d_1 \sim d_2$, if they define the same element of $\mathrm{Cl}(V)$ — i.e., if

$$d_1 - d_2 = (f)$$

for some $f \in k(V)$.

Let us consider some examples:

EXAMPLE 5.9.5. If $V = \mathbb{A}^n$, then $\mathrm{Cl}(V) = 0$ because one can construct rational functions with arbitrary factors in the numerator and denominator:

$$\frac{f(X_1, \ldots, X_n)}{g(X_1, \ldots, X_n)}$$

The fact that $\mathrm{Cl}(\mathbb{A}^n) = 0$ implies that there are *many* rational functions on $\mathbb{A}^n$.

REMARK. This also happens if every prime divisor is principal: if

$$d = \sum n_i \langle d_i \rangle - m_j \langle e_j \rangle$$

is an arbitrary divisor, where the $\langle d_i \rangle, \langle e_j \rangle$ are principal and the coefficients are positive, then there exist rational functions f_i, g_j with $\langle d_i \rangle = (f_i)$, $\langle e_j \rangle = (g_j)$ for all i, j and

$$d = \left(\frac{\prod f_i^{n_i}}{\prod g_j^{m_j}} \right)$$

so d is principal.

EXAMPLE 5.9.6. If $V = k\mathbb{P}^n$, every possible divisor can be represented as a "formal quotient"

$$d = \frac{f(X_1, \ldots, X_n)}{g(X_1, \ldots, X_n)}$$

where f and g are arbitrary homogeneous polynomials, each representing unions of hypersurfaces in $k\mathbb{P}^n$. Each such divisor has a *degree* equal to

$$\deg f - \deg g$$

and a divisor is *principal* (i.e., defines a rational function on $k\mathbb{P}^n$) if and only if $\deg d = 0$ — see example 5.4.3 on page 243. If two divisors,

$$d_1 = \frac{f_1(X_1, \ldots, X_n)}{g_1(X_1, \ldots, X_n)}$$

$$d_2 = \frac{f_2(X_1, \ldots, X_n)}{g_2(X_1, \ldots, X_n)}$$

have the same degree, then

$$d_1 - d_2 = \left(\frac{f_1(X_1, \ldots, X_n)}{g_1(X_1, \ldots, X_n)} \cdot \frac{g_2(X_1, \ldots, X_n)}{f_2(X_1, \ldots, X_n)} \right)$$

is principal. It follows that the element of $\mathrm{Cl}(k\mathbb{P}^n)$ that a divisor represents is determined by its degree, so $\mathrm{Cl}(k\mathbb{P}^n) = \mathbb{Z}$, and $k\mathbb{P}^n$ has "fewer" rational functions than $\mathbb{A}^n$.

This implies that:

EXAMPLE 5.9.7. If $V = k\mathbb{P}^n$, then $\text{Cl}(V)$ is generated by a hyperplane (in fact, *any* hyperplane), $\langle H \rangle$. If $W = \mathcal{P}(F) \subset k\mathbb{P}^n$ is a hypersurface, where F is a homogeneous polynomial of degree d, then $\langle W \rangle = d \cdot \langle H \rangle \in \text{Cl}(V)$.

We can use Bézout's theorem to get a partial generalization of what was established in example 5.9.6 on the previous page:

PROPOSITION 5.9.8. *If $V \subset k\mathbb{P}^2$ is a smooth one-dimensional projective variety (i.e., a curve) and $x \in k(V)$ is a rational function, then $\deg(x) = 0$.*

REMARK. It follows that, on one-dimensional varieties, divisor-*classes* have a well-defined *degree* — i.e., we get a homomorphism

$$\deg\colon \text{Cl}(V) \to \mathbb{Z}$$

It is *not* necessarily true that $\text{Cl}(V) = \text{Cl}(k\mathbb{P}^1) = \mathbb{Z}$, though. For instance, section 6.2 on page 313 discusses a class of smooth curves whose divisor class groups are *uncountable.*

PROOF. Suppose $V \subset k\mathbb{P}^t$ is of degree n and $x = f/g$, where f and g are regular functions that extend to homogeneous polynomials F and G of degree m on $k\mathbb{P}^t$. If $p_1, \ldots, p_t$ are the zeros of f, they are precisely points of $V \cap W \subset k\mathbb{P}^t$, where W is the hypersurface defined by $F = 0$. Furthermore, the coefficients of these points in the expression for the divisor are the intersection-multiplicities. Bézout's theorem (5.8.6 on page 279) implies that $\deg(f) = n \cdot m$. A similar argument implies that $\deg(g) = n \cdot m$ so

$$\deg(x) = n \cdot m - n \cdot m = 0$$

$\square$

DEFINITION 5.9.9. If V is a one-dimensional projective variety, define $\text{Cl}^0(V)$ to be the kernel of

$$\deg\colon \text{Cl}(V) \to \mathbb{Z}$$

i.e., the subgroup $\text{Cl}^0(V) \subset \text{Cl}(V)$ of degree-0 divisors.

We can generalize the computation in example 5.9.6 on the previous page to *products* of projective spaces:

EXAMPLE 5.9.10. If $V = k\mathbb{P}^1 \times \cdots \times k\mathbb{P}^{n_r}$, then corollary 5.5.2 on page 247 shows that a hypersurface, $W \subset V$ is specified by a polynomial

$$f(X_{0,1} \ldots, X_{n_1,1}, \ldots, X_{0,r}, \ldots, X_{n_r,r})$$

that is homogeneous in each *set* of variables $\{X_{0,i}, \ldots, X_{n_i,i}\}$ — with distinct sets possibly having different degrees. Let $\deg f_{\{X_{0,i},\ldots,X_{n_i,i}\}}$ denote the degree of the ith set of homogeneous variables. A divisor is a "formal quotient"

$$\frac{f(X_{0,1} \ldots, X_{n_1,1}, \ldots, X_{0,r}, \ldots, X_{n_r,r})}{g(X_{0,1} \ldots, X_{n_1,1}, \ldots, X_{0,r}, \ldots, X_{n_r,r})}$$

and this defines a rational function on V if and only if $\deg f_{\{X_{0,i},\ldots,X_{n_i,i}\}} = \deg g_{\{X_{0,i},\ldots,X_{n_i,i}\}}$ for all $i = 1, \ldots, r$. Reasoning like that used in example 5.9.6 on the previous page shows that

$$\text{Cl}(V) = \mathbb{Z}^r$$

The divisor class group contains some interesting information about a variety:

PROPOSITION 5.9.11. *Let V be an irreducible normal affine variety. Then $k[V]$ is a unique factorization domain if and only $\text{Cl}(V) = 0$.*

PROOF. Lemma 2.8.39 on page 115 shows that $k[V]$ is factorial if and only if all height 1 primes are principal. Prime divisors correspond to minimal nonzero primes $\mathfrak{p} \subset k[V]$. Since $k[V]$ is factorial, it is a domain and these primes contain *the* minimal prime, (0) so they are of height 1. Since these are principal, it follows that $\text{Cl}(V) = 0$. The converse is also clear.   □

There are cases where computing the divisor class group is not difficult:

PROPOSITION 5.9.12. *If V is an irreducible normal variety and $W \subset V$ is a proper closed subvariety with $U = V \setminus W$, then:*

(1) there is a surjective homomorphism

$$\text{Cl}(V) \xrightarrow{r} \text{Cl}(U)$$

that sends $\sum n_i \cdot P_i$ to $\sum n_i \cdot (P_i \cap U)$, where we ignore the intersections that are empty and the P_i are prime divisors,

(2) if W is of codimension ≥ 2, this homomorphism is an isomorphism.

(3) if W is irreducible of codimension 1, then there is an exact sequence

$$\mathbb{Z} \to \text{Cl}(V) \to \text{Cl}(U) \to 0$$

where the kernel is the image of $\mathbb{Z} \cdot \langle W \rangle$.

PROOF. The intersections $P_i \cap U$ are either prime divisors of U or empty, so the map given is a homomorphism. It is surjective because every prime divisor $Q \subset U$ has a closure $\bar{Q} \subset V$ with the property that

$$Q = \bar{Q} \cap U$$

The second statement follows from the fact that the divisor class group is only influenced by subvarieties in codimension 1.

The third statement follows from the fact that W is a prime divisor and a divisor is in the kernel of r if and only if its support is in W.   □

Weil divisors, as in definition 5.9.1 on page 284 have the problem that they are not well-behaved on varieties that are not normal. For instance, determining the divisor of a function requires normality. It is possible to extend the definition of divisors to general varieties, giving so-called Cartier divisors. We begin by looking at some of the properties of Weil divisors.

In a neighborhood of each point $p \in V$ of a variety, a prime Weil divisor is given by an equation $f = 0$ (given by the local parameter in lemma 3.4.3 on page 146, which extends to a small neighborhood of a point) where f is a rational function — so that it is *locally principal.* Suppose we cover V by such open sets, so that our divisor is given by $f_i = 0$ on U_i

$$V = \bigcup U_i$$

On the overlap $U_i \cap U_j$ the divisors defined by $f_i = 0$ and $f_j = 0$ must coincide. This will happen if and only if

$$\frac{f_i}{f_j}$$

is a regular function on $U_i \cap U_j$ that is *never zero* there. We usually express this by saying

$$\frac{f_i}{f_j} \text{ and } \frac{f_j}{f_i}$$

are regular functions on $U_i \cap U_j$. This leads to the definition:

DEFINITION 5.9.13. If V is an irreducible variety, a *Cartier divisor* on V is

  (1) a covering of V by open affines

$$V = \bigcup U_i$$

  (2) a set of rational functions $f_i \in \mathcal{K}_V^*$ defined on U_i for all i that are not identically 0 and with the property that

$$\frac{f_i}{f_j} \text{ and } \frac{f_j}{f_i}$$

   are both regular on the overlap $U_i \cap U_j$, for all pairs i, j.

  (3) The *product* of two Cartier divisors $\{(U_i, f_i\}$ and $\{V_j, g_j\}$ is the divisor $\{U_i \cap V_j, f_i \cdot g_j\}$.

  (4) Two systems of functions $\{(U_i, f_i\}$ and $\{(U_i, g_i\}$ define the *same* Cartier divisor if f_i/g_i is regular and nonzero on U_i.

  (5) If $f \in k(V)$ is a globally defined rational function, it defines a Cartier divisor via $f_i = f|U_i$. This is called a *principal divisor*.

The group of Cartier divisors (under multiplication) of V is denoted $\mathrm{Cart}(V)$.

REMARK. If V is *normal,* the argument above shows that every prime Weil divisor gives rise to a Cartier divisor. This is extended to the group of divisors by mapping the formal sum of Weil divisors to the *product* of the corresponding Cartier divisors.

Condition 2 above implies that $f_i \in \mathcal{K}_V^*(U_i)$ for all i.

EXAMPLE 5.9.14. Recall the "blowing-up" operation in section 5.5.3 on page 248. Proposition 5.5.9 on page 252 shows that the exceptional fiber is actually a Cartier divisor — for instance, on the overlap (in the notation of proposition 5.5.9 on page 252)

$$\left(E \cap \mathbb{A}^n \times \mathbb{A}_i^{t-1} \right) \cap \left(E \cap \mathbb{A}^n \times \mathbb{A}_j^{t-1} \right)$$

with $i \neq j$, we have $Y_i = Y_j = 1$ so that $f_i = f_j$, satisfying condition 2 of definition 5.9.13 on page 290. This is why the exceptional *fiber* of a blowup is often called its *exceptional divisor*.

PROPOSITION 5.9.15. *If V is a normal variety and $d = \{U_i, f_i\}$ is a Cartier divisor on V, then d induces a unique Weil divisor w.*

PROOF. In a neighborhood of each point of U_i, define w to be $(f_i) \in \mathrm{Div}(V)$. Condition 2 in definition 5.9.13 on page 5.9.13 implies that f_i and f_j will induce the same divisor on $U_i \cap U_j$.   □

Cartier divisors are well-behaved under regular maps:

LEMMA 5.9.16. *A surjective regular map* $g: V \to W$ *of irreducible varieties induces a natural homomorphism*

$$g^*: \text{Cart}(W) \to \text{Cart}(V)$$

that sends principle divisors to principle divisors. If V and W are both normal, g induces natural homomorphisms

$$g^*: \text{Div}(W) \to \text{Div}(V)$$
$$g^{\#}: \text{Cl}(W) \to \text{Cl}(V)$$

REMARK. This shows the "correct" way to map Weil divisors: if w is a prime Weil divisor, it is a codimension-1 subvariety and its image in $\text{Div}(V)$ is not merely $g^{-1}(w)$; it might have a multiplicity other than 1.

PROOF. If $d = \{(U_i, f_i\}$ is a Cartier divisor on W, define $g^*(d)$ to be $\{(g^{-1}(U_i), g \circ f_i)\}$. If both varieties are normal, then Cartier divisors correspond to Weil divisors. □

The reader might notice similarities between the patching condition of a Cartier divisor (statement 2 in definition 5.9.13 on page 290) and the *transition functions* of a line bundle. This is no coincidence and implies

PROPOSITION 5.9.17. *If V is a variety and $P \subset \text{Cart}(V)$ is the subgroup of principal divisors, there exists an exact sequence*

(5.9.3) $$0 \to P \to \text{Cart}(V) \xrightarrow{b} \text{Pic}(V)$$

(see definition C.1.13 on page 513).

REMARK. It follows that

$$\frac{\text{Cart}(V)}{P} \subset \text{Pic}(V)$$

PROOF. Given a Cartier divisor, $\{f_i\}$, we can construct a line-bundle by identifying $U_i \times \mathbb{A}^1$ with $U_j \times \mathbb{A}^1$ by the transition function $\varphi_{i,j} = f_i/f_j$ (on U_i). It is also not hard to see that *products* of Cartier divisors map to *tensor-products* of line-bundles (the transition functions are multiplied in each case). Consequently, we get a homomorphism

(5.9.4) $$b: \text{Cart}(V) \to \text{Pic}(V)$$

whose kernel is precisely $P \subset \text{Cart}(V)$. □

The map

$$b: \text{Cart}(V) \to \text{Pic}(V)$$

is *often* surjective. On normal varieties, this gives us a way to compute Picard groups because this quotient is equal to $\text{Cl}(V)$. For instance example 5.9.6 on page 287 would imply that $\text{Pic}(k\mathbb{P}^n) = \mathbb{Z}$.

To characterize when this happens, we need

DEFINITION 5.9.18. Let ξ be a line-bundle on a variety, V, with sheaf of sections $\mathscr{S}_\xi$. If $\mathcal{K}_V$ is the sheaf of meromorphic functions (see definition 4.3.21 on page 178) then

$$\mathscr{S}_\xi \otimes_{\mathcal{O}_V} \mathcal{K}_V$$

is called the *sheaf of meromorphic sections* of ξ. There is a canonical inclusion

$$\mathscr{S}_\xi \hookrightarrow \mathscr{S}_\xi \otimes_{\mathcal{O}_V} \mathcal{K}_V$$

REMARK. Despite the verbiage involving sheaves, meromorphic sections are fairly easy to describe:

> Let $\{U_\alpha\}$ be an open cover of V such that $\xi|U_\alpha$ is trivial (see definition C.1.3 on page 507), with transition functions $\varphi_{\alpha,\beta}: U_\alpha \cap U_\beta \to k$ — functions that never vanish. A *meromorphic section* of ξ is a set of meromorphic functions $\{f_\alpha \in \mathcal{K}_V^*(U_\alpha)\}$ such that

(5.9.5)
$$f_\alpha \cdot f_\beta^{-1} = \varphi_{\alpha,\beta}$$

> or $f_\alpha = \varphi_{\alpha,\beta} \cdot f_\beta$ (see proposition C.1.11 on page 512).

So a "meromorphic section" is a set of meromorphic functions that *mimic* the properties of a section. It is not a *real* section unless those functions are *regular*. If $\mathcal{K}_V(V) = F$ is a field, we can regard meromorphic sections as sections of the line bundle over F given by $\xi \otimes_k F$.

Equation 5.9.5 implies that the functions $\{f_\alpha\}$ define a Cartier divisor whose class (in $\mathrm{Cart}(V)/P$) is uniquely determined by ξ. This implies that:

PROPOSITION 5.9.19. *The image of the map b in diagram 5.9.3 on the preceding page consists of line-bundles that have a meromorphic section.*

Given the identification of line-bundles with invertible sheaves, we also get:

PROPOSITION 5.9.20. *If V is a variety and $D = \{(U_i, f_i)\}$ is a Cartier divisor then the subsheaf generated by*

$$\bar{D} = \{f_i^{-1} \subset \mathcal{K}_V(U_i)\}$$

defines an invertible sheaf, $\mathbf{s}(D) \subset \mathcal{K}_C$ on V corresponding to the line-bundle $b(D)$ in 5.9.4 on the previous page.

PROOF. The transition functions $\varphi_{\alpha,\beta} = f_\alpha \cdot f_\beta^{-1}$ in equation 5.9.5 are compatible with the definition above. □

COROLLARY 5.9.21. *If V is a variety and $D = \{(U_i, f_i)\}$ is an effective Cartier divisor defining a codimension-1 subscheme $W \subset V$ then the ideal $\mathscr{I}$ defining W is given by*

$$\mathscr{I} = \mathbf{s}(-D)$$

PROOF. If $D = \{(U_i, f_i)\}$, the sheaf $\mathbf{s}(-D)$ is just the subsheaf of $\mathcal{K}_V$ generated by the functions $f_i \cdot \mathcal{O}_V$ on U_i. This is clearly a quasi-coherent ideal in $\mathcal{O}_V$ that defines the subscheme W. □

Our main result characterizes the Picard group in terms of Cartier divisors:

THEOREM 5.9.22. *If V is an irreducible projective variety, the map*

$$b \colon \operatorname{Cart}(V) \to \operatorname{Pic}(V)$$

is surjective so that

$$\frac{\operatorname{Cart}(V)}{P} \cong \operatorname{Pic}(V)$$

REMARK. This is a simplified version of proposition 21.3.4 of [**67**, p. 264] — which proved the result in enormous generality (and showed that the irreducibility of V is not a *necessary* condition). Our treatment follows that in [**70**].

PROOF. Let ξ be a line-bundle on V and pick an open cover $\{U_\alpha\}$ with the property that $\xi|U_\alpha$ is trivial and with transition functions $\varphi_{\alpha,\beta} \colon U_\alpha \cap U_\beta \to k$. If U_α is one such open set,

$$\mathscr{S}_\xi(U_\alpha) = \mathcal{O}_V(U_\alpha)$$

so

$$\mathscr{S}_\xi|U_\alpha \otimes_{\mathcal{O}_V|U_\alpha} \mathcal{K}_V|U_\alpha = \underline{\mathcal{K}_V}|U_\alpha$$

— a constant sheaf of the meromorphic functions on U_α.

CLAIM. Since $U_\alpha \subset V$ is *dense,* we claim that the sheaf of meromorphic sections is *globally* a constant sheaf equal to the global meromorphic functions.

This follows by an argument due to Serre in [**150**]: Consider the maps to the stalks $f_x \colon \mathcal{K}_V(U) \to \mathcal{K}_{V,x}$ for $x \in U$ (see B.1.1 on page 495) and suppose $c \in \mathcal{K}_V(U)$ is in the kernel of f_x. Since $\mathcal{K}_{V,x} = \varinjlim \mathcal{K}_V(\hat{U})$ as $\hat{U}$ runs over all open sets with $x \in \hat{U}$, it follows that exists a maximal open set $U' \subset U$ with $p_{U'}^U(c) = 0$. Since the sheaf is locally constant, U' must also be *closed.* This means it is *empty* since V is irreducible. So the maps $\mathcal{K}_V(U) \to \mathcal{K}_{V,x}$ are all *injective*.

On the other hand, if $c \in \mathcal{K}_{V,x}$ there exists a neighborhood U with $c \in \mathcal{K}_V(U)$ and with $\mathcal{K}_V|U$ a constant sheaf. Since V is irreducible, *every other* open set U' intersects U and $U \cap U' \neq \varnothing$. Since $\mathcal{K}_V|U'$ is a constant sheaf, it follows that $c \in \mathcal{K}_V(U')$ and $c \in \mathcal{K}_V(V)$. So the maps $\mathcal{K}_V(U) \to \mathcal{K}_{V,x}$ are also *surjective*.

We get an isomorphism

$$\mathscr{S}_\xi(U_\alpha) \otimes_{\mathcal{O}_V(U_\alpha)} \mathcal{K}_V(U_\alpha) = \mathcal{O}_V(U_\alpha) \otimes_{\mathcal{O}_V(U_\alpha)} \mathcal{K}_V(U_\alpha) = \mathcal{K}_V(U_\alpha) \xrightarrow{\cong} \underline{\mathcal{K}_V(V)}$$

for all α. If $f_\alpha \in \mathcal{K}_V(U_\alpha)$, for all α, are the inverse images of $1 \in \mathcal{K}_V(V)$ under this map, then $f_\alpha = \varphi_{\alpha,\beta} \cdot f_\beta$ so that the $\{(f_\alpha, U_\alpha)\}$ constitute a meromorphic section of ξ. $\qquad \square$

The following result is due to Serre, in [**150**]:

COROLLARY 5.9.23. *If $V = k\mathbb{P}^n$, then $\operatorname{Pic}(V) = \mathbb{Z}$ with the element $n \in \operatorname{Pic}(V)$ given by the Serre twist, $\mathcal{O}_V(n)$ (see definition 5.3.1 on page 234).*

REMARK. Serre's original proof in [**150**] was constructive and direct.

PROOF. Example 5.9.6 on page 287, proposition 5.9.17 on page 291, and theorem 5.9.22 imply that $\operatorname{Pic}(V) = \mathbb{Z}$. To see that $\operatorname{Pic}(V)$ is generated by $\mathcal{O}_V(1)$, note that proposition 5.3.3 on page 235 shows that $\mathcal{O}_V(1)$ has a section given by X_0 (in homogeneous coordinates $[X_0 : \cdots : X_n]$), which defines a hyperplane ($X_0 = 0$) in V. This generates $\operatorname{Cl}(V)$ and all other hyperplanes are equivalent

to it (i.e. other sections include those defined by X_i, which give hyperplanes $X_i = 0$, but these are linearly equivalent to the one defined by X_0 since they are multiples by rational functions, X_i/X_0). □

Here's an important example of a line-bundle on $k\mathbb{P}^1$:

EXAMPLE 5.9.24. Let $V = k\mathbb{P}^n$ and define the "tautological line-bundle," η, on V as the subspace of $k\mathbb{P}^n \times \mathbb{A}^{n+1}$ given by

$$[X_0: \cdots : X_n] \times (tX_0, \ldots, tX_n)$$

for all $t \in k$. Every point of $k\mathbb{P}^n$ represents a line through the origin of $\mathbb{A}^{n+1}$ and this line-bundle simply pairs a point with the line it represents (hence the term "tautological"). On $\mathbb{A}_i^n$ we have

$$X_j = x_j \cdot X_i$$

where $x_j = X_j/X_i$ are the non-homogeneous coordinates, so the bundle is given by

$$(x_0, \ldots, x_{i-1}, 1, x_i, \ldots, x_n) \times (tx_0 \cdot X_i, \ldots, tX_i, \ldots, tx_n \cdot X_i)$$

On the overlap $\mathbb{A}_i^n \cap \mathbb{A}_j^n$ (coming from $\mathbb{A}_i^n$) we have

$$(x_0, \ldots, x_{i-1}, 1, x_i, \ldots, x_{j-1}, 1, x_j \ldots, x_n)$$
$$\times (x_0 \cdot X_i, \ldots, X_i, \ldots, X_i, \ldots, x_n \cdot X_i)$$

and coming from $\mathbb{A}_j^n$ we have

$$(x_0, \ldots, x_{i-1}, 1, x_i, \ldots, x_{j-1}, 1, x_j \ldots, x_n)$$
$$\times (x_0 \cdot X_j, \ldots, X_j, \ldots, X_j, \ldots, x_n \cdot X_j)$$

so the transition function is $\varphi_{j,i} = X_j/X_i$. This means that $\eta = \mathcal{O}_V(-1)$ (see definition 5.3.1 on page 234).

In general the Grassmannian, $\mathbb{G}_{n,r}$ (see definition 5.2.10 on page 230) has a "tautological" r-plane bundle — a subspace of

$$\xi \subset \mathbb{G}_{n,r} \times \mathbb{A}^n$$

pairing a point of $\mathbb{G}_{n,r}$ with the r-dimensional subspace of $\mathbb{A}^n$ that the point represents.

EXERCISES.

1. If V is a variety, recall its sheaf of meromorphic functions, $\mathcal{K}_V$ and its sheaf of regular meromorphic functions, $\mathcal{K}_V^*$ — see section 4.3.2 on page 4.3.2. If $\mathcal{O}_V^*$ denotes the sheaf defined on affine open sets as the *nonzero* elements of $\mathcal{O}_V$, show that

$$\text{Cart}(V) = \left(\frac{\mathcal{K}_V^*}{\mathcal{O}_V^*} \right)(V)$$

— the *quotient sheaf* (see definition B.2.6 on page 501) evaluated on V, and that a Cartier divisor is principal if it is in the image of $\mathcal{K}_V^*(V)$.

2. Suppose $V = \mathbb{C}P^1$ with finite open set $\mathbb{C}_0^1$. Compute the divisor, (f), if:

    a. $f(x) = x$
    b. $f(x) = (x-1)/x$
    c. $f(x) = 1/x^2$
    d. $f(x) = (x-1)(x-4)$

3. Suppose $V = k\mathbb{P}^n$ and $W = \mathcal{P}(F)$ is a smooth degree-d hypersurface where F is a homogeneous polynomial of degree d. Show that $\mathrm{Pic}(V \setminus W) = \mathbb{Z}/d \cdot \mathbb{Z}$, generated by $\mathcal{O}_V(1)|V \setminus W$.

4. If $x = f/g \in k(V)$, show that $(x)_0 \equiv (x)_\infty \in \mathrm{Cl}(V)$.

5. If V is a variety and W is a blow-up show that

$$\mathrm{Cl}(W) \cong \mathrm{Cl}(V) \oplus \mathbb{Z}$$

5.9.2. Divisors and rational maps. We can use divisors to analyze rational maps. Although rational maps can have singularities, in some cases these can be relatively "small":

PROPOSITION 5.9.25. *If V is a smooth variety and $f: V \to k\mathbb{P}^n$ is a rational map, let*

$$(f_i) = \sum_{j=1}^m a_{i,j} C_j = D_i$$

where C_j are prime divisors and set

$$D = \gcd(\{D_i\}) = \sum_{j=1}^m \min_i(a_{i,j}) C_j$$

Then the divisors $D_i' = D_i - D$ have no common components and the map f is non-regular at the points of

$$\bigcap_{i=0}^n \mathrm{supp} D_i'$$

which has codimension ≥ 2.

PROOF. In a neighborhood of any point $p \in V$, we have

$$f_i = \left(\prod_{j=1}^m \pi_j^{a_{i,j}} \right) u_i \in \mathcal{O}_{V,p}$$

where the $\{\pi_j\}$ are local uniformizers (see definition 3.3.21 on page 132) and the u_i are units. Since $\mathcal{O}_{V,p}$ is a unique factorization domain (see theorem 3.3.41 on page 144), there exists a well-defined greatest common divisor, g, of the f_i.

Since the target of f is a projective space,

$$(f_0: \cdots : f_n) \sim (f_0 g^{-1} : \cdots : f_n g^{-1}) \in k\mathbb{P}^n$$

so the map f is regular at point p if not all of the $f_i g^{-1}$ are zero at p. The points where they all vanish are precisely the intersection of the support of the divisors D_i'. $\qquad \square$

This has the remarkable consequence that *rational maps* induce homomorphisms of differential forms:

COROLLARY 5.9.26. *If X and Y are smooth varieties with Y projective, and $f: X \to Y$ is a rational mapping such that $f(X) \subset Y$ is dense, then f induces an injection of sheaves*

$$f^*: \Omega_Y^i \to \Omega_X^i$$

REMARK. Smoothness plays a vital role here: pullbacks of differential forms on Y may have singularities, but proposition 5.9.25 on the preceding page implies that these will be of codimension ≥ 2. Since these pullbacks are essentially *functions* (multiplying differentials), and singularities of *functions* are codimension 1, they must be regular.

PROOF. Most of the proof has been outlined above. The rational map f will be regular on $X \setminus Z$ where Z is of codimension ≥ 2. If $\omega \in \Omega_Y^i$, then $f^*(\omega)$ is regular on $X \setminus Z$. If $V \subset Y$ is an open set with $\omega|V$ a linear combination of basic differentials and $U \subset f^{-1}(V)$ is an open set in X, then

$$f^*\omega = \sum g_{n_1,\dots,n_i} du_{n_1} \wedge \cdots du_{n_i}$$

where the $g_{n_1,\dots,n_i}$ are functions that are regular on $U \setminus (U \cap Z)$. Since $U \cap Z$ is of codimension ≥ 2 and singularities of a function are codimension 1, we conclude that the $g_{n_1,\dots,n_i}$ are regular on *all* of U.

Injectivity of f^* follows from the fact that the image of f is dense in Y. $\quad\square$

5.9.3. Linear spaces of divisors. We can use divisors to carry out a much more "fine-grained" analysis of polynomials and rational functions by specifying the maximum degree and the locations and types of zeros and singularities:

DEFINITION 5.9.27. If $D \in \text{Div}(V)$ is a divisor defined on a variety, V, the *space of D*, denoted $\mathcal{L}(D)$ is the set of meromorphic functions $f \in k(V)$ such that

$$(f) + D \geq 0$$

or, if $D = \sum_{i=1}^m n_i \cdot P_i$, then

$$\text{ord}_{P_i}(f) \geq -n_i$$

for all $i = 1, \dots, n$. The *dimension of D*, denoted $\ell(D)$ is the dimension of this vector-space.

REMARK. A polynomial of degree n has a singularity of order n at ∞, so the space of $n \cdot \langle \infty \rangle$ is precisely the set of polynomials of degree $\leq n$ and $\ell(n \cdot \langle \infty \rangle) = n + 1$.

The dimension of $\mathcal{L}(D)$ only depends on its *class* in $\text{Cl}(V)$:

PROPOSITION 5.9.28. *If V is a variety and $D_1 \equiv D_2 \in \text{Cl}(V)$ are divisors, then there is an isomorphism*

$$\iota: \mathcal{L}(D_2) \to \mathcal{L}(D_1)$$

PROOF. If $D_2 = D_1 + (g)$, where g is a meromorphic function on V then

$$(fg) + D_1 \geq 0 \Leftrightarrow (f) + (g) + D_1 \geq 0$$

so $f \in \mathcal{L}(D_2)$ if and only if $fg \in \mathcal{L}(D_1)$. The isomorphism is multiplication by g. $\quad\square$

In the case of plane *curves,* it is easy to draw conclusions about divisors:

PROPOSITION 5.9.29. *If D_1, D_2 are divisors on a smooth curve, V then*

(1) $D_1 \leq D_2$ *implies that* $\mathcal{L}(D_1) \subseteq \mathcal{L}(D_2)$ *and* $\ell(D_1) \leq \ell(D_2)$,
(2) *if* $D_1 \leq D_2$, *then* $\dim_k \mathcal{L}(D_2)/\mathcal{L}(D_1) \leq \deg D_2 - \deg D_1$,
(3) $\mathcal{L}(0) = k$,
(4) *if* $D \geq 0$, *then* $\ell(D) \leq \deg D + 1$,

REMARK. Since $\text{Cl}(k\mathbb{P}^1) = \mathbb{Z}$, a divisor, D, is equivalent to $\deg(D) \cdot \langle \infty \rangle$. If $D \geq 0$, it follows that

$$\ell(D) = \deg D + 1$$

on $k\mathbb{P}^1$.

PROOF. The first statement follows from the transitivity of an order-relation: if $(f) + D_1 \geq 0$, then $(f) + D_2 \geq (f) + D_1 \geq 0$.

To see the second statement, consider $\mathcal{L}(D_1 + P)$, where P is a prime divisor. Let π be a uniformizing parameter of $\mathcal{O}_{V,P}$. If n is the coefficient of P in D_1, and $f \in \mathcal{L}(D_1 + P)$, note that $\pi^{n+1} \cdot f \in \mathcal{O}_{V,P}$ since $\text{ord}_P(f) \geq -n-1$. Consequently, we can define a homomorphism

$$g: \mathcal{L}(D_1 + P) \quad \rightarrow \quad k$$
$$f \quad \mapsto \quad (\pi^{n+1} \cdot f)(x)$$

— or, equivalently, the image of $\pi^{n+1} \cdot f$ under the projection

$$\mathcal{O}_{V,P} \rightarrow \mathcal{O}_{V,P}/\pi \cdot \mathcal{O}_{V,P} = k$$

The kernel of g is $\mathcal{L}(D_1)$ so we conclude

$$\dim_k \mathcal{L}(D_1 + P)/\mathcal{L}(D_1) \leq 1$$

with equality if $\mathcal{L}(D_1) \subsetneq \mathcal{L}(D_1 + P)$. Statement 2 follows by induction after writing $D_2 = D_1 + P_1 + \cdots + P_t$.

Statement 3 follows from the Nullstellensatz — $f \in \mathcal{L}(0)$ has no poles, so it is a global regular function — a constant.

Statement 4 follows from statements 3 and 2, which imply that

$$\dim \mathcal{L}(D)/\mathcal{L}(0) = \ell(D) - 1 \leq \deg D$$

The final statement follows from the fact that, if $D = D_1 - D_2$ with $D_1, D_2 \geq 0$ then $D \leq D_1$ so

$$\mathcal{L}(D) \subset \mathcal{L}(D_1)$$

$\square$

EXAMPLE 5.9.30. Suppose $V = k\mathbb{P}^1$ for $k = \mathbb{F}_3$ and $P_1 = [1{:}0]$, $P_2 = [1{:}2]$, $\odot = [0{:}1]$. Then we have

D	Generators of $\mathcal{L}(D)$	deg D	$\ell(D)$
$\langle \odot \rangle - \langle P_1 \rangle$	X	0	1
$2\langle \odot \rangle - 2\langle P_2 \rangle$	$(X+1)^2$	0	1
$3\langle \odot \rangle - 2\langle P_2 \rangle$	$(X+1)^2, X(X+1)^2$	1	2
$\langle \odot \rangle - 2\langle P_1 \rangle$	$\varnothing$	-1	0

LEMMA 5.9.31. *Suppose D is a Cartier divisor on a irreducible, normal variety, V over a field k, with associated line-bundle, ξ. Then there exists an isomorphism*

$$\mathcal{L}(D) \to \Gamma(V, \xi) = \mathscr{S}_{\xi}(V)$$

The set of effective divisors linearly equivalent to D is $\mathbb{P}(\Gamma(V, \xi))$. In addition, $\mathcal{L}(D)$ is a finite-dimensional vector space over k.

PROOF. If f is any function on V and D is locally defined by meromorphic functions f_α on open affines U_α, then

$$\{f \cdot f_\alpha\}$$

defines a *meromorphic section, m*, of ξ (see definition 5.9.18 on page 292 and the remark following it). If $f \in \mathcal{L}(D)$, then $f \cdot f_\alpha$ is a regular function for all α so m is an *actual* section of ξ. Conversely, given an actual section s of ξ, the functions

$$(s|U_\alpha) \cdot f_\alpha^{-1}$$

define a global meromorphic function g such that $g \in \mathcal{L}(D)$.

Every global section of ξ defines a meromorphic function f such that $(f) + D \geq 0$ so that $(f) + D$ is an effective divisor linearly equivalent to D — and any effective divisor linearly equivalent to D arises this way.

Suppose two sections, s and s', of ξ give rise to the same divisor. This implies that divisors of

$$(s|U_\alpha) \cdot f_\alpha^{-1} \text{ and } (s'|U_\alpha) \cdot f_\alpha^{-1}$$

are identical. This means that

$$(s|U_\alpha) \cdot f_\alpha^{-1} \cdot \left((s'|U_\alpha) \cdot f_\alpha^{-1}\right)^{-1}$$

are regular functions on the affine sets U_α that never vanish. The Nullstellensatz implies that they are *constants*. The fact that these constants must agree on overlaps implies that there is a "global" constant, $c \in k$, such that

$$s = c \cdot s'$$

But this is precisely the condition that s and s' define the same point in $\mathbb{P}(\Gamma(V, \xi))$.

The final statement follows from theorem 5.5.24 on page 260. □

The vector space $\ell(D)$ can *vanish* under some circumstances:

COROLLARY 5.9.32. *If D is a divisor on a smooth one-dimensional variety with $\ell(D) > 0$, then $\deg D \geq 0$.*

REMARK. We'll be more interested in the contrapositive: if $\deg D < 0$, then $\ell(D) = 0$.

PROOF. If $\ell(D) > 0$, there exists a rational function f such that $(f) + D \geq 0$. Since proposition 5.9.8 on page 288 shows that $\deg(f) = 0$, we have $\deg(D + (f)) = \deg D \geq 0$. □

EXERCISES.

6. Suppose V is a variety and $D \in \mathrm{Div}(V)$. If $f \in \mathcal{L}(D)$ show that $f \notin \mathcal{L}(D - P)$ for all but a finite number of prime divisors, P.

5.9.4. Differential forms and divisors. In this section, we will study divisors associated with differential forms (see section 4.6.2 on page 210).
This has an interesting consequence:

COROLLARY 5.9.33. *If V is an irreducible smooth projective variety over a field of characteristic* 0, *then*
$$h^i = \dim_k \Omega^i_V(V)$$
is finite for all i.

REMARK. For a proof of this that doesn't make use of sheaf-cohomology, see [**153**].
The fact that the maps in a birational equivalence have dense images and corollary 5.9.26 on page 296 imply that the $\{h^i\}$ are *birational invariants* of V.

PROOF. Since Ω^i_V is clearly coherent (see definition 3.5.5 on page 155), the conclusion follows from theorem 5.5.24 on page 260. $\qquad\square$

We can also define the module of *rational* forms on V:

DEFINITION 5.9.34. If V is an irreducible variety, then the module of *rational forms*, $\Omega^t(V)$ is the set of equivalence classes of pairs $(U, \omega \in \Omega^t_U)$, where
$$(U_1, \omega_1) \sim (U_2, \omega_2)$$
if $\omega_1 | U_1 \cap U_2 = \omega_2 | U_1 \cap U_2$.

REMARK. Note the subtle distinction in notation: Ω^t_V verses $\Omega^t(V)$: the former is a locally-free sheaf and the latter is a vector space.

We can immediately conclude

COROLLARY 5.9.35. *If V is an n-dimensional irreducible variety, then $\Omega^t(V)$ is a vector space over $k(V)$ of dimension $\binom{n}{t}$.*

PROOF. If $x \in \Omega^t(V)$, then there exists an open set U such that $x \in \Omega^t_U$ and is a linear combination of $\binom{n}{t}$ basis elements:
$$x = \sum f_{i_1,\dots,i_t} du_{i_1} \wedge \cdots \wedge du_{i_t}$$
where the $\{u_j\}$ are local parameters in U. Since the $\{f_{i_1,\dots,i_t}\} \subset k[U]$, they define elements of $k(X)$, as do the $\{u_j\}$. $\qquad\square$

5.9.5. The canonical divisor class. If V is an n-dimensional smooth irreducible variety, then $\Omega^n(V)$ is *one-dimensional* over $k(V)$ — in other words

$$\Omega^n(V) = \omega \cdot k(V)$$

In an open affine, U_1

$$\omega = du_1 \wedge \cdots \wedge du_n$$

where the u_i are local parameters. If U_2 is another open affine with local parameters $\{w_1, \ldots, w_n\}$ then, on $U_1 \cap U_2$ there exists an invertible matrix, M with entries in $k[V]$ such that

$$\mathbf{w} = M \cdot \mathbf{u}$$

where

$$\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

and lemma A.6.14 on page 480 implies that

$$dw_1 \wedge \cdots \wedge dw_n = \det M \cdot du_1 \wedge \cdots \wedge du_n$$

where $\det M$ is nonvanishing on $U_1 \cap U_2$. This implies that ω defines a *Cartier divisor* on V — and that they are all in the same *class* (i.e. they differ by globally defined rational functions).

DEFINITION 5.9.36. The divisor class, K_V, described above is called the *canonical class* of V.

REMARK. The canonical class encapsulates deep algebraic and topological information about a projective variety.

PROPOSITION 5.9.37. *Let V be a smooth n-dimensional projective variety and let K_V be the canonical class. If K_V is represented by the rational differential form $\omega \cdot du_1 \wedge \cdots \wedge du_n$ then*

$$(f) + K_V \geq 0$$

if and only if $f\omega \cdot du_1 \wedge \cdots \wedge du_n \in \Omega_V^n$. It follows that

$$\dim_k \Omega_V^n(V) = \ell(K_V)$$

In some cases, computing the canonical class can be straightforward:

EXAMPLE 5.9.38. If $V = k\mathbb{P}^1$, $\mathrm{Cl}(V) = \mathbb{Z}$ so a divisor-class is completely determined by its degree. If

$$k\mathbb{P}^1 = \mathbb{A}_0^1 \cup \mathbb{A}_1^1$$

let x_i be the single coordinate of $\mathbb{A}_i^1$. Then $\Omega_{k[x_0]/k} = \Omega_{\mathbb{A}_0^1}^1 = k[x_0] \cdot dx_0$, so $\Omega^1(\mathbb{A}_0^1) = k(x_0) \cdot dx_0$. In gluing together $\mathbb{A}_0^1$ and $\mathbb{A}_1^1$, we identify x_1 with x_0^{-1} so that dx_0 becomes

$$dx_0 = -\frac{1}{x_1^2} \cdot dx_1$$

making the canonical class on $k\mathbb{P}^1$ the element $-2 \in \mathbb{Z} = \mathrm{Cl}(k\mathbb{P}^1)$, or $-2 \cdot \langle p \rangle$ where $p \in k\mathbb{P}^1$ is *any* point.

We will conclude this section with a more complex and instructive example — a *hypersurface*. We will follow the treatment of this problem by Shafarevich in [**153**]:

EXAMPLE 5.9.39. Let $H \subset k\mathbb{P}^n$ be a smooth degree-s hypersurface defined by

$$p(X_0, \ldots, X_n) = 0$$

where p is a degree-s homogeneous polynomial with $p_i = p(X_0, \ldots, X_{i-1}, 1, X_{i+1}, \ldots, X_n)$ defining $H_i = H \cap \mathbb{A}_i^n$.

If we focus on H_0, let $x_i = X_i/X_0$ and let U_t be the (open, affine) set of points where $\partial p_0 / \partial x_t = p_{0,x_t} \neq 0$ on H_0. The fact that H_0 is smooth implies

$$H_0 = \bigcup_{j=1}^{n} U_j$$

— i.e., at least one of the derivatives is nonvanishing at each point. The $n - 1$-dimensional tangent plane at a point is defined by

(5.9.6) $$\sum_{t=1}^{n} p_{0,x_t} \cdot dx_t = 0$$

so that $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ is a set of local parameters on U_i — see 3.3.21 on page 132.

We can define

$$\omega_i = \frac{dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_n}{p_{0,x_i}}$$

on U_i. If we take the $\wedge$-product of

$$dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n$$

with equation 5.9.6, we get

$$(-1)^{i-1} p_{0,x_i} \cdot dx_1 \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n$$
$$+ (-1)^{j-2} p_{0,x_j} dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_n = 0$$

or

$$(-1)^{i-1} p_{0,x_i} \cdot dx_1 \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n$$
$$= (-1)^{j-1} p_{0,x_j} dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_n$$

which implies that $\omega_j = (-1)^{i-j} \omega_i$ so ω_n is regular everywhere on H_0 and generates $\Omega_{H_0}^{n-1}$. By definition 5.9.34 on page 299, it defines a unique element of $\Omega^{n-1}(H)$.

Now we consider what the differential form ω_n looks like on H_1. Suppose the coordinates of H_1 are

$$y_i = \begin{cases} \dfrac{X_0}{X_1} = \dfrac{1}{x_1} & \text{if } i = 1 \\[2mm] \dfrac{X_i}{X_1} = \dfrac{x_i}{x_1} & \text{if } i > 1 \end{cases}$$

This implies that

$$dx_1 = -\frac{dy_1}{y_1^2}$$

and

$$dx_i = \frac{y_1 \cdot dy_i - y_i \cdot dy_1}{y_1^2}$$

so

$$dx_1 \wedge \cdots \wedge dx_{n-1} = -\frac{dy_1 \wedge \cdots \wedge dy_{n-1}}{y_1^n}$$

which implies that

$$\omega_n = -\frac{dy_1 \wedge \cdots \wedge dy_{n-1}}{y_1^n \cdot p_{0,x_n}}$$

It remains to convert p_{0,x_n} to the y-coordinates. We have

$$p_1 = y_1^s \cdot p_0 \left(\frac{1}{y_1}, \ldots, \frac{y_n}{y_1} \right)$$

so

$$
\begin{aligned}
p_{1,y_n} = \frac{\partial p_1}{\partial y_n} &= y_1^s \cdot \frac{\partial x_n}{\partial y_n} \cdot p_{0,x_n} \left(\frac{1}{y_1}, \ldots, \frac{y_n}{y_1} \right) \\
&= y_1^{s-1} \cdot p_{0,x_n}(x_1, \ldots, x_n)
\end{aligned}
$$

which gives the conversion

$$p_{0,x_n}(x_1, \ldots, x_n) = \frac{p_{1,y_n}}{y_1^{s-1}}$$

We conclude that

$$\omega_n = -\frac{dy_1 \wedge \cdots \wedge dy_{n-1}}{y_1^{n-s+1} \cdot p_{1,y_n}} \in \Omega^{n-1}(H_1)$$

which defines the Cartier divisor

$$-\frac{1}{y_1^{n-s+1} \cdot p_{1,y_n}}$$

We conclude that the canonical class of H is

$$K_H = -(n - s + 1) \cdot \langle Y \rangle$$

where Y is the hypersurface in H defined by $y_1 = 0$.

If we had picked some other H_i to do these computations, we would have gotten a similar result — and the resulting divisor would have differed from $-(n - s + 1) \cdot \langle Y \rangle$ by a rational function so the *class* would have been the same.

The upshot of this is that the canonical class is equivalent to

(5.9.7) $$K_H = -(n - s + 1)\langle L \rangle$$

where L is any linear function.

Now we will compute the *dimension* of $\mathcal{L}(K_H)$, or $\ell(K_H)$. Suppose $f(x_1, \ldots, x_n)$ is a degree-d polynomial on H_0 defining the differential form

$$f \cdot \omega_n \in \Omega_{H_0}^{n-1}$$
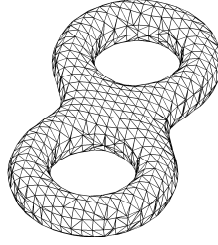
FIGURE 5.9.1.   Topological genus 2

On H_1, we get

$$
\begin{aligned}
f &= f\left(\frac{1}{y_1}, \ldots, \frac{y_n}{y_1}\right) \\
&= y_1^{-d} \cdot y_1^{d} \cdot f\left(\frac{1}{y_1}, \ldots, \frac{y_n}{y_1}\right) \\
&= \frac{\bar{f}(y_1, \ldots, y_n)}{y_1^{d}}
\end{aligned}
$$

where $\bar{f}(y_1, \ldots, y_n) \in k[y_1, \ldots, y_n]$ is a polynomial, which means that

$$
f \cdot \omega_n = -\frac{\bar{f}(y_1, \ldots, y_n) dy_1 \wedge \cdots \wedge dy_{n-1}}{y_1^{d+n-s+1} \cdot p_{1,y_n}}
$$

This corresponds to the divisor

$$
-\frac{\bar{f}(y_1, \ldots, y_n)}{y_1^{d+n-s+1} \cdot p_{1,y_n}}
$$

and is *effective* if and only if $d + n - s + 1 \leq 0$ or $d \leq s - n - 1$. If m is the dimension of H, then $m = n - 1$ and we have $d \leq s - m - 2$. The set of such polynomials form a vector space of dimension

$$
\ell(K_H) = \begin{cases} 0 & \text{if } s - m - 2 < 0 \\ \dbinom{s-1}{m+1} & \text{otherwise} \end{cases}
$$

— see proposition 5.4.6 on page 244. If $m = 1$ — the case of projective plane curves studied in chapter 6.1 on page 305 — we get

(5.9.8) $$\ell(K_H) = \binom{s-1}{2} = \frac{(s-1)(s-2)}{2}$$

DEFINITION 5.9.40. If V is an irreducible smooth projective variety, the number $p_g(V) = \ell(K_V)$ is called the *geometric genus* of V.

REMARK. We've just shown that the genus of a hypersurface of degree d in $k\mathbb{P}^2$ is $(d-1)(d-2)/2$.

If $k = \mathbb{C}$, our *curve* is topologically a two-dimensional *surface* and $\ell(K_V)$ turns out to equal the *topological* genus of the curve — essentially the number

of handles one must sew onto a sphere to get the surface. See figure 5.9.1. Proving this is beyond the scope of this book.

If $k \neq \mathbb{C}$, V is not a surface and the topological genus is undefined.

The geometric genus is a birational invariant:

THEOREM 5.9.41. *If X and Y are smooth, birationally equivalent projective varieties, then*

$$p_g(X) = p_g(Y)$$

REMARK. The computations in example 5.9.39 on page 301 implies that there are an infinite number of birationally inequivalent hypersurfaces.

PROOF. This follows immediately from corollary 5.9.26 on page 296, the fact that birational equivalences have dense images, and proposition 5.9.37 on page 300. □

CHAPTER 6

Curves

"Everyone knows what a curve is, until he has studied enough mathematics to become confused through the countless number of possible exceptions."
—Felix Klein, quoted by Carl B. Boyer — *Scientific American,* "The Invention of Analytic Geometry", 1949.

6.1. Basic properties

In this chapter, we will focus on smooth, one-dimensional projective varieties or *curves* which, according to theorem 5.6.9 on page 265, naturally embed in $k\mathbb{P}^3$. In fact, throughout most of this chapter, we will deal with a special type of curve:

DEFINITION 6.1.1. A one-dimensional projective variety, V, is a *projective plane curve* if it has an embedding in $k\mathbb{P}^2$.

REMARK. A projective plane curve is clearly defined by a single equation.

Despite their apparent simplicity, projective plane curves have a rich array of properties — even in the case of *elliptic* curves (the subject of the next section). To begin, we can try to write down exactly what the canonical class is:

PROPOSITION 6.1.2. *If $V \subset k\mathbb{P}^2$ is a smooth projective plane curve of degree d and $H = \mathcal{P}(G) \subset k\mathbb{P}^2$ is a hyperplane that intersects V at points $\{p_1, \ldots, p_t\}$ with multiplicities $\{m_1, \ldots, m_t\}$, respectively, then*

(1) $\sum_{i=1}^{t} m_i = d$
(2) *if $D = \sum_{i=1}^{t} m_i \cdot \langle p_i \rangle$ then $K_V = (d-3) \cdot D$*
(3) $\deg K_V = d \cdot (d-3) = 2g - 2$, *where g is the genus in definition 5.9.40 on page 303.*

REMARK. If H intersects V in points $\{q_1, \ldots, q_s\}$ with multiplicities $\{n_1, \ldots, n_s\}$, respectively, then

$$(G) = \sum_{j=1}^{s} n_j \cdot \langle p_j \rangle$$

PROOF. The first statement follows from Bézout's theorem (5.8.6 on page 279). The second statement follows from equation 5.9.7 on page 302 and the fact that all hyperplanes define equivalent divisors. The third statement follows immediately from the first two (and simple arithmetic). □

One interesting consequence of curves being one-dimensional is:

PROPOSITION 6.1.3. *If $f: V \to W$ be a rational map of curves, then f is regular at every nonsingular point of V.*

PROOF. In a neighborhood of a nonsingular point of V, we can use proposition 5.9.25 on page 295 to conclude that the sets of points where f fails to be regular is of codimension ≥ 2, which in this case means it is empty.   □

Despite our restriction to projective plane curves:

COROLLARY 6.1.4. *All irreducible smooth curves are birationally equivalent to projective plane curves. Furthermore, if V is a smooth curve and*

$$f: V \to W \subset k\mathbb{P}^2$$

is a birational equivalence, we can assume that f is regular everywhere.

REMARK. The plane curve, W, may have singularities, in which case the *inverse* map will not be regular.

PROOF. This follows immediately from theorem 2.8.37 on page 114. If $V \subset k\mathbb{P}^n$ is a curve, and V_0 is an open affine, then there exists a birational equivalence

$$f: V_0 \to W \subset \mathbb{A}^2$$

It follows that an open subset of V is isomorphic to an open subset of the projective completion of W.   □

COROLLARY 6.1.5. *A birational equivalence between smooth curves is an isomorphism.*

REMARK. Normally, birational equivalences are isomorphisms on the open sets where they are regular (see theorem 2.5.27 on page 86). In the present situation, this is everywhere.

This allows us to formulate a criterion for a smooth plane curve to be *rational* (see definition 2.5.28 on page 86):

COROLLARY 6.1.6. *If P is a smooth plane curve that is rational and $f: P \to P$ is a birational equivalence that is not the identity map, then f has at most two fixed points.*

PROOF. In light of 6.1.5, it suffices to prove this for $k\mathbb{P}^1$. The map f will have the form

$$f(x_0 : x_1) = (p_0(x_0, x_1) : p_1(x_0 : x_1))$$

where p_0 and p_1 are homogeneous polynomials of the same degree. Since f must be regular and one-to-one, it follows that the p_i must be *linear*, so

$$f(x_0 : x_1) = (a \cdot x_0 + b \cdot x_1 : c \cdot x_0 + d \cdot x_1)$$

A point $(x_0 : x_1) \in k\mathbb{P}^1$ is a fixed point of f if $f(x_0 : x_1) = (k \cdot x_0 : k \cdot x_1)$ for a nonzero k, or

$$\begin{aligned}
a \cdot x_0 + b \cdot x_1 &= k \cdot x_0 \\
c \cdot x_0 + d \cdot x_1 &= k \cdot x_1
\end{aligned}$$

It follows that the allowable values of k are just the ≤ 2 nonzero *eigenvalues* of the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and the *fixed points* are the corresponding *eigenvectors*.   □

This allows us to prove:

PROPOSITION 6.1.7. *An elliptic curve*

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

in $k\mathbb{P}^2$ with $4A^3 + 27B^2 \neq 0$ is not rational.

PROOF. We extend this curve to $k\mathbb{P}^2$ by rewriting it in homogeneous coordinates

$$\left(\frac{x_2}{x_3}\right)^2 = \left(\frac{x_1}{x_3}\right)^3 + A\frac{x_1}{x_3} + B$$

or

$$x_2^2 x_3 = x_1^3 + Ax_1 x_3^2 + Bx_3^3$$

If we set $x_3 = 0$, we get $x_1 = 0$ so x_2 must be nonzero, giving a point at infinity, $(0\!:\!1\!:\!0)$. The map

$$f\colon (x_1\!:\!x_2\!:\!x_3) \mapsto (x_1\!:\!-x_2\!:\!x_3)$$

is a birational equivalence from the elliptic curve to itself. The point at infinity is clearly a fixed point of this map. In the finite part of the plane ($x_3 \neq 0$), the fixed points are the points with $x_2 = 0$, i. e., solutions of

$$x_1^3 + Ax_1 x_3^2 + Bx_3^3 = 0$$

or

$$\left(\frac{x_1}{x_3}\right)^3 + a\frac{x_1}{x_3} + b = 0$$

Because $4A^3 + 27B^2 \neq 0$, there are *three* of these (see definition A.1.55 on page 366 and the discussion following corollary A.1.56 on page 366). It follows that f has *four* fixed points. The conclusion follows from corollary 6.1.6 on the facing page.   □

We get another criterion for rationality:

LEMMA 6.1.8. *Suppose V is a smooth projective curve and $p_1, p_2 \in V$ are two distinct points. If the divisor $\langle p_1 \rangle - \langle p_2 \rangle$ is principal, then V is rational.*

PROOF. Suppose f is the rational function whose divisor is $\langle p_1 \rangle - \langle p_2 \rangle$. Then we may regard f as a map

$$f\colon V \to k\mathbb{P}^1$$

that sends p_1 to 0 and p_2 to the point at infinity and is regular at all points of V. Since $k\mathbb{P}^1$ is irreducible and theorem 5.5.14 on page 255 implies that the image of f is *closed*, f must be *surjective*. Since V and $k\mathbb{P}^1$ are both one-dimensional, the induced extension of fields

(6.1.1) $$f^*\colon k(k\mathbb{P}^1) \hookrightarrow k(V)$$

must be *algebraic,* hence a finite extension. Let $\mathcal{O}_{k\mathbb{P}^1,0}$ be the coordinate ring of $k\mathbb{P}^1$ at 0 and let $\mathcal{O}_{V,p_1}$ be the coordinate ring of V at p_1. The map f^* defines an inclusion

$$\mathcal{O}_{k\mathbb{P}^1,0} \hookrightarrow \mathcal{O}_{V,p_1}$$

Let t be a local parameter at $0 \in k\mathbb{P}^1$ and let s be a local parameter at $p_1 \in V$ — these are localizations of the coordinate rings of open affines $U \subset k\mathbb{P}^1$ and $U' \subset V$. The *fiber* of f at 0 consists only of the point p_1 (see lemma 2.5.8 on page 75) so that

$$\frac{\mathcal{O}_{V,p_1}}{(t)} \cong \frac{\mathcal{O}_{k\mathbb{P}^1,0}}{(t)} \cong k$$

so it follows that $\mathcal{O}_{V,p_1} \cong \mathcal{O}_{k\mathbb{P}^1,0}$ and the field-extension in equation 6.1.1 on the previous page is of degree 1, so f is a birational equivalence. □

Sometimes it is advantageous to focus on a set of prime divisors to the exclusion of all others.

For instance:

PROPOSITION 6.1.9. *Let $V \subset k\mathbb{P}^2$ be a smooth irreducible projective plane curve and let $\{p_1,\ldots,p_t\} \in V$ be a finite collection of prime divisors (i.e., points). If $m_1,\ldots,m_t$ is a sequence of integers, there exists a meromorphic function $f \in k(V)$ with the property that*

$$\operatorname{ord}_{p_i}(f) = m_i$$

REMARK. It will most emphatically *not* be true that

$$(f) = \sum_{i=1}^{t} m_i \cdot \langle p_i \rangle$$

In fact (f) will contain *many* more nonzero terms, but *among the divisors* $\{p_1,\ldots,p_t\}$ it will have the required property.

PROOF. Let $\ell_i \subset k\mathbb{P}^2$ be a line that passes through p_i but not through any p_j for $j \neq i$: in homogeneous coordinates if $p_i = (x_i,y_i,z_i) \in k\mathbb{P}^2$, set

$$\ell_i(x,y,z) = a_i(x-x_i) + b_i(y-y_i) + c_i(z-z_i)$$

for $a_i,b_i,c_i \in k$. This is possible because k is infinite — in each ℓ_i we only have to avoid a finite set of values of (a_i,b_i,c_i). We can even ensure that ℓ_i is not tangent to V at p_i: if V is defined by $F(x,y,z) = 0$ we simply avoid values for (a_i,b_i,c_i) that make

$$\left(\frac{\partial F}{\partial x}\right)_{p_i} \cdot a_i + \left(\frac{\partial F}{\partial y}\right)_{p_i} \cdot b_i + \left(\frac{\partial F}{\partial z}\right)_{p_i} \cdot c_i = 0$$

— a finite set of forbidden ratios $a_i/c_i, b_i/c_i$. Smoothness implies that at least one derivative is nonzero at each point.

Now define:

$$f = \prod_{i=1}^{t} \left(\ell_i^{m_i}(x,y,z) \cdot \bar{\ell}^{-m_i}(x,y,z)\right)$$

where $\bar{\ell}(x,y,z)$ is a linear function that does *not* pass through *any* of the p_i. This induces a meromorphic function on V with the required properties. □

We will also need a kind of *relative* version of definition 5.9.27 on page 296:

DEFINITION 6.1.10. If V is a variety and S is a set of prime divisors of V the relative space of this divisor with respect to S, denoted $\mathcal{L}^S(D)$ is the set of meromorphic functions f such that

$$\operatorname{ord}_P(f) \geq -\operatorname{ord}_P(D)$$

for all $P \in S$. We define $\deg_S D$ to be the sum of the coefficients of the terms in S.

REMARK. If the set S does not include any of the prime divisors of D, then $\deg_S D = 0$ — even if D is an effective divisor.

Note that, in general, $\mathcal{L}^S(0) \neq k$: the argument used in proposition 5.9.29 on page 297 no longer works because we limit our focus to prime divisors in S.

We can also prove relative versions of proposition 5.9.29 on page 297:

PROPOSITION 6.1.11. If $D_1 \leq D_2$, then $\mathcal{L}^S(D_1) \subseteq \mathcal{L}^S(D_2)$. If S is finite, then

$$\dim\left(\mathcal{L}^S(D_2)/\mathcal{L}^S(D_1)\right) = \deg_S(D_2 - D_1)$$

REMARK. Note that this is a "sharper" statement than the corresponding one in proposition 5.9.29 on page 297: we have equality rather than an inequality.

PROOF. We prove this like the way we proved statement 2 in proposition 5.9.29 on page 297. We set $D_2 = D_1 + P$ and define a homomorphism

$$\varphi\colon \mathcal{L}^S(D_1 + P) \to k$$

whose kernel is $\mathcal{L}^S(D_1)$. We can use proposition 6.1.9 on the preceding page to show that this map is always *surjective:* we can always find a meromorphic function that has prescribed orders at the prime divisors in S. $\square$

> Max Noether (1844 – 1921) was a German mathematician who made contributions to algebraic geometry and the theory of algebraic functions. He was also the father of the more famous Emmy Noether.

One of his Max Noether's notable achievements was the oddly-named "AF+BG Theorem:"

DEFINITION 6.1.12. Let $p = \mathbb{A}_0^2 \in k\mathbb{P}^2$ be a point and let $V = \mathcal{P}(F) \subset k\mathbb{P}^2$ and $W = \mathcal{P}(G) \subset k\mathbb{P}^2$ be projective plane curves that have no common component with $p \in V \cap W$. If $X = \mathcal{P}(H) \subset k\mathbb{P}^2$ is another curve, we say that *Noether's Conditions are satisfied at p* (with respect to F, G, and H) if

$$H(x,y,1) \in (F(x,y,1), G(x,y,1)) \subset \mathcal{O}_{k\mathbb{P}^2,p} = \mathcal{O}_{\mathbb{A}^2,p}$$

REMARK. If $p \notin \mathbb{A}_0^2$, we can substitute any of the other open affines (and make slight changes to the other statements).

For instance, if the multiplicity of the intersection $p \in V \cap W$ is 1, then lemma 3.3.34 on page 140 implies that Noether's conditions are satisfied at p.

Another case where Noether's conditions are satisfied is given by:

PROPOSITION 6.1.13. *Let $V = \mathcal{P}(F) \subset k\mathbb{P}^2$ be an irreducible smooth projective plane curve let $p \in V$ be a point, and suppose $W = \mathcal{P}(G) \subset k\mathbb{P}^2$ and $X = \mathcal{P}(H) \subset k\mathbb{P}^2$ are plane curves. Noether's conditions are satisfied at the point p if*

$$\mathrm{ord}_p(H) \geq \mathrm{ord}_p(G)$$

PROOF. We need to show that $\mathrm{im}\, H \in \mathrm{im}\,(F,G) \subset \mathcal{O}_{\mathbb{A}^2_0,p}$. This is equivalent to $\mathrm{im}\, H \in (G) \subset \mathcal{O}_{V,p}$. Since $\mathcal{O}_{V,p}$ is a discrete valuation ring (see section A.4.3 on page 431) this is completely determined by the condition on valuations given above. The conclusion follows. $\qquad\square$

THEOREM 6.1.14 (Max Noether's AF+BG theorem). *Let $V = \mathcal{P}(F) \subset k\mathbb{P}^2$ and $W = \mathcal{P}(G) \subset k\mathbb{P}^2$ be projective plane curves that have no common component and let $X = \mathcal{P}(H) \subset k\mathbb{P}^2$ be another curve with $\deg H > \deg F$ and $\deg H > \deg G$. Then there exist forms $A, B \in k[X, Y, Z]$ with*

$$\begin{aligned}
\deg A &= \deg H - \deg F \\
\deg B &= \deg H - \deg G
\end{aligned}$$

such that

$$H = AF + BG$$

if and only if Noether's conditions are satisfied at all points in $V \cap W$.

REMARK. This is similar to Bézout's lemma, A.1.12 on page 345. The special features of the present result are that *local* properties imply a *global* one.

PROOF. If $H = AF + BG$ then the same is true at all of the localizations of $k[X, Y, Z]$ so that Noether's conditions will be satisfied at all points of $k\mathbb{P}^2$.

Conversely, suppose Noether's conditions are satisfied at all points of $V \cap W$. Without loss of generality, assume $V \cap W = \{p_1, \ldots, p_t\} \in \mathbb{A}^2_0 \subset k\mathbb{P}^2$. If

$$R = \frac{k[X, Y]}{(F(X, Y, 1), G(X, Y, 1))}$$

then $\mathrm{Spec}\, R$ consists of a finite number of points and R is an *Artinian* ring (see section A.1.7 on page 381). The solution to exercise 3 on page 177 implies that

$$R \cong \prod_{j=1}^{t} \mathcal{O}_{\mathbb{A}^2, p_j}$$

so that Noether's conditions being satisfied at all of the points p_j implies that

$$\mathrm{im}\, H(X, Y, 1) = 0 \in R$$

or

$$(H(X, Y, 1)) \in (F(X, Y, 1), G(X, Y, 1)) \subset k[X, Y]$$

In other words, there exist polynomials, $a(X, Y)$ and $b(X, Y)$ such that

$$H(X, Y, 1) = a(X, Y) \cdot F(X, Y, 1) + b(X, Y) \cdot G(X, Y, 1)$$

Setting $X = \bar{X}/Z$ and $Y = \bar{Y}/Z$ gives

$$H(\bar{X}, \bar{Y}, Z) = Z^r a(\bar{X}/Z, \bar{Y}/Z) \cdot F(\bar{X}, \bar{Y}, Z) + Z^r b(\bar{X}/Z, \bar{Y}/Z) \cdot G(\bar{X}, \bar{Y}, Z)$$

in $k[\bar{X}, \bar{Y}, Z]$, where $r = \deg H$. The conclusion follows. $\qquad\square$

DEFINITION 6.1.15. Suppose $V \subset k\mathbb{P}^n$ is a variety and $W = \mathcal{P}(F) \subset k\mathbb{P}^n$ is a hypersurface sharing no components with V. If $V \cap W = \{Z_1, \dots, Z_t\}$, the *intersection-cycle* is the expression

$$V \bullet W = \sum_{i=1}^{t} \mu_i \langle Z_i \rangle \in \mathrm{Div}(V)$$

where $\mu_i = \mathrm{ord}_{Z_i}(F)$.

REMARK. Note that this divisor describes the intersection of V and W. It is also the divisor of the function F on V because the intersections of W with V occur at precisely the points where F vanishes on V.

Bézout's Theorem implies that

$$\deg(V \bullet W) = (\deg V) \cdot (\deg W)$$

where degrees on the *right* are as defined in 5.7.12 on page 272. Some other properties of these intersection cycles are given by:

PROPOSITION 6.1.16. *If $V = \mathcal{P}(F) \subset k\mathbb{P}^2$, $W = \mathcal{P}(G) \subset k\mathbb{P}^2$, $X = \mathcal{P}(H) \subset k\mathbb{P}^2$ share no components then*

(1) $V \bullet \mathcal{P}(GH) = V \bullet \mathcal{P}(G) + V \bullet \mathcal{P}(H) = V \bullet W + V \bullet X$
(2) $V \bullet \mathcal{P}(G + AF) = V \bullet \mathcal{P}(G) = V \bullet W$

REMARK. The first property makes intuitive sense because $\mathcal{P}(GH) = \mathcal{P}(G) \cup \mathcal{P}(H)$.

PROOF. The first property follows from the definition of order as a *discrete valuation* (see definition A.4.31 on page 431): $\mathrm{ord}_p(GH) = \mathrm{ord}_p(G) + \mathrm{ord}_p(H)$. The second follows from the fact that:

(1) On $V = \mathcal{P}(F)$, $G + AF \equiv G$ so all of the points of $V \cap W$ are points of $V \cap \mathcal{P}(G + AF)$,
(2) $\mathrm{ord}_p(G + AF) = \mathrm{ord}_p(G)$ since order is computed in a localized coordinate ring of V and AF vanishes identically there.

$\square$

COROLLARY 6.1.17. *Let $V = \mathcal{P}(F) \subset k\mathbb{P}^2$ and $W = \mathcal{P}(G) \subset k\mathbb{P}^2$ be curves with no common component and which intersect in $(\deg V) \cdot (\deg W)$ distinct points $\{p_i\}$. If $X = \mathcal{P}(H) \subset k\mathbb{P}^2$ is another curve with $\deg H > \deg F$ and $\deg H > \deg G$ passing through all of the p_i, then there exists a curve $U = \mathcal{P}(B)$ such that*

$$V \bullet U = V \bullet X - V \bullet W$$

REMARK. This is one of the main applications of the AF+BG theorem. It allows one to perform subtle geometric manipulations — creating a curve that intersects V only in points where W does *not*.

PROOF. Bézout's theorem implies that the intersections of V and W, counted with multiplicities, is $(\deg V) \cdot (\deg W)$. It follows that all of the intersections must be of multiplicity 1, which means that Noether's conditions are satisfied (see the remark following definition 6.1.12 on page 309. Noether's AF+BG theorem (6.1.14 on the facing page) implies that

$$H = AF + BG$$

Proposition 6.1.16 on the previous page implies that

$$
\begin{aligned}
V \bullet X &= V \bullet \mathcal{P}(AF + GB) \\
&= V \bullet \mathcal{P}(GB) \\
&= V \bullet W + V \bullet \mathcal{P}(B)
\end{aligned}
$$

and the conclusion follows.   □

Another application of the AF+BG theorem is:

COROLLARY 6.1.18. *Let V be a smooth curve and let D, D' be effective divisors of V with $D \equiv D' \in \mathrm{Cl}(V)$. If $G \in k[X, Y, Z]$ is a form defining $W = \mathcal{P}(G) \subset k\mathbb{P}^2$ with $(G) = D + A$ and $A \geq 0$, then there exists a form G' such that $(G') = D' + A$.*

REMARK. Since $D - D' = (f)$ for some rational function $f \in k(V)$, this just says that we can modify a smooth curve, W, by an arbitrary rational function to get another curve.

PROOF. As remarked above, $D - D' = (f)$, where

$$
f = \frac{H'}{H}
$$

where H and H' are forms of the *same* degree (see proposition 5.9.8 on page 288). We get

$$
D + H = D' + H'
$$

and

$$
(GH) = D' + (H') + A \geq (H')
$$

Proposition 6.1.13 on page 310 implies that Noether's conditions are satisfied for $V, \mathcal{P}(H')$ and $\mathcal{P}(GH)$. The AF+BG theorem gives an expression

$$
GH = \bar{A}F + G'H'
$$

where the degree of G' is the same as that of G and

$$
\begin{aligned}
(GH) &= (G'H') \\
(G) + (H) &= (G') + (H')
\end{aligned}
$$

because AF vanishes identically on V. We conclude that

$$
(G') = D' + A
$$

as claimed.   □

EXERCISES.

1. Use corollary 6.1.17 on page 311 to prove the statement:
If V and W are cubic curves with

$$V \bullet W = \sum_{i=1}^{9} p_i$$

(9 distinct points) and X is a quadratic curve with

$$V \bullet X = \sum_{i=1}^{6} p_i$$

then the points p_7, p_8, p_9 lie on a straight line.

2. Prove Pascal's Mystic Hexagram (see section 1.5 on page 25).

3. Prove a "converse" of Pascal's theorem:
If we have a hexagon and extend opposite sides until they intersect, and these intersections lie on a straight line, then the vertices of the original hexagon were on a *quadratic curve* (i.e. a conic section).

4. What is the genus of an elliptic curve?

5. If a smooth projective plane curve is of genus 1, show that it must be elliptic.

6.2. Elliptic curves

6.2.1. Divisors. Proposition 6.1.2 on page 305 implies that a projective plane curve is of genus 1 if and only if it is defined by a cubic form. In [**123**], Nagell showed (among other things) that every cubic form can be transformed into an equation like the projective closure of the affine curve V defined by

(6.2.1) $$Y^2 = s(X) = X^3 + AX + B \subset k\mathbb{P}^2$$

where $4A^3 + 27B^2 \neq 0$. We assume that the characteristic of the field k is not 2 or 3. See Milne's excellent treatise, [**113**] for the details.

It follows that we can define an elliptic curve to be a smooth plane curve of genus 1.

This section's main result is a theorem (6.2.11 on page 320) of Jacobi and Abel computing the structure of $\mathrm{Cl}^0(E)$, for an elliptic curve, E, and defining a group operation on it. The theory of elliptic curves is a very rich one — Andrew Wiles's proof of Fermat's Last Theorem makes extensive use of them (see [**174**]).

We loosely follow the excellent surveys [**30, 29**]. Since our elliptic curve is not rational, lemma 6.1.8 on page 307 implies that

$$\langle p \rangle - \langle p_0 \rangle \neq 0 \in \mathrm{Cl}^0(V) \subset \mathrm{Cl}(V)$$

(see definition 5.9.9 on page 288) where $p \neq p_0$. If we fix a point $p_0 \in V$, we get an injective map

$$f_{p_0} : V \hookrightarrow \mathrm{Cl}^0(V)$$
$$p \mapsto \langle p \rangle - \langle p_0 \rangle$$

This implies, for instance, that $\mathrm{Cl}^0(V)$ and $\mathrm{Pic}(V)$ are quite large. To study this map and its image, we need to analyze the coordinate ring of V:

PROPOSITION 6.2.1. *Let $k[x,y] = k[V]$ be the coordinate ring of the affine cone of V — where x and y are the images of X and Y, respectively, under the projection*

$$k[X,Y] \to k[X,Y]/(Y^2 - s(X)) = k[V]$$

Then every regular function $f(x,y)$ can be uniquely written as

$$f(x,y) = g(x) + y \cdot h(x)$$

Given such a representation, we define the conjugate of $f(x,y)$ to be

$$\overline{f(x,y)} = g(x) - y \cdot h(x)$$

and the norm of f to be

$$N(f) = f \cdot \bar{f}$$

The degree of a polynomial, $f(x,y)$, is defined to be the degree of its norm, as a polynomial in x.

REMARK. It is not hard to see that the map

$$k[X] \to k[V]$$

is an inclusion. The norm of f can be regarded as a polynomial in x alone because it is equal to

$$g(x)^2 - s(x) \cdot h(x)^2$$

It is easily verified that $N(f_1 \cdot f_2) = N(f_1) \cdot N(f_2)$.

Multiplying a rational function

$$\frac{f(x,y)}{g(x,y)} = \frac{a(x) + y \cdot b(x)}{c(x) + y \cdot d(x)}$$

by $\bar{g}(x,y)/\bar{g}(x,y)$ puts it into the form

$$r_1(x) + y \cdot r_2(x)$$

where r_1 and r_2 are rational functions of x alone.

If $f(x,y) = g(x) + y \cdot h(x)$, then $\deg f = \max(2 \cdot \deg a, 3 + 2 \cdot \deg h)$.

We will regard the point, $(0,1,0)$ at ∞, as the identity element — or the point, p_0 used to define the correspondence between divisors and points in 6.2.11 on page 320.

We will use two coordinate-systems on V: coordinates (x,y) in the finite portion of V and homogeneous coordinates $(x_1 : x_2 : x_3)$ on the whole of it. They are related via

$$x = x_1/x_3$$
$$y = x_2/x_3$$

In homogeneous coordinates, the defining equation of V is

$$x_2^2 x_3 = x_1^3 + A x_1 x_3^2 + B x_3^3$$

DEFINITION 6.2.2. We distinguish three kinds of points on V:
  (1) The point $(0:1:0)$ *at infinity*, denoted $\odot$.
  (2) The points $\eta_i = (\omega_i:0:1)$ where $\omega_1, \omega_2, \omega_3$ are the three solutions of

$$X^3 + AX + B = 0$$

  These are called the *points of degree 2*.
  (3) All of the rest — called *generic* points.

We need to analyze the behavior of rational functions at $\odot$:

PROPOSITION 6.2.3. *If*

$$r = \frac{f(x,y)}{g(x,y)}$$

is a rational function:
  (1) $r(\odot) = 0$ if and only if $\deg g > \deg f$.
  (2) if $\deg f = \deg g$, then $r(\odot)$ is the quotient of the leading terms and is nonzero.

PROOF. Suppose $f = a(x) + y \cdot b(x)$ and $g = c(x) + y \cdot d(x)$. The hypothesis is that

$$\max(2 \cdot \deg a, 3 + 2 \cdot \deg b) < \max(2 \cdot \deg c, 3 + 2 \cdot \deg d)$$

This gives rise to four cases:
  (1) The a and c terms dominate (i.e., determine degrees). Since $3 + 2 \cdot \deg b \le 2 \cdot \deg a$, we conclude that $\deg b < \deg a - 1$ and similar reasoning implies that $\deg d < \deg c - 1$. After substituting $x = x_1/x_3$ and $y = x_2/x_3$, we get

$$r = x_3^{\deg c - \deg a} \frac{x_3^{\deg a} \cdot [a(x_1/x_3) + (x_2/x_3)b(x_1/x_3)]}{x_3^{\deg c} \cdot [c(x_1/x_3) + (x_2/x_3)d(x_1/x_3)]}$$

  where the numerator and denominator are now homogeneous polynomials in x_1, x_2, x_3 (as in 1.2.6 on page 5). It follows that $r(\odot) = 0$, and that this is the only way this can happen when the a and c terms dominate.
  If $\deg f = \deg g$, then the leading term of $x_3^{\deg a} \cdot a(x_1/x_3)$ will be $k_1 \cdot x_1^{\deg a}$ and the leading term of $x_3^{\deg c} \cdot c(x_1/x_3)$ will be $k_2 \cdot x_1^{\deg c}$, so the value of r at $\odot$ will be $k_1/k_2 \ne 0$.
  (2) The b and c terms dominate. In this case $\deg b \ge \deg a - 1$ and $\deg d < \deg c - 1$. We get

$$r = x_3^{\deg c - \deg b - 1} \frac{x_3^{\deg b + 1} \cdot [a(x_1/x_3) + (x_2/x_3)b(x_1/x_3)]}{x_3^{\deg c} \cdot [c(x_1/x_3) + (x_2/x_3)d(x_1/x_3)]}$$

  and the initial exponent of x_3 is positive because $2 \cdot \deg c > 3 + 2 \cdot \deg b$ or $\deg c > \deg b + 1$.

The remaining cases are left as an exercise to the reader. In the case where $\deg f = \deg g$ is *odd* (so b and d dominate), the leading term of the numerator is the leading term of $x_3^{\deg b + 1} \cdot (x_2/x_3) b(x_1/x_3) = x_3^{\deg b} \cdot x_2 \cdot b(x_1/x_3) = k_1 x_2 x_1^{\deg b}$ and that of the denominator will be $k_2 x_2 x_1^{\deg d}$ so the value of r at $\circledcirc$ will, again, be $k_1/k_2 \neq 0$. $\qquad\square$

In order to understand divisors on V we must define local parameters or *uniformizers* at each type of point, as in section 3.3.2 on page 132.

PROPOSITION 6.2.4. *Local parameters on points of V are:*
  (1) *At a generic point $(a\!:\!b\!:\!1)$, the function $x - a$ serves as a local parameter.*
  (2) *At a point of degree 2, the function y is a local parameter.*
  (3) *At the point $\circledcirc$, the function x/y is a parameter. If $p(x,y)$ is a polynomial then*
$$\operatorname{ord}_\circledcirc\left(\frac{1}{p(x,y)}\right) = -\deg p(x,y)$$

PROOF. Let $r(x,y)$ be a rational function that vanishes at (a,b). Then $r(x,y) = f(x,y)/d(x,y)$ where $f(x,y) = g(x) + yh(x)$ vanishes at the point (a,b) and d does not. If $\bar{f}$ also vanishes at this point, then the equations
$$\begin{aligned} g(a) + bh(a) &= 0 \\ g(a) - bh(a) &= 0 \end{aligned}$$
(and the fact that the characteristic of k is $\neq 2$) implies that $g(a) = h(a) = 0$ so $x - a | f(x,y)$. If $\bar{f}(a,b) \neq 0$, we multiply $f(x,y)$ by $\bar{f}(x,y/\bar{f}(x,y))$ to get
$$f(x,y) = \frac{g(x)^2 - s(x)h(x)^2}{\bar{f}}$$
The numerator is a polynomial in x that vanishes at $x - a$ so, *again*, $x - a | f(x,y)$. It follows that $x - a$ generates the ideal of rational functions that vanish at (a,b).

If $r(x,y)$ vanishes at a point, ω_i, of degree 2, and $r(x,y) = f(x,y)/d(x,y)$ then assume $i = 1$. Then
$$N(f)(\omega_1) = g(\omega_1)^2 = 0$$
so $(X - \omega_1) | g(X)$ and $f(x,y)$ satisfies
$$\begin{aligned} f(x,y) &= a(x)(x - \omega_1) + yh(x) \\ &= \frac{a(x)(x - \omega_1)(x - \omega_2)(x - \omega_3) + yh(x)(x - \omega_2)(x - \omega_3)}{(x - \omega_2)(x - \omega_3)} \\ &= \frac{a(x)s(x) + yh(x)(x - \omega_2)(x - \omega_3)}{(x - \omega_2)(x - \omega_3)} \\ &= \frac{a(x)y^2 + yh(x)(x - \omega_2)(x - \omega_3)}{(x - \omega_2)(x - \omega_3)} \\ &= y \cdot \frac{a(x)y + h(x)(x - \omega_2)(x - \omega_3)}{(x - \omega_2)(x - \omega_3)} \end{aligned}$$
so $y | r(x,y)$.

If $r(x_1\!:\!x_2\!:\!x_3)$ vanishes at the point $(0\!:\!1\!:\!0) = \circledcirc \in V$, it is of the form
$$\frac{f(x,y)}{d(x,y)}$$

with $\deg d(x,y) - \deg f(x,y) = k$ as in proposition 6.2.1 on page 314. Since $\deg y = 3$ and $\deg x = 2$,

$$\frac{y^k \cdot f(x,y)}{x^k \cdot g(x,y)}$$

has a nonzero value at $\circledcirc$ and

$$r(x,y) = \left(\frac{x}{y}\right)^k \cdot \frac{y^k \cdot f(x,y)}{x^k \cdot g(x,y)}$$

$\square$

Given these local parameters, we can compute *divisors* of functions:

EXAMPLE 6.2.5. Consider the function $x - a$, where $a \in V$ is a generic point. Its coefficient at $(a, \pm b)$ is clearly 1 and at a finite point whose x-coordinate is not equal to a, it is 0. Since it has positive degree, its *order* at $\circledcirc$ (see definition 5.9.3 on page 286) will be negative — i.e., we must compute

$$\mathrm{ord}_\circledcirc \left(\frac{1}{x-a}\right)$$

Since $x - a$ has degree 2 (via proposition 6.2.1 on page 314), it follows $\mathrm{ord}_\circledcirc (1/(x-a)) = -2$ and

$$(x - a) = \langle (a,b) \rangle + \langle (a,-b) \rangle - 2\langle \circledcirc \rangle \in \mathrm{Cl}(V)$$

The function y has order 1 at each of the points of degree 2 and, since y is of degree 3, we get

$$(y) = \langle \eta_1 \rangle + \langle \eta_2 \rangle + \langle \eta_3 \rangle - 3\langle \circledcirc \rangle$$

We will use our local parameters to geometrically describe the group-operation on $\mathrm{Cl}^0(V)$.

DEFINITION 6.2.6. A *line on V* is a linear function

$$\ell(x,y) = c_1 x + c_2 y + c_3$$

in affine coordinates. We say that a line *passes through a point* (x,y) if $\ell(x,y) = 0$.

REMARK. Since an elliptic curve involves a polynomial of degree 3, Bézout's theorem (1.4.7 on page 25 or 5.8.6 on page 279) implies that a line can pass through at most 3 points of the curve.

Computing the *divisor* of a line (as a rational function on V) has interesting consequences. We begin with

PROPOSITION 6.2.7. *If $p_1 = (a,b) \in V$ is any point not of degree 2 and $p_2 = (a,-b)$, then there exists a line ℓ whose divisor is*

$$(\ell) = \langle p_1 \rangle + \langle p_2 \rangle - 2\langle \circledcirc \rangle$$

If $p = \eta_i$ for any i, then there exists a line ℓ with

$$(\ell) = 2\langle p \rangle - 2\langle \circledcirc \rangle$$

PROOF. If we set $\ell = x - a$, the conclusion follows from example 6.2.5 on the preceding page. In the second case, without loss of generality assume that $i = 1$. We have $x - \omega_1$ and $y = 0$, and

$$x - \omega_1 = \frac{(x - \omega_1)(x - \omega_2)(x - \omega_3)}{(x - \omega_2)(x - \omega_3)} = \frac{y^2}{(x - \omega_2)(x - \omega_3)}$$

so $y^2 | x - \omega_1$ and $\mathrm{ord}_{\eta_1}(x - \omega_1) = 2$. In both cases, the order at $\odot$ is -2.   □

When a line is tangent to V at the point of intersection we get

PROPOSITION 6.2.8. *If $p_1 = (a, b) \in V$ is a point that is not of degree 2, and for which*

$$3a^4 + 6a^2 A + 12 aB - A^2 \neq 0$$

there exists another point $p_2 \in V$ and a line $\ell(x, y)$ such that

$$(\ell) = 2\langle p_1 \rangle + \langle p_2 \rangle - 3\langle \odot \rangle$$

If

(6.2.2) $$3a^4 + 6a^2 A + 12 aB - A^2 = 0$$

$$(\ell) = 3\langle p_1 \rangle - 3\langle \odot \rangle$$

PROOF. The general line through p_1 is $m(x - a) + y - b$ and we compute m to make ℓ tangent to V. We differentiate equation 6.2.1 on page 313

$$2y \frac{dy}{dx} = 3x^2 + A$$

or

$$\frac{dy}{dx} = \frac{3a^2 + A}{2b} = m$$

so we get

$$\ell(x, y) = \frac{3a^2 + A}{2b}(x - a) - b + y$$

If we solve for y and plug into equation 6.2.1 on page 313 and simplify, we get

$$P(x) = -x^3 + x^2 \frac{A^2 + 9a^4 + 12a^2 B + 6a^2 A}{4(a^3 + Aa + B)}$$
$$- x \frac{6a^5 + 12a^2 B + 2a^2 A}{4(a^3 + Aa + B)} + \frac{A^2 a^2 - 2a^4 A + a^6 - 8a^3 B}{4(a^3 + Aa + B)}$$

The denominator is nonzero because p_1 is not of degree 2. This is twice-divisible by $(x - a)$ because

$$P(x) = (x - a)^2 \cdot \left(-x + \frac{a^4 - 2a^2 A - 8aB + A^2}{4(a^3 + Aa + B)} \right)$$

which gives the x-coordinate of p_2 — the y-coordinate is computed by setting $\ell(x, y) = 0$. The function $\ell(x, y)$ has an order of 2 at p_1 because

$$\ell(x, y) = \ell(x, y) \frac{\overline{\ell(x, y)}}{\ell(x, y)} = \frac{P(x)}{\ell(x, y)}$$

Equation 6.2.2 on the preceding page is the result of setting

$$a = \frac{a^4 - 2a^2 A - 8aB + A^2}{4(a^3 + Aa + B)}$$

$\square$

If we have three distinct points that sum to zero in the group defined by V, we get:

PROPOSITION 6.2.9. *If $p_1, p_2, p_3 \in V$ are three distinct points with*

$$\langle p_1 \rangle + \langle p_2 \rangle + \langle p_3 \rangle - 3\langle \circledcirc \rangle = 0 \in \mathrm{Cl}^0(V)$$

there exists a line $\ell(x, y)$ with

$$(\ell) = \langle p_1 \rangle + \langle p_2 \rangle + \langle p_3 \rangle - 3\langle \circledcirc \rangle$$

PROOF. The fact that

$$\langle p_1 \rangle + \langle p_2 \rangle + \langle p_3 \rangle - 3\langle \circledcirc \rangle = 0 \in \mathrm{Cl}^0(V)$$

implies that there exists a function, f, such that

$$(f) = \langle p_1 \rangle + \langle p_2 \rangle + \langle p_3 \rangle - 3\langle \circledcirc \rangle$$

and Bézout's theorem implies that this function must be linear. $\square$

Finally, we need to know that *all* elements of $\mathrm{Cl}^0(V)$ are of the right form:

PROPOSITION 6.2.10. *If $x \in \mathrm{Cl}^0(V)$, there exists a unique point, $p \in V$ such that*

$$x = \langle p \rangle - \langle \circledcirc \rangle$$

PROOF. We will begin by showing that the point p *exists.* Suppose

(6.2.3)
$$x = \sum_{i=1}^{k} n_i \langle p_i \rangle - n \langle \circledcirc \rangle$$

with $\sum n_i = n$.

CLAIM. 1. We can assume that all points of degree 2 have a coefficient of 1.

If p_i is of degree 2 and n_i is even, we can use proposition 6.2.8 on the facing page to eliminate it entirely — by adding a suitable multiple of $2\langle p_i \rangle - 2\langle \circledcirc \rangle$ to x. If p_i is of degree 2 and n_i is odd, we can make $n_i = 1$ in the same way. We can assume that all points of degree 2 in equation 6.2.3 have coefficient $+1$.

CLAIM. 2. We can assume that the coefficients of all points in equation 6.2.3 are *positive.*

If a point p_i has a *negative* coefficient, we can use proposition 6.2.8 on the preceding page to add copies of $\langle p_i \rangle + \langle \bar{p}_i \rangle - 2\langle \circledcirc \rangle$ to x (where, if $p_i = (a, b)$ then $\bar{p}_i = (a, -b)$) to eventually make $n_i = 0$.

CLAIM. 3. We can assume that all coefficients in equation 6.2.3 are 1, i.e.,

$$x = \sum_{i=1}^{k} \langle p_i \rangle - k \langle \circledcirc \rangle$$

Suppose that N is the maximum of the positive coefficients in x. If $N > 1$, suppose, without loss of generality, $N = n_1$. Since $n_1 > 1$, we can use proposition 6.2.8 on page 318 to subtract a multiple of $2\langle p_1 \rangle + \langle p_1' \rangle - 2\langle \circledcirc \rangle$ to make the coefficient of $\langle p_1 \rangle \leq 1$. This introduces a term $-\lfloor n_1/2 \rfloor \langle p_1' \rangle$, which we can make positive by adding suitable copies of $\langle p_1' \rangle + \langle \bar{p}_1' \rangle - 2\langle \circledcirc \rangle$. This introduces new elements $\lfloor n_1/2 \rfloor \langle \bar{p}_1' \rangle$, but the maximum of all positive coefficients will have decreased. We can clearly do this until all coefficients are 1.

CLAIM. 4. We can assume $k = 1$ or $x = 0$.

If $k > 2$, we can reduce it by referring to proposition 6.2.9 on the preceding page and subtracting $\langle p_1 \rangle + \langle p_2 \rangle + \langle p_3 \rangle - 3\langle \circledcirc \rangle$ to reduce it.

If $k = 2$ we can still reduce it by subtracting $\langle p_1 \rangle + \langle p_2 \rangle + \langle p_3 \rangle - 3\langle \circledcirc \rangle$ to get

$$x = -\langle p_3 \rangle + \langle \circledcirc \rangle$$

and then make the coefficient of $\langle p_3 \rangle$ positive via the steps outlined above. It follows that

$$x = \langle p \rangle - \langle \circledcirc \rangle$$

If there was another point q with $x = \langle q \rangle - \langle \circledcirc \rangle$, we would have

$$\langle p \rangle - \langle q \rangle = (f)$$

for some rational function, f and lemma 6.1.8 on page 307 would imply that V is rational, which is a contradiction. $\qquad \square$

This allows us to conclude something remarkable about the divisors on an elliptic curve:

THEOREM 6.2.11. *If V is an elliptic curve, then there is a one-to-one correspondence between the elements of* $\mathrm{Cl}^0(V)$ *and the points of V given by*

$$p \in V \leftrightarrow \langle p \rangle - \langle \circledcirc \rangle \in \mathrm{Cl}^0(V)$$

REMARK. This implies that the points of V constitute an *abelian group.* This is one of the first examples of an *abelian variety:*

DEFINITION 6.2.12. An algebraic group that is an irreducible projective variety is called an *abelian variety*.

REMARK. Abelian varieties were first studied by Abel in connection with his work on elliptic functions. There is an extensive theory of abelian varieties — see [**120**] and [**111**].

> Niels Abel, (1802–1829) was a brilliant Norwegian mathematician whose many contributions to algebra include the proof of the nonexistence of a formula for solving degree-5 polynomials and the theory of elliptic functions. Abel's original approach to the preceding theorem was wildly different from the one given here — involving elliptic functions and elliptic integrals (see [**1**]).

Theorem 6.2.11, proposition 5.9.17 on page 291, and the fact that V is smooth imply that V is a *moduli space* for isomorphism classes of *line bundles* over V (compare with the discussion following definition 5.2.10 on page 230).
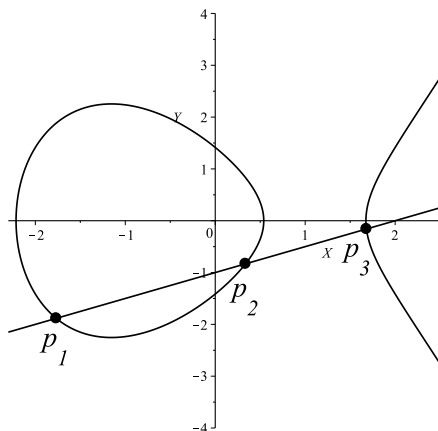
FIGURE 6.2.1.   Three points that sum to zero

Propositions 6.2.7 on page 317, 6.2.9 on page 319, and 6.2.9 on page 319 define the group-structure of V and its relation to geometry.

COROLLARY 6.2.13. *If $p_1, p_2, p_2 \in V$ are three points of V, then*

$$p_1 + p_2 + p_3 = 0$$

if and only if

$$(\langle p_1 \rangle - \langle \circledcirc \rangle) + (\langle p_2 \rangle - \langle \circledcirc \rangle) + (\langle p_3 \rangle - \langle \circledcirc \rangle) = 0 \in \mathrm{Cl}^0(V)$$

which happens if and only if p_1, p_2, p_3 lie on the same line, as in figure 6.2.1. The additive inverse of (a, b) is $(-a, b)$.

REMARK. Proposition 6.2.8 on page 318 considers degenerate cases where two of the points coincide (and the line is tangent to V) or where all three points coincide at *inflection* points of V.

It is easy to check that this group-law makes V into an algebraic (abelian) group whenever k is a field — we never required k to be algebraically closed.

When $k = \mathbb{Q}$, Mordell proved that V is finitely generated in [**119**], answering a question posed by Poincaré in 1908. In 1929, André Weil generalized this by proving that the group of *any* abelian variety over a number-field (i.e., a finite extension of $\mathbb{Q}$) is finitely generated, in [**171**]. The general result is called the Mordell-Weil theorem.

EXERCISES.

1. Show that the condition $4A^3 + 27B^2 \neq 0$ guarantees that the elliptic curve defined by $Y^2 = X^3 + AX + B$ is *smooth*.

2. Explicitly write $x - 1$ (assuming 1 is not a root of the polynomial $X^3 + aX + b$) and y in terms of the local parameter x/y at $\circledcirc$.

3. Why doesn't proposition 6.2.7 on page 317 violate Bézout's theorem? After all, it involves a $x - a$ that defines a line intersecting V in only *two* points.

4. Given points, $(a_1, b_1), (a_2, b_2) \in V$ find an explicit formula for $(a_1, b_1) + (a_2, b_2) = (a_3, b_3)$.

5. Use lemma 5.5.19 on page 258 to show that the group-operation of an abelian variety is commutative (making it "abelian" in two different senses!).

6.2.2. Elliptic curve cryptography. Elliptic curve cryptography was independently discovered by Neal Koblitz (see [**90**]) and Victor Miller (in [**110**]) in 1985. It is a cryptographic system that uses elliptic curves over a finite field — usually $\mathbb{F}_p$ or $\mathbb{F}_{2^n}$ and is based on the computational difficulty of deducing $n \in \mathbb{Z}$, given $v \in V$ and

$$(6.2.4) \qquad\qquad n \cdot v = \sum_{i=1}^{n} v \in V$$

for some very large value of n (the so-called *discrete logarithm problem*). Although other discrete-logarithm systems are in use, ones based on elliptic curves appear to offer greater security for a given value of n.

Suppose Joe Blow and Sally Slow want to communicate confidentially across a channel that is easily monitored by others. The standard way to do this is with an agreed-upon cryptosystem — involving well-known algorithms. Secrecy is provided by coupling this system with a *key* known only to Joe and Sally:

$$\text{Message} + \text{Key} \to \text{Encrypted message}$$

They could agree on a key beforehand, but this has the problem that it might fall into the wrong hands or be deduced by a good cryptologist (after analyzing many messages[1]). The most secure approach (therefore the one used in practice) is to randomly generate a new key for each message. This raises the obvious question: how do Joe and Sally exchange this key without it falling into the wrong hands?

This is where elliptic curves come into play. Joe and Sally have an agreed-upon elliptic curve, V, over $\mathbb{F}_p$ for some large (i.e. 100 digits or more) prime p or over $\mathbb{F}_{2^n}$ for a large value of n, given by

$$Y^2 = X^3 + AX + B$$

and some fixed point $v \in V$. So they initially agree upon p, A, B, v (or $2^n, A, B, v$) — and must assume that their adversaries know these parameters. Suppose the order of the group, V, is N — we assume this is a very large number.

Joe picks a random large number, n_{Joe}, from 1 to $N - 1$ and publicly sends Sally $K_{\text{Joe}} = n_{\text{Joe}} \cdot v$ — where "multiplication" is done as in equation 6.2.4. Similarly, Sally picks a random n_{Sally} and sends Joe $K_{\text{Sally}} = n_{\text{Sally}} \cdot v$.

When Sally receives K_{Joe}, she multiplies it by n_{Sally} to get

$$x = n_{\text{Sally}} \cdot K_{\text{Joe}} = n_{\text{Sally}} n_{\text{Joe}} \cdot v$$

Joe does a similar thing with K_{Sally} to get

$$x = n_{\text{Joe}} \cdot K_{\text{Sally}} = n_{\text{Sally}} n_{\text{Joe}} \cdot v$$

---

[1]The book [**161**] gives a vivid account of this happening in the Second World War.

At the end of this exchange, Joe and Sally share a secret, x, that is unknown to the world at large — even though the defining parameters were publicly broadcast. They can pass secret messages to each other, using x as the key.

This subject has motivated a great deal of research into elliptic curves — for instance, the two monographs [**30**] and [**29**] were sponsored by the US Institute for Defense Analysis. There are published guidelines for "good" values for the parameters p, A, B, v, and ones to avoid — see [**28**]. There are also efficient algorithms for computing $n \cdot v$ for large values of n — see the exercises. For more on this subject, see [**14**].

EXERCISES.

6. If $n = 2^k$ and $x \in V$ is a point in an elliptic curve, describe a fast procedure for computing $n \cdot x$.

7. How can the procedure in the previous problem be used to compute $n \cdot x$ for an arbitrary natural number n.

6.3. The Riemann-Roch Theorem

6.3.1. Introduction. The material in this section is true for arbitrary smooth curves — not just projective plane curves.

Georg Friedrich Bernhard Riemann (1826–1866) was an influential German mathematician who made contributions to many fields including analysis, number theory and differential geometry. Riemann's work in differential geometry provided the mathematical foundation for Einstein's Theory of General Relativity (see [**116**]).

The Riemann-Roch theorem has a long and complex history. In its original form, it was a theorem about meromorphic functions on a Riemann surface. Riemann proved a limited form of it known as Riemann's Inequality in [**142**] and Roch expanded on this to create an *equation* in [**143**].

In 1931, Friedrich Karl Schmidt proved a version for algebraic curves in [**148**] over finite fields and applied this to number-theoretic problems. In [**80**], Hirzebruch extended the Riemann-Roch theorem still further, developing a topological version that led to many other powerful theorems including the Atiyah-Singer Index theorem (see [**134**]).

In [**47**], Gerd Faltings proved a version of the Riemann-Roch theorem (and Noether's AF+BG theorem) for schemes over algebraic number fields (*finite* extensions of $\mathbb{Q}$).

Today the phrase "Riemann-Roch" has evolved to become a generic term for a vast array of results.

6.3.2. Riemann's Inequality. In this section we will consider the original result Riemann proved in 1857 (in [**142**]):

THEOREM 6.3.1 (Riemann's Inequality). *If V is an irreducible smooth projective curve, there exists an integer $\xi(V)$ such that for any divisor, D*

$$\ell(D) \geq \deg D + 1 - \xi(V)$$

REMARK. This number, $\xi(V)$, will turn out to be the *genus* of V (see definition 5.9.40 on page 303).

Riemann originally proved this inequality for meromorphic functions on a Riemann surface — in which case, g was the *topological* genus. Such a surface can be "flattened out" by $2g$ "cuts" to form a polygon in $\mathbb{C}$ with $4g$ sides. Riemann reasoned that a meromorphic function on the surface is equivalent to one on the $4g$-sided polygon whose values *agree* across the cuts — i.e., ones that satisfy $2g$ conditions.

Using a construction called the Dirichlet Principle, Riemann was able to construct g basic linearly independent regular functions and add functions of the form $1/(z-a)$ to give d singularities. He reasoned that the resulting function is specified by $g + d + 1$ constants. Since it must satisfy $2g$ constraints (when we glue the cuts together again), we get at least $g + d + 1 - 2g = d + 1 - g$ degrees of freedom.

Riemann's use of the Dirichlet Principle was refuted by his student Emil Prym in 1870, and it cast doubt on Riemann's proof of the inequality and Roch's later enhancements to it. The first rigorous proof of the Riemann-Roch theorem (and the first time it was called that) is due to Alexander Brill and Max Noether in 1870 ([**19**]).

We follow the elementary and very concise treatment of the subject in Fulton's classic, [**49**].

We begin by proving it for a special type of divisor:

LEMMA 6.3.2. *Suppose V is a smooth projective curve with field of meromorphic functions $k(V)$, suppose $x \in k(V)$ and $x \notin k$. If $n = [k(V):k(x)]$ and $D = (x)_0$ is the divisor of the zeros of x, then*

(1) *D is an effective divisor and $\deg D = n$.*
(2) *There exists a number τ such that $\ell(r \cdot D) \geq rn - \tau$ for all $r \in \mathbb{Z}$.*

PROOF. Effectiveness of D follows from its definition. If $m = \deg D$, we begin by showing that $m \leq n$. If

$$D = \sum_{i=1}^{t} n_{p_i} \cdot \langle p_i \rangle$$

Let $S = \{ \langle p \rangle \in \mathrm{Div}(V) | n_p > 0 \}$. Proposition 6.1.11 on page 309 implies that

$$\dim \mathcal{L}^S(0) / \mathcal{L}^S(-D) = m$$

Let $v_i \in \mathcal{L}^S(0)$, $i = 1, \ldots, m$ be elements that map to a basis $\{\bar{v}_1, \ldots, \bar{v}_m\}$ of $\mathcal{L}^S(0)/\mathcal{L}^S(-D)$.

We claim that the v_i are linearly independent over $k(x)$. If not, there is an equation

$$\sum_{i=1}^{m} f_i v_i = 0$$

with not all $f_i \in k(x)$ equal to zero. After clearing denominators, we get an equation

$$(6.3.1) \qquad\qquad \sum_{i=1}^{m} g_i v_i = 0$$

where $g_i \in k[x]$. If we write $g_i = \lambda_i + x h_i$ where $\lambda_i \in k$ and $h_i \in k[x]$, not all of the λ_i are 0 (since we could multiply by x^{-n} for suitable n to ensure this). We get

$$\sum_{i=1}^{m} \lambda_i v_i = -x \sum_{i=1}^{m} h_i v_i \in \mathcal{L}^S(-D)$$

so the image of $\sum_{i=1}^{m} g_i v_i$ in $\mathcal{L}^S(0)/\mathcal{L}^S(-D)$ is $\sum_{i=1}^{m} \lambda_i \bar{v}_i = 0$, which contradicts the assumption that $\bar{v}_i$ are a basis.

To prove the second statement (and complete the proof of the first), we will explicitly construct linearly independent elements of $\mathcal{L}(r \cdot D)$. Let $\{w_1, \dots, w_n\} \in k(V)$ be a basis for $k(V)$ over $k(x)$ — and regard $k(X)$ as the field of fractions of $k[x^{-1}]$. Proposition A.2.12 on page 390 implies that each of the w_i satisfies a polynomial

$$w_i^{n_i} + a_{i,n_i-1} w_i^{n_i-1} + \cdots + a_{i,0} = 0$$

with $a_{i,j} \in k(x)$ and the solution to exercise 2 on page 427 implies that, without loss of generality, we can assume that the $a_{i,j}$ lie in $k[x^{-1}]$. This implies that $\text{ord}_P\left(a_{i,j}\right) \geq 0$ for all prime divisors P such that $P \notin S$. We will use the w_i to construct elements of $\mathcal{L}(r \cdot D)$.

Claim: If $P \notin S$, then $\text{ord}_P\left(w_i\right) \geq 0$, so the elements we construct will be effective at prime divisors not in S.

If $\text{ord}_P\left(w_i\right) < 0$, we would have $b = \text{ord}_P\left(w_i^{n_i}\right) < \text{ord}_P\left(a_{i,j} w_i^j\right)$ which, by statement 4 of proposition A.4.32 on page 431 would imply that

$$\text{ord}_P\left(w_i^{n_i} + a_{i,n_i-1} w_i^{n_i-1} + \cdots + a_{i,0}\right) = b$$

a *contradiction*.

It follows that, for some integer $t > 0$, $\left(w_i\right) + t \cdot D \geq 0$ for $i = 1, \dots, n$. If $j = 0, \dots, r$, then

$$w_i x^{-j} \in \mathcal{L}((r+t)D)$$

Since the w_i are linearly independent and the $\{x^{-j}\}$ are linearly independent, it follows that the set

$$\{w_i x^{-j}\}$$

is linearly independent (see proposition A.2.7 on page 388). It follows that $\ell((r+t)D) \geq n(r+1)$. But

$$\ell((r+t) \cdot D) = \ell(r \cdot D) + \dim\left(\frac{\mathcal{L}((r+t) \cdot D)}{\mathcal{L}(r \cdot D)}\right) \leq \ell(r \cdot D) + t \cdot m$$

by statement 2 of proposition 5.9.29 on page 297. We conclude that

$$\ell(r \cdot D) \geq n(r+1) - t \cdot m = r \cdot n - \tau$$

Statement 4 of proposition 5.9.29 on page 297 implies that

$$\mathcal{L}(r \cdot D) \leq r \cdot m + 1$$

so we get the inequality $r \cdot n - \tau \leq r \cdot m + 1$. Taking the limit as $r \to \infty$ shows that $n \leq m$, which completes the proof of the first statement. $\qquad \square$

Although divisors of the form $D = (x)_0$ are specialized, we can show that arbitrary divisors are related to them:

PROPOSITION 6.3.3. *Let V be a smooth projective curve and let $B \in \mathrm{Div}(V)$ be an arbitrary divisor. If $x \in k(V) \setminus k$ is an arbitrary element such that $[k(V) : k(x)] < \infty$ and $D = (x)_0$, then there exists a divisor $B' \equiv B \in \mathrm{Cl}(V)$ such that*

$$B' \leq r \cdot D$$

for all sufficiently large r.

PROOF. If $D = \sum n_p \langle p \rangle$ and $B = \sum m_p \langle p \rangle$, we will construct a meromorphic function f such that $B' = B - (f)$ and $m_p - \mathrm{ord}_p(f) \leq r \cdot n_p$ for all $p \in V$. Let $y = x^{-1}$ and set $T = \{ p \in V \mid m_p > 0, \mathrm{ord}_p(y) \geq 0 \}$. Set

$$f = \prod_{p \in T} (y - y(p))^{m_p}$$

If $p \in T$ (i.e. $\mathrm{ord}_p(y) \geq 0$), then $m_p - \mathrm{ord}_p(f) \leq 0$, so the conclusion holds at these points. If $\mathrm{ord}_p(y) < 0$, then the corresponding $n_p > 0$ so a sufficiently large value of r will ensure the conclusion. $\qquad \square$

Proof of theorem 6.3.1 on page 324:

We will show that there exists some integer ξ such that, for any divisor, D, on V

$$\ell(D) \geq \deg d + 1 - \xi$$

and later (when we prove Roch's enhancement of the theorem) show that ξ is the genus of V.

If $s(D) = \deg D + 1 - \ell(D)$, we will show that there is an upper bound for $s(D)$ as D runs over all divisors and:

(1) If $D = 0$, statement 3 of proposition 5.9.29 on page 297 implies that $s(0) = 0$, so $\xi(V) \geq 0$ if it exists.
(2) Propositions 5.9.8 on page 288 and 5.9.28 on page 296 show that if $D_1 \equiv D_2 \in \mathrm{Cl}(V)$ then $s(D_1) = s(D_2)$.
(3) Statement 2 of proposition 5.9.29 on page 297 shows that $D_1 \leq D_2$ implies that $s(D_2) \leq s(D_1)$.
(4) If $x \in k(V) \setminus k$ and $D = (x)_0$ as in lemma 6.3.2 on page 324, then $s(r \cdot D) \leq \tau + 1$ for all r. If τ is the smallest value that works in lemma 6.3.2 on page 324, we can set $\xi(V) = \tau + 1$. Since this must be a value taken on by $s(r \cdot D)$ and since $r \cdot D \leq (r + 1) \cdot D$, it follows that $s(r \cdot D) = \tau + 1 = \xi(V)$ for all sufficiently large values of r.

At this point, proposition 6.3.3 finishes the proof.

6.3.3. The full Riemann-Roch Theorem. In this section, we will state the theorem in what is essentially its modern form.

> Gustav Roch (1839 – 1866), a student of Riemann, was a German mathematician who made important contributions to the theory of Riemann surfaces. What promised to be a brilliant career was cut short by tuberculosis, when Roch was 26.

In [**143**], he enhanced the Riemann inequality to an *equation:*

THEOREM 6.3.4 (Riemann-Roch). *If V is a smooth projective plane curve with canonical class K_V (see definition 5.9.36 on page 300) and D is any divisor, then*

$$(6.3.2) \qquad \ell(D) = \ell(K_V - D) + 1 + \deg D - \xi(V)$$

REMARK. Setting $D = 0$ gives $\xi(V) = \ell(K_V)$, so it is actually the genus in definition 5.9.40 on page 303.

The proof presented here follows that of Alexander Brill and Max Noether in [**19**] — which only works for projective plane curves (since they were studying Riemann surfaces).

Because of statements 2 on the preceding page and 3 on the facing page, we already know this theorem for D "sufficiently large:"

PROPOSITION 6.3.5. *In the notation of theorem 6.3.4 and statement 3 on the facing page, if D is a divisor with the property that*

$$\ell(D) = \deg D + 1 - \xi(V)$$

and $D' \geq D$ and D'' is any divisor with $D'' \equiv D' \in \mathrm{Cl}(V)$, then

$$\ell(D'') = \deg D'' + 1 - \xi(V)$$

Statement 4 on the preceding page implies that:

PROPOSITION 6.3.6. *If V is a smooth projective curve with field of meromorphic functions $k(V)$, and $x \in k(V) \setminus k$ and $D = (x)_0$ is the divisor of the zeros of x, then*

$$\ell(r \cdot (x)_0) = \deg(r \cdot (x)_0) + 1 - \xi(V)$$

for r sufficiently large.

This allows us to refine our statement about a divisor being sufficiently large:

COROLLARY 6.3.7. *Under the hypotheses of theorem 6.3.4, there exists an integer N such that any divisor D with $\deg D \geq N$ satisfies*

$$(6.3.3) \qquad \ell(D) = \deg D + 1 - \xi(V)$$

REMARK. If D is "large" then $K_V - D$ is "small" and $\ell(K_V - D) = 0$. It follows that $\ell(K_V - D)$ is a "correction factor" that only comes into play for "small" divisors.

PROOF. Let $r \cdot (x)_0$ be a divisor that satisfies proposition 6.3.6 and set $N = \deg(r \cdot (x)_0) + g$. If $\deg D \geq N$, then

$$\deg (D - r \cdot (x)_0) \geq \xi(V)$$

and Riemann's inequality (6.3.1 on page 324) implies that

$$\ell(D - r \cdot (x)_0) > 0$$

so there exists a meromorphic function f with

$$D - r \cdot (x)_0 + (f) \geq 0$$

which implies that

$$D + (f) \geq r \cdot (x)_0$$

and the conclusion follows from proposition 6.3.5. □

We can use this result to compute $\xi(V)$:

PROPOSITION 6.3.8. *If $V = \mathcal{P}(F) \subset k\mathbb{P}^2$ is a smooth curve of degree d, then*

$$\xi(V) = \frac{(d-1)(d-2)}{2} = g$$

— *the geometric genus in definition 5.9.40 on page 303.*

REMARK. Henceforth, we will dispense with the term $\xi(V)$ and use the genus.

PROOF. Suppose, without loss of generality, that V intersects the line $T = \mathcal{P}(Z)$ in d distinct points. We will explicitly construct $\ell(n \cdot (Z))$ for large values of n and compute $\xi(V)$ as

$$\xi(V) = \deg(n \cdot (Z)) + 1 - \ell(n \cdot (Z))$$

First of all, $\deg(n \cdot (Z)) = nd$, and if $V_n \subset k[X, Y, Z]$ is the vector space of forms of degree n, we get

$$\dim V_n = \binom{n+2}{2} = \frac{(n+1)(n+2)}{2}$$

— see proposition 5.4.6 on page 244. We can define a map

$$\begin{aligned} (6.3.4) \qquad\qquad p_n \colon V_n &\;\to\; \mathcal{L}(n \cdot (Z)) \\ F &\;\mapsto\; \frac{F}{Z^n} \end{aligned}$$

We claim that this is surjective. If $f = R/S \in \mathcal{L}(n \cdot (Z))$, where R and S are forms of the same degree, then

$$(RZ^n) \geq (S)$$

Proposition 6.1.13 on page 310 implies that Noether's conditions are satisfied, so theorem 6.1.14 on page 310 provides an expression

$$RZ^n = AS + BF$$

so $f = R/S = A/Z^n \in k(V)$. The kernel of p_n in 6.3.4 consists of all multiplies of F. We get an exact sequence

$$0 \to V_{n-d} \xrightarrow{\times F} V_n \xrightarrow{p_n} \mathcal{L}(n \cdot (Z)) \to 0$$

It follows that

$$\begin{aligned} \ell(n \cdot (Z)) &= \binom{n+2}{2} - \binom{n-d+2}{2} \\ &= nd - \frac{d(d-3)}{2} \end{aligned}$$

From which we conclude that

$$\begin{aligned} \xi(V) &= nd + 1 - \ell(n \cdot (Z)) \\ &= 1 + \frac{d(d-3)}{2} \\ &= \frac{(d-1)(d-2)}{2} \end{aligned}$$

$\square$

It follows that we know the Riemann-Roch theorem for divisors of high degrees. To prove it in general, we must "work our way down" to lower degrees. The key to this is

LEMMA 6.3.9 (Max Noether's Reduction Lemma). *Let V be a smooth projective curve with canonical class K_V, let D be a divisor, and let $\langle p \rangle$ is a prime divisor (point). If $\ell(D) > 0$ and $\ell(K_V - D - \langle p \rangle) \neq \ell(K_V - D)$ then $\ell(D + \langle p \rangle) = \ell(D)$.*

REMARK. We can regard this as the main structural property of K_V that makes theorem 6.3.4 on page 327 work.

PROOF. Using proposition 6.1.2 on page 305, set $K_V = (d - 3) \cdot (Z)$ where Z is the third homogeneous coordinate in $k\mathbb{P}^2$. We may assume $\mathcal{L}(K_V - D) \subset \mathcal{L}((d - 3) \cdot (Z))$.

If $v \in \mathcal{L}(K_V - D) \setminus \mathcal{L}(K_V - D - \langle p \rangle)$, it follows that $v \in \mathcal{L}((d - 3) \cdot (Z))$ and
$$(Z^{d-3}v) \geq 0$$
so write
$$v = \frac{G}{Z^{d-3}}$$
where G is a function of degree $d - 3$. We conclude that
$$K_V - D + (v) \geq 0 \qquad \text{which implies that}$$
$$K_V - D + (G) - (Z^{d-3}) \geq 0 \qquad \text{and, since } K_V = (Z^{d-3}), \text{ we get}$$
$$-D + (G) \geq 0$$
We conclude that $(G) \geq D$ or $(G) = D + A$ where
$$(6.3.5) \qquad\qquad A \not\geq \langle p \rangle$$
since $v \notin \mathcal{L}(K_V - D - \langle p \rangle)$.

Let $\mathcal{P}(L) \subset k\mathbb{P}^2$ be a line whose intersection with V consists of p and $d - 1$ other points, all distinct from p, so
$$(L) = \langle p \rangle + B$$
and
$$(L \cdot G) = D + A + \langle p \rangle + B$$
where $L \cdot G$ is a form of degree $d - 2$.

If $f \in \mathcal{L}(D + \langle p \rangle)$, we must show that $f \in \mathcal{L}(D)$, let $D' = D + (f)$ so that $D \equiv D'$. Corollary 6.1.18 on page 312 implies the existence of a smooth curve $\mathcal{P}(H)$ of degree $d - 2$ such that
$$(H) = D' + \langle p \rangle + A + B$$
Then $\mathcal{P}(H)$ contains the $d - 1$ collinear points of B. Bézout's Theorem implies that $\mathcal{P}(H)$ contains $\mathcal{P}(L)$ as a component, so that $H(p) = 0$. Since $\langle p \rangle$ does not appear in $A + B$ (by 6.3.5 and by the construction of L), we get
$$D' + \langle p \rangle \geq \langle p \rangle$$
— i.e. none of the terms in D' *cancel* $\langle p \rangle$. It follows that $D + (f) = D' \geq 0$, so $f \in \mathcal{L}(D)$. $\square$

We are ready to prove theorem 6.3.4 on page 327.

PROOF. We consider several cases:

Case $\mathcal{L}(K_V - D) = 0$**:** In this case, we do induction on $\ell(D)$:

(1) If $\ell(D) = 0$, applying the Riemann inequality to D and $K_V - D$ gives

$$\begin{aligned} g &= \deg D + 1 \\ &= \deg(K_V - D) + 1 \end{aligned}$$

and equation 6.3.2 on page 327 is satisfied.

(2) If $\ell(D) = 1$, $D \geq 0$ and statement 2 of proposition 5.9.29 on page 297 implies that $g = \ell(K_V) \leq \ell(K_V - D) + \deg D$. Proposition 6.1.2 on page 305 so $\deg D \geq g$ and Riemann's inequality implies that

$$1 \geq \deg D + 1 - g$$

or $\deg D \leq g$. It follows that $\deg D = g$ and equation 6.3.3 on page 327 is satisfied.

(3) If $\ell(D) > 1$, suppose the result is known for all *smaller* values of $\ell(D)$ and choose $p \in V$ such that $\ell(D - \langle p \rangle) = \ell(D) - 1$ (see exercise 6 on page 299). The *contrapositive* of Noether's reduction lemma (6.3.9 on the previous page) implies that

$$\ell(K_V - (D - \langle p \rangle)) = \ell((K_V - D) = 0$$

so the inductive hypothesis implies that equation 6.3.3 on page 327 is satisfied.

Case $\ell(K_V - D) > 0$**:** If $\ell(D) = 0$, we can *swap* D and $K_V - D$ (i.e. $K_V - (K_V - D) = D$) and use the case where $\ell(K_V - D) = 0$ to prove the result. Consequently, we assume $\ell(D) > 0$. Statement 2 of proposition 5.9.29 on page 297 implies that

$$\deg D \leq \deg K_V = 2g - 2$$

— see proposition 6.1.2 on page 305. If equation 6.3.2 on page 327 is false, let D be chosen to be a divisor with *maximal* $\ell(D)$ that violates it. This exists, because if $\ell(D)$ is large enough, $\ell(K_V - D) = 0$ and we fall back to the previous case. So D has the property that equation 6.3.2 on page 327 is true for $D + \langle p \rangle$ for any $p \in V$. Select point $p \in V$ with the property that

$$\ell(K_V - D - \langle p \rangle) = \ell(K_V - D) - 1$$

(see exercise 6 on page 299). The Noether Reduction theorem (6.3.9 on the preceding page) implies that $\ell(D + \langle p \rangle) = \ell(D)$. Since equation 6.3.2 on page 327 is true for $D + \langle p \rangle$ we have

$$\begin{aligned} \ell(D) = \ell(D + \langle p \rangle) &= \deg(D + \langle p \rangle) + 1 + \ell(K_V - D - \langle p \rangle) - g \\ &= \deg D + 1 + \ell(K_V - D) - g \end{aligned}$$

and the result is proved.

$\square$

6.4. The modern approach to Riemann-Roch

In this section, we will prove the Riemann-Roch theorem by more modern methods, using sheaf cohomology (see appendix D on page 519). This approach is valid for arbitrary 1-dimensional projective varieties (i.e. they do not have to be projective plane curves) — indeed, it generalizes to higher-dimensional varieties and leads to a vast array of results.

The modern approach, developed by Serre in [**149**], is to surround the problem with the heavy machinery of Serre Duality — and *gently squeeze*.

THEOREM 6.4.1 (Riemann-Roch). *If V is a smooth projective curve with canonical class K_V (see definition 5.9.36 on page 300) and D is any divisor, then*

$$(6.4.1) \qquad \ell(D) = \ell(K_V - D) + 1 + \deg D - \xi(V)$$

REMARK. This statement is essentially identical to that of theorem 6.3.4 on page 327 except that we have omitted the word "plane". It turns out that our modern sheaf-cohomology tools do not require V to be embedded in $k\mathbb{P}^2$.

PROOF. We rewrite this as

$$\ell(D) - \ell(K_V - D) = 1 + \deg D - \xi(V)$$

where $\ell(D) = \dim_k H^0(V, \mathbf{s}(D))$ where $\mathbf{s}(D)$ is the invertible sheaf corresponding to D in proposition 5.9.20 on page 292 (or line-bundle — see remark C.2.2 on page 514) — and $\ell(K_V - D) = \dim_k H^0(V, \omega_V \otimes_{\mathcal{O}_V} \mathbf{s}(D)^\vee)$, where ω_V is the canonical sheaf of V (see definition 4.6.24 on page 212).

Serre Duality (corollary D.4.20 on page 573) implies that

$$H^0(V, \omega_V \otimes_{\mathcal{O}_V} \mathbf{s}(D)^\vee) \cong H^1(V, \mathbf{s}(D))^*$$

so that $\dim_k H^0(V, \omega_V \otimes_{\mathcal{O}_V} \mathbf{s}(D)^\vee) = \dim_k H^1(V, \mathbf{s}(D))$ and our formula becomes

$$\begin{aligned} \dim_k H^0(V, \mathbf{s}(D)) - \dim_k H^1(V, \mathbf{s}(D)) &= 1 + \deg D - \xi(V) \\ &= \chi(V, \mathbf{s}(D)) \end{aligned}$$

where $\chi(V, \mathbf{s}(D))$ is the Euler Characteristic in definition D.3.23 on page 559.

If we set $D = 0$, we get $1 - \ell(K_V) = 1 - \xi(V)$ so the formula is true, since we have defined $\xi(V) = \ell(K_V)$. We will show that, if it is true for a divisor D and $p \in V$ is a point, then it must be true for $D + \langle p \rangle$. A simple induction implies the full result.

If $p \in V$ is a point, its structure sheaf, $\mathcal{O}_p$, is the skyscraper sheaf (see example B.1.7 on page 498), k_p, at the point p. If f is a function defining the Cartier divisor of $\langle p \rangle$, then the defining ideal, $\mathscr{I} \subset \mathcal{O}_V$, of p is given by $\mathbf{s}(-\langle p \rangle)$ — see 5.9.21 on page 292 — and we get an exact sequence

$$0 \to \mathscr{I} = \mathbf{s}(-\langle p \rangle) \to \mathcal{O}_V \to k_p \to 0$$

If we take the tensor product of this with $\mathbf{s}(D + \langle p \rangle)$, we get

$$0 \to \mathbf{s}(D) \to \mathbf{s}(D + \langle p \rangle) \to k_p \to 0$$

The sequence remains exact since $\mathbf{s}(D + \langle p \rangle)$ is a locally free sheaf of rank 1. and $\mathbf{s}(D + \langle p \rangle) \otimes_{\mathcal{O}_V} k_p = k_p$. Proposition D.3.24 on page 560 implies that

$$\chi(V, \mathbf{s}(D + \langle p \rangle)) = \chi(V, \mathbf{s}(D)) + \chi(V, k_p)$$

and $H^0(V, k_p) = k$, and $H^1(V, k_p) = H^1(p, k) = 0$ by lemma D.3.18 on page 554 so

$$\chi(V, \mathbf{s}(D + \langle p \rangle)) = \chi(V, \mathbf{s}(D)) + 1$$

Since $\deg(D + \langle p \rangle) = \deg D + 1$ the Riemann-Roch formula remains true. $\square$

We get a number of applications of the Riemann-Roch theorem.

COROLLARY 6.4.2. *If V is a smooth projective curve of genus g and D is a divisor with $\deg D \geq 2g - 1$, then*

$$\ell(D) = \deg D + 1 - g$$

PROOF. If $\deg D \geq 2g - 1$ then $\deg(K_V - D) \leq 2g - 1 - \deg D < 0$ by propositions 6.1.2 on page 305 and 5.9.29 on page 297. Corollary 5.9.32 on page 298 implies that $\ell(K_V - D) = 0$. $\square$

This immediately leads to a kind of converse to corollary 5.9.32 on page 298 and allows us to classify curves of low genus:

COROLLARY 6.4.3. *If V is a smooth curve of genus 0, then every divisor of degree 0 is of the form (f) for $f \in k(V)$. Consequently, V is isomorphic to $k\mathbb{P}^1$.*

PROOF. If D is a divisor of degree 0, corollary 6.4.2 implies that $\ell(D) = 1$. Suppose $p_1 p_2 \in V$ are two distinct points. Then the divisor $\langle p_1 \rangle - \langle p_2 \rangle$ is of degree 0, hence $\langle p_1 \rangle - \langle p_2 \rangle = (f)$ for some $f \in k(V)$. The conclusion follows from lemma 6.1.8 on page 307. $\square$

We can also conclude:

COROLLARY 6.4.4. *If V is a smooth one-dimensional variety of genus 1, then V is isomorphic to a cubic curve embedded in $k\mathbb{P}^2$.*

REMARK. It follows that a smooth one-dimensional variety of genus 1 can be defined by a single equation.

PROOF. The genus being 1 implies that, for any divisor D of degree > 0, $\ell(D) = \deg D$ (see corollary 6.4.2). If $x \in V$ is any point, then we conclude that

$$\ell(3\langle x \rangle) = 3$$

If $f_0, f_1, f_2 \in k(V)$ are a basis for $\mathcal{L}(3\langle x \rangle)$, then we get a rational map

$$F = (f_0 : f_1 : f_2) : V \to k\mathbb{P}^2$$

Since V is smooth, proposition 6.1.3 on page 306 implies that this map is regular, i.e., an embedding. If its image is W, then the divisor (f_i) can be regarded as the inverse image of the hyperplane $H_i = \mathcal{P}(x_i) \subset k\mathbb{P}^2$. Since at least one of the f_i is of degree 3 (f_0, say), it follows that $H_0 \cap F(V)$ is of degree 3 and Bézout's Theorem implies that $F(V)$ is of degree 3 as well. $\square$

Using sheaf-cohomology, one can give another common definition of genus:

DEFINITION 6.4.5. Let $V \subset k\mathbb{P}^n$ is a variety of dimension t, define the *arithmetic genus* of V by

$$g_a(V) = (-1)^t(\chi(V, \mathcal{O}_V) - 1)$$

where $\chi(V, \mathcal{O}_V)$ is the Euler characteristic in definition D.3.23 on page 559.

REMARK. If V is a projective curve, $\dim_k H^0(V, \mathcal{O}_V) = 1$, and $\dim_k H^1(V, \mathcal{O}_V) = p_g$ by exercise 1 so $g_a(V) = g_p(V)$. Although these genera agree for curves, they do not agree in general.

PROPOSITION 6.4.6. *If V is a smooth irreducible projective curve over an algebraically closed field with canonical class K_V and geometric genus g, then*

$$\deg K_V = 2g - 2$$

REMARK. This generalizes 6.1.2 on page 305 to curves that are not immersed in $k\mathbb{P}^2$.

PROOF. This follows immediately from the Riemann-Roch formula, setting $D = K_V$:

$$\ell(K_V) = \ell(0) + 1 + \deg K_V - g$$

where $\ell(K_V) = g$, from definition 5.9.40 on page 303, and $\ell(0) = 1$, by proposition 5.9.29 on page 297, statement 3. $\qquad\square$

EXERCISES.

1. Show that the geometric genus of a projective curve, V, over an algebraically closed field, k, is given by

$$p_g(V) = \dim_k H^1(V, \mathcal{O}_V)$$

This is closely related to the topological definition of genus.

6.5. The Hurwitz-Riemann Formula

Recall the existence of local parameters as in definition 3.3.21 on page 132. In the case of a smooth curve, V, we have a *single* local parameter and the local rings $\mathcal{O}_{V,p}$ are *discrete valuation rings* for all $p \in V$ — see section A.4.3 on page 431. This means

(1) $\mathcal{O}_{V,p}$ is a local ring with unique maximal ideal $\mathfrak{m}_p = (g)$ for a uniformizing parameter $g \in \mathcal{O}_{V,p}$. Given any other uniformizing parameter g', there exists a unit, $u \in \mathcal{O}_{V,p}$ such that $g = u \cdot g'$.

(2) $\mathcal{O}_{V,p}$ is a PID with a *valuation-function* $v : \mathcal{O}_{V,p} \to \mathbb{Z}$ and, for *any* element $x \in \mathcal{O}_{V,p}$, $x = g^{v(x)} \cdot u$, where u is a unit.

This rigid structure of the $\mathcal{O}_{V,p}$-rings allows us to simplify some results:

PROPOSITION 6.5.1. *Let $f : V \to W$ be a finite morphism of smooth curves, let $y \in W$ be a point, and let $f^{-1}(y) = \{x_1, \ldots, x_k\}$. If g_y is the uniformizing parameter of $\mathcal{O}_{W,y}$, and v_i is the valuation of $\mathcal{O}_{W,x_i}$, then the ramification index of f at x_i is equal to*

$$e_i = v_i(f^* g_y)$$

where $f^* : k[W] \hookrightarrow k[V]$ is the induced morphism.

REMARK. This implies that, in a neighborhood of a ramification point, x_i, the map f "looks like"

$$v_i \mapsto v_i^{e_i}$$

where v_i is a uniformizing parameter of $\mathcal{O}_{V,x_i}$. In other words, it is like example 3.4.8 on page 148.

PROOF. Without loss of generality, assume $k[W] \subset k[V]$. Simply examine the proof of theorem 3.4.9 on page 148. In one step, we form the factor by $\mathfrak{m}_y$. In the present context, $\mathfrak{m}_y = (g_y)$. Equation 3.4.8 on page 148 becomes

$$k[V]_{(g_y)}/(g_y) \cong \left(k[W]_{(g_y)}[T]/(p(T)) \right)/(g_y) = \mathcal{O}_{W,y}/(g_y) \cong k[T]/(p_y(T))$$

and equation 3.4.6 on page 149 implies that

$$\mathcal{O}_{V,x_i}/(g_y) = k^{e_i} = \mathcal{O}_{V,x_i}/(v_i^{e_i})$$

so that

$$g_y = v_i^{e_i} \cdot u_i \in \mathcal{O}_{V,x_i}$$

where v_i is a local uniformizing parameter, and u_i is a unit. It follows that $e_i = v_i(g_y)$, where v_i is the valuation associated with $\mathcal{O}_{V,x_i}$. $\square$

DEFINITION 6.5.2. If $f: V \to W$ is a finite morphism of smooth curves, we can define a homomorphism

$$f^*: \mathrm{Div}(W) \to \mathrm{Div}(V)$$

as follows:

For any point, $y \in W$, let $t \in \mathcal{O}_{W,y}$ be a local parameter at y, i.e. an element of $k(W)$ with $v_y(t) = 1$, where v_y is the valuation associated with the discrete valuation ring $\mathcal{O}_{W,y}$. We define

$$f^*y = \sum_{f(x)=y} v_x(t) \cdot x$$

Since f is a finite map, this is a finite sum, so we get a divisor on V. Note that f^*y does not depend on the local parameter, t, since other local parameters differ from t by units.

Since f^* preserves linear equivalence, it induces a homomorphism

$$f^*: \mathrm{Cl}(Y) \to \mathrm{Cl}(X)$$

PROPOSITION 6.5.3. *Let $f: V \to W$ be a finite morphism of smooth curves. Then, for any divisor, D, on W*

$$\deg f^*D = \deg f \cdot \deg D$$

PROOF. It suffices to prove this for a point $w \in W$. The fact that $\deg f^*w = \deg f$ follows from definition 6.5.2, 6.5.1 on the preceding page and the fact that the sum of the ramification indices of the inverse images of a point sum up to the degree of the map (see 3.4.9 on page 148). $\square$

PROPOSITION 6.5.4. *If $f: V \to W$ is a finite morphism of smooth curves over characteristic 0, then there is an exact sequence of sheaves on V:*

(6.5.1) $$0 \to f^*\Omega_W \xrightarrow{df} \Omega_V \to \Omega_{V/W} \to 0$$

PROOF. We get most of this in proposition 4.6.22 on page 211, where we have set $Z = \operatorname{Spec} k$. We need only verify that

$$df: f^*\Omega_W \to \Omega_V$$

is injective. We note that $f^*\Omega_W$ and Ω_V are invertible sheaves — i.e. they are coherent sheaves representing line bundles on V. The kernel of df will also be a coherent sheaf representing a line-bundle over V. Since the only sub-bundle of a line bundle is 0 or the entire line-bundle, we conclude that df is either the zero-map or an inclusion.

To see that df is an inclusion, it suffices to localize at the generic point (prime ideal (0)). We get $\Omega_{k(V)/k(W)} = 0$ by corollary A.7.16 on page 490 and the remark following it. $\qquad\square$

Now we explore the structure of $\Omega_{V/W} = \Omega_V/f^*\Omega_W$. If $x \in V$, let $f(x) = y \in W$, and let $t \in \mathcal{O}_{W,y}$ and $s \in \mathcal{O}_{V,x}$ be local parameters. Then dt generates the free $\mathcal{O}_{W,y}$-module $\Omega_{W,y}$ and ds generates the free $\mathcal{O}_{V,x}$-module $\Omega_{V,x}$ (see corollary A.7.19 on page 492). In particular, there exists a unique element $u \in \mathcal{O}_{V,x}$ such that $df(dt) = u \cdot ds$. We will write $u = dt/ds$.

PROPOSITION 6.5.5. *If $f: V \to W$ is a finite morphism of smooth curves over characteristic 0, then*

(1) *$\Omega_{V/W}$ is a torsion sheaf with support equal to the set of ramification points of f*
(2) *for each $x \in V$, the stalk, $(\Omega_{V/W})_x$ is a principal $\mathcal{O}_{V,x}$-module of length $v_x(dt/ds)$*
(3) *if f is ramified at x with ramification index e_x, then $\operatorname{length}((\Omega_{V/W})_x) = e_x - 1$.*

PROOF. The quotient, $\Omega_{V/W}$, is actually a linear combination of skyscraper-sheaves — see example B.1.7 on page 498. If f is unramified at x, then $f^*(t) = s \cdot u$, where u is a unit, and $(\Omega_{V/W})_x = \mathcal{O}_{V,x}/(dt/ds) = 0$.

If f is ramified at x with index e_x, then $f^*(t) = s^{e_x} \cdot u$, where $u \in \mathcal{O}_{V,x}$ is a unit and

$$dt = e_x s^{e_x-1} \cdot u \cdot ds$$

where u and e_x are units. It follows that

$$(\Omega_{V/W})_x = \mathcal{O}_{V,x}/(dt/ds) = \mathcal{O}_{V,x}/(s^{e_x-1})$$

It is left as an exercise to the reader to verify that

$$\operatorname{length}(\mathcal{O}_{V,x}/(s^{e_x-1})) = e_x - 1$$

$\qquad\square$

DEFINITION 6.5.6. If $f: V \to W$ is a finite morphism of smooth curves over characteristic 0, then the *ramification divisor*, R, of f is defined by

$$R = \sum_{x \in V} \operatorname{length}((\Omega_{V/W})_x) \cdot x$$

LEMMA 6.5.7. *Let $f: V \to W$ be a finite morphism of smooth curves over characteristic 0. If K_V and K_W are the canonical classes (see section 5.9.5 on page 300), respectively of V and W, then*

$$K_V \sim f^*K_W + R$$

PROOF. If we regard R as a closed subscheme of V, its structure-sheaf, $\mathcal{O}_R$ is isomorphic to $\Omega_{V/W}$. Let Ω_V^{-1} denote the invertible sheaf whose transition-functions are the *inverses* of those of Ω_V. Then

$$\Omega_V \otimes_{\mathcal{O}_V} \Omega_V^{-1} \cong \mathcal{O}_V$$

so the sheaves "cancel out" (this is another definition of invertible sheaf!).

If we take the tensor product of the exact sequence in 6.5.1 on page 334 with Ω_V^{-1}, we get

(6.5.2) $$0 \to f^*\Omega_W \otimes_{\mathcal{O}_V} \Omega_V^{-1} \to \mathcal{O}_V \to \mathcal{O}_R \to 0$$

which implies that $f^*\Omega_W \otimes_{\mathcal{O}_V} \Omega_V^{-1}$ is the sheaf of ideals defining R as a sub-scheme of V. But corollary 5.9.21 on page 292 implies that

$$f^*\Omega_W \otimes_{\mathcal{O}_V} \Omega_V^{-1} \cong \mathbf{s}(-R)$$

Taking associated divisors gives

$$f^*K_W - K_V = -R$$

or the formula given above.   □

This leads to this section's primary result:

THEOREM 6.5.8 (Hurwitz-Riemann formula). *If $f\colon V \to W$ is a degree-d finite morphism of smooth curves over characteristic 0, then*

$$2g(V) - 2 = d \cdot (2g(W) - 2) + \sum_{x \in V} (e_x - 1)$$

PROOF. Just take the degrees of the divisors in lemma 6.5.7 on the previous page — see proposition 6.4.6 on page 333.   □

⚠⚠ *Note for the topologically-inclined.* Modulo several well-known topological facts, there's a very intuitive argument for the Hurwitz-Riemann formula (which may have inspired it in the first place):

Figure 5.9.1 on page 303 shows a genus-2 surface, S, subdivided into triangles, also known as 2-simplices. The line-segments that form the sides of these triangles are called 1-simplices, and the vertices of the triangles are points, called 0-simplices. *It turns out[2] that*

$$\chi(S) = n_0 - n_1 + n_2$$

where n_i is the number of i-simplices for $i = 0, 1, 2$, and $\chi(S)$ is the *Euler characteristic*, mentioned in definition 6.4.5 on page 332, satisfying the equation

(6.5.3) $$\chi(S) = 2 - 2g$$

If $f\colon R \to S$ is an unramified degree-d map of smooth surfaces, every simplex of S has d simplices of R mapping to it, so

$$\chi(R) = d \cdot \chi(S)$$

If f is ramified at a 0-simplex, σ, of R with ramification-index e, then fewer than d 0-simplices of R map to $f(\sigma)$: in fact, there are $d - (e - 1)$ such simplices. It follows that, for a ramified map, $f\colon R \to S$, we have

$$\chi(R) = d \cdot \chi(S) - \sum (e_j - 1)$$

---

[2]We're not proving it here!

summed over all the ramified points of f. This (and equation 6.5.3 on the preceding page) immediately implies the Hurwitz-Riemann formula.

EXERCISES.

1. In proposition 6.5.4 on page 334, why isn't the map df an *isomorphism* (i.e., how can it have a nontrivial cokernel)?

2. In equation 6.5.2 on the preceding page, why is $\mathcal{O}_R \otimes_{\mathcal{O}_V} \Omega_V^{-1} = \mathcal{O}_R$?

3. If U and V are curves with $g(U) < g(V)$, then there does not exist a surjective map from U to V.

6.6. The *j*-invariant

In this section, we will develop an exhaustive classification of elliptic curves using what Felix Klein called the *j-invariant*.

Christian Felix Klein (1849 – 1925) was a German mathematician known for his work with group theory, complex analysis, non-Euclidean geometry, and on the associations between geometry and group theory. His 1872 Erlangen program, classifying geometries by their basic symmetry groups, was a synthesis of much of the mathematics of the time. His work in complex analysis includes study of elliptic modular functions (related to *our* elliptic curves and the reason they are called 'elliptic.').

We begin with:

PROPOSITION 6.6.1. *Given an elliptic curve defined by the affine curve, E,*

$$Y^2 = X^3 + aX + b \subset k\mathbb{P}^2$$

where $4a^3 + 27b^2 \neq 0$, *we have*

- *projection,* $E \to \mathbb{A}^1$, *to the X-axis is a map of degree 2*
- *projection,* $E \to \mathbb{A}^1$, *to the Y-axis is a map of degree 3*

PROOF. These projections are induced, respectively, by the inclusions

$$k[X] \hookrightarrow k[X, Y]/(Y^2 - X^3 - aX - b)$$
$$k[Y] \hookrightarrow k[X, Y]/(Y^2 - X^3 - aX - b)$$

These are inclusions of integral domains of degree 2 and 3, respectively, so they induce corresponding inclusions of fraction-fields of the same degrees.

If P is any point of an elliptic curve, E, then see corollary 6.4.2 on page 332 implies that $\ell(n \cdot P) = n$. It follows that

$$\ell(2\circledcirc) = 2$$

and $\mathcal{L}(2\circledcirc) = \{1, X\}$ and $\mathcal{L}(3\circledcirc) = \{1, X, Y\}$. $\qquad\square$

LEMMA 6.6.2. *If $E \subset k\mathbb{P}^2$ is an elliptic curve and $q \in E$ is an arbitrary point, there exists an isomorphism $f \colon E \to E$ that sends q to $\circledcirc = (0\colon 1\colon 0)$.*

PROOF. Simply use the abelian group-operation on E — see theorem 6.2.11 on page 320 — and define

$$f(x) = x + (\circledcirc - q) \colon E \to E$$

$\square$

Now we have this section's main result:

THEOREM 6.6.3. *Let E and $\bar{E}$ be two elliptic curves given by*

$$Y^2 = X^3 + aX + b$$
$$\bar{Y}^2 = \bar{X}^3 + \bar{a}\bar{X} + \bar{b}$$

respectively, over a field, k, that is algebraically closed and of characteristic $\neq 2, 3$. If $\varphi \colon \bar{E} \to E$ is an isomorphism that (without loss of generality, by lemma 6.6.2) we assume maps $\bar{\circledcirc} \in \bar{E}$ to $\circledcirc \in E$. Then there exists a $c \in k^\times$ such that $\bar{a} = c^4 a$, $\bar{b} = c^6 b$ and φ is given by

$$\varphi(X\colon Y\colon Z) = (c^2 X \colon c^3 Y \colon Z)$$

If we define

$$j(E) = \frac{1728(4a^3)}{4a^3 + 27b^2}$$

then $j(E)$ only depends on E — and two elliptic curves, E and $\bar{E}$, are isomorphic if and only if $j(E) = j(\bar{E})$.

REMARK. We make heavy use of the fact that k is algebraically closed. The statement isn't even true otherwise.

PROOF. The regular function $X \circ \varphi$ has a pole of degree 2 (see proposition 6.6.1 on the previous page) at $\bar{\circledcirc}$, so $X \circ \varphi \in \mathcal{L}(2\bar{\circledcirc}) = \{1, \bar{X}\}$. It follows that $X \circ \varphi = r \cdot \bar{X} + s$ for $r, s \in k$. Similar reasoning shows that $Y \circ \varphi = u \cdot \bar{Y} + v \cdot \bar{X} + w$.

We must have

$$(u \cdot \bar{Y} + v \cdot \bar{X} + w)^2 = (r \cdot \bar{X} + s)^3 + a(r \cdot \bar{X} + s) + b$$
$$\bar{Y}^2 = \bar{X}^3 + \bar{a}\bar{X} + \bar{b}$$

from which we conclude $v = w = s = 0$, and $u^2 = r^3$. If $c = u/r$, we get $\bar{a} = c^4 a$ and $\bar{b} = c^6 b$, and φ is as described.

If $\bar{E}$ is isomorphic to E (as computed above), then $j(\bar{E}) = j(E)$. To prove the converse, we consider two cases:

- $a = 0$, $j(E) = 0$, $\bar{a} = 0$. Our elliptic curves are

$$Y^2 = X^3 + b$$
$$\bar{Y}^2 = \bar{X}^3 + \bar{b}$$

  which are isomorphic via $\bar{Y} = Y\sqrt{b/\bar{b}}$, $\bar{X} = X\sqrt[3]{b/\bar{b}}$.
- If $a, \bar{a} \neq 0$, let $c = \sqrt[4]{\bar{a}/a}$ and replace (a, b) by $(c^4 a, c^6 b)$. We now have $\bar{a} = a$. If $j(\bar{E}) = j(E)$, we get $\bar{b} = \pm b$. If necessary, we can correct the sign by replacing c by $c\sqrt{-1}$.

□

REMARK 6.6.4. Every element of k occurs as the j-invariant of some elliptic curve.

- $Y^2 = X^3 + 1, j = 0$
- $Y^2 = X^3 + X, j = 1728$
- $Y^2 = X^3 - \frac{27}{4} \frac{j}{j-1728} X - \frac{27}{4} \frac{j}{j-1728}, j \neq 0, 1728$

We conclude that $\mathbb{A}^1$ is the *moduli space* for elliptic curves over k (assuming k is is algebraically closed and of characteristic $\neq 2, 3$).

The interested reader is referred to the excellent treatise on elliptic curves by Milne, [**113**].

EXERCISES.

1. If we write the equation of an elliptic curve in the form

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

with $\lambda \neq 0, 1$, what is the equation of the j-invariant in terms of λ?

APPENDIX A

Algebra

"L'algèbre n'est qu'une géométrie écrite; la géométrie n'est qu'une algèbre figurée." (Algebra is merely geometry in words; geometry is merely algebra in pictures)
— Sophie Germain, [**56**]

A.1. Rings

In order to proceed further, we need a fair amount of algebraic machinery. We begin with a mathematical structure that we first meet in elementary school:

DEFINITION A.1.1. A ring, R, is a set equipped with two binary operations, denoted $+$ and $\cdot$ such that, for all $r_1, r_2, r_2 \in R$,

(1) $(r_1 + r_2) + r_3 = r_1 + (r_2 + r_3)$
(2) $(r_1 \cdot r_2) \cdot r_3 = r_1 \cdot (r_2 \cdot r_3)$
(3) $r_1 \cdot (r_2 + r_3) = r_1 \cdot r_2 + r_1 \cdot r_3$
(4) $(r_1 + r_2) \cdot r_3 = r_1 \cdot r_3 + r_1 \cdot r_3$
(5) there exists elements $0, 1 \in R$ such that $r + 0 = 0 + r = r$ and $r \cdot 1 = 1 \cdot r = r$ for all $r \in R$.
(6) For every $r \in R$, there exists an element $s \in R$ such that $r + s = 0$.

The ring R will be called *commutative* if $r_1 \cdot r_2 = r_2 \cdot r_1$ for all $r_1, r_2 \in R$.

REMARK. In algebraic geometry we will deal with commutative rings exclusively. We will also regard the set containing only the number 0 as a ring with $0 + 0 = 0 = 0 \cdot 0$ — the *trivial ring* (the multiplicative and additive identities are the same). When an operation is written with a '$+$' sign it is implicitly assumed to be commutative.

We also have an important variation on this:

DEFINITION A.1.2. Given a ring k, we will call an abelian group, R, a *k-algebra* if there exists:

(1) A multiplication-operation
$$\mu: R \times R \to R$$
with an identity element $1 \in R$,
(2) a homomorphism of abelian groups $f: k \to R$ that is also a ring-homomorphism with respect to μ.

We use the notation
$$x \cdot y = \mu(f(x), y)$$
$$\mu(r, s) = rs$$

for all $x \in k$ and $y \in R$ and $r, s \in R$. The action of k on R is bilinear in the sense that

$$(\alpha \cdot x + \beta \cdot y)z = \alpha \cdot xz + \beta \cdot yz$$
$$x(\alpha \cdot y + \beta \cdot z) = \alpha \cdot xy + \beta \cdot xz$$

for all $\alpha, \beta \in k$ and $x, y, z \in R$.

REMARK. This is a slight extension of the usual definition of an algebra, where k is a field and R is a k-vector space.

Note that all commutative rings are $\mathbb{Z}$-algebras in a unique way.

Algebras are both more and less general than rings. They are not required to be commutative or even associative — although all algebras in this book will be. For examples of non commutative and non associative algebras, see [**155**].

We can classify elements of a ring by certain basic properties:

DEFINITION A.1.3. An element $u \in R$ of a ring will be called a *unit* if there exists another element $v \in R$ such that $u \cdot v = v \cdot u = 1$. The set of units of a ring, R, form a group, denoted $R^{\times}$. A ring for which every nonzero element is a unit is called a *field.*

An element $u \in R$ is called a *zero-divisor* if it is nonzero and if there exists a nonzero element $v \in R$ such that $u \cdot v = 0$.

EXAMPLE. Perhaps the simplest example of a ring is the integers, $\mathbb{Z}$. This is simple in terms of familiarity to the reader but a detailed analysis of the integers is a very deep field of mathematics in itself (number theory). Its only units are ± 1, and it has no zero-divisors.

We can use the integers to construct:

EXAMPLE. If m is an integer, the numbers modulo m, $\mathbb{Z}_m$ is a ring under addition and multiplication modulo m.

We also have *polynomial* rings:

DEFINITION A.1.4. If R is a ring, rings of polynomials $R[X]$ is the ring of polynomials where addition and multiplication are defined

$$\left(\sum_{i=0}^{n} a_i X^i \right) + \left(\sum_{i=0}^{m} b_i X^i \right) = \sum_{i=0}^{\max(n,m)} (a_i + b_i) X^i$$
$$\left(\sum_{i=0}^{n} a_i X^i \right) \left(\sum_{i=0}^{m} b_i X^i \right) = \sum_{i=1}^{n+m} \left(\sum_{j+k=i} a_j b_k \right) X^i$$

with $a_i, b_j \in R$ and $a_i = 0$ if $i > n$ and $b_i = 0$ if $i > m$.

More formally, one can define $R[X]$ as the set of infinite sequences

(A.1.1) $$(r_0, \ldots, r_i, \ldots)$$

with the property that all but a *finite* number of the r_i vanish, and with addition defined by

$$(r_0, \ldots, r_i, \ldots) + (s_0, \ldots, s_i, \ldots) = (r_0 + s_0, \ldots, r_i + s_i, \ldots)$$

and multiplication defined by

$$(r_0, \ldots, r_i, \ldots)(s_0, \ldots, s_i, \ldots) = (t_0, \ldots, t_i, \ldots)$$

with

$$t_n = \sum_{\substack{i+j=n \\ i \geq 0, j \geq 0}} r_i s_j$$

In this case,

$$\sum_{i=0}^{k} r_i X^i$$

becomes the *notation* for the sequence $(r_0, \ldots, r_i, \ldots, r_k, 0 \cdots)$.

EXAMPLE. $\mathbb{Z}[X]$ is the ring of polynomials with integer coefficients.

We can also define power-series ring

DEFINITION A.1.5. If R is a ring, the *ring of power-series $R[[X]]$* over R is the ring of formal power series

$$\sum_{i=1}^{\infty} a_i X^i$$

with addition and multiplication defined as for $R[X]$. As with polynomial-rings, one can formally define the elements of $R[[X]]$ as infinite sequences like those in A.1.1 on the facing page where we allow an *infinite* number of the r_i to be nonzero.

REMARK. Note that these power-series are like infinite polynomials. If we impose a metric on R the ring of power-series that *converge* with respect to that metric can be very different from $R[[X]]$.

CLAIM A.1.6. We can define a metric on $R[[X]]$ that makes power-series convergent in the usual sense.

Let $p, q \in R[[X]]$ and define the distance between them by

$$d(p, q) = \left(\frac{1}{2} \right)^{v(p-q)}$$

where $X^{v(p-q)} | (p - q)$ but $X^{v(p-q)+1} \nmid (p - q)$, i.e. the function $v(x)$ is equal to the degree of the lowest-degree term of x. In this metric all formal power-series converge and we can define Cauchy-sequences, etc.

Power series rings can have very different properties than polynomial rings. For instance

PROPOSITION A.1.7. *In the ring $R[[X]]$, any element*

$$\alpha = \sum_{k=0}^{\infty} a_k X^k$$

where $a_0 \in R$ is a unit (see definition A.1.3 on the preceding page) has a multiplicative inverse.

REMARK. The inverse of α is

$$\frac{1}{a_0} - \frac{a_1}{a_0^2}X + \frac{a_0 a_2 - a_1^2}{a_0^3}X^2 + \cdots$$

PROOF. Suppose the inverse is

$$\sum_{j=0}^{\infty} b_j X^j$$

and multiply α by this to get

$$\sum_{n=0}^{\infty} c_n X^n$$

with

$$c_n = \sum_{j=0}^{n} a_j b_{n-j}$$

$$\begin{aligned} c_0 &= a_0 b_0 \\ b_0 &= a_0^{-1} \end{aligned}$$

In general, we get a recursive equation

$$b_n = -a_0^{-1} \sum_{k=0}^{n-1} b_k a_{n-k}$$

that computes b_n for any n.   □

We also have extension rings

DEFINITION A.1.8. Suppose we have an embedding of rings $R \subset \Omega$ and $\alpha \in \Omega$ is some element. Then $R[\alpha] \subset \Omega$ is the subring of all possible polynomials

$$\sum_{i=1}^{n} c_i \alpha^i$$

with $c_i \in R$.

EXAMPLE. In the extension $\mathbb{Q}[\sqrt{2}]$, the fact that $(\sqrt{2})^2 \in \mathbb{Q}$ implies that all elements of $\mathbb{Q}[\sqrt{2}]$ will actually be of the form $a + b\sqrt{2}$, with $a, b \in \mathbb{Q}$.

Although not used much in algebraic geometry, non-commutative rings are widespread:

EXAMPLE. Let $\mathbf{M}(R, n)$ be the set of $n \times n$ matrices with entries in a ring R. Then $\mathbf{M}(R, n)$ is also a ring under matrix-addition and multiplication.

We conclude this section by focusing on the *integers* — and properties of them that extend to some more general rings.

PROPOSITION A.1.9. *Let n and d be positive integers. Then it is possible to write*

$$n = q \cdot d + r$$

where $0 \le r < d$. If $r = 0$, we say that $d|n$ — stated "d divides n".

REMARK. So $u \in R$ is a unit if and only if $u|1$.

The division algorithm mentioned above gives rise to the concept of greatest common divisor in the integers (and some other rings):

DEFINITION A.1.10. Let n and m be positive integers. The *greatest common divisor* of n and m, denoted $\gcd(n, m)$, is the largest integer d such that $d \mid n$ and $d \mid m$. The *least common multiple* of n and m, denoted $\operatorname{lcm}(n, m)$, is the smallest positive integer k such that $n \mid k$ and $m \mid k$.

Since 0 is divisible by any integer, $\gcd(n, 0) = \gcd(0, n) = n$.

There is a very fast algorithm for computing the greatest common divisor due to Euclid — see [**43, 44**].

ALGORITHM A.1.11. *Given positive integers n and m with $n > m$, use the division algorithm to set*

$$
\begin{aligned}
n &= q_0 \cdot m + r_0 \\
m &= q_1 \cdot r_0 + r_1 \\
r_0 &= q_2 \cdot r_1 + r_2 \\
&\vdots \\
r_{k-2} &= q_k \cdot r_{k-1} + r_k
\end{aligned}
$$

with $m > r_0 > r_1 > \cdots > r_k$. At some point $r_N = 0$ and we claim that $r_{N-1} = \gcd(n, m)$.

REMARK. Euclid's original formulation was geometric, involving line-segments. Given two line-segments of lengths r_1 and r_2, it found a real number r such that

$$
\frac{r_1}{r}, \frac{r_2}{r} \in \mathbb{Z}
$$

An ancient proof of the irrationality of $\sqrt{2}$ showed that this process never terminates if one of the line-segments is of unit length and the other is the diagonal of a unit square.

PROOF. To see that this works, note that $r_{N-1} | r_{N-2}$ since $r_N = 0$. A simple induction shows that $r_{N-1} | r_i$ for all $i < N$ and $r_{N-1} | m$ and $r_{N-1} | n$. Consequently $r_{N-1} | \gcd(m, n)$. On the other hand, another simple induction shows that r_{N-1} is an integer linear combination of m and n so $\gcd(m, n) | r_{N-1}$, so $r_{N-1} \geq \gcd(m, n)$. To summarize:

(1) r_{n-1} is a divisor of n and m
(2) $r_{N-1} \geq \gcd(m, n)$

Since $\gcd(m, n)$ is the *greatest* common divisor, we must have $r_{N-1} = \gcd(m, n)$. $\square$

As trivial as proposition A.1.9 on the facing page appears to be, it allows us to prove Bézout's Identity:

LEMMA A.1.12. *Let n and m be positive integers. Then there exist integers u and v such that*

$$
\gcd(n, m) = u \cdot n + v \cdot m
$$

REMARK. Bézout proved this identity for polynomials — see [**13**]. However, this statement for integers can be found in the earlier work of Claude Gaspard Bachet de Méziriac (1581–1638) — see [**165**].

PROOF. Let z be the smallest positive value taken on by the expression

$$(A.1.2) \qquad\qquad z = u \cdot n + v \cdot m$$

as u and v run over all possible integers. Clearly, $\gcd(n, m)|z$ since it divides any possible linear combination of m and n. It follows that $\gcd(n, m) \leq z$.

We claim that $z|n$. If not, then proposition A.1.9 on page 344 implies that $n = q \cdot z + r$, where $0 < r < z$, or $r = n - q \cdot z$. Plugging that into equation A.1.2 gives

$$
\begin{aligned}
r &= n - q \cdot (u \cdot n + v \cdot m) \\
&= (1 - q \cdot u) \cdot n - q \cdot v \cdot m
\end{aligned}
$$

which is a linear combination of n and m *smaller* than z — a contradiction. Similar reasoning shows that $z|m$ so z is a common divisor of m and $n \geq \gcd(m, n)$ so it must *equal* $\gcd(m, n)$. $\qquad\square$

Recall that a prime number is an integer that is not divisible by any integer other than 1 or ($\pm$)itself. The ring-theoretic analogue is an *irreducible element*.

Bézout's Identity immediately implies:

PROPOSITION A.1.13. *Let p be a prime number and let n and m be integers. Then*

$$p|m \cdot n \implies p|m \text{ or } p|n$$

PROOF. Suppose $p \nmid m$. We will show that $p|n$. Since p is prime and $p \nmid m$, we have $\gcd(p, m) = 1$. Lemma A.1.12 on the previous page implies that there exist integers u and v such that

$$1 = u \cdot m + v \cdot p$$

Now multiply this by n to get

$$n = u \cdot mn + v \cdot n \cdot p$$

Since p divides each of the terms on the right, we get $p|n$. A similar argument show that $p \nmid n \implies p|m$. $\qquad\square$

A simple induction shows that:

COROLLARY A.1.14. *If p is a prime number, $k_i \in \mathbb{Z}$ for $i = 1, \dots, n$ and*

$$p \Big| \prod_{i=1}^{n} k_i$$

then $p|k_j$ for at least one value of $1 \leq j \leq n$. If p and q are both primes and

$$q|p^i$$

for some integer $i \geq 1$, then $p = q$.

PROOF. We do induction on n. Proposition A.1.13 on the facing page proves the result for $n = 2$.

Suppose the result is known for $n - 1$ factors, and we have n factors. Write

$$\prod_{i=1}^{n} k_i = k_1 \cdot \left(\prod_{i=2}^{n} k_i \right)$$

Since

$$p | k_i \cdot \left(\prod_{i=2}^{n} k_i \right)$$

we either have $p | k_1$ or

$$p | \prod_{i=2}^{n} k_i$$

The inductive hypothesis proves the result. If the k_j are all copies of a prime,p, we must have $q | p$, which only happens if $q = p$. $\qquad \square$

This immediately implies the well-known result:

LEMMA A.1.15. *Let n be a positive integer and let*

$$\begin{aligned} n &= p_1^{\alpha_1} \cdot \dots \cdot p_k^{\alpha_k} \\ (\text{A.1.3}) \qquad &= q_1^{\beta_1} \cdot \dots \cdot q_\ell^{\beta_\ell} \end{aligned}$$

be factorizations into powers of distinct primes. Then $k = \ell$ and there is a reordering of indices $f \colon \{1, \dots, k\} \to \{1, \dots, k\}$ such that $q_i = p_{f(i)}$ and $\beta_i = \alpha_{f(i)}$ for all i from 1 to k.

PROOF. First of all, it is easy to see that a number can be factored into a product of primes. We do induction on k. If $k = 1$ we have

$$p_1^{\alpha_1} = q_1^{\beta_1} \cdot \dots \cdot q_\ell^{\beta_\ell}$$

Since $q_1 | p_1^{\alpha_1}$, corollary A.1.14 on the preceding page implies that $q_1 = p_1$, $\beta_1 = \alpha_1$ and that the primes $q_i \neq p_1$ cannot exist in the product. So $\ell = 1$ and the conclusion follows.

Assume the result for numbers with $k - 1$ distinct prime factors. Equation A.1.3 implies that

$$q_1 | p_1^{\alpha_1} \cdot \dots \cdot p_k^{\alpha_k}$$

and corollary A.1.14 on the preceding page implies that $q_1 | p_j^{\alpha_j}$ for some value of j. It also implies that $p_j = q_1$ and $\alpha_j = \beta_1$. We define $f(1) = j$ and take the quotient of n by $q_1^{\beta_1} = p_j^{\alpha_j}$ to get a number with $k - 1$ distinct prime factors. The inductive hypothesis implies the conclusion. $\qquad \square$

The Extended Euclid algorithm explicitly calculates the factors that appear in the Bézout Identity:

ALGORITHM A.1.16. *Suppose n, m are positive integers with $n > n$ and we use Euclid's algorithm (A.1.11 on page 345) to compute $\gcd(n, m)$. Let q_i, r_i for $0 < i \le$*

N (in the notation of A.1.11 on page 345) denote the quotients and remainders used. Now define

$$\begin{aligned} x_0 &= 0 \\ y_0 &= 1 \\ x_1 &= 1 \\ y_1 &= -q_1 \end{aligned}$$

(A.1.4)

and recursively define

$$\begin{aligned} x_k &= x_{k-2} - q_k x_{k-1} \\ y_k &= y_{k-2} - q_k y_{k-1} \end{aligned}$$

(A.1.5)

for all $2 \le k \le N$. Then

$$r_i = x_i \cdot n + y_i \cdot m$$

so that, in particular,

$$\gcd(n, m) = x_{N-1} \cdot n + y_{N-1} \cdot m$$

PROOF. If $r_i = x_i \cdot n + y_i \cdot m$ then

$$\begin{aligned} r_k &= r_{k-2} - q_k r_{k-1} \\ &= x_{k-2} \cdot n + y_{k-2} \cdot m - q_k(x_{k-1} \cdot n + y_{k-1}m) \\ &= (x_{k-2} - q_k x_{k-1}) \cdot n + (y_{k-2} - q_k y_{k-1}) \cdot m \end{aligned}$$

This implies the inductive formula A.1.5, and to get the correct values for r_1 and r_2:

$$\begin{aligned} r_1 &= n - m \cdot q_1 \\ r_2 &= m - r_1 \cdot q_2 \\ &= m - q_2 \cdot (n - m \cdot q_1) \\ &= -q_2 \cdot n + (1 + q_1 q_2) \cdot m \end{aligned}$$

we must set x_0, x_1, y_0, y_1 to the values in equation A.1.4. $\square$

EXERCISES.

1. Show that additive and multiplicative inverses in rings are *unique*, i.e., if $r + s_1 = r + s_2 = 0 \in R$, then $s_1 = s_2$.

2. What are the units of $\mathbb{Z}$?

3. Find the units of $\mathbb{Z}_m$, where $m > 1$ is some integer.

4. Find the greatest common divisor of 123 and 27 and find integers a and b such that

$$\gcd(123, 27) = a \cdot 123 + b \cdot 27$$

5. If $x > 0$ is a rational number that is not an integer, show that x^x is irrational.

A.1.1. Homomorphisms and ideals. Now that we have defined rings, we can define mappings of them:

DEFINITION A.1.17. Given two rings, R and S, a function $f\colon R \to S$ is called a *homomorphism* if, for all $r_1, r_2 \in R$:

(1) $f(r_1 + r_2) = f(r_1) + f(r_2) \in S$
(2) $f(r_1 \cdot r_2) = f(r_1) \cdot f(r_2) \in S$ and $f(1) = 1$.

The set of elements $r \in R$ with the property that $f(r) = 0$ is called the *kernel* of the homomorphism, or ker f. If the homomorphism is *surjective* and its kernel vanishes, it is called an *isomorphism*. An isomorphism from a ring to itself is called an *automorphism*.

PROPOSITION A.1.18. *Let K be the kernel of a homomorphism $f\colon R \to S$ of rings. If $k \in K$ and $r \in R$, then $r \cdot k, k \cdot r \in K$.*

PROOF. The defining property of a homomorphism implies that $f(r \cdot k) = f(r) \cdot f(k) = f(r) \cdot 0 = 0$. □

We can abstract out the important property of kernels:

DEFINITION A.1.19. If R is a ring, an *ideal $\mathfrak{I}$ in R* is an abelian subgroup of R with the property that, for all $x \in \mathfrak{I}, r \in R, r \cdot x, x \cdot r \in \mathfrak{I}$.

(1) An ideal, $\mathfrak{I} \subset R$ is *prime* if $a \cdot b \in \mathfrak{I}$ implies that $a \in \mathfrak{I}$ or $b \in \mathfrak{I}$ (or both).
(2) The *ideal generated by* $\alpha_1, \ldots, \alpha_n \in R$, denoted $(\alpha_1, \ldots \alpha_n) \subseteq R$, is the set of all linear combinations

$$\sum_{k=1}^{n} r_k \cdot \alpha_k$$

where the r_i run over all elements of R. The number 0 is an ideal, as well as the whole ring.
(3) An ideal $\mathfrak{I} \subset R$ is *maximal* if $\mathfrak{I} \subset \mathfrak{K}$, where $\mathfrak{K}$ is an ideal, implies that $\mathfrak{K} = R$. This is equivalent to saying that for any $r \in R$ with $r \notin \mathfrak{I}$,

$$\mathfrak{I} + (r) = R$$

(4) An ideal generated by a *single element* of R is called a *principal ideal.*
(5) Given two ideals $\mathfrak{a}$ and $\mathfrak{b}$, their product is the ideal generated by all products $\{(a \cdot b) | \forall a \in \mathfrak{a}, b \in \mathfrak{b}\}$.

REMARK. Following a convention in algebraic geometry, we will usually denote ideals by *Fraktur letters.* An ideal would be a subring except that most interesting ideals do not contain 1.

Julius Wilhelm Richard Dedekind (1831 – 1916) was a German mathematician who worked in abstract algebra, algebraic number theory and analysis (he gave one of the first rigorous definitions of the real numbers). The concept of an ideal originated in Dedekind's research on Fermat's last theorem — see [**35**].

EXAMPLE. We claim that the ideals of $\mathbb{Z}$ are just the sets

$$
\begin{aligned}
(2) &= \{\ldots, -4, -2, 0, 2, 4, 6, 8, \ldots\} \\
(3) &= \{\ldots, -6, -3, 0, 3, 6, 9, 12, \ldots\} \\
&\quad\vdots \\
(n) &= \{n \cdot \mathbb{Z}\}
\end{aligned}
$$

for various values of n. Proposition A.1.12 on page 345 shows that $(n, m) = (\gcd(m, n))$ and a simple induction shows that all ideals of $\mathbb{Z}$ are generated by a single element. Note that the ideal $(1) = \mathbb{Z}$. An ideal $(n) \subset \mathbb{Z}$ is *prime* if and only if n is a prime number.

Maximal ideals are prime:

PROPOSITION A.1.20. *If R is a ring with maximal ideal $\mathfrak{I}$, then $\mathfrak{I}$ is also prime.*

PROOF. This is similar to the proof of proposition A.1.13 on page 346. Suppose $r, s \in R$, $r \cdot s \in \mathfrak{I}$ but $r \notin \mathfrak{I}$. Then $\mathfrak{I} + (r) = R$ so that there exists a $t \in R$ such that

$$ a + t \cdot r = 1 $$

where $a \in \mathfrak{I}$. If we multiply this by s, we get

$$ a \cdot s + t \cdot r \cdot s = s $$

Since both terms on the left are in $\mathfrak{I}$, it follows that $s \in \mathfrak{I}$. $\qquad\square$

Proposition A.1.18 on the previous page shows that the kernel of a homomorphism is an ideal. The following is a converse to that:

PROPOSITION A.1.21. *Let R be a ring and let $\mathfrak{I} \subset R$ be an ideal. For all $r_1, r_2 \in R$ define*

$$ r_1 \equiv r_2 \pmod{\mathfrak{I}} $$

if $r_1 - r_2 \in \mathfrak{I}$. Then $\equiv$ is an equivalence relation. If we denote the set of equivalence-classes by $R/\mathfrak{I}$, then the ring-operations of R induce corresponding operations on $R/\mathfrak{I}$ making it into a ring (called the quotient ring of R by $\mathfrak{I}$). The canonical map

$$ R \to R/\mathfrak{I} $$

that sends an element to its equivalence class is a homomorphism with kernel $\mathfrak{I}$.

REMARK A.1.22. We can also think of the elements of $R/\mathfrak{I}$ as disjoint sets of elements of R, namely sets of the form

$$ r + \mathfrak{I} $$

These are all of the elements of R equivalent to $r \in R$.

PROOF. It is not hard to see that

$$ r_1 \equiv r_2 \pmod{\mathfrak{I}} $$

and

$$ s_1 \equiv s_2 \pmod{\mathfrak{I}} $$

implies that

$$ r_1 + s_1 \equiv r_2 + s_2 \pmod{\mathfrak{I}} $$

so that addition is well-defined in $R/\mathfrak{I}$. To see that multiplication is also well-defined note that

$$r_1 s_1 - r_2 s_2 = (r_1 - r_2)s_1 + r_2(s_1 - s_2) \in \mathfrak{I}$$

due to the closure property in definition A.1.19 on page 349. The final statement follows from the fact that $\mathfrak{I}$ is just the set of elements of R equivalent to 0. □

EXAMPLE. Here are examples of quotient rings:
(1) For instance, $\mathbb{Z}/(n) = \mathbb{Z}_n$, the integers modulo n, where $\mathbb{Z}/(1)$ is the trivial ring.
(2) In the example given earlier, $\mathbb{Q}[X,Y]/(X) = \mathbb{Q}[Y]$ and $\mathbb{Q}[X,Y]/(X,Y) = \mathbb{Q}$.
(3) We can think of the ring $\mathbb{Q}[\sqrt{2}]$ two ways: as an extension or as a quotient

$$\mathbb{Q}[X]/(X^2 - 2)$$

There's a homomorphism

$$\begin{aligned} \mathbb{Q}[X] &\to \mathbb{Q}[\sqrt{2}] \\ X &\mapsto \sqrt{2} \end{aligned}$$

whose kernel is exactly $(X^2 - 2)$. This induces an isomorphism $\mathbb{Q}[X]/(X^2 - 2) \cong \mathbb{Q}[\sqrt{2}]$.

Complementing the concept of kernel, we have the cokernel:

DEFINITION A.1.23. If $f\colon R \to S$ is a homomorphism of rings and if $f(R) \subset S$ is an ideal in S, the quotient

$$\frac{S}{f(R)}$$

is called the cokernel of f.

REMARK. Cokernels for homomorphisms of rings do not always exist because one cannot "divide" a ring by an arbitrary subring.

Here is an algebraic term motivated *entirely* by algebraic geometry (see definition 3.2.1 on page 117):

DEFINITION A.1.24. A ring R is called a *local ring* if it has a *unique* maximal ideal.

For instance, let R be the subring of $\mathbb{Q}$ of fractions

$$\frac{p}{q}$$

where q is an odd number. Then $2 \cdot R \subset R$ is the only ideal not equal to all of R. It follows that R is a local ring.

We could also have defined R by

$$R = \mathbb{Z}[\frac{1}{3}, \frac{1}{5}, \ldots, \frac{1}{p}, \ldots]$$

where p runs over all odd primes.

Here is how the projection to a quotient ring affects ideals:

LEMMA A.1.25. *Let R be a ring and let $\mathfrak{a} \subset R$ be an ideal and let*

$$p\colon R \to R/\mathfrak{a}$$

Then p induces a one-to-one correspondence between ideals of $R/\mathfrak{a}$ and ideals $\mathfrak{b} \subset R$ that contain $\mathfrak{a}$. In addition,

- *$p(\mathfrak{b})$ is prime or maximal in $R/\mathfrak{a}$ if and only if $\mathfrak{b}$ is prime or maximal in R*
- *$p^{-1}(\mathfrak{c})$ is prime or maximal in R if and only if $\mathfrak{c}$ is prime or maximal in $R/\mathfrak{a}$.*

PROOF. Let $\mathfrak{b} \subset R$ be an ideal containing $\mathfrak{a}$ and let $y \in R$ with $p(y) = x \in R/\mathfrak{a}$. Then $x \cdot p(\mathfrak{b}) = p(y \cdot \mathfrak{b}) \subset p(\mathfrak{b})$ so that $p(\mathfrak{b}) \subset R/\mathfrak{a}$ is an ideal.

Suppose $\mathfrak{b}$ is maximal in R. Then $(x) + p(\mathfrak{b}) = p((y) + \mathfrak{b}) = p(R) = R/\mathfrak{a}$ so $p(\mathfrak{b})$ is maximal in $R/\mathfrak{a}$.

If $\mathfrak{b}$ is prime, $X_1 \cdot x_2 \in p(\mathfrak{b})$ implies that $y_1 \cdot y_2 \in \mathfrak{b}$, where $p(y_i) = x_i$, and either $y_1 \in \mathfrak{b}$ or $y_2 \in \mathfrak{b}$, which implies that $x_1 \in p(\mathfrak{b})$ or $x_2 \in \mathfrak{b}$. This means that $p(\mathfrak{b}) \subset R/\mathfrak{a}$ is prime.

Now suppose $\mathfrak{c} \subset R/\mathfrak{a}$. Then $\mathfrak{a} \subset p^{-1}(\mathfrak{a})$ (since $\mathfrak{a} = p^{-1}(0)$). If $x \in R$, then $x \cdot p^{-1}(\mathfrak{c})$ has the property that its image under p is equal to $\mathfrak{c}$, i.e., it is contained in $p^{-1}(\mathfrak{c})$. It follows that $p^{-1}(\mathfrak{c})$ is an ideal of R.

Suppose $\mathfrak{c}$ is maximal in $R/\mathfrak{a}$, and suppose that $x \in R$ has the property that $x \notin p^{-1}(\mathfrak{c})$. Then $p(x) \notin \mathfrak{c}$ and $\mathfrak{I} = (x) + p^{-1}(\mathfrak{c})$ is an ideal of R that has the property that $p(\mathfrak{I}) = (p(x)) + \mathfrak{c} = R/\mathfrak{a}$. So $\mathfrak{I} = R$ and $p^{-1}(\mathfrak{c})$ is maximal.

We leave the final statement that p^{-1} of a prime ideal is prime as an exercise. □

We will also need to know the effect of multiple quotients:

LEMMA A.1.26. *Let R be a ring with ideals $\mathfrak{a} \subset \mathfrak{b} \subset R$. Let*

(1) *$f\colon R \to R/\mathfrak{a}$,*
(2) *$g\colon R \to R/\mathfrak{b}$ and*
(3) *$h\colon R/\mathfrak{a} \to (R/\mathfrak{a})/f(\mathfrak{b})$*

be projections to the quotients. Then $(R/\mathfrak{a})/f(\mathfrak{b}) = R/\mathfrak{b}$ and the diagram

$$
\begin{array}{ccc}
R & \xrightarrow{\;\;f\;\;} & R/\mathfrak{a} \\
{\scriptstyle g}\downarrow & & \downarrow{\scriptstyle h} \\
R/\mathfrak{b} & =\!\!=\!\!= & (R/\mathfrak{a})/f(\mathfrak{b})
\end{array}
$$

commutes.

PROOF. Elements of $R/\mathfrak{a}$ are equivalence classes of the equivalence relation

$$r_1 \sim_{\mathfrak{a}} r_2 \text{ if } r_1 - r_2 \in \mathfrak{a}$$

or *sets* of the form (see remark A.1.22 on page 350)

$$r + \mathfrak{a} \subset R$$

and elements of $R/\mathfrak{b}$ are sets of the form

$$r + \mathfrak{b} \subset R$$

Elements of $(R/\mathfrak{a})/f(\mathfrak{b})$ are sets of the form

$$q + f(\mathfrak{b})$$

where $q \in R/\mathfrak{a}$, or sets of the form

$$r + \mathfrak{a} + \mathfrak{b} = r + \mathfrak{b}$$

This shows that $(R/\mathfrak{a})/f(\mathfrak{b}) = R/\mathfrak{b}$. The commutativity of the diagram follows from the fact that the image of $r \in R$ under the maps going down either side of the diagram is the set $r + \mathfrak{b}$. □

For the next result, we will use Zorn's Lemma, a classic result found in [**100, 178**]:

LEMMA A.1.27. *If S is a partially-ordered set with the property that every increasing sequence of elements*

$$e_1 \prec e_2 \prec \cdots$$

has an upper bound, then S contains a maximal element.
    If every decreasing sequence of elements

$$e_1 \succ e_2 \succ \cdots$$

has a lower bound, then S has a minimal element.

REMARK. In this context, "maximal" means "there is an element $e \in S$ such that there does *not* exist an element $e' \in S$ with $e \prec e'$."
    Zorn's lemma is equivalent to the axiom of choice in set theory.

PROPOSITION A.1.28. *If $\mathfrak{I} \subset R$ is an ideal in a ring with $1 \notin \mathfrak{I}$, then there exists a maximal ideal $\mathfrak{M} \subset R$ such that*

$$\mathfrak{I} \subset \mathfrak{M}$$

PROOF. The ideals of R that contain $\mathfrak{I}$ can be ordered by inclusion. Every ascending chain of such ideals has an upper bound, namely the union. Zorn's Lemma implies (A.1.27) that there is a maximal such ideal. □

EXERCISES.

6. Suppose A is an algebra over a field, k, that is also a commutative ring (i.e., it is commutative and associative). If $\mathfrak{g} \subset A$ is an ideal, show that

$$A/\mathfrak{g}$$

is also an algebra over k.

7. If $f \colon R \to S$ is a homomorphism of rings, and $\mathfrak{a} \subset S$ is an ideal, show that $f^{-1}(\mathfrak{a}) \subset R$ is an ideal such that there's an induced map

$$\frac{R}{f^{-1}(\mathfrak{a})} \to \frac{S}{\mathfrak{a}}$$

that is a monomorphism.

8. If $x, y \in R$ are two elements with the property that $(x, y) = R$, show that $(x^n, y^m) = R$ for positive integers n, m.

9. Show that the converse of proposition A.1.7 on page 343 is also true: if

$$\alpha = \sum_{i=0}^{\infty} a_i X^i \in R[[X]]$$

is a unit, so is a_0.

10. If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in a ring, show that $\mathfrak{a} \cdot \mathfrak{b} \subset \mathfrak{a} \cap \mathfrak{b}$.

11. Suppose $\mathfrak{a}, \mathfrak{b}, \mathfrak{p} \subset R$ are ideals in a commutative ring. If $\mathfrak{p}$ is a prime ideal and

$$\mathfrak{a} \cdot \mathfrak{b} \subset \mathfrak{p}$$

(for instance, if $\mathfrak{a} \cap \mathfrak{b} \subset \mathfrak{p}$) prove that either $\mathfrak{a} \subset \mathfrak{p}$ or $\mathfrak{b} \subset \mathfrak{p}$

12. If

$$\mathfrak{p}_1 \supset \mathfrak{p}_2 \supset \cdots$$

is a decreasing sequence of prime ideals in a ring, show that

$$\mathfrak{p} = \bigcap \mathfrak{p}_i$$

is also a prime ideal.

13. In the ring $R = \mathbb{Q}[X, Y]$, show that the ideal (X) is prime but not maximal.

14. In the ring $R = \mathbb{Q}[\sqrt{2}]$, show that the map that leaves $\mathbb{Q}$ fixed and is defined by

$$\begin{aligned} f: \mathbb{Q}[\sqrt{2}] &\rightarrow \mathbb{Q}[\sqrt{2}] \\ \sqrt{2} &\mapsto -\sqrt{2} \end{aligned}$$

is an isomorphism of rings (so it is an *automorphism* of $\mathbb{Q}[\sqrt{2}]$).

15. Show that the ring $R = \mathbb{Q}[\sqrt{2}]$ is a field by finding a multiplicative inverse for any nonzero element.

16. Suppose R is a ring and $\mathfrak{J}$ is the intersection of all *maximal ideals* of R, i.e.

$$\mathfrak{J} = \bigcap_{\mathfrak{m} \text{ maximal in } R} \mathfrak{m}$$

If $r \in R$ has the property that $r \equiv 1 \pmod{\mathfrak{J}}$, show that r is a *unit* (i.e., has a multiplicative inverse).

17. If $\mathfrak{a}_1, \ldots, \mathfrak{a}_n \subset R$ are distinct ideals with the property that $\mathfrak{a}_i + \mathfrak{a}_j = R$ for any $i \neq j$, show that

$$\mathfrak{a}_i + \prod_{j \neq i} \mathfrak{a}_j = R$$

for any i where the product is take over all the integers $1, \ldots, n$ except i.

18. If $\mathfrak{a}_1, \ldots, \mathfrak{a}_n \subset R$ are distinct ideals with the property that $\mathfrak{a}_i + \mathfrak{a}_j = R$ for any $i \neq j$, and

$$\mathfrak{a} = \bigcap_{i=1}^{n} \mathfrak{a}_i$$

show that

$$\frac{R}{\mathfrak{a}} = \prod_{i=1}^{n} \frac{R}{\mathfrak{a}_i}$$

This is a generalization of the Chinese Remainder Theorem in number theory.

A.1.2. Integral domains and Euclidean Rings. Now we are in a position to define classes of rings with properties like those of the integers. An integral domain is a ring without zero-divisors (see definition A.1.3 on page 342), and a Euclidean ring is one in which a version of the division algorithm (proposition A.1.9 on page 344) applies.

DEFINITION A.1.29. Let R be a ring. Then R is an *integral domain* (or just a *domain*) if, for all $r_1, r_2 \in R$, $r_1 \cdot r_2 = 0$ implies that at least one of r_1 or r_2 is 0.

An element, x, of an integral domain is called *irreducible* if $x = a \cdot b$ implies that $x = u \cdot a$ or $x = u \cdot b$ where u is some unit of the ring (see definition A.1.3 on page 342).

An element, x, is called *prime* if the principal ideal, (x), is prime (see definition A.1.19 on page 349).

REMARK. For instance, $\mathbb{Z}$ is an integral domain but $\mathbb{Z}_6$ is not since $2 \cdot 3 \equiv 0$ (mod 6).

When we discussed the integers, we defined prime numbers as positive. In a general ring, the concept of "> 0" is not well-defined so we have to define irreducible elements "up to multiplication by a unit." It is as if we regarded 2 and -2 as essentially the same prime.

LEMMA A.1.30. *Let $\mathfrak{a} \subset R$ be an ideal in a ring. Then:*
(1) $\mathfrak{a}$ is prime if and only if $R/\mathfrak{a}$ is an integral domain.
(2) $\mathfrak{a}$ is maximal if and only if $R/\mathfrak{a}$ is a field.

PROOF. Let $a, b \in R/\mathfrak{a}$ be the images of $x, y \in R$ under the standard projection

$$R \to R/\mathfrak{a}$$

(see proposition A.1.21 on page 350) Then $a \cdot b = 0 \in R/\mathfrak{a}$ if and only if

$$x \cdot y = 0 \quad (\mathrm{mod}\ \mathfrak{a})$$

which is equivalent to saying that $x \cdot y \in \mathfrak{a}$. If $\mathfrak{a}$ is prime, $x \cdot y \in \mathfrak{a}$ implies that $x \in \mathfrak{a}$ or $y \in \mathfrak{a}$, which means that $a = 0$ or $b = 0$. Conversely, if $a \cdot b = 0 \in R/\mathfrak{a}$ always implies $a = 0$ or $b = 0$, then $x \cdot y \in \mathfrak{a}$ would always imply that $x \in \mathfrak{a}$ or $y \in \mathfrak{a}$.

If $\mathfrak{a}$ is maximal, then it is also prime (see proposition A.1.20 on page 350) so we know that $R/\mathfrak{a}$ is an integral domain. Suppose $x \in R$ projects to $a \neq 0 \in R/\mathfrak{a}$. Since $a \neq 0$, we know that $x \notin \mathfrak{a}$, and since $\mathfrak{a}$ is maximal,

$$\mathfrak{a} + (x) = R$$

so $1 \in \mathfrak{a} + (x)$ and

$$y \cdot x + z = 1$$

for some $z \in \mathfrak{a}$ and $y \cdot x = 1 \pmod{\mathfrak{a}}$ so the image of y in $R/\mathfrak{a}$ is a multiplicative inverse of a.

The converse is left to the reader as an exercise. $\square$

DEFINITION A.1.31. A Euclidean domain, R, is an integral domain that has a function called the *norm*, $N: R \to \mathbb{N}$ that measures the "size" of an element, and such that a version of the division algorithm holds (see proposition A.1.9 on page 344):

> Given elements $a, b \in R$ with $b \nmid a$, there exist elements $q, r \in R$
> such that
>
> $$a = b \cdot q + r$$
>
> with $r \neq 0$ and $N(r) < N(b)$.

REMARK. The term "norm" has at least two unrelated meanings in commutative algebra: the meaning above (which is like the degree of a polynomial) and norms of *field extensions* in section A.2.3 on page 396.

EXAMPLE A.1.32. The ring of rational polynomials, $\mathbb{Q}[X]$ is a Euclidean domain, where the norm is the degree of a polynomial. Any irreducible polynomial generates a prime ideal.

Many basic properties of the integers immediately carry over to Euclidean rings — for instance, we have Bézout's Identity (that he originally proved for the Euclidean ring $\mathbb{R}[X]$):

PROPOSITION A.1.33. *If R is a Euclidean ring and $a, b \in R$, and we define the greatest common divisor, $\gcd(a, b)$ of a and b to be the largest in terms of the norm, then there exist elements $u, v \in R$ such that*

$$\gcd(a, b) = u \cdot a + v \cdot b$$

If a and b have no common divisors (other than 1) then we can find $u, v \in R$ such that

$$1 = u \cdot a + v \cdot b$$

PROOF. Exactly the same as the proof of lemma A.1.12 on page 345, but we replace every occurrence of "minimal" with "nonzero elements with minimal N(*)". $\square$

In fact, we can also prove this for a principal ideal domain:

PROPOSITION A.1.34. *If R is a principal ideal domain, the concept of greatest common divisor is well-defined and, for any two elements $x, y \in R$, there exist elements $u, v \in R$ such that*

$$\gcd(x, y) = u \cdot x + v \cdot y$$

PROOF. If $x, y \in R$, then the ideal $(x, y) \subset R$ is generated by a single element (g), i.e. $(x, y) = (g)$. It follows that $g | x$ and $g | y$ — and $g = u \cdot x + v \cdot y$, which implies that any common divisor of x and y must divide g. We define g to be the greatest common divisor of x and y. $\square$

In rings with greatest common divisor, we can prove:

COROLLARY A.1.35. *Let R be a Euclidean domain or a principal ideal domain, let $r \in R$ be some element, and let*

$$\begin{aligned} r &= p_1^{\alpha_1} \cdot \cdots \cdot p_k^{\alpha_k} \\ &= q_1^{\beta_1} \cdot \cdots \cdot q_\ell^{\beta_\ell} \end{aligned}$$

be factorizations into powers of irreducible elements. Then $k = \ell$ and there is a reordering of indices $f\colon \{1, \ldots, k\} \to \{1, \ldots, k\}$ such that $q_i = u_{f(i)} \cdot p_{f(i)}$ for some units, $u_{f(i)}$, and $\beta_i = \alpha_{f(i)}$ for all i from 1 to k.

PROOF. Simply repeat the proof of proposition A.1.13 on page 346. □

Note that all ideals of R are principal, i.e., generated by a single element. We will be interested in general rings that share this property:

DEFINITION A.1.36. A *principal ideal domain* is an integral domain in which all ideals are principal.

PROPOSITION A.1.37. *All Euclidean domains are principal ideal domains.*

PROOF. Let R be a Euclidean domain with norm $N\colon R \to \mathbb{Z}$ and $\mathfrak{a} \subset R$ be an ideal. If $\mathfrak{a}' = \mathfrak{a} \setminus \{0\}$, let $x \in \mathfrak{a}'$ be a minimal element in the sense that there does not exist any element $y \in \mathfrak{a}'$ with $N(y) < N(x)$. We claim that $\mathfrak{a} = (x)$. If $y \in \mathfrak{a}$ is not a multiple of x, then we can divide y by x to get

$$y = x \cdot q + r$$

Because $\mathfrak{a}$ is an ideal, $x \cdot q \in \mathfrak{a}$. Since $y \in \mathfrak{a}$, it follows that $r \in \mathfrak{a}'$ and $N(r) < N(x)$, which contradicts the minimality of x. □

Another important class of rings are unique factorization domains:

DEFINITION A.1.38. A ring, R, is a *unique factorization domain* if it is a domain whose elements satisfy the conclusion of corollary A.1.35 on page 357, i.e., if factorization of elements into irreducibles is unique up to units.

REMARK A.1.39. Since Bézout's identity was used to prove unique factorization of integers (see proposition A.1.13 on page 346), it follows that any principal ideal domain has unique factorization.

We have already seen several examples of unique factorization domains: the integers, polynomials over the rational numbers.

It is useful to give an example of a ring that is *not* a unique factorization domain. It shows that such examples are fairly common:

EXAMPLE A.1.40. Consider the extension ring $\mathbb{Z}[\sqrt{-5}] \subset \mathbb{C}$. It is the set of all numbers

$$a + b\sqrt{-5}$$

with $a, b \in \mathbb{Z}$. These elements satisfy the multiplication law

(A.1.6) $\quad (a_1 + b_1\sqrt{-5}) \cdot (a_2 + b_2\sqrt{-5}) = a_1 a_2 - 5 b_1 b_2 + (a_1 b_2 + a_2 b_1)\sqrt{-5}$

It is not hard to see that the map $f\colon \mathbb{Z}[\sqrt{-5}] \to \mathbb{Z}[\sqrt{-5}]$ that sends $\sqrt{-5}$ to $-\sqrt{-5}$ is an automorphism (see definition A.1.17 on page 349) — just plug it into equation A.1.6.

If $x = a + b\sqrt{-5} \in \mathbb{Z}[\sqrt{-5}]$, then define

$$N(x) = x \cdot f(x) = a^2 + 5b^2 \in \mathbb{Z}$$

and

(1) $N(x) = 0$ if and only if $x = 0$.
(2) for all $x, y \in \mathbb{Z}[\sqrt{-5}]$,

$$N(x \cdot y) = x \cdot y \cdot f(x \cdot y) = x \cdot y \cdot f(x) \cdot f(y) = N(x) \cdot N(y)$$

since f is a homomorphism. This means that $a|b \in \mathbb{Z}[\sqrt{-5}]$ implies that $N(a)|N(b) \in \mathbb{Z}$.

Now note that $N(2) = 4$ and $N(3) = 9$. The only elements $z = a + b\sqrt{-5}$ with $N(z) \leq 9$ are $1 \pm \sqrt{-5}$. Both have $N(z) = 6$ which does not divide 4 or 9. It follows that the four elements 2, 3, $1 \pm \sqrt{-5} \in \mathbb{Z}[\sqrt{-5}]$ are *irreducible* — i.e., primes.

The formula

$$6 = 2 \cdot 3 = (1 - \sqrt{-5}) \cdot (1 + \sqrt{-5})$$

gives an example of non-unique factorization. So the ring $\mathbb{Z}[\sqrt{-5}]$ is not a unique factorization domain. The function, N, is an example of a *norm of a field-extension*, a topic covered in more detail in section A.2.3 on page 396.

EXERCISES.

19. If $\mathbb{F}$ is a field, show that the equation $x^n = 1$ in $\mathbb{F}$ has at most n solutions.

20. Let $C[0,1]$ be the ring of all real-valued continuous functions on the unit interval, $[0,1]$. If $a \in [0,1]$, let $\mathfrak{f}_a = \{f \in C[0,1] | f(a) = 0\}$. Show that $\mathfrak{f}_a \subset C[0,1]$ is a maximal ideal.

21. Find the greatest common divisor of

$$a(X) = X^4 + 3X^3 - 2X^2 + X + 1$$

and

$$b(X) = X^5 - X^3 + X + 5$$

in $\mathbb{Q}[X]$.

22. Show that there exists integral domains with pairs of elements that have no greatest common divisor. Hint: consider the subring $R \subset \mathbb{Q}[X]$ of polynomials with *no linear term* — i.e., polynomials of the form

$$f(x) = a_0 + a_2 X^2 + \cdots$$

A.1.3. Radicals. We conclude this section by discussing radicals of an ideal. We begin by defining multiplicative sets:

DEFINITION A.1.41. A *multiplicative set, S,* is a set of elements of a ring, *R* that:

(1) contains 1
(2) is closed under multiplication.

Our main application of multiplicative sets will be in constructing rings and modules of fractions in section A.1.8 on page 383. We need this concept here, to prove A.1.47 on the following page.

EXAMPLE. For instance, if $\mathfrak{p} \subset R$ is a prime ideal, then $S = R \setminus \mathfrak{p}$ is a multiplicative set.

We have a kind of converse to this:

PROPOSITION A.1.42. *If $S \subset R$ is a multiplicative set in a commutative ring with $S^{-1}R \neq 0$, then any ideal $\mathfrak{I} \subset R$ with $\mathfrak{I} \cap S = \varnothing$ that is maximal with respect to this property is prime.*

REMARK. "Maximal with respect to this property" means that, given any other ideal $\mathfrak{J}$ with $\mathfrak{J} \cap S = \varnothing$ and $\mathfrak{I} \subset \mathfrak{J}$, then $\mathfrak{I} = \mathfrak{J}$ — i.e. $\mathfrak{I}$ is not *properly* contained in $\mathfrak{J}$.
Such a maximal ideal always exists, by Zorn's Lemma (A.1.27 on page 353).

PROOF. Let $\mathfrak{I}$ be such a maximal ideal and assume it is not prime. The there exist $a, b \in R$ such that $ab \in \mathfrak{I}$ and $a \notin \mathfrak{I}$ and $b \notin \mathfrak{I}$. Then $(a + \mathfrak{I}) \cap S \neq \varnothing$ and $(b + \mathfrak{I}) \cap S \neq \varnothing$. Let $s_1 \in (a + \mathfrak{I}) \cap S$ and $s_2 \in (b + \mathfrak{I}) \cap S$. Then

$$s_1 s_2 \in (a + \mathfrak{I})(b + \mathfrak{I}) \subset ab + a\mathfrak{I} + b\mathfrak{I} + \mathfrak{I}^2 \subset \mathfrak{I}$$

which is a contradiction. $\square$

DEFINITION A.1.43. If $\mathfrak{a}$ is an ideal in a ring *K*, define the *radical* of $\mathfrak{a}$, $\sqrt{\mathfrak{a}}$ to be

$$\{f | f^r \in \mathfrak{a}, \text{ for some } r > 0\}$$

PROPOSITION A.1.44. *The radical of an ideal has the following properties*

- $\sqrt{\mathfrak{a}}$ *is an ideal*
- $\sqrt{\sqrt{\mathfrak{a}}} = \sqrt{\mathfrak{a}}$

PROOF. If $a \in \sqrt{\mathfrak{a}}$, then $a^r \in \mathfrak{a}$ so $f^r a^r = (fa)^r \in \mathfrak{a}$ so $fa \in \sqrt{\mathfrak{a}}$ for all $f \in K$. If $a, b \in \sqrt{\mathfrak{a}}$ and $a^r, b^s \in \mathfrak{a}$. The binomial theorem expands $(a + b)^{r+s}$ to a polynomial in which every term has a factor of a^r or b^s.
If $a^r \in \sqrt{\mathfrak{a}}$ then $a^{rs} \in \mathfrak{a}$. $\square$

DEFINITION A.1.45. An ideal is called *radical* if it equals its own radical.
Equivalently, $\mathfrak{a}$ is radical if and only if $K/\mathfrak{a}$ is a *reduced ring* — a ring without nonzero nilpotent elements. Since integral domains are reduced, prime ideals (and maximal ideals) are radical.

It is not hard to see that intersections of radical ideals are radical. Since $f^r(P) = (f(P))^r$, f^r vanishes wherever f vanishes. It follows that $\mathcal{IV}(\mathfrak{a}) \supset \sqrt{\mathfrak{a}}$. We conclude this section with study of the *nilpotent elements* of a ring.

DEFINITION A.1.46. An element $x \in R$ of a ring is called *nilpotent* if $x^k = 0$ for some integer k. The set of all nilpotent elements of a ring forms an ideal, $\mathfrak{N}(R) = \sqrt{(0)}$, called the *nilradical*.

REMARK. We leave the proof that the set of all nilpotent element forms an ideal as an exercise.

THEOREM A.1.47. *If $\mathfrak{I} \subset R$ is an ideal in a ring, then*

$$\sqrt{\mathfrak{I}} = \bigcap \mathfrak{p}_i$$

where the intersection is taken over all prime ideals that contain $\mathfrak{I}$. Consequently, $\mathfrak{N}(R)$ is equal to the intersection of all prime ideals.

REMARK. Every ideal is contained in a maximal ideal (see proposition A.1.28 on page 353) which is prime by proposition A.1.20 on page 350, so there is always at least one prime in this intersection.

PROOF. Suppose $x \in \sqrt{\mathfrak{I}}$ and $\mathfrak{I} \subset \mathfrak{p}$ where $\mathfrak{p}$ is prime. Then $x^n = x \cdot x^{n-1} \in \mathfrak{I} \subset \mathfrak{p}$. If $x^{n-1} \notin \mathfrak{p}$ then $x \in \mathfrak{p}$. Otherwise, a simple downward induction on n proves that $x \in \mathfrak{p}$. It follows that

$$\sqrt{\mathfrak{I}} \subseteq \bigcap \mathfrak{p}_i$$

where we take the intersection over all prime ideals of R.

If $x \in R \setminus \sqrt{\mathfrak{I}}$, we will construct a prime ideal that *does not* contain x. Note that $S = \{x^n, n = 1, \dots\}$ is a multiplicative set. Proposition A.1.42 on the preceding page show that the maximal ideal that does *not* intersect S is prime. □

We can define a related concept:

A.1.4. Noetherian rings. We fill our menagerie of ring-types (see figure A.1.1 on the next page) with

DEFINITION A.1.48. A ring R is *noetherian* if all of its ideals are finitely generated.

REMARK. This is a generalization of principal ideal domain. The term noetherian is in honor of the mathematician Emmy Noether (1882-1935) whose contributions extend to many fields, including physics (see [126]).

The definition given above is equivalent to the statement:

> All increasing sequences of ideals in R eventually become constant, i.e., if
> $$\mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$$
> then there exists a number n such that $\mathfrak{a}_i = \mathfrak{a}_{i+1}$ for all $i \geq n$.
> This is called the *ascending chain condition* or ACC.

The similar-looking *descending* chain condition leads to a class of rings called Artinian rings — see definition A.1.83 on page 381.

The following result (due to Emmy Noether — see [128]) shows that noetherian rings are extremely common:

LEMMA A.1.49. *If R is noetherian, then so is $R[X]$.*

PROOF. Recall that, for a polynomial

$$f(X) = a_k X^k + \cdots + a_0$$

k is called the degree and a_k is called the leading coefficients. If $\mathfrak{a} \subseteq R[X]$ is an ideal, let $\mathfrak{c}_i$ be the set of all leading coefficients of polynomials in $\mathfrak{a}$ of degree $\leq i$.

Then $\mathfrak{c}_i \subseteq R$ is an ideal and

$$\mathfrak{c}_1 \subseteq \mathfrak{c}_2 \subseteq \cdots \subseteq \mathfrak{c}_i \subseteq \cdots$$

Because R is noetherian, this sequence eventually becomes constant, say $\mathfrak{c}_d = \mathfrak{c}_{d+1} = \cdots$. For each $i \leq d$, let

$$\mathfrak{c}_i = (a_{i,1}, \ldots, a_{i,n(i)}) \subset R$$

and let $f_{i,j} \in \mathfrak{a} \subset R[X]$ be a polynomial whose leading coefficient is $a_{i,j}$. If $f \in \mathfrak{a}$, we will show by induction on the degree of f that it lies in the ideal generated by the (finite) set of $f_{i,j}$.

When f has degree 0, the result is clear. If f has degree $s < d$ then

$$f = aX^s + \cdots$$

with $a \in \mathfrak{c}_s$, and

$$a = \sum_{j=1}^{n(s)} b_j \cdot a_{s,j}$$

for some $b_j \in R$, so

$$f - \sum_{j=1}^{n(s)} b_j \cdot f_{s,j}$$

is a polynomial of degree $s - 1$ and induction implies the conclusion.

If f has degree $s \geq d$, then

$$f = aX^s + \cdots$$

with $a \in \mathfrak{c}_d$. It follows that

$$a = \sum b_j \cdot a_{d,j}$$

for some $b_j \in R$ and that

$$f - \sum_j b_j \cdot f_{d,j} X^{s-d}$$

has degree $< \deg f$, and so lies in the ideal generated by the $\{f_{i,j}\}$ (by induction). $\qquad\square$

Some relations between classes of rings is illustrated in figure A.1.1.
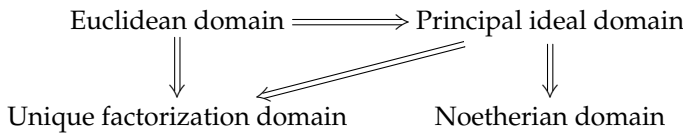
FIGURE A.1.1.   Relations between classes of rings

THEOREM A.1.50 (Hilbert Basis Theorem). *If R is noetherian, then so is $R[X_1, \ldots, X_n]$, i.e., every ideal is finitely generated.*

REMARK. Technically this is Noether's generalization of the Hilbert Basis Theorem. Hilbert originally proved it for R a field.

PROOF. Since R is noetherian, and

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

the theorem follows by an easy induction from lemma A.1.49 on page 360. □

A variation of this argument even shows that *power-series rings* are noetherian.

LEMMA A.1.51. *If R is a noetherian ring, so is $R[[X]]$.*

REMARK. As above, a simple induction shows that

$$R[[X_1, \ldots, X_k]]$$

is noetherian for any finite k.

PROOF. Given a power series $s \in R[[X]]$, let $\ell(s)$ denote its lowest nonzero coefficient.

Let $\mathfrak{a} \subset R[[X]]$ be an ideal whose elements are

$$v = \sum_{i=0}^{\infty} c(v)_i X^i$$

with $c(V)_i \in R$. Define $\mathfrak{h}_n \subset R$ to be composed of the nth coefficients of elements of $\mathfrak{a}$ whose lower coefficients vanish, i.e.

$$\mathfrak{h}_n = \{c(v)_n | v \in \mathfrak{a}, c(v)_0 = \cdots = c(v)_{n-1} = 0\}$$

We claim that the $\mathfrak{h}_n$ are all ideals of R (from the way elements of $R[[X]]$ are multiplied). Now set

$$\mathfrak{H}_n = \bigcup_{i=0}^{n} \mathfrak{h}_n$$

Then we get an ascending chain of ideals in R

$$\mathfrak{H}_0 \subset \cdots$$

which must eventually become constant with some finitely generated ideal (since R is noetherian)

$$\mathfrak{H}_{m+1} = \mathfrak{H}_m = (r_0, \ldots, r_t)$$

So the $\{r_i\}$ generate all of the coefficients of all elements of $\mathfrak{a}$. Each of the r_i is the lowest nonzero coefficient of some power series $f_i \in \mathfrak{a}$. We claim that

$$\mathfrak{a} = (f_0, \ldots, f_t)$$

Given $z = \sum_{i=0}^{\infty} c_i X^i \in \mathfrak{a}$, we must show that there exist power-series $d_i = \sum_{j=0}^{\infty} d_{i,j} X^j$ such that

(A.1.7) $$z = d_1 f_1 + \cdots + d_t f_t$$

If s is the highest degree of the lowest nonzero term that occurs in the f_i, we can subtract R-linear combinations of the f_i from z that will kill off all of its terms of degree $\leq s$ — giving z_1. This gives the *constant terms* of the d_i, i.e. $\{d_{i,0}\}$.

To cancel the lowest term of z_1, we know that its coefficient, c_{s+1}, is also a linear combination of the r_i. We must multiply suitable f_i by X to reach it, thus defining the $d_{1,i}$. Subtracting this linear combination gives z_2.

Continuing this indefinitely results in equation A.1.7 on the preceding page. Despite the seeming "infinite complexity" of z, we express it as a finite linear combination because we have infinite series available to us as coefficients. □

We conclude this section with a result due to Emmy Noether:

LEMMA A.1.52. *Let* $\mathfrak{I} \subset R$ *be an ideal in a noetherian ring. Then:*

(1) in the set of prime ideals $\mathfrak{p}$ *such that* $\mathfrak{I} \subset \mathfrak{p}$, *there is a minimal element*
(2) the set of minimal prime ideals containing $\mathfrak{I}$ *is finite.*

REMARK. Coupled with theorem A.1.47 on page 360, this implies that $\sqrt{\mathfrak{I}}$ is equal to the intersection of a finite number of prime ideals.

PROOF. The first statement follows from:

(1) every ideal is contained in a maximal ideal (see proposition A.1.28 on page 353),
(2) maximal ideals are prime (see proposition A.1.20 on page 350) so every ideal is contained in at least one prime ideal,
(3) the intersection of a decreasing sequence of prime ideals is prime (see exercise 12 on page 354).

We prove the second statement by contradiction. Let $\mathscr{I}$ denote the set of ideals with an infinite number of minimal primes that contain them. Every ascending chain of ideals in $\mathscr{I}$

$$\mathfrak{I}_1 \subset \mathfrak{I}_2 \subset \cdots$$

has an upper bound since the sequence stabilizes after a finite number of terms (this is the only place where we use the noetherian property of R). Zorn's Lemma (A.1.27 on page 353) implies that $\mathscr{I}$ has a maximal member, $\mathfrak{M}$.

Clearly, $\mathfrak{M}$ is not prime because it would be the (one and only) minimal prime containing it. It follows that

(1) there exist $a, b \in R$ such that $a \cdot b \in \mathfrak{M}$ and $a \notin \mathfrak{M}$ and $b \notin \mathfrak{M}$.
(2) if $\mathfrak{A} = (a, \mathfrak{M})$ and $\mathfrak{B} = (b, \mathfrak{M})$, then $\mathfrak{M} \subset \mathfrak{A}$, $\mathfrak{M} \subset \mathfrak{B}$, and $\mathfrak{A} \cdot \mathfrak{B} \subset \mathfrak{M}$

If $\{\mathfrak{p}_i\}$ is the infinite set of minimal primes that contain $\mathfrak{I}$, exercise 11 on page 354 implies that, for each i, $\mathfrak{A} \subset \mathfrak{p}_i$ or $\mathfrak{B} \subset \mathfrak{p}_i$. It follows that $\mathfrak{A}$ or $\mathfrak{B}$ (or both) is contained in an infinite number of the $\mathfrak{p}_i$ — without loss of generality, we will say it is $\mathfrak{A}$. Since $\mathfrak{M} \subsetneq \mathfrak{A}$, it follows that $\mathfrak{A}$ can only have a *finite* number of minimal primes containing it. This is the contradiction. □

EXERCISES.

23. Show that every finite integral domain is a field.

24. Prove the statement following definition A.1.48 on page 360, i.e. that a ring is noetherian if and only if every increasing sequence of ideals is eventually constant.

25. Use the results of exercise 24 to show that any quotient of a noetherian ring is noetherian.

26. Show that $\mathbb{Q}[X, Y]$ is *not* a Euclidean domain.

27. if R is a principal ideal domain and $x = p_1^{n_1} \cdots p_k^{n_k}$ is a factorization into primes, show that

$$\sqrt{(x)} = (p_1 \cdots p_k)$$

28. Find all the maximal ideals of $\mathbb{Z}[X]$.

29. Show that an element of $R[[X_1, \ldots, X_n]]$ is a unit if and only if its constant term is a unit in R.

30. If R is a noetherian ring, show that the nilradical is nilpotent, i.e. that there exists a integer $k > 0$ such that $\mathfrak{N}(R)^k = 0$.

31. Suppose $\mathfrak{p} \subset R$ is a minimal prime ideal in a noetherian (commutative) ring. Show that all of the elements of $\mathfrak{p}$ are zero-divisors.

A.1.5. Polynomial rings and elementary symmetric functions. If R is a commutative ring, consider the polynomial ring

$$P = R[X_1, \ldots, X_n]$$

The symmetric group, S_n, acts on this ring by permuting the variables. Each such permutation of the variables defines an automorphism of P so the set of elements

$$S = R[X_1, \ldots, X_n]^{S_n}$$

fixed by the action of S_n is a *subring* of P. It is interesting that the structure of this subring is completely understood — and was in the time of Isaac Newton. The actual description of this subring will be important in the sequel and is used in several areas of algebraic geometry.

In order to give this description, we will need to define the elementary symmetric functions. The quickest (if not the simplest) way to describe them is to consider a polynomial in the ring $P[t]$ where t is a new indeterminate:

$$(A.1.8) \qquad \prod_{i=1}^{n}(t - X_i) \;\; = \;\; q(t)$$

$$= \;\; t^n - \sigma_1 \cdot t^{n-1} + \cdots + (-1)^n \sigma_n$$

Since $q(t)$ is unchanged when the X_i are permuted, the coefficients of $q(t)$ must be functions of the X_i that are also unchanged by permuting the X_i. They are

$$\sigma_0(X_1, \ldots, X_n) \;\; = \;\; 1$$

$$\sigma_1(X_1, \ldots, X_n) \;\; = \;\; \sum_{i=1}^{n} X_i$$

$$\sigma_2(X_1, \ldots, X_n) \;\; = \;\; \sum_{1 \leq i < j \leq n} X_i X_j$$

$$\vdots$$

$$\sigma_n(X_1, \ldots, X_n) \;\; = \;\; \prod_{i=1}^{n} X_i$$

where $\sigma_i(X_1, \ldots, X_n)$ is $(-1)^i \times$ the coefficient of t^{n-i}.

If we consider the ring $R[\sigma_1, \ldots, \sigma_n]$ of polynomials of the σ_i, it is clear that

$$R[\sigma_1, \ldots, \sigma_n] \subset R[X_1, \ldots, X_n]$$

and even that

$$R[\sigma_1, \ldots, \sigma_n] \subset R[X_1, \ldots, X_n]^{S_n} = S$$

since the σ_i are unchanged by permutations of the X_i. It is remarkable that:

THEOREM A.1.53. *The subring of polynomials of*

$$R[X_1, \ldots, X_n]$$

that are invariant under all permutations of the X_i is precisely the polynomial ring of elementary symmetric functions, i.e.

$$R[\sigma_1, \ldots, \sigma_n] = R[X_1, \ldots, X_n]^{S_n}$$

PROOF. Let $p(X_1, \ldots, X_n) \in R[X_1, \ldots, X_n]^{S_n}$. We will express this as a polynomial of the elementary symmetric functions. Suppose

$$m = r \cdot X_1^{\alpha_1} \cdots X_n^{\alpha_n}$$

is a monomial of p, where $r \in R$. Since p is invariant under permutations of the X_i, p also contains a ordered-monomial that is equivalent to m under the action of S_n, where an ordered monomial is of the form

$$r \cdot X_1^{\beta_1} \cdots X_n^{\beta_n}$$

where $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$ where the β_i's are some permutation of the α_i's. We may focus our attention entirely on ordered-monomials of this type, since every monomial will be equivalent to one of these. The (unique) ordered monomial of $\sigma_i(X_1, \ldots, X_n)$ is

(A.1.9) $$X_1 \cdots X_i$$

Now we order the ordered-monomials of $p(X_1, \ldots, X_n)$ lexicographically by exponents, i.e. so

$$X_1^{\alpha_1} \cdots X_n^{\alpha_n} \succ X^{\beta_1} \cdots X_n^{\beta_n}$$

if $\alpha_j > \beta_j$ and $\alpha_i = \beta_i$, for $i = 1 \ldots j - 1$.

The polynomial p will contain a unique maximal ordered-monomial, say

$$r \cdot X_1^{\beta_1} \cdots X_n^{\beta_n}$$

and this agrees with the unique maximal ordered monomial of

(A.1.10) $$r \cdot \sigma_n^{\beta_n} \cdot \sigma_{n-1}^{\beta_{n-1} - \beta_n} \cdots \sigma_1^{\beta_1 - \beta_2}$$

by equation A.1.9. It follows that the unique maximal ordered monomial of

$$p - r \cdot \sigma_n^{\beta_n} \cdot \sigma_{n-1}^{\beta_{n-1} - \beta_n} \cdots \sigma_1^{\beta_1 - \beta_2}$$

is *strictly* $\prec r \cdot X_1^{\beta_1} \cdots X_n^{\beta_n}$. Since there are only a finite number of monomials $\prec r \cdot X_1^{\beta_1} \cdots X_n^{\beta_n}$, repeating this procedure over and over again must terminate after a finite number of steps. The polynomial p is equal to the sum of the symmetric polynomials we subtracted from p.   □

The proof gives us an algorithm for computing the expression of symmetric polynomials in terms of symmetric functions:

EXAMPLE. Consider
$$X^2 + Y^2 \in \mathbb{Q}[X, Y]$$
The maximal ordered monomial of this is X^2 — which corresponds to σ_1^2 in equation A.1.10 on the previous page. The difference is
$$
\begin{aligned}
X^2 + Y^2 - \sigma_1^2 &= X^2 + Y^2 - (X + Y)^2 \\
&= -2XY
\end{aligned}
$$
which is equal to $-2\sigma_2$. So we get
$$X^2 + Y^2 = \sigma_1^2 - 2\sigma_2$$

An interesting consequence of formula A.1.8 on page 364 and theorem A.1.53 on the previous page is:

PROPOSITION A.1.54. *Let* $X^n + a_{n-1}X^{n-1} + \cdots + a_0 = p(X) \in \mathbb{Q}[X]$ *and suppose* $q(X_1, \dots, X_n) \in \mathbb{Q}[X_1, \dots, X_n]$ *is invariant under permutations of the* X_i. *If* $\mu_1, \dots, u_n \in \mathbb{C}$ *are the roots of* $p(X)$, *then there exists a polynomial* $z(X_1, \dots, X_n)$ *such that*
$$q(\mu_1, \dots, \mu_n) = z(a_0, \dots, a_{n-1})$$

PROOF. Theorem A.1.53 on the preceding page implies that $q(X_1, \dots, X_n) = z(\sigma_1, \dots, \sigma_n)$. Equation A.1.8 on page 364 shows that $\sigma_i(\mu_1, \dots, \mu_n) = (-1)^i a_{n-i}$ and the result follows. $\square$

This has an interesting application in the definition of discriminants of polynomials:

DEFINITION A.1.55. *Let* $p(x) \in \mathbb{Q}[x]$ *be of degree* n *with roots* $\alpha_1, \dots, \alpha_n \in \mathbb{C}$. *The* discriminant, D, *of* $p(X)$ *is defined to be*
$$D = \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j)^2$$

REMARK. The discriminant is nonzero if and only if $p(X)$ has n distinct roots (so it *discriminates* between roots).

Since the discriminant is unchanged by a permutation of the roots, proposition A.1.54 implies that

COROLLARY A.1.56. *If*
$$p(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$$
there is a polynomial function $z(a_0, \dots a_{n-1})$ *equal to the discriminant of* $p(X)$.

For instance, the discriminant of $X^2 + aX + b$ is
$$
\begin{aligned}
(\alpha_1 - \alpha_2)^2 &= \alpha_1^2 - 2\alpha_1\alpha_2 + \alpha_2^2 \\
&= \sigma_1^2(\alpha_1, \alpha_2) - 4\alpha_1\alpha_2 \\
&= \sigma_1^2(\alpha_1, \alpha_2) - 4\sigma_2(\alpha_1, \alpha_2) \\
&= a^2 - 4b
\end{aligned}
$$

A lengthy calculation shows that the discriminant of $X^3 + aX^2 + bX + c$ is

$$D = a^2b^2 - 4b^3 - 4a^3c - 27c^2 + 18abc$$

EXERCISES.

32. Express $X^3 + Y^3 + Z^3 \in \mathbb{Q}[X, Y, Z]$ in terms of elementary symmetric functions.

33. Let $p_1, p_2 \in \mathbb{Q}[t]$ are monic polynomials with roots $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ and $\beta_1, \dots, \beta_m \in \mathbb{C}$, respectively. Let

$$\Delta = \prod(\alpha_i - \beta_j)$$

with i running from 1 to n and j running from 1 to m. Show that Δ is a polynomial function of the coefficients of p_1 and p_2 — equal to the *resultant* of the polynomials — see proposition 1.3.3 on page 13.

34. Another way of defining the discriminant involves using the Vandermonde matrix. Given elements $\alpha_1, \dots, \alpha_n$, we define the corresponding *Vandermonde matrix* as

$$V = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ 1 & \alpha_3 & \alpha_3^2 & \cdots & \alpha_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{bmatrix}$$

Show that

$$\det V = \prod_{1 \le i < j \le n} (\alpha_j - \alpha_i)$$

35. Suppose k is a field and

$$\mathfrak{I} \subset k[X_1, \dots, X_n]$$

is an ideal. If $f(X_1, \dots, X_n) \in k[X_1, \dots, X_n]$ is any polynomial, show that

$$\frac{k[X_1, \dots, X_n]}{\mathfrak{I}} \cong \frac{k[X_1, \dots, X_{n+1}]}{\mathfrak{I} + (X_{n+1} - f(X_1, \dots, X_n))}$$

In other words, show that a variable, like X_{n+1}, that can be expressed in terms of the others is superfluous.

A.1.6. Modules. Modules are a kind of generalization of vector-spaces, or "vector-spaces over a ring:"

DEFINITION A.1.57. If R is a commutative ring, a *module* over R is

(1) an abelian group, A,

(2) an action of R on A, i.e. a map

$$f\colon R \times A \to A$$

such that

$$f(r, *)\colon r \times A \to A$$

is a homomorphism of abelian groups, for all $r \in R$, and

$$f(r_1, f(r_2, a)) = f(r_1 r_2, a)$$

and

$$f(r_1 + r_2, a) = f(r_1, a) + f(r_2, a)$$

This action is usually written with a product-notation, i.e. $f(r, a) = r \cdot a$ (in analogy with multiplication by scalars in a vector space).

If $B \subset A$ is a subgroup with the property that $r \cdot B \subset B$ for all $r \in R$, then B is called a *submodule* of A.

EXAMPLE. We can regard a ring, R, as a module over *itself.* Its *submodules* are precisely its ideals.

If $g\colon R \to S$ is a homomorphism of rings, S naturally becomes a module over R by defining $r \cdot s = g(r)s$ for all $r \in R$ and $s \in S$.

EXAMPLE A.1.58. If R is a ring and $R^n = \bigoplus_{i=1}^n R$, then R^n is a module over R with the action defined by multiplication in R. This is called the *free module of rank n over R*. An n-dimensional vector space over a field k is a free module of rank n over that field.

PROPOSITION A.1.59. *Every R-module is a quotient of a free R-module.*

PROOF. If M is an R-module with generating set $\{m_1, \dots\}$, let F be free on generators $\{e_1, \dots\}$ — one for each of the m_i. We define a map

$$f\colon F \to M$$
$$\sum r_j e_j \mapsto \sum r_j m_j$$

This is clearly surjective, so that

$$\frac{F}{\ker f} \cong M$$

$\square$

It is possible to come up with more "exotic" examples of modules:

EXAMPLE A.1.60. Let V be an n-dimensional vector space over a field F and let M be an $n \times n$ matrix over F. Then V is a module over the polynomial ring $k[X]$, where a polynomial, $p(X) \in k[X]$ acts via

$$p(M)\colon V \to V$$

In other words, we plug M into $p(X)$ to get a matrix and then act on V via that matrix.

Note that a vector-subspace $W \subset V$ is a *submodule* if and only if $M(W) \subset W$. It follows that the module-structure of V over $k[X]$ depends strongly on the matrix M.

DEFINITION A.1.61. Let M_1 and M_2 be modules over the same ring, R. A homomorphism of modules is a map of their underlying abelian groups

$$f\colon M_1 \to M_2$$

such that $f(r \cdot m) = r \cdot f(m)$ for all $m \in M$ and $r \in R$. The set of elements $m \in M_1$ with $f(m) = 0$ is called *the kernel of f* and denoted ker f. The set of elements $m \in M_2$ of the form $f(n)$ for some $n \in M_1$ is called the *image of f* and denoted im f.

If ker $f = 0$, the homomorphism f is said to be *injective*. If im $f = M_2$, the homomorphism is said to be *surjective*. If f is both injective and surjective, it is called an *isomorphism*.

REMARK. Note that, if f above is injective, we can regard M_1 as a submodule of M_2. Isomorphic modules are algebraically equivalent.

The corresponding statements about abelian groups imply that

PROPOSITION A.1.62. *Let M be a module over a ring R and let A and B be submodules of M. Then:*

(1) *we can define the quotient M/A as the set of equivalence classes of the equivalence relation*

$$m_1 \equiv m_2 \pmod{A}$$

if $m_1 - m_2 \in A$, for all $m_1, m_2 \in M$. We can also define M/A as the set of cosets $\{m + A\}$ for $m \in M$.

(2) *the map*

$$p\colon M \to M/A$$

sending an element to its equivalence class, is a homomorphism of modules.

(3) *the map p defines a 1-1 correspondence between submodules of M containing A and submodules of M/A*

(4) *there is a canonical isomorphism*

$$\frac{A+B}{A} \cong \frac{B}{A \cap B}$$

DEFINITION A.1.63. If $f\colon M_1 \to M_2$ is a homomorphism of modules, the quotient

$$\frac{M_2}{f(M_1)}$$

is called the *cokernel* of f.

REMARK. Since one can form quotients of modules with respect to arbitrary submodules, cokernels always exist for module-homomorphisms.

DEFINITION A.1.64. A sequence of modules and homomorphisms (all over the same ring)

$$\cdots \xrightarrow{f_{n+1}} M_{n+1} \xrightarrow{f_n} M_n \xrightarrow{f_{n-1}} M_{n-1} \to \cdots$$

is said to be *exact* if im $f_{n+1} = $ ker f_n for all n. An exact sequence with five terms like

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

is called a *short exact sequence*.

REMARK. In the *short* exact sequence above, the kernel of f must be 0, so A can be identified with a submodule of B, and the map g must be surjective (since the kernel of the rightmost map is all of C).

The exactness of the (long) sequence above is equivalent to saying that the short sequences

$$0 \to \operatorname{im} f_n \to M_n \to \operatorname{im} f_{n-1} \to 0$$

are exact for all n.

Exact sequences are widely used in homological algebra and algebraic topology, facilitating many types of computations.

DEFINITION A.1.65. If M is a module over a ring R, a set of elements $S = \{m_1, \dots\} \in M$ will be called a *generating set* if every element $m \in M$ can be expressed in terms of the elements of S

$$m = \sum_{m_i \in S} r_i \cdot m_i$$

with the $r_i \in R$.

A module is said to be *finitely generated* if it has a finite generating set.

EXAMPLE A.1.66. As in example A.1.60 on page 368Let V be an n–dimensional vector space over a field F and let M be the $n \times n$ permutation matrix

$$M = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Then, as a module over $k[X]$ (defined as in example A.1.60 on page 368), V has a *single generator,* namely

$$g = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

This is because

$$Mg = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

and $M^k g = g_k$, the kth basis element of V. So all of the basis-elements of V are in the orbit of powers of M and of $k[X]$.

It is interesting to consider what properties of vector-spaces carry over to modules, or whether we can do a kind of "linear algebra" over a general ring. This is a deep field of mathematics that includes several areas, such as group-representations (see [50]), homological algebra (see [173]) and algebraic K-theory (see [114]).

Even a simple question like

> "Is a submodule of a finitely generated module finitely generated?"

can have a complex answer. For instance, let $R = k[X_1, \dots]$ — a polynomial ring over an infinite number of variables. It is finitely generated as a module over itself (generated by 1). The submodule of polynomials with vanishing constant term is *not* finitely generated since every polynomial has a finite number of variables.

We need to find a class of modules that is better-behaved.

DEFINITION A.1.67. A module M over a ring R will be called *noetherian* if all of its submodules are finitely generated — this is equivalent to saying that all ascending chains of submodules of M

$$(A.1.11) \qquad M_1 \subset M_2 \subset \cdots \subset M_i \subset \cdots$$

becomes constant from some finite point on, i.e. $M_t = M_{t+i}$ for all $i > 0$. A module will be said to be *Artinian* if every *descending* chain of submodules

$$M_1 \supset M_2 \supset \cdots \subset M_i \supset \cdots$$

becomes constant from some finite point on.

The *length* of a noetherian module, M, denoted length(M) is the length of the longest possible ascending chain of submodules as in A.1.11.

REMARK. A ring is noetherian if and only if it is noetherian as a module over itself. The length of a module is a direct generalization of the dimension of a vector space.

PROPOSITION A.1.68. *Let*

$$0 \to M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3 \to 0$$

be a short exact sequence (see definition A.1.64 on page 369) of modules over a ring. Then M_2 is noetherian or Artinian if and only if M_1 and M_3 are both noetherian or Artinian, respectively.

PROOF. We will prove this in the noetherian case; the Artinian case is almost identical. Clearly, if M_2 is noetherian, M_1 will inherit this property since it is a submodule. Any increasing chain of submodules of M_3 will lift to one in M_2, which becomes constant from some finite point on. It follows that M_2 being noetherian implies that M_1 and M_3 are also noetherian.

Conversely, suppose that M_1 and M_3 are noetherian and

$$N_1 \subset N_2 \subset \cdots \subset N_i \subset \cdots$$

is an increasing sequence of submodules of M_2. Since M_3 is noetherian, this image of this sequence in M_3 will become constant from some finite point on, say k. Then

$$(A.1.12) \qquad \frac{N_i}{N_i \cap M_1} = \frac{N_{i+1}}{N_{i+1} \cap M_1}$$

for $i > k$. Since M_1 is noetherian, the sequence

$$N_j \cap M_1 \subset N_{j+1} \cap M_2 \subset \cdots$$

will become constant from some finite point on — say $j = t$. Then, for $i > \max(k, t)$, equation A.1.12 and

$$N_i \cap M_1 = N_{i+1} \cap M_1$$

imply that $N_i = N_{i+1}$.   □

COROLLARY A.1.69. *If R is a noetherian ring, then R^n is a noetherian module.*

PROOF. This is a simple induction on n. If $n = 1$, $R^1 = R$ is noetherian over itself since it is a noetherian ring. For $n > 1$ we use proposition A.1.68 on the previous page with the short exact sequence

$$0 \to R \xrightarrow{f} R^{n+1} \xrightarrow{g} R^n \to 0$$

where

$$\begin{aligned} g(r_1, \ldots, r_{n+1}) &= (r_1, \ldots, r_n) \\ f(r) &= (0, \ldots 0, r) \end{aligned}$$

where the image of f is the $n + 1$st entry in R^{n+1}.   □

Now we can define a large class of well-behaved modules:

LEMMA A.1.70. *If R is a noetherian ring, a module, M over R is noetherian if and only if it is finitely generated.*

PROOF. If M is noetherian it must be finitely generated (since its submodules, including itself, are). Suppose M is finitely generated, say by generators $(a_1, \ldots, a_n)$. Then there exists a surjective homomorphism of modules

$$\begin{aligned} R^n &\xrightarrow{f} M \\ (r_1, \ldots, r_n) &\mapsto r_1 \cdot a_1 + \cdots + r_n \cdot a_n \end{aligned}$$

This map fits into a short exact sequence

$$0 \to \ker f \to R^n \xrightarrow{f} M \to 0$$

and proposition A.1.68 on the preceding page and corollary A.1.69 imply the conclusion.   □

Although noetherian modules are somewhat "well-behaved," they still are more complex than vector-spaces. For instance, a subspace of a vector space of dimension n must have dimension $< n$. The ring $k[X, Y, Z]$ is a module over itself with *one* generator: 1. On the other hand, the ideal $(X, Y, Z) \subset k[X, Y, Z]$, is a proper submodule that requires *three* generators.

The most "straightforward" modules are the *free* ones like R^n above. They are closely related to projective modules:

DEFINITION A.1.71. If R is a ring and P is an R-module, then P is said to be *projective* if it is a direct summand of a free module.

REMARK. In other words, P is projective if there exists an R-module, Q, such that $P \oplus Q = R^n$ for some n. All free modules are (trivially) projective but not all projective modules are free. For instance, if $R = \mathbb{Z}_6$, note that $\mathbb{Z}_2 \oplus \mathbb{Z}_3 = \mathbb{Z}_6$ as rings so $\mathbb{Z}_2$ and $\mathbb{Z}_3$ are projective modules that are not free.

Projective modules have an interesting property that is often (usually?) used to define them:

PROPOSITION A.1.72. *Let R be a ring and let P be a projective module over R. If $\alpha\colon M \to N$ is a surjective homomorphism of R-modules and $\beta\colon P \to N$ is any homomorphism, then a homomorphism, $\gamma\colon P \to M$ exists that makes the diagram*

(A.1.13)
$$
\begin{array}{ccc}
 & & P \\
 & \swarrow^{\gamma} & \downarrow{\scriptstyle\beta} \\
M & \xrightarrow[\alpha]{} & N
\end{array}
$$

commute.

PROOF. Since P is projective, there exists a module Q such that $P \oplus Q = R^n$ for some value of n. Consider the diagram

$$
\begin{array}{ccc}
 & & P \oplus Q \\
 & & \downarrow{\scriptstyle\beta\oplus 1} \\
M \oplus Q & \xrightarrow[\alpha\oplus 1]{} & N \oplus Q
\end{array}
$$

and note that

$$
P \oplus Q = R^n = \bigoplus_{i=1}^{n} x_i \cdot R
$$

where the $x_i \in R^n$ are its generators. Since α is surjective, $\alpha \oplus 1$ will also be, and we can choose $y_i \in M \oplus Q$ such that $(\alpha \oplus 1)(y_i) = (\beta \oplus 1)(x_i)$ for $i = 1, \ldots, n$. Then we can define a homomorphism

$$
\begin{array}{ccc}
G\colon R^n & \to & M \oplus Q \\
x_i & \mapsto & y_i
\end{array}
$$

making the diagram

$$
\begin{array}{ccc}
 & & P \oplus Q \\
 & \swarrow^{G} & \downarrow{\scriptstyle\beta\oplus 1} \\
M \oplus Q & \xrightarrow[\alpha\oplus 1]{} & N \oplus Q
\end{array}
$$

commute. Since $(\alpha \oplus 1)(Q) = Q$, we can extend diagram this to a commutative diagram

$$
\begin{array}{ccc}
 & & P \oplus Q \\
 & \swarrow^{G} & \downarrow{\scriptstyle\beta\oplus 1} \\
M \oplus Q & \xrightarrow[\alpha\oplus 1]{} & N \oplus Q \\
{\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle p_2} \\
M & \xrightarrow[\alpha]{} & N
\end{array}
$$

where the p_i are projections to the factors. If γ is the composite $P \hookrightarrow P \oplus Q \xrightarrow{G} M \oplus Q \xrightarrow{p_1} M$, it will have the desired properties. $\qquad\square$

Once we know a module is finitely generated, we can prove many other interesting results. It is possible to represent modules over a noetherian ring in terms of prime ideals of that ring.

DEFINITION A.1.73. Let $m \in M$ be an element of a module over a ring, R. Then the *annihilator* of m, denoted ann(m), is defined by

$$\text{ann}(m) = \{r \in R | r \cdot m = 0\}$$

The annihilator of M is defined by

$$\text{Ann}(M) = \{r \in R | \forall_{m \in M} \ r \cdot m = 0\} = \bigcap_{m \in M} \text{ann}(m)$$

A prime ideal $\mathfrak{p} \subset R$ is *associated* to M if it annihilates an element $m \in M$. The set of associated primes is denoted Assoc(M).

REMARK. It is not hard to see that ann(m), Ann$(M) \subset R$ are always ideals. The following properties are also easy to verify:

(1) If $m \in M$ then ann$(m) = R$ if and only if $m = 0$.
(2) If $m \in M$ and $s \in R$, then ann$(m) \subseteq \text{ann}(s \cdot m)$.

It is not at all obvious that *any* associated primes exist, since Ann(M) is usually not prime.

EXAMPLE A.1.74. Let $R = \mathbb{C}[X]$ and let A be an $n \times n$ matrix over $\mathbb{C}$. If $V = \mathbb{C}^n$ is a vector-space, we can make it a module over R by defining

$$X \cdot v = Av$$

for all $v \in V$. We know that some element of R annihilates V because the Cayley-Hamilton theorem (see chapter 6 of [**155**]) states that A "satisfies" its characteristic polynomial, i.e., $\chi_A(A) = 0$. Because R is a principal ideal domain, the annihilator of V is a principal ideal $(p(X))$ such that $\chi_A(X) \in (p(X))$, i.e. $p(X)|\chi_A(X)$. The polynomial, $p(X)$, is called the *minimal polynomial* of A.

In general, the minimal polynomial of a matrix is not equal to its characteristic polynomial. For instance, if $A = 3 \cdot I$, where I is the identity matrix then $p(X) = X - 3$ and $\chi_A(X) = (X - 3)^n$.

In studying the structure of ideals that annihilate elements of a module, we begin with:

LEMMA A.1.75. *Let M be a finitely generated module over a noetherian ring R. If $\mathfrak{I} \subset R$ is an ideal that is maximal among ideals of R of the form* ann(m) *for $m \neq 0 \in R$, then $\mathfrak{I}$ is prime.*

REMARK. We will construct an ascending chain of ideals — the fact that R is noetherian implies that this chain has a maximal element.

This shows that, for a finitely generated module over a noetherian ring, at least *one* associated prime exists.

PROOF. Suppose $r, s \in R$ and $rs \in \mathfrak{I}$ but $s \notin \mathfrak{I}$. Then we will show that $r \in \mathfrak{I}$. We have

$$rs \cdot m = 0$$

but $s \cdot m \neq 0$. It follows that $(r) + \mathfrak{I}$ annihilates $sm \in M$. Since $\mathfrak{I}$ is maximal among ideals that annihilate elements of M, we have $(r) + \mathfrak{I} \subseteq \mathfrak{I}$ so $r \in \mathfrak{I}$. $\square$

We go from this to show that *many* associated primes exist:

COROLLARY A.1.76. *Let M be a finitely generated module over a noetherian ring, R. If $Z \subset R$ is the set of elements that annihilate nonzero elements of M, then*

$$Z = \bigcup_{\mathfrak{p} \in \mathrm{Assoc}(M)} \mathfrak{p}$$

PROOF. The definition of $\mathrm{Assoc}(M)$ implies that

$$\bigcup_{\mathfrak{p} \in \mathrm{Assoc}(M)} \mathfrak{p} \subset Z$$

If $x \in Z$, then $x \cdot m = 0$ for $m \in M$, $m \neq 0$. The submodule $R \cdot m \subset M$ has an associated prime, $\mathfrak{p} = \mathrm{ann}(y \cdot m)$, by lemma A.1.75 on page 374, which is also an associated prime to M. Since $x \cdot m = 0$, it follows that $xy \cdot m = 0$ so that $x \in \mathrm{ann}(y \cdot m) = \mathfrak{p}$. $\qquad \square$

Our main result classifying the structure of modules is

THEOREM A.1.77. *Let M be a finitely generated module over a noetherian ring R. Then there exist a finite filtration*

$$0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_n = M$$

such that each

$$\frac{M_{i+1}}{M_i} \cong \frac{R}{\mathfrak{p}_i}$$

for prime ideals $\mathfrak{p}_i \subset R$.

REMARK. This sequence $\{M_i\}$ is called the *prime filtration* of M, and the primes $\{\mathfrak{p}_i\}$ that occur here are called the *prime factors* of M. Note that the $\{\mathfrak{p}_i\}$ might not all be distinct — a given prime ideal may occur more than once (see examples A.1.78 on the following page and A.1.79 on the next page).

The associated primes occur among the primes that appear in this decomposition, so that $\mathrm{Assoc}(M)$ is *finite* (for finitely generated modules over a noetherian ring)

PROOF. Lemma A.1.75 on the facing page states that the maximal ideal, $\mathfrak{p}$, that annihilates an element $m \in M$ is prime. Consider the submodule, M_0, generated by this $m \in M$, i.e., $R \cdot m \subset M$. We get a homomorphism of modules

$$\begin{aligned} R &\rightarrow R \cdot m \\ r &\mapsto r \cdot m \end{aligned}$$

Since $\mathfrak{p}$ is in the kernel, we get a homomorphism of R-modules

$$g \colon R/\mathfrak{p} \rightarrow R \cdot m$$

Since $\mathfrak{p}$ is *maximal,* any element in the kernel of g must lie in $\mathfrak{p}$, so

$$\frac{R}{\mathfrak{p}} \cong R \cdot m = M_1$$

Now, form the quotient M/M_1 and carry out the same argument, forming a submodule $M_2' \subset M/M_1$ and its inverse image over the projection

$$M \rightarrow M/M_1$$

is M_2. We continue this process over and over until we get to 0. It must terminate after a finite number of steps because M is finitely generated over a noetherian ring (see definition A.1.67 on page 371). $\qquad\square$

EXAMPLE. If $M = \mathbb{Z}_{60}$ is regarded as a module over $\mathbb{Z}$, then a maximal ideal is of the form $(p) \subset \mathbb{Z}$ for some prime p. For instance, (2) annihilates $M_1 = 30 \cdot \mathbb{Z}_{60}$ and $M/M_1 = \mathbb{Z}_{30}$. The ideal (2) annihilates $M_2 = 15 \cdot \mathbb{Z}_{30}$ and we get $M_1/M_2 = \mathbb{Z}_{15}$. The ideal (3) annihilates $5 \cdot \mathbb{Z}_{15}$ and we are done (the final quotient is $\mathbb{Z}_5$). We can lift these modules up into M to get

$$0 \subset 30 \cdot \mathbb{Z}_{60} \subset 15 \cdot \mathbb{Z}_{60} \subset 5 \cdot \mathbb{Z}_{60} \subset \mathbb{Z}_{60}$$

with prime factors, $\mathbb{Z}_2$, $\mathbb{Z}_2$, $\mathbb{Z}_3$ and $\mathbb{Z}_5$, respectively.

EXAMPLE A.1.78. Returning to example A.1.74 on page 374, let $\mathfrak{p} \subset R$ be a prime ideal that annihilates an element $v \in V$. Then $\mathfrak{p} = (X - \lambda)$ and λ must be an *eigenvalue.* The element, v, annihilated by $\mathfrak{p}$ is the corresponding *eigenvector.* A simple induction shows that all of the prime ideals we get in the prime decomposition of V are of the form $(X - \lambda_i)$ where the λ_i run through the eigenvalues of A.

Here's a much more detailed example:

EXAMPLE A.1.79. Let $R = \mathbb{C}[X, Y]/\mathfrak{a}$ where

(A.1.14) $\qquad\qquad \mathfrak{a} = (Y^3, XY + Y, X^2 + Y^2 - 1) \subset \mathbb{C}[X, Y]$

Lemma A.1.25 on page 352 implies that the prime ideals of R are images under the projection

$$p: \mathbb{C}[X, Y] \to \mathbb{C}[X, Y]/\mathfrak{a} = R$$

of the prime ideals of $\mathbb{C}[X, Y]$ that contain $\mathfrak{a}$. We will skip ahead and use theorem 2.2.3 on page 40 to conclude that the prime ideals of $\mathbb{C}[X, Y]$ are either of the form $(f(X, Y))$ for some irreducible polynomial, f, or of the form $(X - \alpha, Y - \beta)$ for $\alpha, \beta \in \mathbb{C}$. We reject the possibility of a principal ideal because $f(X, Y)|Y^3$ implies that $f(X, Y) = Y$ but that does not divide $X^2 + Y^2 - 1$.

If $\mathfrak{a} \subset (X - \alpha, Y - \beta)$, then equations like

$$\begin{array}{rcl}
p_1(X, Y)(X - \alpha) + q_1(X, Y)(Y - \beta) & = & Y^3 \\
p_2(X, Y)(X - \alpha) + q_2(X, Y)(Y - \beta) & = & XY + Y \\
p_3(X, Y)(X - \alpha) + q_3(X, Y)(Y - \beta) & = & X^2 + Y^2 - 1
\end{array}$$

must hold, for some $p_i, q_i \in k[X, Y]$. The top equation forces $\beta = 0$ (i.e., set $X = \alpha$). This also satisfies the second equation since we can set $p_2 = 0$ and $q_2 = X + 1$. The bottom equation becomes

$$p_3(X, Y)(X - \alpha) + q_3(X, Y)Y = X^2 + Y^2 - 1$$

We simplify this by setting $q_3 = Y$ and subtracting to get

$$p_3(X, Y)(X - \alpha) = X^2 - 1 = (X - 1)(X + 1)$$

which implies that $\alpha = \pm 1$. It follows that $\mathfrak{P}_1 = (X - 1, Y)$ and $\mathfrak{P}_2 = (X + 1, Y)$ are the only two ideals of the form $(X - \alpha, Y - \beta)$ that contain $\mathfrak{a}$.

Let $x, y \in R$ be the images of X and Y, respectively under the projection, p above — they clearly generate R as a ring. Then the prime ideals of R are

$\mathfrak{p}_1 = (x - 1, y)$ and $\mathfrak{p}_2 = (x + 1, y)$. In addition, lemma A.1.26 on page 352 implies that

$$\frac{R}{\mathfrak{p}_i} = \frac{\mathbb{C}[X, Y]}{\mathfrak{P}_i} = \mathbb{C}$$

for $i = 1, 2$.

Now we will compute a prime filtration of R as a module over itself. Since $y^2 \cdot \mathfrak{p}_2 = 0$, we regard $y^2 \cdot R$ as a candidate for R_1. Let us compute what happens to $\{1, x, y\}$ when we multiply by y^2

$$y^2 \cdot 1 = y^2$$
$$y^2 \cdot x = y \cdot (-y) = -y^2 \quad \text{because of the relation } xy + y = 0$$
$$\text{in equation A.1.14}$$

$$y^2 \cdot y = 0 \qquad\qquad\qquad \text{because of the relation } y^3 = 0$$
$$\text{in equation A.1.14}$$

It follows that $y^2 \cdot R = \mathbb{C} \cdot y^2 = \mathbb{C}$ and we get an isomorphism

$$R_1 = y^2 \cdot R = \frac{R}{\mathfrak{p}_2} = \mathbb{C} \cdot y^2$$

Following the proof of theorem A.1.77 on page 375, we form the quotient

$$\text{(A.1.15)} \quad R' = \frac{R}{R_1} = \frac{k[X, Y]}{(Y^3, Y^2, XY + Y, X^2 + Y^2 - 1)}$$

$$= \frac{k[X, Y]}{(Y^2, XY + Y, X^2 - 1)}$$

— where we eliminated Y^3 since $(Y^3) \subset (Y^2)$ and eliminated the Y^2 term from $X^2 + Y^2 - 1$. Notice that $\mathfrak{p}_2 \cdot y = 0$. This suggests using $y \cdot R'$ as our second prime quotient. As before, we enumerate the effect of $y \cdot$ on the generators of R':

$$y \cdot 1 = y$$
$$y \cdot x = -y \quad \text{because of the relation } xy + y = 0$$
$$\text{in equation A.1.15}$$

$$y \cdot y = 0 \qquad \text{because of the relation } y^2 = 0$$
$$\text{in equation A.1.15}$$

Again, we conclude that $y \cdot R' = \mathbb{C}$, generated by y, and get

$$\frac{R}{\mathfrak{p}} = y \cdot R' = \mathbb{C} \cdot y$$

and we take the inverse image of $y \cdot R'$ over the projection $R \to R/R_1 = R'$ to get $R_2 = y \cdot R$ and a *partial* prime filtration

$$0 \subsetneq y^2 \cdot R \subsetneq y \cdot R$$

Continuing, we form the quotient again

$$\text{(A.1.16)} \qquad R'' = \frac{R'}{\mathbb{C} \cdot y} = \frac{k[X, Y]}{(Y^2, Y, XY + Y, X^2 - 1)} = \frac{k[X, Y]}{(Y, X^2 - 1)}$$

and notice that $\mathfrak{p}_2 \cdot (x - 1) = 0$. Computing $(x - 1) \cdot R''$ gives

$$(x - 1) \cdot 1 = (x - 1)$$
$$(x - 1) \cdot x = x^2 - x = -(x - 1) \quad \text{because of the relation}$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad x^2 - 1 = 0 \text{ in equation A.1.16}$$
$$(x - 1) \cdot y = 0 \qquad\qquad\qquad \text{because of the relation } y = 0$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in equation A.1.16}$$

so $(x - 1) \cdot R'' = \mathbb{C}$, generated by $x - 1$ and

$$\frac{R}{\mathfrak{p}_2} = (x - 1) \cdot R'' = \mathbb{C} \cdot (x - 1)$$

and this lifts to $(x - 1) \cdot R \subset R$. In the final step, we get

$$R''' = \frac{R''}{(x - 1) \cdot R''} = \frac{k[X, Y]}{(Y, X^2 - 1, X - 1)} = \frac{k[X, Y]}{(Y, X - 1)} = \frac{R}{\mathfrak{p}_1}$$

(since $(X^2 - 1) \subset (X - 1)$), so we get our complete prime filtration

$$0 \subsetneq y^2 \cdot R \subsetneq y \cdot R \subsetneq (x - 1) \cdot R \subsetneq R$$

The prime $\mathfrak{p}_2$ occurs three times, and the last factor involves the prime $\mathfrak{p}_1$.

Another interesting and useful result is called *Nakayama's Lemma* — it has a number of applications to algebraic geometry and other areas of algebra:

LEMMA A.1.80 (Nakayama's Lemma). *Let M be a finitely-generated module over a commutative ring, R. If $\mathfrak{a} \subset R$ is an ideal with the property that*

$$\mathfrak{a} \cdot M = M$$

then there exists an element $r \in R$ such that $r \equiv 1 \pmod{\mathfrak{a}}$ and

$$r \cdot M = 0$$

REMARK. This result is named after Tadashi Nakayama who introduced it in [**124**]. Special cases of it had been discovered earlier by Krull, and Azumaya had published the general case in [**9**] before Nakayama's paper. The version for noncommutative rings is called the Krull-Azumaya Theorem.

PROOF. Let $m_1 \ldots, m_k$ denote the generators of M over R so

$$M = R \cdot m_1 + \cdots + R \cdot m_k$$

Since $\mathfrak{a} \cdot M = M$, we have

$$m_i = \sum_{j=1}^{k} A_{i,j} m_j$$

for some $k \times k$ matrix $A = [A_{i,j}]$ with entries in $\mathfrak{a}$. Subtracting gives

$$\sum_{j=1}^{n} (\delta_{i,j} - A_{i,j}) m_j = 0$$

where $\delta_{i,j}$ is the $(i, j)^{\text{th}}$ entry of the identity matrix, or

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

or

$$(I - A) \begin{bmatrix} m_1 \\ \vdots \\ m_k \end{bmatrix} = 0$$

Cramer's Rule implies that

$$\det(I - A)m_i = C_i = 0$$

for all i, where C_i is the determinant of the matrix one gets by replacing the ith column by 0's. So $r \in R$ in the statement of the lemma is just $\det(I - A)$.

We claim that $\det(I - A) = 1 + a$ for some $a \in \mathfrak{a}$. The determinant of $I - A$ is what one gets from the characteristic polynomial $p_A(x)$ by setting $x = 1$. Since the characteristic polynomial is monic, one term is equal to 1 and the remaining terms are linear combinations of elements of $\mathfrak{a}$. $\square$

Here's a consequence of Nakayama's lemma (a special case of the Krull Intersection Theorem — see [**97**]):

LEMMA A.1.81. *Let* $\mathfrak{m} \subset R$ *be a maximal ideal of a noetherian ring, R or an arbitrary ideal of a noetherian domain. Then*

$$\bigcap_{j=1}^{\infty} \mathfrak{m}^j = (0)$$

PROOF. Call this infinite intersection $\mathfrak{b}$. Since R is noetherian, $\mathfrak{b}$ is finitely generated as a module over R. Since

$$\mathfrak{m} \cdot \mathfrak{b} = \mathfrak{b}$$

Nakayama's Lemma (A.1.80 on the preceding page) implies that $\mathfrak{b}$ is annihilated by an element $x \in R$ such that $x \equiv 1 \pmod{\mathfrak{m}}$ and such an element is a unit so $\mathfrak{b} = (0)$.

If R is an integral domain and $\mathfrak{m}$ is an arbitrary ideal, $x \neq 0$ and $x \cdot \mathfrak{b} = 0$ implies $\mathfrak{b} = 0$ since R has no zero-divisors. $\square$

We also get a result for local rings:

COROLLARY A.1.82. *Let R be a local ring (see definition A.1.24 on page 351) with unique maximal ideal $\mathfrak{m} \subset R$. If M is an R-module with the property that*

$$\mathfrak{m} \cdot M = M$$

then $M = 0$.

PROOF. Nakayama's lemma implies that there exists $r \in R$ such that $r \equiv 1 \pmod{\mathfrak{m}}$ and $r \cdot M = 0$. Since R is a *local* ring, $\mathfrak{m}$ is the *only* maximal ideal and therefore equal to the intersection of *all* maximal ideals. Exercise 16 on page 354 implies that this r is a unit, i.e., has a multiplicative inverse, $s \in R$. Consequently

$$r \cdot M = 0 \implies s \cdot r \cdot M = 0 \implies 1 \cdot M = 0$$

and the conclusion follows. $\square$

EXERCISES.

36. If

$$0 \to U \to V \to W \to 0$$

is a short exact sequence of vector-spaces, show that

$$\dim V = \dim U + \dim W$$

37. Prove this basic result in linear algebra:

> A vector-space over an infinite field cannot be a finite union of proper subspaces.

38. Give a counterexample to statement in exercise 37 if the field of definition is *finite.*

39. If P and M are R-modules, with P projective and

$$f: M \to P$$

is a surjective homomorphism, show that there exists a homomorphism $g: P \to M$ such that $f \circ g = 1: P \to P$.

40. If

$$0 \to M_1 \to M_2 \to M_3 \to 0$$

is a short exact sequences of modules over a ring R, show that

$$\mathrm{Ann}(M_1) \cdot \mathrm{Ann}(M_3) \subset \mathrm{Ann}(M_2) \subset \mathrm{Ann}(M_1) \cap \mathrm{Ann}(M_3)$$

41. Let

$$0 \to U \xrightarrow{q} V \xrightarrow{p} W \to 0$$

be a short exact sequence of modules, and suppose there exists a homomorphism

$$h: W \to V$$

such that $p \circ h = 1: W \to W$ (such short exact sequences are said to be *split* and h is called a *splitting map).* Show that there exists an isomorphism

$$V \cong U \oplus W$$

42. Let

$$0 \to U \xrightarrow{q} V \xrightarrow{p} P \to 0$$

be a short exact sequence of modules, and suppose that P is a projective module. Show that there exists an isomorphism

$$V \cong U \oplus P$$

A.1.7. Artinian rings. Artinian rings are an example of the effect of slightly changing the defining property of noetherian rings. It turns out (theorem A.1.87 on the next page) that Artinian rings *are* noetherian rings with a special property.

DEFINITION A.1.83. A ring, R, will be called *Artinian* if every descending sequence of ideals becomes constant from some finite point on — i.e., if

$$\mathfrak{a}_1 \supseteq \mathfrak{a}_2 \supseteq \cdots$$

is a descending chain of ideals, there exists an integer n such that $\mathfrak{a}_i = \mathfrak{a}_{i+1}$ for all $i \geq n$. A ring is Artinian if and only if it is an Artinian module over itself.

REMARK. Emil Artin (1898-1962) introduced these rings in the papers [**4**] and [**3**]. At first glance, this definition appears very similar to the definition of noetherian ring in definition A.1.48 on page 360 (at least if you look at the remark following the definition).

For instance $\mathbb{Z}$ is noetherian but not Artinian since we have an infinite descending sequence of ideals that does *not* become constant

$$(2) \supset (4) \supset \cdots \supset (2^k) \supset \cdots$$

Artinian rings have some unusual properties:

LEMMA A.1.84. *If R is an Artinian ring:*

   (1) every quotient of R is Artinian
   (2) if R is an integral domain, it is a field
   (3) every prime ideal of R is maximal
   (4) the number of maximal ideals in R is finite.

REMARK. Statement 3 implies that all Artinian rings are Jacobson rings.

PROOF. The first statement follows immediately from the definition of Artinian ring and lemma A.1.25 on page 352.

To prove the second statement, suppose R is an integral domain and $x \neq 0 \in R$. Then the descending chain of ideals

$$(x) \supset (x^2) \subset \cdots \supset (x^n) \supset \cdots$$

must stabilize after a finite number of steps, so $(x^t) = (x^{t+1})$ and $x^t = r \cdot x^{t+1}$ for some $r \in R$, or $x^t - r \cdot x^{t+1} = 0$. Since R is an integral domain $x^t \cdot (1 - r \cdot x) = 0$ implies $1 - r \cdot x = 0$ so $r = x^{-1}$.

The third statement follows from the first two: if $\mathfrak{p} \subset R$ is a prime ideal, then $R/\mathfrak{p}$ is an Artinian integral domain, hence a field. This implies that $\mathfrak{p}$ is maximal.

Suppose we have an infinite set of distinct maximal ideals, $\{\mathfrak{m}_i\}$ and consider the following descending sequence of ideals

$$\mathfrak{m}_1 \supset \mathfrak{m}_1 \cdot \mathfrak{m}_2 \supset \cdots \supset \mathfrak{m}_1 \cdots \mathfrak{m}_k \supset \ldots$$

The Artinian property implies that this becomes constant at some point, i.e.,

$$\mathfrak{m}_1 \cdots \mathfrak{m}_n \subset \mathfrak{m}_1 \cdots \mathfrak{m}_{n+1} \subset \mathfrak{m}_{n+1}$$

The fact that maximal ideals are prime (see proposition A.1.20 on page 350) and exercise 11 on page 354 implies that either

$$\mathfrak{m}_1 \subset \mathfrak{m}_{n+1}$$

a contradiction, *or*

$$\mathfrak{m}_2 \cdots \mathfrak{m}_n \subset \mathfrak{m}_{n+1}$$

In the latter case, a simple induction shows that *one* of the $\mathfrak{m}_i \subset \mathfrak{m}_{n+1}$, so a contradiction cannot be avoided.

If R is an Artinian ring and $\square$

We can completely characterize Artinian rings. The first step to doing this is:

LEMMA A.1.85. *Let R be a ring in which there exists a finite product of maximal ideals equal to zero, i.e.*

$$\mathfrak{m}_1 \cdots \mathfrak{m}_k = 0$$

Then R is Artinian if and only if it is noetherian.

PROOF. We have a descending chain of ideals

$$R \supset \mathfrak{m}_1 \supset \mathfrak{m}_1\mathfrak{m}_2 \supset \cdots \supset \mathfrak{m}_1 \cdots \mathfrak{m}_{k-1} \supset \mathfrak{m}_1 \cdots \mathfrak{m}_k = 0$$

Let for $1 \leq i \leq k$, let $M_i = \mathfrak{m}_1 \cdots \mathfrak{m}_{i-1}/\mathfrak{m}_1 \cdots \mathfrak{m}_i$, a module over $R/\mathfrak{m}_i$ i.e., a vector space over $R/\mathfrak{m}_i$. Then M_i is Artinian if and only if it is noetherian — if and only if it is finite-dimensional. The conclusion follows from induction on k, proposition A.1.68 on page 371 and the short exact sequences

$$0 \rightarrow \mathfrak{m}_1 \cdots \mathfrak{m}_i \rightarrow \mathfrak{m}_1 \cdots \mathfrak{m}_{i-1} \rightarrow M_i \rightarrow 0$$

□

PROPOSITION A.1.86. *If R is an Artinian ring, the nilradical, $\mathfrak{N}(R)$, is nilpotent, i.e. there exists an integer k such that $\mathfrak{N}(R)^k = 0$.*

REMARK. We have already seen this for noetherian rings — see exercise 30 on page 364.

PROOF. Since R is Artinian, the sequence of ideals

$$\mathfrak{N}(R) \supset \mathfrak{N}(R)^2 \supset \cdots$$

becomes constant after a finite number of steps. Suppose $\mathfrak{n} = \mathfrak{N}(R)^k = \mathfrak{N}(R)^{k+1}$. We claim that $\mathfrak{n} = 0$.

If not, consider the set, $\mathscr{I}$ of ideals, $\mathfrak{a}$, in R such that $\mathfrak{a} \cdot \mathfrak{n} \neq 0$. Since all descending sequences of such ideals have a lower bound (because R is Artinian), Zorn's Lemma (A.1.27 on page 353) implies that $\mathscr{I}$ has a minimal element, $\mathfrak{b}$. There exists an element $x \in \mathfrak{b}$ such that $x \cdot \mathfrak{n} \neq 0$, and the minimality of $\mathfrak{b}$ implies that $\mathfrak{b} = (x)$. The fact that $\mathfrak{n}^2 = \mathfrak{n}$ implies that $(x \cdot \mathfrak{n}) \cdot \mathfrak{n} = x \cdot \mathfrak{n}^2 = x \cdot \mathfrak{n}$ so $x \cdot \mathfrak{n} \subset (x)$. The minimality of $\mathfrak{b} = (x)$ implies that $x \cdot \mathfrak{n} = (x)$ so that there is an element $y \in \mathfrak{n}$ such that

$$x \cdot y = x = x \cdot y^2 = \cdots = x \cdot y^m$$

Since $y \in \mathfrak{N}(R)$, we have $y^n = 0$ for some $n > 0$, which implies that $x = 0$, which in turn contradicts the requirement that $x \cdot \mathfrak{n} \neq 0$. This contradiction is the result of assuming that $\mathfrak{n} \neq 0$.   □

We are finally ready to characterize Artinian rings:

THEOREM A.1.87. *A ring is Artinian if and only if it is noetherian and all of its prime ideals are maximal.*

REMARK. The reader may wonder whether this contradicts our statement that $\mathbb{Z}$ is not Artinian. After all, all of its prime ideals of the form (p) for a prime number $p \in \mathbb{Z}$ are maximal. The one exception is $(0) \subset (p)$ which is a prime ideal that is proper subset of another prime ideal.

PROOF. If R is Artinian, then lemma A.1.84 on the preceding page implies that all of its prime ideals are maximal. PropositionA.1.86 implies that $\mathfrak{N}(R)^k = 0$ for some $k > 0$ and the proof of A.1.47 on page 360 implies that

$$\mathfrak{m}_1 \cap \cdots \cap \mathfrak{m}_n \subset \mathfrak{N}(R)$$

where $\mathfrak{m}_1, \ldots, \mathfrak{m}_n$ are the finite set of maximal ideals of R. Since

$$\mathfrak{m}_1 \cdots \mathfrak{m}_n \subset \mathfrak{m}_1 \cap \cdots \cap \mathfrak{m}_n$$

it follows that a finite product of maximal ideals is equal to 0. Lemma A.1.85 on the preceding page then implies that R is noetherian.

Conversely, if R is noetherian and all of its prime ideals are maximal, lemma A.1.52 on page 363 implies that the number of these will be finite. Since the nilradical is nilpotent (see exercise 30 on page 364), the argument above implies that R is Artinian. □

Another interesting property of Artinian rings is:

THEOREM A.1.88. *An Artinian ring decomposes (uniquely) into a product of finitely many local Artinian rings.*

PROOF. Let A be an Artinian ring with maximal ideals $\{\mathfrak{m}_1, \ldots, \mathfrak{m}_n\}$. Then

$$\mathfrak{N}(R) = \mathfrak{m}_1 \cap \cdots \cap \mathfrak{m}_n = \mathfrak{m}_1 \cdots \mathfrak{m}_n$$

Let k be a value for which $\mathfrak{N}(R)^k = 0$ (this exists by proposition A.1.86 on the facing page). Then

$$(\mathfrak{m}_1 \cdots \mathfrak{m}_n)^k = \mathfrak{m}_1^k \cdots \mathfrak{m}_n^k$$

and the Chinese Remainder Theorem (see exercise 18 on page 354) implies that

$$R = \frac{R}{(\mathfrak{m}_1 \cdots \mathfrak{m}_n)^k} = \prod_{i=1}^{n} \frac{R}{\mathfrak{m}_i^k}$$

Each of the quotients $R/\mathfrak{m}_i^k$ has a unique maximal ideal, namely the image of $\mathfrak{m}_i$ so it is a local ring.

Suppose we have an expression

$$A = A_1 \times \cdots \times A_t$$

where the A_i are Artinian local rings. Then every ideal, $\mathfrak{I} \subset A$ is of the form

$$\mathfrak{I} = \mathfrak{I}_1 \times \cdots \times \mathfrak{I}_t$$

and the maximal ideals of A are of the form

$$\mathfrak{m}_i = A_1 \times \cdots \times \mathfrak{M}_i \times \cdots \times A_t$$

where $\mathfrak{M}_i \subset A_i$ is a maximal ideal. This implies that $t = n$ — i.e., the number of factors is uniquely determined by A. We also conclude that

$$\mathfrak{m}_1^k \cdots \mathfrak{m}_n^k = \mathfrak{M}_1^k \times \cdots \times \mathfrak{M}_n^k = 0$$

so that $\mathfrak{M}_i^k = 0$ for all i. We finally note that

$$\frac{R}{\mathfrak{m}_i^k} = \frac{A_1 \times \cdots \times A_n}{A_1 \times \cdots \times A_{i-1} \times 0 \times A_{i+1} \times \cdots \times A_n} = A_i$$

so the decomposition is unique. □

A.1.8. Rings and modules of fractions. Given a multiplicative set, we can define the corresponding ring of fractions:

DEFINITION A.1.89. Let M be a module over a ring, R, and let $S \subset R$ be a multiplicative set. Then the module $S^{-1}M$ consists of pairs $(s, m) \in S \times M$, usually written m/s, subject to the relation

$$\frac{m_1}{s_1} \equiv \frac{m_2}{s_2}$$

if $u \cdot (s_2 \cdot m_1 - s_1 \cdot m_2) = 0$ for some $u \in S$ and $m_1, m_2 \in M$. We make $S^{-1}M$ a module by defining:

$$\frac{m_1}{s_1} + \frac{m_2}{s_2} = \frac{s_2 \cdot m_1 + s_1 \cdot m_2}{s_1 \cdot s_2}$$
$$r \cdot \frac{m_1}{s_1} = \frac{r \cdot m_1}{s_1}$$

for all $m_1, m_2 \in M$, $s_1, s_2 \in S$, and $r \in R$.

There exists a canonical homomorphism $f \colon M \to S^{-1}M$ that sends $m \in M$ to $m/1 \in S^{-1}M$.

If $M = R$ as a module over itself, then $S^{-1}R$ is a ring with multiplication defined by

$$\frac{r_1}{s_1} \cdot \frac{r_2}{s_2} = \frac{r_1 \cdot r_2}{s_1 \cdot s_2}$$

for all $r_1, r_2 \in R$ and $s_1, s_2 \in S$.

REMARK. The kernel of the canonical map $M \to S^{-1}M$ consists of elements of M that are annihilated by elements of S. If R is an integral domain, the map $R \to S^{-1}R$ is injective.

This construction has a universal property described in proposition A.5.24 on page 452.

PROPOSITION A.1.90. *If a multiplicative set $S \subset R$ contains elements s_1, s_2 with the property that $s_1 s_2 = 0$, then $S^{-1}M = 0$, for any R-module, M.*

PROOF. Suppose $m \in M$. We claim that

$$\frac{m}{1} = \frac{0}{1} \in S^{-1}M$$

In order for this to be true, we must have

$$s(m - 0) = 0$$

for some $s \in S$. But the fact that $s_1 s_2 = 0$ implies that $0 \in S$ and we can just set

$$0(m - 0) = 0$$

$\square$

DEFINITION A.1.91. Let R be a ring and let $h \in R$. Then $S_h = \{1, h, h^2, \dots\}$ is a multiplicative subset of A and we define $R_h = S_h^{-1}R$.

REMARK. Every element of R_h can be written in the form a/h^m and

$$\frac{a}{h^m} = \frac{b}{h^n} \Leftrightarrow h^J(ah^n - bh^m) = 0$$

for some integer $J \geq 0$.

LEMMA A.1.92. *For any ring A and $h \in A$, the map*

$$\sum a_i x^i \mapsto \sum \frac{a_i}{h^i}$$

defines an isomorphism

$$A[X]/(1 - hX) \to A_h$$

PROOF. If $h = 0$, both rings are zero, so assume $h \neq 0$. In the ring $A' = A[X]/(1 - hX)$, $1 = hX$ so h is a unit. Let $\alpha: A \to B$ be a homomorphism of rings such that $\alpha(h)$ is a unit in B.

The homomorphism

$$\sum a_i X^i \mapsto \sum \alpha(a_i)\alpha(h)^{-i}: A[X] \to B$$

factors through A' because $1 - hX \mapsto 1 - \alpha(a)\alpha(h)^{-1} = 0$.

Because $\alpha(h)$ is a unit in B, this is the unique extension of α to A'. Therefore A' has the same universal property as A_h so the two are uniquely isomorphic.

When $h|h'$ so $h' = hg$, there is a canonical homomorphism

$$\frac{a}{b} \mapsto \frac{ag}{h'}: A_h \to A_{h'}$$

so the rings A_h form a direct system indexed by the set S. $\qquad\square$

PROPOSITION A.1.93. *Suppose A is a ring and $S \subset A$ is a multiplicative set. Then:*

- *If $S \subset A$ and $\mathfrak{b} \subset A$ is an ideal, then $S^{-1}\mathfrak{b}$ is an ideal in $S^{-1}A$.*
- *If $\mathfrak{b}$ contains any element of S, then $S^{-1}\mathfrak{b} = S^{-1}A$.*

It follows that

COROLLARY A.1.94. *The ideals in $S^{-1}A$ are in a 1-1 correspondence with the ideals of A that are disjoint from S.*

DEFINITION A.1.95. If $\mathfrak{p} \subset A$ is a prime ideal, then $S = A \setminus \mathfrak{p}$ is a multiplicative set. Define $A_{\mathfrak{p}} = S^{-1}A$.

REMARK. Since any ideal $\mathfrak{b} \not\subseteq \mathfrak{p}$ intersects S, it follows that $S^{-1}\mathfrak{p}$ is the *unique* maximal ideal in $S^{-1}A$.

$S^{-1}A$ is, therefore, a *local ring* (a ring with a unique maximal ideal). The word "local" is motivated by geometry — see chapter 3 on page 117.

If a ring is not an integral domain, it has no field of fractions. Nevertheless we can define a "closest approximation" to it

DEFINITION A.1.96. If R is a ring and S is the set of non-zero-divisors of R then

$$Q(R) = S^{-1}R$$

is called the *total quotient ring* of R.

REMARK. If R is an integral domain, $Q(R)$ is just the field of fractions of R.

EXERCISES.

43. Suppose R is a ring with a multiplicative set S and $a \cdot s = 0$ for $a \in R$ and $s \in S$. Show that

$$\frac{a}{1} = 0 \in S^{-1}R$$

44. Use the results of exercise 24 on page 363 to show that if R is noetherian, so is $S^{-1}R$ for any multiplicative set S.

45. If R and S are rings, show that $Q(R \times S) = Q(R) \times Q(S)$. Here, $R \times S$ is the ring of pairs (r,s) with pairwise addition and multiplication:

$$(r_1, s_1) + (r_2, s_2) = (r_1 + r_2, s_1 + s_2)$$
$$(r_1, s_1) \cdot (r_2, s_2) = (r_1 \cdot r_2, s_1 \cdot s_2)$$

46. If R is a ring and M is an R-module, show that an element $m \in M$ goes to 0 in all localizations $M_{\mathfrak{a}}$, where $\mathfrak{a} \subset R$ runs over the maximal ideals of R if and only if $m = 0$.

47. If R is a ring and M is an R-module, show that $M_{\mathfrak{a}} = 0$ for all maximal ideals $\mathfrak{a} \subset R$ if and only if $M = 0$.

48. Suppose k is a field and $R = k[[X]]$ is the ring of power-series in X (see definition A.1.5 on page 343). If $F = k((X))$, the field of fractions of R, show that every element of F can be written in the form

$$X^{\alpha} \cdot r$$

for some $\alpha \in \mathbb{Z}$ and some $r \in R$.

A.2. Fields

A.2.1. Definitions.

DEFINITION A.2.1. A *field* is a commutative integral domain whose nonzero elements have multiplicative inverses. If F is a field, the set of nonzero elements is denoted $F^{\times}$ and is an *abelian group* (under multiplication).

If we define $m \cdot 1$ for $m \in \mathbb{Z}$, $m > 0$ as the sum of m copies of 1, then the smallest positive integral value of m such that $m \cdot 1 = 0$ is called the *characteristic* of the field. If $m \cdot 1 \neq 0$ for all values of m, the field is said to be of *characteristic* 0.

An inclusion of fields $F \subset \Omega$ is called an *extension* and often denoted

$$\Omega$$
$$|$$
$$F$$

REMARK. It is not hard to see that *all* homomorphisms of fields must be inclusions or isomorphisms. Consequently, extensions play a very large part in field theory.

DEFINITION A.2.2. If $k \subset K$ is an extension of fields, then K is a vector space over k. The dimension of K as a vector space is called the *degree of the extension*, denoted $[K:k]$.

PROPOSITION A.2.3. *If it is not 0, the characteristic of a field must be a prime number.*

PROOF. Suppose $0 < m$ is the characteristic of a field, $\mathbb{F}$, and $m = a \cdot b \in \mathbb{Z}$. Then $a, b \neq 0 \in \mathbb{F}$ and $a \cdot b = 0 \in \mathbb{F}$, which contradicts the fact that a field is an integral domain. $\square$

PROPOSITION A.2.4. *Let F be a field and let $F[X]$ be the ring of polynomials over F. Then $F[X]$ is a principal ideal domain.*

PROOF. It's easy to see that $F[X]$ is an integral domain. We claim that it is a Euclidean domain as well — see definition A.1.31 on page 356. This is because we can divide polynomials as we do integers: given two polynomials $p(X)$, $q(X)$ we can write
$$p(X) = a(X)q(X) + r(X)$$
with $a(X)$ as the quotient and $r(X)$ as the remainder where $\deg r(X) < \deg q(X)$. So the conclusion follows from proposition A.1.37 on page 357. $\square$

DEFINITION A.2.5. If k is a field, an *algebra over k* is a vector space F over k that has a multiplication-operation
$$F \times F \to F$$
that makes it a commutative ring. The identity element $1 \in F$ defines an inclusion
$$\begin{aligned} k &\to F \\ x &\mapsto x \cdot 1 \end{aligned}$$

REMARK. For instance the polynomials rings $k[X_1, \ldots, X_n]$ are algebras over k.

Strictly speaking, algebras over fields are not required to be commutative or even associative. For instance, the quaternions and Cayley numbers are regarded as algebras over $\mathbb{R}$. In our applications, algebras over a field will always be commutative rings.

An immediate consequence of definition A.2.2 is:

PROPOSITION A.2.6. *If F is a finite field, the number of elements in F must be p^n, where p is some prime number and n is a positive integer.*

PROOF. The characteristic of F must be some prime, p, by proposition A.2.3. It follows that $\mathbb{Z}_p \subset F$ so F is a vector space over $\mathbb{Z}_p$. If $n = [F:\mathbb{Z}_p]$, is the dimension of that vector space, then F has p^n elements. $\square$

Examples of fields are easy to find:
- The familiar examples: $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. They are fields of characteristic 0.
- If R is any integral domain and $S = R \setminus \{0\}$, then S is a multiplicative set in the sense of definition A.1.41 on page 359 and $S^{-1}R$ is a field. This is called the field of fractions of R.

- If $\mathbb{F}$ is a field, the set of *rational functions* with coefficients in $\mathbb{F}$, denoted $\mathbb{F}(X)$, (with round rather than square brackets) is a field. This is the field of fractions of the polynomial ring, $\mathbb{F}[X]$.
- If p is a prime, $\mathbb{Z}_p$ is a field of characteristic p.

The following innocuous-looking result solved a great mystery of ancient Greece:

PROPOSITION A.2.7. *Let $E \subset F$ and $F \subset G$ be finite extensions of fields. Then*

$$[G:E] = [G:F] \cdot [F:E]$$

PROOF. Let $\{x_1, \ldots, x_n\} \in G$ be a basis for it over F and let $\{y_1, \ldots, y_m\} \in F$ be a basis for it over E. So every element of G can be expressed as a linear combination of the x_i

$$(A.2.1) \qquad\qquad\qquad g = \sum_{i=1}^{n} f_i x_i$$

with the $f_i \in F$. Each of the f_i is given by

$$(A.2.2) \qquad\qquad\qquad f_i = \sum_{j=1}^{m} e_{i,j} y_j$$

which means that

$$g = \sum_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} e_{i,j} x_i y_j$$

which shows that the $n \cdot m$ elements $\{x_i \cdot y_j\}$ span G over F. To see that they are linearly independent, set $g = 0$ in equation A.2.1 on page 388. The linear independence of the x_i implies that $f_i = 0$, $i = 1, \ldots, n$. These, and the linear independence of the y_j imply that $e_{i,j} = 0$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$ which proves the result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

DEFINITION A.2.8. If $k \subset K$ is an inclusion of fields, then $\alpha \in K$ is said to be *algebraic* over k if it is a root of a polynomial with coefficients in k. If $\alpha \in K$ is not algebraic, it is said to be *transcendental*.

The notation, $k(\alpha) \subset K$, represents the field of rational functions of α.

REMARK. For instance, if we think of

$$\mathbb{Q} \subset \mathbb{C}$$

then $\sqrt{2}$ is algebraic over $\mathbb{Q}$, but e is not.

In comparing $k(\alpha)$ with the ring $k[\alpha]$, it is not hard to see that:

(1) $k(\alpha)$ is the smallest subfield of K containing k and α.
(2) $k(\alpha)$ is the field of fractions of $k[\alpha]$.

PROPOSITION A.2.9. *Let $k \subset K$ be an inclusion of fields and let $f: k[X] \to K$ be the unique homomorphism that sends X to $\alpha \in K$. Then α is algebraic over k if and only if $\ker f \neq 0$, in which case $\ker f = (p(X))$ and $p(X)$ is called the minimal polynomial of α.*

The minimal polynomial is always irreducible.

REMARK. The minimal polynomial is the lowest-degree polynomial such that $f(\alpha) = 0$. If $g(X)$ is any polynomial with the property that $g(\alpha) = 0$, then $f(X)|g(X)$. See example 2.3.20 on page 60 for techniques for computing it.

This result implies that $\alpha \in K$ is transcendental if and only if the homomorphism f is injective.

The numbers π and e are well-known to be transcendental — see [**55**].

PROOF. The kernel of f is just the polynomials that vanish when evaluated at α. This kernel is a principal ideal because $k[X]$ is a principal ideal domain — see proposition A.2.4 on page 387.

If $f(X) = p(X) \cdot q(X)$ then

$$p(\alpha)q(\alpha) = 0$$

implies that $p(\alpha) = 0$ or $q(\alpha) = 0$. If p and q are of lower degree than f, it would contradict the minimality of $f(X)$. □

Consider

$$\mathbb{Q} \subset \mathbb{C}$$

and form the extension field

$$\mathbb{Q}(\sqrt{2})$$

which is the field of all possible rational functions

$$\frac{\sum_{i=1}^{m} p_i(\sqrt{2})^i}{\sum_{j=1}^{n} q_j(\sqrt{2})^j}$$

where the $p_i, q_j \in \mathbb{Q}$ — or the smallest subfield of $\mathbb{C}$ containing $\mathbb{Q}$ and $\sqrt{2}$.

Upon reflection, it becomes clear that we can always have $n, m \leq 1$ since $(\sqrt{2})^2 \in \mathbb{Q}$, so every element of $\mathbb{Q}(\sqrt{2})$ is really of the form

$$\frac{a + b\sqrt{2}}{c + d\sqrt{2}}$$

with $a, b, c, d \in \mathbb{Q}$.

We can even clear out the denominator because

$$\frac{a + b\sqrt{2}}{c + d\sqrt{2}} = \frac{a + b\sqrt{2}}{c + d\sqrt{2}} \cdot \frac{c - d\sqrt{d}}{c - d\sqrt{2}} = \frac{ac - 2bd + \sqrt{2}(bc - ad)}{c^2 - 2d^2}$$

$$= \frac{ac - 2bd}{c^2 - 2d^2} + \frac{bc - ad}{c^2 - 2d^2}\sqrt{2}$$

We have just proved that

$$\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[\sqrt{2}]$$

This is no accident — it is true for *all* algebraic extensions:

LEMMA A.2.10. *Let $F \subset H$ be an extension of fields and suppose $\alpha \in H$ is algebraic over F. Then*

$$F(\alpha) = F[\alpha]$$

and $[F(\alpha):F]$ is equal to the degree of the minimal polynomial of α.

PROOF. If α is algebraic over F, it has a minimal polynomial $p(X) \in F[X]$ (see definition A.2.9 on page 388) which is irreducible so the ideal $(p(X))$ is prime. Since the ring $F[X]$ is a principal ideal domain (see proposition A.1.37 on page 357) the ideal $(p(X))$ is also *maximal* and

$$F[\alpha] = F[X]/(p(X))$$

is a field (see proposition A.1.30 on page 355), so it is equal to $F(\alpha)$. $\qquad\square$

One famous problem the ancient Greek geometers puzzled over is that of *doubling the cube* — using straightedge and compass constructions. In other words, they wanted to construct $\sqrt[3]{2}$ via their geometric techniques. It can be shown that ancient Greek compass-and-straightedge techniques can construct

(1) all integers
(2) the square root of any number previously constructed (by drawing a suitable circle).
(3) the sum, difference, product and quotient of any two numbers previously constructed.

Consequently, the numbers they constructed all lay in fields of the form

$$F_n = \mathbb{Q}(\sqrt{\alpha_1})(\sqrt{\alpha_2})\cdots(\sqrt{\alpha_n})$$

where each α_i is contained in the field to the left of it. Since the minimal polynomial of $\sqrt{\alpha_{i+1}}$ is $X^2 - \alpha_{i+1} \in F_i[X]$, lemma A.2.10 on the preceding page implies that $[F_{i+1}:F_i] = 2$ and proposition A.2.7 on page 388 implies that $[F_n:\mathbb{Q}] = 2^n$. But $[\mathbb{Q}(\sqrt[3]{2}):\mathbb{Q}] = 3$ and $3 \nmid 2^n$ for any n, so $\sqrt[3]{2} \notin F_n$ for any n.

So the problem of constructing $\sqrt[3]{2}$ is literally unsolvable by ancient Greek techniques.

EXERCISES.

1. Suppose $F \subset H$ is a finite extension of fields and $\alpha \in H$. If n is the degree of the minimum polynomial of α, show that $n|[H:F]$.

A.2.2. Algebraic extensions of fields.

DEFINITION A.2.11. An extension of fields, $E \subset F$, is said to be *algebraic* if every element $x \in F$ is algebraic over E. If an extension is not algebraic, it is *transcendental*.

REMARK. For instance, $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2})$ is algebraic and $\mathbb{Q} \subset \mathbb{R}$ and $\mathbb{F} \subset \mathbb{F}(X)$ are transcendental extensions, where $\mathbb{F}$ is any field.

PROPOSITION A.2.12. *If $E \subset F$ is a finite field extension, then it is algebraic.*

PROOF. Suppose $[F:E] = n$ and let $x \in F$. Then the powers

$$\{1, x, x^2, \ldots, x^n\}$$

must be linearly dependent over E so we get a nontrivial algebraic equation

$$a_1 + a_1 x + \cdots + a_n x^n = 0$$

with $a_i \in E$. $\qquad\square$

Extensions containing roots of polynomials always exist:

COROLLARY A.2.13. *Let F be a field and let $f(X) \in F[X]$ be a polynomial. Then there exists an extension $F \subset \Omega$ such that Ω contains a root of f.*

PROOF. Factor f as

$$f(X) = p_1(X)^{\alpha_1} \cdots p_k(X)^{\alpha_k}$$

where the $p_i(X)$ are irreducible. This can be done (and is even unique) by corollary A.1.35 on page 357. As in the proof of A.2.10 on page 389, the quotient

$$E = F[X]/(p_1(X))$$

is a field containing F.

The image, α, of X under the quotient-mapping

$$F[X] \to F[X]/(p_1(X)) = E$$

has the property that $p_1(\alpha) = f(\alpha) = 0$. $\qquad\square$

COROLLARY A.2.14. *Let F be a field and let $f(X) \in F[X]$ be a polynomial. Then there exists an extension $F \subset \Omega$ such that*

$$f(X) = \prod_{k=1}^{\deg(f)} (X - \alpha_k) \in \Omega[X]$$

REMARK. This extension, Ω, is called a *splitting field* for $f(X)$. We can write

$$\Omega = F[\alpha_1, \ldots, \alpha_d]$$

where $d = \deg f$.

The solution to exercise 8 on page 403 shows that these splitting fields are *unique* up to isomorphism.

PROOF. This follows by an inductive application of corollary A.2.13. We construct a field Ω_1 that contains a root, α, of $f(X)$. If $f(X)$ splits into linear factors in Ω_1, we are done. Otherwise, factor $f(X)$ as

$$f(X) = (X - \alpha)^k \cdot g(X) \in \Omega_1[X]$$

where $g(X)$ is relatively prime to $X - \alpha$, and construct an extension Ω_2 of Ω_1 that contains a root of $g(X)$. Eventually this process terminates with a field Ω that contains all of the roots of $f(X)$. $\qquad\square$

Since a polynomial splits into linear factors in its splitting field, one might expect the greatest common divisor of two polynomials to depend on the field in which one computes it. It is interesting that this does *not* happen:

PROPOSITION A.2.15. *Let $F \subset \Omega$ be an inclusion of fields and let $f(X), g(X) \in F[X]$ be polynomials. Then*

$$g_F(X) = \gcd(f(X), g(X)) \in F[X]$$

is also their greatest common divisor in $\Omega[X]$.

REMARK. Since Ω could be the splitting field of $f(X)$ and $g(X)$, the greatest common divisor of these polynomials (up to units) is

$$\prod_{i=1}^{n}(X - \alpha_i)$$

where the α_i are *all* of the roots that $f(X)$ and $g(X)$ have in common. Somehow this product always defines an element of $F[X]$ (even though the α_i are *not* in F).

PROOF. Let us pass to a field $K \supset \Omega$ that is a splitting field for f and g. Suppose $f(X)$ and $g(X)$ have the following common roots in K:

$$\alpha_1, \ldots, \alpha_n$$

Then $g_F(X)$ also splits into linear factors and

(A.2.3) $$g_F(X) = \prod_{k=1}^{t}(X - \alpha_{j_k})$$

where $\{\alpha_{j_1}, \ldots, \alpha_{j_t}\}$ is, possibly, a subset of $\{\alpha_1, \ldots, \alpha_n\}$ such that the product in equation A.2.3 lies in $F[X]$. If this product lies in $F[X]$, it is also in $\Omega[X]$, so the greatest common divisor calculated in this larger field will have these factors, at *least*. We conclude that

$$g_F(X)|g_\Omega(X)$$

where $g_\Omega(X)$ is the greatest common divisor calculated in $\Omega[X]$.

On the other hand, the Euclidean algorithm (proposition A.1.33 on page 356) implies that there exist $a(X), b(X) \in F[X]$ such that

$$a(X) \cdot f(X) + b(X) \cdot g(X) = g_F(X)$$

so $g_\Omega(X)|g_F(X)$.   □

There is an interesting result regarding repeated roots of a polynomial:

LEMMA A.2.16. *Let F be a field with a splitting field Ω and let $f(X) \in F[X]$ be a polynomial. Then $f(X)$ has a repeated root in Ω if and only if*

$$f(X), f'(X)$$

have a common root. This occurs if and only if $\mathrm{Res}(f, f', X) = 0$ *(in the notation of definition 1.3.2 on page 13) which happens if and only if*

$$\gcd(f(X), f'(X)) \neq 1$$

REMARK. This is interesting because the criteria, $\mathrm{Res}(f, f', X) = 0$ or $\gcd(f(X), f'(X)) \neq 1$, make no direct reference to Ω.

Note that, in characteristic $p \neq 0$, the derivative of X^p is 0.

PROOF. The first statement follows by the chain-rule and the product-formula for derivatives. If

$$f(X) = (X - \alpha)^k g(X)$$

with $\alpha > 1$, then

$$f'(X) = k(X - \alpha)^{k-1}g(X) + (X - \alpha)^k g'(X)$$

which will share a root with $f(X)$ regardless of whether the characteristic of F divides k (for instance, the characteristic of F (and, therefore, Ω) might be p and k might be a multiple of p).

The second statement follows from proposition 1.3.3 on page 13 and the statement about the greatest common divisor follows from proposition A.2.15 on page 391. $\square$

This result tells us something important about irreducible polynomials:

LEMMA A.2.17. *Let $f(X) \in F[X]$ be an irreducible polynomial of degree n with splitting field Ω, and suppose that F is of characteristic 0. Then, in the factorization*

$$f(X) = \prod_{i=1}^{n}(X - \alpha_i) \in \Omega[X]$$

the α_i are all distinct.

REMARK. This argument *fails* if the characteristic of F is $p \neq 0$. In this case, we can have an irreducible polynomial, $f(X^p)$, that has repeated roots.

PROOF. Since the characteristic of F is 0, and $f(X)$ is not constant, $f'(X) \neq 0$.

If $f(X)$ had a repeated factor, we would have

$$\gcd(f(X), f'(X)) = p(X) \neq 1$$

with $\deg p(X) < \deg f(X)$ since $\deg f'(X) < \deg f(X)$, which would contradict the irreducibility of $f(X)$. $\square$

In characteristic p, it is possible to say *exactly* what may prevent an irreducible polynomial from having distinct roots:

PROPOSITION A.2.18. *Let F be a field of characteristic p and suppose $f(X) \in F[X]$ is an irreducible nonconstant polynomial with repeated roots. Then there exists an irreducible polynomial $g(X) \in F[X]$ whose roots are all distinct such that*

$$f(X) = g(X^{p^t})$$

for some integer $t > 0$.

PROOF. If $f(X)$ has repeated roots, then it has roots in common with $f'(X)$. If $f'(X) \neq 0$, the greatest common divisor of $f(X)$ and $f'(X)$ would be a lower-degree polynomial that divides $f(X)$ — contradicting its irreducibility. It follows that $f'(X) = 0$, which is only possible if all of the exponents in $f(X)$ are multiples of p (since p is a prime and the coefficients of f are relatively prime to p). In this case,

$$g_1(X) = f(X^{1/p})$$

is a well-defined polynomial that is *still* irreducible: any nontrivial factorization of $g_1(X)$ implies one for $f(X)$. If $g_1'(X) = 0$, repeat this process. Since each iteration lowers degree by a factor of p, after a finite number of steps we arrive at an irreducible polynomial

$$g_t(X) = g(X) = f(X^{1/p^t})$$

with $g'(X) \neq 0$. $\square$

DEFINITION A.2.19. Let F be a field and $f(X) \in F[X]$ be a polynomial. Then $f(X)$ will be called *separable* if it factors into a product of *distinct* linear factors in a splitting field.

If $F \subset \Omega$ is an inclusion of fields and $\alpha \in \Omega$ is algebraic over F, then α will be called a *separable element* if its minimal polynomial is separable. The field Ω will be called a *separable extension of* F if every element of Ω is separable over F.

REMARK. "Separable" = roots are *separated*. The whole question of separability is moot unless the fields in question have characteristic $p \neq 0$.

DEFINITION A.2.20. A field F is said to be *perfect* if every finite extension of F is separable.

REMARK. This is equivalent to saying that all irreducible polynomials have distinct roots. Most of the fields that we have dealt with have been perfect. Perfect fields include:

- Any field of characteristic 0 (lemma A.2.17 on the previous page).
- Finite fields (theorem A.2.42 on page 405).
- Algebraically closed fields (definition A.2.26 on page 399).

It is interesting that algebraic extensions of fields can always be generated by a single element:

THEOREM A.2.21. *Let $F \subset H$ be an extension of infinite fields and suppose $\alpha_1, \ldots, \alpha_n \in H$ are algebraic over F. In addition, suppose $\alpha_2, \ldots, \alpha_n$ are separable over F.*

Then there exists an element $\beta \in H$ such that

$$F[\alpha_1, \ldots, \alpha_n] = F[\beta]$$

REMARK. The element β is called a *primitive element* and this result is often called the *primitive element theorem*.

The condition that F and H be infinite is *not* necessary — see theorem A.2.46 on page 406 for what happens in the finite case.

PROOF. We will prove it for $n = 2$ — a simple induction proves the general case. We will show that $F[\alpha, \beta] = F[\gamma]$

Let α, β have minimal polynomials $f(X), g(X) \in F[X]$, respectively and let $\Omega \supset H$ be a splitting field for $f(X)$ and $g(X)$. Then $f(X)$ and $g(X)$ have roots

$$\alpha = \alpha_1, \ldots, \alpha_n$$
$$\beta = \beta_1, \ldots, \beta_m$$

respectively, with the β_i all distinct. For $j \neq 1$, the equation

$$\alpha_i + X\beta_j = \alpha_1 + X\beta_1 = \alpha + X\beta$$

has exactly one solution

$$x_{i,j} = \frac{\alpha_i - \alpha}{\beta - \beta_j}$$

If we choose $c \in F$ different from any of these elements (using the fact that F is infinite), we get

$$\alpha_i + c\beta_j \neq \alpha + c\beta$$

unless $i = j = 1$. We claim $\gamma = \alpha + c\beta$ will satisfy the hypotheses of this theorem.

The polynomials $g(X)$ and $h(X) = f(\gamma - cX) \in F[\gamma][X]$ have a β as a common root:

$$
\begin{aligned}
g(\beta) &= 0 \\
h(\beta) &= f(\gamma - c\beta) \\
&= f(\alpha) \\
&= 0
\end{aligned}
$$

By the choice of c above, they will *only* have β as a common root because $\gamma - c\beta_j \neq \alpha_i$ for any $i \neq 1$ or $j \neq 1$. It follows that

$$\gcd(g(X), h(X)) = X - \beta$$

Proposition A.2.15 on page 391 implies the greatest common divisor has its coefficients in the field in which the polynomials have theirs, so

$$\beta \in F[\gamma]$$

On the other hand, we also have $\alpha = \gamma - c\beta \in F[\gamma]$ so

$$F[\alpha, \beta] = F[\gamma]$$

$\square$

EXERCISES.

2. If F is a field, show that

$$
\begin{aligned}
F(Y) &\to F(X) \\
Y &\mapsto X^2
\end{aligned}
$$

makes $F(X)$ algebraic extension of $F(X)$ of degree 2.

3. If $F = \mathbb{Q}(2^{1/3})$, express

$$\frac{1}{2^{2/3} - 2^{1/3} + 1}$$

as a polynomial in $2^{1/3}$ (see lemma A.2.10 on page 389).

4. Find a primitive element for the field $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$ over $\mathbb{Q}$ and find its minimal polynomial.

5. Find the splitting field of $X^3 - 2$ over $\mathbb{Q}$.

A.2.3. Norm and trace. The *norm* of a field element is an important concept that we have seen before in example A.1.40 on page 357.

DEFINITION A.2.22. If $F \subset H$ is a finite extension of fields, then H is a finite-dimensional vector space over F with basis $\{x_1, \ldots, x_n\}$ where $n = [H:F]$. If $\alpha \in H$, then

(1) $m_\alpha = \alpha \cdot *: H \to H$, i.e. the matrix of the linear transformation of H (as a vector-space over F) defined by multiplication by α, and with respect to the basis $\{x_1, \ldots, x_n\}$,
(2) $\chi_\alpha(X) = \det(X \cdot I - m_\alpha) \in F[X]$ is the characteristic polynomial of m_α and called the *characteristic polynomial of α*,
(3) $N_{H/F}(\alpha) = \det m_\alpha \in F$ is the determinant of m_α, and is called the *norm* of α.
(4) $T_{H/F}(\alpha) = \text{Tr}(m_\alpha) \in F$ is the trace of the matrix m_α (i.e. the sum of its diagonal elements), and is called the *trace* of α.

REMARK. The terms are closely related

(A.2.4) $$\chi_\alpha(0) = (-1)^n N_{H/F}(\alpha)$$

where $n = [H:F]$. To see this, just plug $X = 0$ into $I \cdot X - m_\alpha$ and take the determinant. If the characteristic polynomial is of degree n, the trace is $-a_{n-1}$, where a_{n-1} is the coefficient of X^{n-1}.

For instance, suppose $F = \mathbb{Q}$ and $H = \mathbb{Q}[\sqrt{2}]$ with basis $\{1, \sqrt{2}\}$. Then the effect of $\alpha = a + b\sqrt{2}$ on this basis is

$$\begin{aligned} \alpha \cdot 1 &= a + b\sqrt{2} \\ \alpha \cdot \sqrt{2} &= 2b + a\sqrt{2} \end{aligned}$$

so the matrix m_α is

$$m_\alpha = \begin{bmatrix} a & 2b \\ b & a \end{bmatrix}$$

with characteristic polynomial

$$\chi_\alpha(X) = X^2 - 2aX + a^2 - 2b^2$$

and norm

$$N_{H/F}(\alpha) = a^2 - 2b^2$$

and trace

$$T_{H/F}(\alpha) = 2a$$

The basic properties of matrices imply that

LEMMA A.2.23. *Under the assumptions of definition A.2.22, we have*

(1) *the characteristic polynomial, norm, and trace of an element do not depend on the basis used to compute them,*
(2) $N_{H/F}(1) = 1$
(3) *for all $\alpha, \beta \in H$, $N_{H/F}(\alpha \cdot \beta) = N_{H/F}(\alpha) \cdot N_{H/F}(\beta)$*
(4) *for all $\alpha, \beta \in H$, $T_{H/F}(\alpha + \beta) = T_{H/F}(\alpha) + T_{H/F}(\beta$*
(5) *for all $\alpha \in H$, $N_{H/F}(\alpha) = 0$ if and only if $\alpha = 0$*

PROOF. In a splitting field for $\chi_\alpha(X)$, the characteristic polynomial satisfies

$$\chi_\alpha(X) = \prod_{j=1}^{n}(X - \lambda_j) = X^n + c_{n-1}X^{n-1} + \cdots + c_0$$

where the λ_j are the eigenvalues of m_α, which do not depend on the basis. The determinant is equal to $(-1)^n c_0$, so it is also independent of basis. The same is true for the trace, since it is equal to $-c_{n-1}$.

The second statement follows from the fact that m_1 is the identity matrix.

The third statement follows from the basic properties of determinants: the composite of α and β, as linear transformations, is $m_\alpha \cdot m_\beta = m_{\alpha\cdot\beta}$, and

$$\det(m_\alpha \cdot m_\beta) = \det(m_\alpha) \cdot \det(m_\beta)$$

The fourth statement follows from the fact that

$$m_{\alpha+\beta} = m_\alpha + m_\beta$$

And the fifth follows from the third and the fact that any nonzero element $\alpha \in H$ has a multiplicative inverse α^{-1} so that

$$N_{H/F}(\alpha) \cdot N_{H/F}(\alpha^{-1}) = 1$$

$\square$

We can also say something about the characteristic polynomial of an element

PROPOSITION A.2.24. *Under the assumptions of definition A.2.22 on the preceding page,*

$$m_*\colon H \to Mat(F, n)$$

is a homomorphism into the ring of $n \times n$ matrices with coefficients in F. It follows that an element $\alpha \in H$ satisfies its characteristic polynomial, i.e.

$$\chi_\alpha(\alpha) = 0$$

PROOF. We have already seen that $m_\alpha \cdot m_\beta = m_{\alpha\cdot\beta}$ and it is not hard to see that $m_{\alpha+\beta} = m_\alpha + m_\beta$, which proves the first statement. The second follows from the first and the Cayley-Hamilton theorem (see chapter 6 of [**155**]), which states that $\chi_\alpha(m_\alpha) = 0$. $\square$

Clearly, if $F \subset H$ is an extension of fields and $\alpha \in F$, $N_{H/F}(\alpha) = \alpha^{[H:F]}$.

Here's another example of norms of field extensions:

Let $F = \mathbb{Q}$ and let $H = \mathbb{Q}[\gamma]$ where γ is a root of the polynomial

$$p(X) = X^3 + X^2 + X + 2$$

Eisenstein's Criterion (see theorem A.3.8 on page 414) shows that $p(X)$ is irreducible over $\mathbb{Q}$ so we can construct $\mathbb{Q}[\gamma]$ as the quotient

$$\mathbb{Q}[X]/(p(X))$$

and γ is the element that X maps to under the projection to the quotient. Our basis for $\mathbb{Q}[\gamma]$ is $\{1, \gamma, \gamma^2\}$, where $\gamma^3 = -2 - \gamma - \gamma^2$, and

$$\gamma^4 = -2\gamma - \gamma^2 - \gamma^3 = 2 - \gamma$$

A general element of this field is

$$\alpha = a + b\gamma + c\gamma^2$$

and the effect of this element on the basis is

$$
\begin{aligned}
\alpha \cdot 1 &= a + b\gamma + c\gamma^2 \\
\alpha \cdot \gamma &= a\gamma + b\gamma^2 + c\gamma^3 \\
&= -2c + (a - c)\gamma + (b - c)\gamma^2 \\
\alpha \cdot \gamma^2 &= a\gamma^2 + b\gamma^3 + c\gamma^4 \\
&= 2(c - b) - (c + b)\gamma + (a - b)\gamma^2
\end{aligned}
$$

so we get the matrix

$$
m_\alpha = \begin{bmatrix}
a & -2c & 2(c - b) \\
b & a - c & -(c + b) \\
c & b - c & a - b
\end{bmatrix}
$$

with determinant

$$N_{H/F}(\alpha) = a^3 - a^2 b - ca^2$$
$$+ 5\,acb - 3\,ac^2 + ab^2 + 2\,cb^2 - 2\,b^3 - 2\,bc^2 + 4\,c^3$$

and characteristic polynomial

$$\chi_\alpha(X) =$$
$$X^3 - (3\,a - b - c)\,X^2 - \left(3\,c^2 - 5\,cb - b^2 - 3\,a^2 + 2\,ca + 2\,ab\right) X$$
$$- a^3 + a^2 b + ca^2 - 5\,acb + 3\,ac^2 - ab^2 - 2\,cb^2 + 2\,b^3 + 2\,bc^2 - 4\,c^3$$

and trace

$$T_{H/F}(\alpha) = 3a - b - c$$

Although an element of a field *satisfies* its characteristic polynomial, this does not mean the characteristic polynomial *is* its minimal polynomial.

In fact:

LEMMA A.2.25. *Let $F \subset H$ be a finite field-extension and let $\alpha \in H$ have minimal polynomial $p(X) \in F[X]$. Then*

$$\chi_\alpha(X) = p(X)^{[H:F[\alpha]]}$$

PROOF. Let $\{x_i\}$ be a basis for H over $F[\alpha]$ and let $\{y_j\}$ be a basis for $F[\alpha]$ over F. Then $\{x_i y_j\}$ is a basis for H over F (see proposition A.2.7 on page 388). The effect of α on this basis is to act on the $\{y_j\}$ and leave the $\{x_i\}$ fixed. This means

$$
m_\alpha = \begin{bmatrix}
A & 0 & \cdots & 0 \\
0 & A & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & A
\end{bmatrix}
$$

where $A = m_\alpha$ computed *in $F[\alpha]$*, and this block-matrix has $[H: f[\alpha]]$ rows and columns.

In $F[\alpha]$, the characteristic polynomial is a polynomial that α satisfies, hence is contained in the principal ideal $(p(X)) \subset F[X]$ and is of the same degree as $p(X)$ so it is a multiple of $p(X)$ by a unit $u \in F$. Since both polynomials are monic, we must have $u = 1$.

The conclusion follows from the properties of a determinant of a block matrix. $\square$

EXERCISES.

6. If $H = \mathbb{Q}[2^{1/3}]$ compute the norm and characteristic polynomial of a general element.

7. If $H = \mathbb{Q}[\sqrt{2}, \sqrt{3}]$ compute the norm and characteristic polynomial of a general element.

A.2.4. Algebraically closed fields. These fields play an important part in algebraic geometry.

DEFINITION A.2.26. A field Ω is said to be *algebraically closed* if any polynomial $p(X) \in \Omega[X]$ can be factored into linear terms

$$p(X) = f_0 \cdot \prod_{k=1}^{\deg p} (X - \alpha_k) \in \Omega[X]$$

with $f_0 \in \Omega$.

REMARK. This is equivalent to saying that $p(x)$ has $\deg p$ roots in Ω.

EXAMPLE. The Fundamental Theorem of Algebra implies that the field $\mathbb{C}$ is algebraically closed.

DEFINITION A.2.27. Let $F \subset \Omega$ be an extension of fields. Then Ω is defined to be an algebraic closure of F if

(1) Ω is algebraically closed.
(2) given any extension $F \subset G$ with G algebraically closed, Ω is isomorphic to a subfield, Ω' of G that contains F.

REMARK. The field of complex numbers is clearly the algebraic closure of $\mathbb{R}$.

If they exist, algebraic closures are essentially unique:

THEOREM A.2.28. *Let F be a field and let Ω_1 and Ω_2 be algebraic closures of F. Then there exists an isomorphism*

$$f: \Omega_1 \to \Omega_2$$

such that $f|F = 1: F \to F$.

PROOF. Define a pair (E, τ) to consist of a subfield $E \subset \Omega_1$ such that $F \subset E$ and a monomorphism $\tau : E \to \Omega_2$, such that $\tau | F = 1$. At least one such pair exists because we can simply define $E = F \hookrightarrow \Omega_2$.

Define $(E_1, \tau_1) \prec (E_2, \tau_2)$ if $E_1 \subset E_2$ and $\tau_2 | E_1 = \tau_1$. Then every chain

$$(E_1, \tau_1) \prec (E_2, \tau_2) \prec \cdots$$

has a maximal element, (E, τ): Simply define

$$E = \bigcup_i E_i$$

and define $\tau | E_i = \tau_i$. It follows, by Zorn's lemma that we can find a maximal element among all of the (E, τ). Call this $(\bar{E}, \bar{\tau})$. We claim that $\bar{E} = \Omega_1$. If not, we could find a nontrivial algebraic extension $\bar{E}[\alpha]$ with minimal polynomial $p(x)$ and extend $\bar{\tau}$ to a map

$$\begin{aligned} g : \bar{E}[\alpha] &\to \Omega_2 \\ \alpha &\mapsto \beta \end{aligned}$$

where $\beta \in \Omega_2$ is a root of $\bar{\tau}(p(x))$. This is a contradiction. We also claim that $\bar{\tau}(\Omega_1) = \Omega_2$ since its image will be an algebraically closed subfield of Ω_2. $\square$

It turns out that *every* field has an algebraic closure. We will fix a field F and explicitly construct its algebraic closure using a construction due to Artin (see [**102**]).

We need a lemma first:

LEMMA A.2.29. *Let F be a field and let $f_1(X), \ldots, f_k(X) \in F[X]$ be polynomials. Then there exists an extension $F \subset \Omega$ such that Ω contains a root of each of the $f_i(X)$.*

PROOF. This follows from corollary A.2.13 on page 391 and induction. $\square$

DEFINITION A.2.30. Let S denote the set of all monic, irreducible polynomials in $F[x]$ — this is infinite (just mimic Euclid's proof that the number of primes is infinite).

Form the polynomial ring $F[\{S_f\}]$ with an indeterminate, S_f, for each $f \in S$ and form the ideal $\mathfrak{M} = (\{f(S_f)\})$ — generated by indeterminates representing monic irreducible polynomials *plugged into those very polynomials.*

PROPOSITION A.2.31. *The ideal, $\mathfrak{M} \subset F[\{S_f\}]$, defined in A.2.30 is proper.*

PROOF. We will show that $1 \notin \mathfrak{M}$. Let

$$x = \sum_{k=1}^{n} a_k \cdot f_k(S_{f_k}) \in \mathfrak{M}$$

be some element, where $f_k \in S$. We will set $x = 1$ and get a contradiction.

Let Ω denote an extension of F containing one root, α_k, of each of the n polynomials $f_k(S_{f_k})$. Now define a homomorphism

$$\begin{aligned} \mathbb{F}[\{S_f\}] &\to \Omega \\ S_{f_k} &\mapsto \alpha_k \\ S_{f'} &\mapsto 0 \end{aligned}$$

(A.2.5)

(A.2.6)

for $k = 1, \ldots, n$, where $f' \notin \{f_1, \ldots, f_n\}$. This is clearly possible since the S_{f_k} are all indeterminates. The equation $x = 1$ maps to $0 = 1$, a contradiction. $\square$

REMARK. This argument is delicate: The existence of the mapping in equation A.2.5 on the preceding page requires a *separate* indeterminate for each monic irreducible polynomial.

THEOREM A.2.32. *An algebraic closure, Ω, exists for F. If the cardinality of F is infinite, then the cardinality of Ω is equal to that of F.*

PROOF. Let $\mathfrak{M} \subset F[\{S_f\}]$ be as in proposition A.2.31 on the facing page. Since $\mathfrak{M}$ is proper, proposition A.1.28 on page 353 implies that it is contained in some maximal ideal $\mathfrak{M}'$. Define

$$\Omega_1 = F[\{S_f\}]/\mathfrak{M}'$$

This will be a field, by lemma A.1.30 on page 355. This field will contain roots of all monic irreducible polynomials in $F[X]$. If it is algebraically closed, we are done. Otherwise, continue this construction to form a field Ω_2 containing Ω_1 and all roots of monic irreducible polynomials in $\Omega_1[X]$.

We obtain a (possibly infinite) chain of fields

$$F \subset \Omega_1 \subset \Omega_2 \subset \cdots$$

If any of the Ω_k are algebraically closed, then

$$\Omega_n = \Omega_k$$

for all $n > k$ since the only monic irreducible polynomials in Ω_k will be linear ones.

Define

$$\Omega = \bigcup_{i=1}^{\infty} \Omega_i$$

We claim that this is algebraically closed. Any polynomial $f(X) \in \Omega[X]$ is actually contained in $\Omega_k[X]$ for some value of k, and its roots will be contained in Ω_{k+1}.

The statement about cardinalities follows from the corresponding property of each of the $\Omega_i[\{S_f\}]$. $\square$

EXAMPLE. The algebraic closure of $\mathbb{R}$ is $\mathbb{C}$. The algebraic closure of $\mathbb{Q}$ is called the *algebraic numbers* and written $\bar{\mathbb{Q}}$. It cannot equal $\mathbb{C}$ because it is *countable,* by theorem A.2.32. The structure of $\bar{\mathbb{Q}}$ is extremely complex and not well understood.

The uniqueness of algebraic closures have some interesting consequences:

DEFINITION A.2.33. Let F be a field and let $\alpha \in \bar{F}$ be an element of the algebraic closure of F. Then the minimal polynomial $f(X)$ of α splits into linear factors

$$f(X) = \prod_{i=1}^{\deg f} (X - \alpha_i)$$

with $\alpha_1 = \alpha$. The $\{\alpha_i\}$ are called the *conjugates* of α.

REMARK. The conjugates of α are uniquely determined by α because $\bar{F}$ is uniquely determined up to an isomorphism.

The characteristic polynomial of α in $F[\alpha]$ is the minimal polynomial (it is of the same degree and α satisfies it) and the discussion following A.2.24 on

page 397 shows that the conjugates of α are just the eigenvalues of the matrix m_α in definition A.2.22 on page 396.

For instance, if $z = a + bi \in \mathbb{C}$, then the minimal polynomial of z is its characteristic polynomial over $\mathbb{R}$ (see A.2.22 on page 396), namely

$$X^2 - 2aX + a^2 + b^2$$

and the other root of this polynomial is $a - bi$, the usual complex conjugate.

The conjugates of an algebraic element are related to its norm:

LEMMA A.2.34. *Let $F \subset H$ be a finite extension of fields and let $\alpha \in H$. Then*

$$N_{H/F}(\alpha) = \left(\prod_{j=1}^{m} \alpha_j \right)^{[H:F[\alpha]]}$$

$$T_{H/F}(\alpha) = [H:F[\alpha]] \cdot \sum_{j=1}^{m} \alpha_j$$

where the $\{\alpha_j\}$ run over the conjugates of α (with $\alpha = \alpha_1$).

PROOF. Let the minimal polynomial of α be $p(X) \in F[X]$, of degree m. Then, in an algebraic closure of F

$$p(X) = X^m + c_{n-1}X^{m-1} + \cdots + c_0 = \prod_{j=1}^{m}(X - \alpha_j)$$

from which it follows that

$$c_0 = (-1)^m \prod_{j=1}^{m} \alpha_j$$

The conclusion follows from lemma A.2.25 on page 398 and equation A.2.4 on page 396. The statement about the trace follows from the fact that the trace of a matrix is the sum of its eigenvalues. $\square$

Here is another interesting property of conjugates of an algebraic element:

LEMMA A.2.35. *If F is a field and $\alpha \in \bar{F}$ is an element of the algebraic closure of F, then there exists isomorphisms of fields*

$$\begin{aligned} F[\alpha] &\to F[\alpha'] \\ f &\mapsto f \text{ for all } f \in F \\ \alpha &\mapsto \alpha' \end{aligned}$$

where α' is any conjugate of α.

REMARK. To make this more precise: regard $F[\alpha]$ as a vector-space with F-basis $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ (if the minimal polynomial of α is of degree n). Then map vector-spaces $F[\alpha] \to F[\alpha']$ via the change of basis

$$\alpha^j \mapsto (\alpha')^j$$

for $j = 0, \ldots, n - 1$. This lemma says that this defines a field-isomorphism.

This elementary result is the basis of a deep field of mathematics called Galois Theory. See **[146]** for more on this.

PROOF. Each conjugate of α satisfies the same minimal polynomial, $p(X) \in F[X]$, as α and we have

$$F[\alpha] = F[X]/(p(X)) = F[\alpha']$$

$\square$

EXERCISES.

8. If F is a field and $f(X) \in F[X]$ is a polynomial, show that any two splitting fields, G_1, G_2 of $f(X)$ are isomorphic via isomorphism

$$g: G_1 \to G_2$$

whose restriction to $F \subset G_1$ is the identity map. Consequently, we can speak of *the* splitting field of $f(X)$.

9. Compute the conjugates of an element $\gamma = a + b\, 2^{1/3} \in \mathbb{Q}[2^{1/3}]$.

A.2.5. Finite fields. Finite fields can be completely classified — one of those rare areas of mathematics that have an exhaustive solution.

We begin with a lemma:

LEMMA A.2.36. *Let F be a ring or field of characteristic p. If $\alpha, \beta \in F$ then*

$$(\alpha + \beta)^p = \alpha^p + \beta^p$$

PROOF. This follows from the binomial theorem

$$(\alpha + \beta)^p = \sum_{ik=1}^{p} \frac{p!}{(p-k)! \cdot k!} \alpha^k \beta^{p-k}$$

so all terms except the first and the last have a factor of p in the numerator that is not canceled by any factor in the denominator. $\square$

We know, from proposition A.2.3 on page 387 that the characteristic of a finite field is a prime p and proposition A.2.6 on page 387 implies that the size of a finite field is p^k for some $k > 0$.

We first show that finite fields of order p^k exist for all primes p and all integers $k \geq 1$:

LEMMA A.2.37. *Let $g_k(X) = X^{p^k} - X \in \mathbb{Z}_p[X]$. Then the roots of $g_k(X)$ in the algebraic closure, $\bar{\mathbb{Z}}_p$, of $\mathbb{Z}_p$ form a field of order p^k.*

PROOF. First note that $g'_k(X) = -1 \in \mathbb{Z}_p[X]$ so it has no repeated roots, by lemma A.2.16 on page 392 — meaning it has p^k roots in $\bar{\mathbb{Z}}_p$. Note that

(1) $0, 1 \in \bar{\mathbb{Z}}_p$ are in the set of roots, and if α and β are two such roots then:
   (a) $\alpha \cdot \beta$ also satisfies $g_k(\alpha \cdot \beta) = 0$ so the set of roots is closed under multiplication,

(b) $(\alpha + \beta)^{p^k} = \alpha^{p^k} + \beta^{p^k}$ by lemma A.2.36 on the previous page, so $g_k(\alpha + \beta) = 0$ and the set of roots is closed under addition.

(2) multiplying all nonzero roots by a fixed root, $\alpha \neq 0$, merely permutes them because

$$\alpha \cdot \beta_1 = \alpha \cdot \beta_2 \implies \beta_1 = \beta_2$$

because $\bar{\mathbb{Z}}_p$ is an integral domain. So there exists a root γ such that $\alpha \cdot \gamma = 1$.

It follows that the set of p^k roots of $g_k(X)$ constitute a field. $\qquad\square$

Now that we know fields of order p^k exist, we prove that they are unique:

LEMMA A.2.38. *Let F be any field of order p^k and let $\alpha \in F$ be any element. Then*

$$\alpha^{p^k} - \alpha = 0$$

It follows that F is isomorphic to the field of order p^k constructed in lemma A.2.37 on the preceding page.

PROOF. This is just Lagrange's theorem, applied to the multiplicative group, F^*, of F. In other words, take the product of all nonzero elements of F

$$\delta = \prod_{i=1}^{p^k-1} \alpha_i$$

and multiply each element by α to get

$$\alpha^{p^k-1} \cdot \delta = \prod_{i=1}^{p^k-1} \alpha \cdot \alpha_i$$

Since multiplication by α simply permutes the elements of F^*, we get

$$\alpha^{p^k-1} \cdot \delta = \delta$$

or $\alpha^{p^k-1} = 1$. $\qquad\square$

DEFINITION A.2.39. The *unique* field of order p^k is denoted $\mathbb{F}_{p^k}$.

REMARK. So $\mathbb{F}_p = \mathbb{Z}_p$. The notation $\mathrm{GF}(p^k)$ ("Galois Field") is sometimes used for $\mathbb{F}_{p^k}$ in honor of Galois who is responsible for all of the material in this section.

> Évariste Galois (1811–1832) first described finite fields in [51]. Galois is also responsible for much of what we know in the theory of equations (Galois Theory) and field-theory: see [146].

DEFINITION A.2.40. The *Frobenius map*, $\mathcal{F}_p \colon \mathbb{F}_{p^k} \to \mathbb{F}_{p^k}$ is defined to send $\alpha \in \mathbb{F}_{p^k}$ to $\alpha^p \in \mathbb{F}_{p^k}$.

PROPOSITION A.2.41. *The Frobenius map is an automorphism of finite fields.*

PROOF. By the definition, it clearly preserves multiplication.

If $\alpha, \beta \in \mathbb{F}_{p^k}$ note that $\mathcal{F}_p^k(\alpha) = \alpha^{p^k} = \alpha$ because α is a root of $X^{p^k} - X \in \mathbb{Z}_p[X]$ in the algebraic closure of $\mathbb{Z}_p$, so $\mathcal{F}_p(\alpha) = \mathcal{F}_p(\beta)$ implies that $\mathcal{F}_p^k(\alpha) = \mathcal{F}_p^k(\beta) = \alpha = \beta$. It follows that $\mathcal{F}_p$ is a injective. In addition,

$$\mathcal{F}_p(\alpha + \beta) = (\alpha + \beta)^p = \alpha^p + \beta^p = \mathcal{F}_p(\alpha) + \mathcal{F}_p(\beta)$$

by lemma A.2.36 on page 403, so it also preserves addition.

Since $\mathbb{F}_{p^k}$ is *finite*, $\mathcal{F}_p$ must be $1 - 1$. □

Note that $\mathbb{F}_{p^k} \subset \mathbb{F}_{p^\ell}$ if and only if $k|\ell$, since $\mathbb{F}_{p^\ell}$ must be a vector-space over $\mathbb{F}_{p^k}$ and both are vector-spaces over $\mathbb{F}_p = \mathbb{Z}_p$. With this in mind, we can explicitly describe the algebraic closure of *all* finite fields of characteristic p:

THEOREM A.2.42. *Let p be a prime number. Then the algebraic closure of all finite fields of characteristic p is*

$$\bar{\mathbb{F}}_p = \bigcup_{k=1}^{\infty} \mathbb{F}_{p^{k!}}$$

The Frobenius map

$$\mathcal{F}_p: \bar{\mathbb{F}}_p \to \bar{\mathbb{F}}_p$$

is an automorphism and the finite field $\mathbb{F}_{p^\ell} \subset \mathbb{F}_{p^{\ell!}} \subset \bar{\mathbb{F}}_p$ is the set of elements of $\bar{\mathbb{F}}_p$ fixed by $\mathcal{F}_p^\ell$ (i.e. elements $x \in \bar{\mathbb{F}}_p$ such that $\mathcal{F}_p^\ell(x) = x$).

REMARK. Note that $\bar{\mathbb{F}}_p$ is an infinite field since it contains subfields of order p^k for all k. This implies:

All algebraically closed fields are infinite.

PROOF. If $f(X) \in \mathbb{F}_{p^k}[X]$ is a polynomial, it splits into linear factors in some finite extension, G, of $\mathbb{F}_{p^k}$, by corollary A.2.14 on page 391. It follows that G is a finite field that contains $\mathbb{F}_{p^k}$ — i.e. $\mathbb{F}_{p^\ell}$ for some ℓ that is a multiple of k. Consequently $f(X)$ splits into linear factors in $\bar{\mathbb{F}}_p$. It follows that $\bar{\mathbb{F}}_p$ is algebraically closed and it is the smallest field containing all of the $\mathbb{F}_{p^k}$, so it must be the algebraic closure of all of them.

Since the Frobenius map is an automorphism of all of the $\mathbb{F}_{p^k}$ it must be an automorphism of $\bar{\mathbb{F}}_p$. The statement that $\mathcal{F}_p^\ell(x) = x$ implies that x is a root of $X^{p^\ell} - X = 0$ so the final statement follows from lemma A.2.37 on page 403. □

It is interesting to look at the multiplicative group of a finite field — that is, $\mathbb{F}_{p^k} \setminus \{0\}$, where the operation is multiplication. We need two lemmas:

DEFINITION A.2.43. If n is a positive integer then

$$\phi(n)$$

is called the Euler *phi-function.* It is equal to the number of *generators* of $\mathbb{Z}_n$, or

$$\phi(n) = \begin{cases} \text{the number of integers } 1 \leq d < n \text{ such that } \gcd(d, n) = 1 & \text{if } n > 1, \text{ or} \\ 1 & \text{if } n = 1 \end{cases}$$

REMARK. If $n > 1$, an element x, of $\mathbb{Z}_n$ has a multiplicative inverse if and only if $\gcd(x, n) = 1$ (see lemma A.1.12 on page 345). It follows that the multiplicative group $\mathbb{Z}_n^{\times}$ has $\phi(n)$ *elements*. If $n = 1$, $\mathbb{Z}_1 = (0)$, which has one generator.

This function has a number of interesting number-theoretic properties, including

LEMMA A.2.44. *If n is a positive integer, then*

$$(A.2.7) \qquad\qquad n = \sum_{d|n} \phi(d)$$

where the sum is taken over all positive divisors, d, of n.

PROOF. If $d|n$, let $\Phi_d \subset \mathbb{Z}_n$ be the set of generators of the unique cyclic subgroup of order d (generated by n/d). Since every element of $\mathbb{Z}_n$ generates *one* of the $\mathbb{Z}_d$, it follows that $\mathbb{Z}_n$ is the disjoint union of all of the Φ_d for all divisors $d|n$. It follows that

$$|\mathbb{Z}_n| = n = \sum_{d|n} |\Phi_d| = \sum_{d|n} \phi(d)$$

$\square$

LEMMA A.2.45. *If G is a finite group of order n with the property that the equation $x^k = 1$ has at most k solutions, then G is cyclic.*

REMARK. In a *non-cyclic* group, the equations $x^k = 1$ can have *more* than k solutions. For instance, in the group $\mathbb{Z}_3 \oplus \mathbb{Z}_3$, the equation $3x = 0$ (written additively) has 9 solutions.

PROOF. If $d|n$ and an element, x, of order d exists, then it generates a cyclic subgroup $(x) = \{1, x, \ldots, x^{d-1}\}$ — which has $\phi(d)$ *distinct* generators. The hypothesis implies that all solutions to the equation $x^d = 1$ are elements of (x). It follows that all elements of *order d* are *generators* of (x) and that there are $\phi(d)$ of them. For each $d|n$ the set of elements of order d is either

- *empty,* if there are *no* elements of order d,
- *nonempty* with $\phi(d)$ members.

Equation A.2.7 implies that the number of elements of G is $< n$ unless elements of order d exist for *all* $d|n$ — including n itself. An element of order n generates G and implies it is cyclic. $\square$

THEOREM A.2.46. *If $\mathbb{F}_{p^n}$ is a finite field, its multiplicative group, $\mathbb{F}_{p^n}^{\times}$, is cyclic of order $p^n - 1$.*

PROOF. If $x \in \mathbb{F}_{p^n}^{\times}$, the solution to exercise 19 on page 358 implies that the equation $x^n = 1$ has, at most n solutions. The conclusion follows immediately from lemma A.2.45. $\square$

Among other things, this implies the *Primitive Element Theorem* (see theorem A.2.21 on page 394) for finite fields.

The minimum polynomial of a generator of $\mathbb{F}_{p^n}^{\times}$ over $\mathbb{Z}_p$ is called a *primitive polynomial* and such polynomials are heavily used in cryptography (see [**132**]).

A.2.6. Transcendental extensions. We will characterize transcendental extensions of fields and show that they have transcendence bases similar to the way vector spaces have bases (see table A.2.1 on the next page). A great deal of this material originated in the work of the German mathematician, Ernst Steinitz (1871–1928) in his seminal paper, [**160**].

DEFINITION A.2.47. Consider an inclusion of fields $F \subset \Omega$. Elements $\alpha_1, \ldots, \alpha_m \in \Omega$ will be called *algebraically independent* over F if the natural map

$$
\begin{aligned}
F[X_1, \ldots X_m] &\rightarrow \Omega \\
X_i &\mapsto \alpha_i
\end{aligned}
$$

is *injective*. If they aren't independent, they are said to be algebraically *dependent* over F.

REMARK. In other words, the $\alpha_1, \ldots, \alpha_m \in \Omega$ are algebraically dependent if there exists a polynomial f with coefficients in F such that

$$f(\alpha_1, \ldots, \alpha_n) = 0$$

in Ω.

They are algebraically *independent* if any equation of the form

$$\sum c_{i_1, \ldots, i_m} \alpha_1^{i_1} \cdots \alpha_m^{i_m} = 0$$

implies that all of the $\{c_{i_1, \ldots, i_m}\}$ vanish. Note the similarity between this condition and the definition of linear independence in linear algebra. As we will see, this is not a coincidence, and the theory of transcendence bases is similar to that of bases of vector spaces.

EXAMPLE.
(1) A single element $\alpha \in \Omega$ is algebraically independent if it is transcendental over F.
(2) The numbers π and e are probably algebraically independent over $\mathbb{Q}$ but this has not been proved.
(3) An infinite set $\{\alpha_i\}$ is independent over F if and only if every finite subset is independent.
(4) If $\alpha_1, \ldots, \alpha_n$ are algebraically independent over F, then

$$
\begin{aligned}
F[X_1, \ldots, X_n] &\rightarrow F[\alpha_1, \ldots, \alpha_n] \\
f(X_1, \ldots, X_n) &\mapsto f(\alpha_1, \ldots, \alpha_n)
\end{aligned}
$$

is injective, hence an isomorphism. This isomorphism extends to the fields of fractions. In this case, $F(\alpha_1, \ldots, \alpha_n)$ is called a *pure transcendental extension* of F.
(5) The Lindemann–Weierstrass theorem (see [**55**] and [**12**]) proves that if $\alpha_1, \ldots, \alpha_n$ are algebraic numbers that are linearly independent over $\mathbb{Q}$, then $e^{\alpha_1}, \ldots, e^{\alpha_n}$ are algebraically independent over $\mathbb{Q}$.

We can characterize algebraic elements of a field extension:

LEMMA A.2.48. *Let $f \subset \Omega$ be an extension of fields with $\gamma \in \Omega$ and let $A \subset \Omega$ be some set of elements. The following conditions are equivalent:*
(1) γ is algebraic over $F(A)$.
(2) There exist $\beta_1, \ldots, \beta_t \in F(A)$ such that $\gamma^t + \beta_1 \gamma^{t-1} + \cdots + \beta_t = 0$.

(3) *There exist $\beta_0, \ldots, \beta_t \in F[A]$ such that $\beta_0 \gamma^t + \beta_1 \gamma^{t-1} + \cdots + \beta_t = 0$.*
(4) *There exists an $f(X_1, \ldots, X_m, Y) \in F[X_1, \ldots, X_m, Y]$ and $\alpha_1 \cdots, \alpha_m \in A$ such that $f(\alpha_1, \ldots, \alpha_m, Y) \neq 0$ but $f(\alpha_1, \ldots, \alpha_m, \gamma) = 0$.*

PROOF. Clearly statement 1 $\implies$ statement 2 $\implies$ statement 3 $\implies$ statement 1 — so those statements are equivalent.

Statement 4 $\implies$ statement 3: Write $f(X_1, \ldots, X_m, Y)$ as a polynomial in Y with coefficients in $F[X_1, \ldots, X_m]$, so

$$f(X_1, \ldots, X_m, Y) = \sum f_i(X_1, \ldots, X_m) Y^i$$

Then statement 3 holds with $\beta_i = f_i(\alpha_1, \ldots, \alpha_m)$.

Statement 3 $\implies$ statement 4: The β_i in statement 3 can be expressed as polynomials in a finite number of elements $\alpha_1, \ldots, \alpha_m \in A$

$$\beta_i = f_i(\alpha_1, \ldots, \alpha_m)$$

and we can use the polynomial

$$f(X_1, \ldots, X_m, Y) = \sum f_i(X_1, \ldots, X_m) Y^i$$

in statement 4. $\qquad \square$

When γ satisfies the conditions in the lemma, it is said to be *algebraically dependent on A over F*.

Table A.2.1 illustrates the many similarities the theory of transcendence bases has with linear algebra.

Linear algebra	Transcendence
linearly independent	algebraically independent
$A \subset \text{Span}(B)$	*A* algebraically dependent on *B*
basis	transcendence basis
dimension	transcendence degree

TABLE A.2.1. Analogy with linear algebra

Continuing our analogy with linear algebra, we have the following result, which shows that we can swap out basis elements:

LEMMA A.2.49. EXCHANGE LEMMA:*Let $\{\alpha_1, \ldots, \alpha_t\}$ be a subset of Ω. If $\beta \in \Omega$ is algebraically dependent on $\{\alpha_1, \ldots, \alpha_t\}$ but not on $\{\alpha_1, \ldots, \alpha_{t-1}\}$, then α_t is algebraically dependent on $\{\alpha_1, \ldots, \alpha_{t-1}, \beta\}$.*

PROOF. Since β is algebraically dependent on $\{\alpha_1, \ldots, \alpha_t\}$, there exists a polynomial $f(X_1, \ldots, X_i, Y)$ with coefficients in F such that

$$f(\alpha_1, \ldots, \alpha_t, Y) \neq 0 \quad f(\alpha_1 \ldots, \alpha_t, \beta) = 0$$

Write f as a polynomial in X_t:

$$f(X_1, \ldots, X_t, Y) = \sum z_i(X_1, \ldots, X_{t-1}, Y) X_t^i$$

Because $f(\alpha_1, \ldots, \alpha_t, Y) \neq 0$ at least one of the z_i, say $z_{i_0}(\alpha_1, \ldots, \alpha_{t-1}, Y)$ is not the zero polynomial.

Because β is not algebraically dependent on $\{\alpha_1, \ldots, \alpha_{t-1}\}$, it follows that $z_{i_0}(\alpha_1, \ldots, \alpha_{t-1}, \beta) \neq 0$. Therefore $f(\alpha_1, \ldots, \alpha_{t-1}, X_t, \beta) \neq 0$.

But, because $f(\alpha_1, \ldots, \alpha_{t-1}, \alpha_t, \beta) = 0$, it follows that α_t is algebraically dependent on $\{\alpha_1, \ldots, \alpha_{t-1}, \beta\}$. $\qquad\square$

LEMMA A.2.50. *If C is algebraically dependent on B and B is algebraically dependent on A, then C is algebraically dependent on A.*

PROOF. If γ is algebraic over a field E that is algebraic over F, then γ is algebraic over F. Apply this with $E = F(A \cup B)$ and $F = F(A)$. $\qquad\square$

Now we are ready to prove the main result

THEOREM A.2.51. *Let $F \subset \Omega$ be an extension of fields, let $A = \{\alpha_1, \ldots, \alpha_t\}$ and $B = \{\beta_1, \ldots, \beta_m\}$ be two subsets of Ω, and suppose*

(1) A is algebraically independent over F.
(2) A is algebraically dependent on B over F.

Then $t \leq m$.

PROOF. Let ℓ be the number of elements A and B have in common. If this is t, the conclusion follows, so assume it is $< t$.

Write
$$B = \{\alpha_1, \ldots, \alpha_\ell, \beta_{\ell+1}, \ldots, \beta_m\}$$

Since $\alpha_{\ell+1}$ is algebraically dependent on B, but not on $\{\alpha_1, \ldots, \alpha_t\}$, there will be a β_j with $\ell + 1 \leq j \leq m$ such that $\alpha_{\ell+1}$ is algebraically dependent on $\{\alpha_1, \ldots, \alpha_\ell, \beta_{\ell+1}, \ldots, \beta_j\}$ but not on $\{\alpha_1, \ldots, \alpha_\ell, \beta_{\ell+1}, \ldots, \beta_{j-1}\}$.

The Exchange lemma A.2.49 on the preceding page shows that β_j is algebraically dependent on
$$B_1 = B \cup \{\alpha_{\ell+1}\} \setminus \{\beta_j\}$$

So B is algebraically dependent on B_1 and A is algebraically dependent on B_1. Now we have $\ell + 1$ elements in common between A and B_1.

If $\ell + 1 < t$ repeat this process, using the Exchange property to swap elements of A for elements of B. We will eventually get $\ell = t$, and $t \leq m$. $\qquad\square$

THEOREM A.2.52. *Let $F \subset \Omega$ be an inclusion of fields. Then there exists a (possibly infinite) set of elements $\{\alpha_1, \ldots, \alpha_k\} \in \Omega$ such that the set $\{\alpha_1, \ldots, \alpha_k\}$ is algebraically independent over F, and Ω is an algebraic extension of $F(\alpha_1, \ldots, \alpha_k)$*

The number k is uniquely determined by Ω and is called the transcendence degree of Ω over F.

PROOF. All chains
$$A_1 \subset A_2 \subset \cdots$$
of sets of algebraically independent elements have an upper bound, namely their union. Zorn's lemma (A.1.27 on page 353) implies that there exists a maximal set of algebraically independent elements. If this set is finite and
$$\{\alpha_1, \ldots, \alpha_s\}$$
and
$$\{\beta_1, \ldots, \beta_t\}$$
are two maximal algebraically independent sets, theorem A.2.51 implies that $s \leq t$ and $t \leq s$ so $s = t$. $\qquad\square$

EXAMPLE. The Lindemann–Weierstrass theorem (see [**55**] and [**12**]) proves that if $\alpha_1, \ldots, \alpha_n$ are algebraic numbers that are linearly independent over $\mathbb{Q}$, then $\mathbb{Q}(e^{\alpha_1}, \ldots, e^{\alpha_n})$ has transcendence degree n over $\mathbb{Q}$.

DEFINITION A.2.53. A *transcendence basis* for Ω over F is an algebraically independent set A, such that Ω is algebraic over $F(A)$.

If there is a *finite set $A \subset \Omega$* such that Ω is algebraic over $F(A)$, then Ω has a *finite* transcendence basis over F. Furthermore, *every* transcendence basis of Ω over F is finite and has the same number of elements.

EXAMPLE A.2.54. Let $p_1, \ldots, p_m$ be the elementary symmetric polynomials in $X_1, \ldots X_m$.

CLAIM. The field $F(X_1, \ldots, X_m)$ is algebraic over $F(p_1, \ldots, p_m)$.

Consider a polynomial $f(X_1, \ldots, X_n) \in F(X_1, \ldots, X_m)$. Theorem A.1.53 on page 365 shows that the product

$$\prod_{\sigma \in S_n} (T - f(X_{\sigma(1)}, \ldots, X_{\sigma(n)}))$$

over all permutations of the variables, is a polynomial with coefficients in $F(p_1, \ldots, p_m)$.

It follows that the set $\{p_1, \ldots, p_m\}$ must contain a transcendence basis for $F(X_1, \ldots, X_m)$ over F.

Since the size of a transcendence basis is unique, the $\{p_1, \ldots, p_m\}$ must *be* a transcendence basis and $F(X_1, \ldots, X_m)$ must be an algebraic extension of $F(p_1, \ldots, p_m)$.

Here's an example from complex analysis:

EXAMPLE A.2.55. Let Ω be the field of meromorphic functions on a compact complex manifold.

The only meromorphic functions on the Riemann sphere are the rational functions in z. It follows that Ω is a pure transcendental extension of $\mathbb{C}$ of transcendence degree 1.

EXERCISES.

10. Use the Lindemann–Weierstrass theorem to prove that π is transcendental.

11. Show that the extension

$$
\begin{array}{c}
\mathbb{C} \\
| \\
\mathbb{Q}
\end{array}
$$

has an uncountable degree of transcendence.

A.3. Unique factorization domains

A.3.1. Introduction. This section is one of the most complex in the chapter on commutative algebra and most of it is flagged with a dangerous bend symbol. We study the important question of when a ring has unique factorization. Aside from any inherent interest, unique factorization has geometric implications that will become apparent later.

A great deal of this material is due to Gauss in his groundbreaking [**52**], and Weierstrass (in [**170**]) in his research on complex analysis in several variables.

Johann Carl Friedrich Gauss (1777 – 1855) was a German mathematician and scientist who contributed to many fields, including number theory, analysis, statistics (the normal distribution curve), differential geometry (he essentially invented it), geophysics, electrostatics, astronomy, and optics.

We can characterize unique factorization domains by

LEMMA A.3.1. *If R is a ring, the following three statements are equivalent*
(1) R is a unique factorization domain
(2) For any $r, p, q \in R$ such that

$$r \mid p \cdot q$$

with r irreducible

$$r \nmid p \implies r \mid q$$

(3) For any $r, p, q \in R$ such that

$$r \mid p \cdot q$$

and r and p have no common factors

$$r \nmid p \implies r \mid q$$

(4) For any irreducible element $r \in R$, the principal ideal $(r) \subset R$ is prime.

REMARK. The final statement describes the property of a unique factorization domain most relevant to algebraic geometry.

Prime ideals play an important part in algebraic geometry (see proposition 2.4.18 on page 68), and statement 4 implies that they can have a particularly simple structure.

PROOF. If statement 1 is true, then R is a unique factorization domain by reasoning like that used in lemma A.1.15 on page A.1.15. Conversely, if R is a unique factorization domain, then

$$p = u_1 \prod_{i=1}^{m} p_i^{\beta_i}$$

(A.3.1)
$$q = u_2 \prod_{i=1}^{m} p_i^{\gamma_i}$$

where $u_1, u_2 \in R$ are units, the $p_i \in R$ are irreducible elements and the $\alpha_i, \beta_i, \gamma_i \in \mathbb{Z}$ are all ≥ 0. If $r \mid p \cdot q$, then r must equal one of the p_i, say p_j and $\alpha_j + \beta_j \geq 1$. Since $r \nmid p$, we get $\alpha_j = 0$, which implies that $\beta_j \geq 1$ and this proves the conclusion.

Statement 3 implies statement 2. Conversely, if statement 2 is true, R is a unique factorization domain and equation A.3.1 on the preceding page holds as well as

$$r = u_0 \prod_{i=1}^{m} p_i^{\alpha_i}$$

Since r has no common factors with p, $\alpha_i > 0$ implies that $\beta_i = 0$, hence $\alpha_i \leq \gamma_i$ for all $i = 1, \ldots, n$. This implies that $r|q$.

To see statement 4, suppose $a, b \in U$ and $a \cdot b \in (r)$ or $ra \cdot b$. The previous statement implies that $r|a$ or $r|b$ which means $a \in (r)$ or $b \in (r)$. The implication clearly goes in the opposite direction too. $\square$

One of the easiest results in this section is:

LEMMA A.3.2. *If k is a field, then $k[X]$ is a unique factorization domain.*

PROOF. It is a Euclidean domain in the sense of definition A.1.31 on page 356 so it is also a unique factorization domain by corollary A.1.35 on page 357. $\square$

Even though the ideals in a unique factorization might not all be principal, we have:

PROPOSITION A.3.3. *In a unique factorization domain, the concept of greatest common divisor and least common multiple are well-defined.*

REMARK. In general, we have no analogue of the Euclidean Algorithm, so it may be impossible to have a formula like equation A.1.2 on page 346.

PROOF. If U is a unique factorization domain with elements x and y, then they have factorizations, unique up to multiplication by units. Let $\{p_1, \ldots, p_k\}$ be all of the irreducible factors that occur in their factorizations:

$$x = u \prod_{i=1}^{k} p_i^{\alpha_i}$$

$$y = u' \prod_{i=1}^{k} p_i^{\beta_i}$$

Now we can define

$$\gcd(x, y) = \prod_{i=1}^{k} p_i^{\min(\alpha_i, \beta_i)}$$

$$\mathrm{lcm}(x, y) = \prod_{i=1}^{k} p_i^{\max(\alpha_i, \beta_i)}$$

$\square$

A.3.2. Polynomial rings. Throughout the rest of this section, we *fix* a unique factorization domain, U. We will show that $U[X]$ is also a unique factorization domain. We use a trick to prove this: embed $U[X]$ in $F[X]$, where F is the field of fractions of U and uniquely factor elements there.

The proof involves several steps.

DEFINITION A.3.4. A polynomial $a_n X^n + \cdots + a_0 \in U[X]$ will be called *primitive*, if the greatest common divisor of its coefficients is 1.

Note that if $f \in U[X]$ is a polynomial, we can write $f = u \cdot f'$ where u is the greatest common divisor of the coefficients of f and f' is primitive.

The following result is called Gauss's Lemma:

LEMMA A.3.5. *If $f, g \in U[X]$ are primitive polynomials, then so is fg.*

REMARK. This and the following lemma were proved by Carl Friedrich Gauss in his treatise [**52**].

PROOF. Suppose

$$
\begin{aligned}
f &= a_n X^n + \cdots + a_0 \\
g &= b_m X^m + \cdots + b_0 \\
fg &= c_{n+m} X^{n+m} + \cdots + c_0
\end{aligned}
$$

If $d \in U$ is irreducible, it suffices to prove that d does not divide all of the c_i. Let a_i and b_j be the first coefficients (i.e. with the lowest subscripts) not divisible by d. We claim that c_{i+j} is not divisible by d. Note that

$$
c_{i+j} = \underbrace{a_0 b_{i+j} + \cdots + a_{i-1} b_{j+1}}_{\text{Group 1}} + a_i b_j + \underbrace{a_{i+1} b_{j-1} + \cdots + a_{i+j} b_0}_{\text{Group 2}}
$$

By construction, d divides all of the terms in Group 1 and Group 2. Since U is a unique factorization domain, $d | c_{i+1}$ *if and only if* $d | a_i b_j$. But, the fact that U is a unique factorization domain also implies that $d | a_i b_j$ if and only if $d | a_i$ or $d | b_j$. □

This leads to the following:

LEMMA A.3.6. *Let $f \in U[X]$ be primitive. Then f is irreducible in $U[X]$ if and only if it is irreducible in $F[X]$, where F is the field of fractions of U.*

PROOF. Suppose f is irreducible in $U[X]$, and that $f = gh \in F[X]$. By clearing denominators, we can assume

$$
\begin{aligned}
g &= u_1^{-1} \bar{g} \\
h &= u_2^{-1} \bar{h}
\end{aligned}
$$

where $\bar{g}, \bar{h}$ are primitive polynomials of $U[X]$, and $u_1, u_2 \in U$. We conclude that

$$
u_1 u_2 f = \bar{g} \bar{h}
$$

where $\bar{g} \bar{h}$ is primitive by A.3.5. Since f is also primitive, the factor $u_1 u_2 \in U$ must be a unit. Since f is irreducible in $U[X]$, $\bar{g}$ or $\bar{h}$ must be a unit in $U[X]$ and also in $F[X]$.

On the other hand, suppose $f \in U[X]$ is irreducible in $F[X]$ and assume $f = gh \in U[X]$. Since f is irreducible in $F[X]$, either g or h must be a unit in $F[X]$, i.e. a constant polynomial. If g is a constant polynomial then the formula $f = gh$ with f primitive implies that $g \in R$ is a unit. □

We are finally ready to prove:

THEOREM A.3.7. *If U be a unique factorization domain, then so is $U[X]$.*

REMARK. This is interesting because U doesn't have to be a euclidean domain or a principal ideal domain.

PROOF. We use a trick to prove this: embed $U[X]$ in $F[X]$, where F is the field of fractions of U and uniquely factor elements there. Suppose $r, f, g \in U[X]$ are polynomials and

(1) r is irreducible.
(2) $r|fg$

We will show that $r|f$ or $r|g$ and lemma A.3.1 on page 411 will show that $U[x]$ has unique factorization.

Lemma A.3.2 on page 412 implies that $F[X]$ is a unique factorization domain because it is Euclidean.

Write

$$
\begin{aligned}
r &= ur' \\
f &= u_1 f' \\
g &= u_2 g'
\end{aligned}
$$

where $u, u_1, u_2 \in U$ are, respectively, the greatest common divisors of the coefficients of r, f, g and $r', f', g' \in U[x]$ are primitive. Since r is irreducible, we can assume $u \in U$ is a unit (otherwise $r = ur'$ would be a nontrivial factorization of r).

Lemma A.3.6 on the previous page implies that r' is irreducible in $F[X]$ so, in $F[X]$, $r'|f'$ or $r'|g'$ in $F[x]$. Without loss of generality, assume $r'|f'$, so that

$$
f' = a \cdot r'
$$

where $a \in F[X]$. We can write $a = v^{-1} a'$, where $v \in U$ and $a' \in U[X]$ is primitive. We get

$$
v \cdot f' = a' \cdot r'
$$

in $U[X]$. Since f' and $a' \cdot r'$ are both primitive (by lemma A.3.5 on the preceding page), $v \in U$ must be a unit and we get $r|f$.   □

REMARK. Actually *finding* factorizations in these rings can be challenging. See chapter 2.3 on page 45.

To actually find a factorization of a polynomial, it is helpful to have a criterion for irreducibility. The following is called Eisenstein's Criterion:

THEOREM A.3.8. *Let U be a unique factorization domain and let*

$$
f(X) = a_n X^n + \cdots + a_0 \in U[X]
$$

be a primitive polynomial and let $p \in U$ be irreducible. If $p|a_i$ for $0 \leq i \leq n-1$, $p \nmid a_n$, and $p^2 \nmid a_0$ then $f(X)$ is irreducible in $U[X]$.

REMARK. If F is the field of fractions of U, lemma A.3.6 on the preceding page shows that this criterion works for polynomials over $F[X]$ too, after clearing the denominators of the coefficients.

Eisenstein originally proved this for $\mathbb{Z}[X]$ but it works for any unique factorization domain.

PROOF. We will reason by contradiction. Suppose there exist polynomials

$$p(X) = b_s X^s + \cdots + b_0$$
$$q(X) = c_t X^t + \cdots + c_0$$

such that $f(X) = p(X) \cdot q(X)$ and $s \geq 1$ and $t \geq 1$. Since $p^2 \nmid a_0$ we must have $p \nmid b_0$ or $p \nmid c_0$. Assume that $p|c_0$ and $p \nmid b_0$. Since f is primitive, not all the c_i are divisible by p. Suppose c_k is the first that is not divisible by p. Then

$$a_k = b_k c_0 + \cdots + b_0 c_k$$

By assumption, $p|a_k$ and $p|c_i$ for $0 \leq i < k$, which implies that $p|b_0 c_k$ and this implies that $p|b_0$, which is a contradiction. $\square$

EXAMPLE. The polynomial $X^3 + 3X^2 + 3X + 1 \in \mathbb{Q}[X]$ is irreducible by Eisenstein's Criterion with respect to the prime $p = 3$.

In some cases, one must first transform the polynomial a bit to use Eisenstein's Criterion. For instance, in the polynomial

$$f(X) = X^2 + X + 1 \in \mathbb{Q}[X]$$

there are no primes that divide any of the coefficients. After substituting $X = U + 1$, $f(X)$ becomes

$$g(U) = U^2 + 3U + 3$$

which satisfies Eisenstein's Criterion with respect to the prime $p = 3$. Since $X \to U + 1$ defines an isomorphism

$$\mathbb{Q}[X] \to \mathbb{Q}[U]$$

$f(X)$ is irreducible if and only if $g(U)$ is.

A.3.3. Power-series rings. Next, we tackle the question of unique factorization in *power series* rings. This appears daunting at first glance because power series seem to have "infinite complexity". For instance, it is *not* true that whenever U is a unique factorization domain, $U[[X]]$ is also — see [**147**].

It is gratifying to see that, in some cases, factorization is actually *easier* in power series rings. The point is that factorizations are only well-defined up to multiplication by a unit — and the power series ring $k[[X]]$ (for k a field) has *many units:* Proposition A.1.7 on page 343 shows that any power series

$$z = \sum_{n=0}^{\infty} c_n X^n$$

with $c_0 \neq 0$ is a unit. If the lowest nonzero coefficient in z is c_r then our unique factorization of z is

(A.3.2) $$X^r \cdot (c_r + c_{r+1} X + \cdots)$$

In other words, X is our "only prime" in $k[[X]]$, and an arbitrary element of $k[[X]]$ is the product of a *polynomial in X* and a *unit.*

This will turn out to be true in *general:* the Weierstrass Preparation Theorem will show that certain elements (every element can be transformed into one of these — see lemma A.3.14 on page 418) of

$$k[[X_1, \ldots, X_n]]$$

will equal units $\times$ polynomials in

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

i. e., polynomials in X_n with coefficients in $k[[X_1, \ldots, X_{n-1}]]$. Unique factorization in $k[[X_1, \ldots, X_{n-1}]]$ and $k[[X_1, \ldots, X_{n-1}]][X_n]$ will imply it in $k[[X_1, \ldots, X_n]]$ (via lemma A.3.1 on page 411).

> *We will fix the following notation throughout the rest of this section:*
> $P_n = k[[X_1, \ldots, X_n]]$, where k is a field.

We need to develop some properties of power-series rings.

PROPOSITION A.3.9. *An element $p \in P_n$ is a unit if and only if its constant term is nonzero.*

PROOF. Straightforward induction using proposition A.1.7 on page 343 and exercise 9 on page 354. $\square$

DEFINITION A.3.10. An element $p(X_1, \ldots, X_n) \in P_n$ will be called X_n-*general* if $p(0, \ldots, 0, X_n) \neq 0$. If $X_n^d | p(0, \ldots, 0, X_n)$ and $X_n^{d+1} \nmid p(0, \ldots, 0, X_n)$ for some integer $d > 0$, we say that p is X_n-*general of degree d*.

REMARK. A power series is X_n-general if it has a term that only involves X_n. For instance $X_1 + X_2$ is X_2-general but $X_1 X_2$ is not.

Next, we have a kind of division algorithm for power-series rings (even though these rings are not Euclidean):

THEOREM A.3.11 (Weierstrass Division Theorem). *Let $p \in P_n$ be X_n-general power-series of degree d that is not a unit of P_n. For every power series $g \in P_n$, there exists a power series $u \in P_n$ and a polynomial $r \in P_{n-1}[X_n]$ of degree $d - 1$ such that*

(A.3.3) $g = u \cdot p + r$

The power-series u and polynomial r are uniquely determined.

REMARK. A shorter way to say this is that

$$\frac{P_n}{(p)} = P_{n-1} \oplus X_n \cdot P_{n-1} \oplus \cdots \oplus X_n^{d-1} P_{n-1}$$

or that it is a module over P_{n-1} generated by $\{1, \ldots, X_n^{d-1}\}$.

PROOF. We will explicitly construct u and r.

For every $f \in P_n$, let $r(f)$ equal the set of terms, T, such that $X_n^d \nmid T$, and let $h(f)$ be the factor of X_n^d in $f - r(f)$. Then

$$f = r(f) + X_n^d h(f)$$

for all power series in P_n. So $r(f), h(f) \in P_n$ and $r(f)$ is a polynomial in $P_{n-1}[X_n]$ of degree $< d$. Note that, regarding P_n as a vector space over k, both $r(*)$ and $h(*)$ are linear maps.

CLAIM A.3.12. In addition, $h(p)$ is a unit (since its constant term is the element of k multiplying X_n^d, and $r(f)$ has no constant terms since f is not a unit.

We claim that equation A.3.3 is equivalent to

(A.3.4) $h(g) = h(u \cdot p)$

for some $u \in P_n$. If equation A.3.3 holds then $h(g - u \cdot p) = 0$ and equation A.3.4 is true. Conversely, if equation A.3.4 on page 416is true, then $h(g - u \cdot p) = 0$ and $g - u \cdot p = r(q - u \cdot p)$, a degree $d - 1$ polynomial in $P_{n-1}[X_n]$.

Since $p = r(p) + X_n^d \cdot h(p)$, equation A.3.4 on the preceding page is equivalent to

(A.3.5) $$u \cdot p = u \cdot r(p) + X_n^d \cdot u \cdot h(p)$$

Since $h(p)$ is a unit (see the claim above), it suffices to compute the power-series $v = u \cdot h(p)$. Set

$$m = -r(f) \cdot h(p)^{-1}$$

Then $u \cdot r(p) = -m \cdot v$ and we can rewrite equation A.3.5 to the equivalent equation

(A.3.6) $$h(g) = -h(m \cdot v) + v$$

or

(A.3.7) $$v = h(g) + s(v)$$

where, for any power series, $f \in P_n$, we have defined $s(f) = h(m \cdot f)$. Note that s is a linear operation on power-series.

Let $m = (X_1, \ldots, X_{n-1}) \subset P_{n-1}$ be the maximal ideal. Note that $r(p) \in m[X_n] \subset P_{n-1}[X_n]$ since it is not a unit (so it has vanishing constant term). This means that, if the coefficients of $f \in P_n = P_{n-1}[[X_n]]$ lie in m^j, then the coefficients of $s(f)$ will lie in m^{j+1}.

Now we plug equation A.3.7 into itself to get

$$
\begin{aligned}
v &= h(g) + s(h(g) + s(v)) \\
&= h(g) + s(h(g)) + s^2(v)
\end{aligned}
$$

We can iterate this any number of times:

$$v = \sum_{j=0}^{t} s^j(h(g)) + s^{t+1}(v)$$

or

$$v - \sum_{j=0}^{t} s^j(h(g)) \in m^{t+1}[[X_n]] \subset P_{n-1}[[X_n]]$$

Since lemma A.1.81 on page 379 implies that

$$\bigcap_{j=1}^{\infty} m^j = (0)$$

we claim that

$$v = \sum_{j=0}^{\infty} s^j(h(g))$$

is the unique solution to our problem. It is easy to verify that it satisfies equation A.3.7. Now all we have to do is set

$$u = v \cdot h(p)^{-1}$$

and $r = r(q - u \cdot p)$. $\square$

The Weierstrass Preparation Theorem is a simple corollary:

THEOREM A.3.13 (Weierstrass Preparation Theorem). *Let $p \in P_n$ be X_n-general power-series of degree d that is not a unit of P_n. Then there exists a unit $q \in P_n$ and a monic polynomial $u \in P_{n-1}[X_n]$ of degree d such that*

(A.3.8) $$p = u \cdot w$$

and u and w are uniquely determined by p.

REMARK. This is the general analogue of A.3.2 on page 415 for power-series of n variables. Weierstrass originally proved it for convergent power series using the methods of complex analysis. It gives valuable information on the behavior of the zero-sets of analytic functions of several complex variables (besides implying that the ring of such functions has unique factorization).

Our proof of the Division Theorem is the "combinatorial" or "algebraic" form — that does not use contour integrals.

The polynomial $w \in P_{n-1}[X_n]$ is called the Weierstrass Polynomial of p.

Case 1.   Apply the division theorem (A.3.11 on page 416) to $g = X_n^d$. It gives u and r such that
$$X_n^d = u \cdot p + r$$
so we get
$$u \cdot p = X_n^d - r = w \in P_{n-1}[X_n]$$
We claim that u must be a unit since the lowest X_n term in p is X_n^d. The only way the product could contain X_n^d is for u to have a nonvanishing constant term.

If A is an $n \times n$ invertible matrix whose entries are in k, then A induces an automorphism
$$\begin{aligned} A^* \colon P_n &\rightarrow P_n \\ p(X_1, \ldots, X_n) &\mapsto p^A = p(A^{-1} \cdot (X_1, \ldots, X_n)) \end{aligned}$$
The inverse is given by the inverse of A.

LEMMA A.3.14. *Let $p \in P_n$ be a power series that is not a unit. If the field k is infinite, then there exists a matrix A such that p^A is X_n-general.*

PROOF. Let L be the leading term of p — this consists of the terms of lowest total degree in the power series and will be a homogeneous polynomial in $X_1, \ldots, X_n$. Let
$$(k_1, \ldots, k_n) \in k^n$$
be a set of values on which L is nonvanishing. Such a set of values exists because we can plug in 1 for all of the X_i except one, and the resulting polynomial of one variable vanishes at a finite number of values. Since the field k is infinite, we can find a value for the remaining variable that makes $L \neq 0$ vanish. Let A be an invertible matrix that transforms this point
$$(k_1, \ldots, k_n)$$
to $(0, \ldots, 0, 1)$. The conclusion follows. $\square$

It is easy (and necessary) to generalize this a bit:

COROLLARY A.3.15. *Let $p_1, \ldots, p_t \in P_n$ be a finite set of power series that are non-units. Then there exists an invertible matrix A such that $p_1^A, \ldots, p_t^A \in P_n$ are all X_n-regular.*

PROOF. Simply apply lemma A.3.14 to the product $p_1 \cdots p_t$. $\square$

We are finally ready to prove the main result:

THEOREM A.3.16. *If k is an infinite field, the ring $P_n = k[[X_1, \ldots, X_n]]$ is a unique factorization domain.*

REMARK. The requirement that k be an infinite field is not really necessary but it simplifies the proof of lemma A.3.14 — and k will be infinite in all of our applications of this result.

Weierstrass originally proved this for $k = \mathbb{C}$ and $P_n = \mathbb{C}\{X_1, \ldots, X_n\}$ — the ring of *convergent* power-series. This is essentially the ring of complex-analytic functions. See [**82**].

PROOF. We prove this by induction on n, the number of indeterminates. The result is almost trivial for $n = 1$ — see A.3.2 on page 415.

Let $p_1 \in P_n$ be irreducible and suppose $p_1 | p_2 \cdot p_2$ and $p_1 \nmid p_2$. We will show that this forces $p_1 | p_3$. Use corollary A.3.15 on the preceding page to transform p_1, p_2, p_3 to X_n-regular power series of degrees d_1, d_2, d_3, respectively. Then the Weierstrass Preparation Theorem (theorem A.3.13 at page 417) implies that

$$
\begin{aligned}
p_1^A &= u_1 \cdot w_1 \\
p_2^A &= u_2 \cdot w_2 \\
p_3^A &= u_3 \cdot w_3
\end{aligned}
$$

where $u_1, u_2, u_3 \in P_n$ are units and $w_1, w_2, w_3 \in P_{n-1}[X_n]$ are the Weierstrass polynomials of the p_i. We claim that the polynomial $w_1 \in P_{n-1}[X_n]$ is irreducible. This is because a nontrivial factorization of it would give a nontrivial factorization of p_1, since A induces an automorphism. Since P_{n-1} is a unique factorization domain by induction and $P_{n-1}[X_n]$ is one by theorem A.3.7 on page 413, we must have

$$w_1 | w_3$$

which implies that

$$p_1^A | p_3^A$$

and

$$p_1 | p_3$$

which means P_n is a unique factorization domain, by lemma A.3.1 on page 411. $\qquad \square$

A.4. Further topics in ring theory

A.4.1. Integral extensions of rings. The theory of integral extensions of rings is crucial to algebraic number theory and algebraic geometry. It considers the question of "generalized integers:"

If $\mathbb{Z} \subset \mathbb{Q}$ is the subring of integers, what subring, $R \subset \mathbb{Q}[\sqrt{2}]$, is like its "ring of integers"?

In algebraic geometry, integral extensions are used to define the *existence* of algebraic sets — see theorem 2.2.2 on page 40 and the concept of *dimension*, see corollary 2.8.23 on page 107.

DEFINITION A.4.1. If $A \subset K$ is the inclusion of an integral domain in a field, $x \in K$ will be called *integral* over A if it satisfies an equation

$$x^j + a_1 x^{j-1} + \cdots + a_k = 0 \in A$$

with the $a_i \in A$ (i.e., is a root of a *monic* polynomial).

REMARK. For instance, consider $\mathbb{Z} \subset \mathbb{Q}$. The only integral elements over $\mathbb{Z}$ are in $\mathbb{Z}$ itself.

In the case of $\mathbb{Z} \subset \mathbb{Q}(i)$, we get integral elements $n_1 + n_2 \cdot i$ where $n_1, n_2 \in \mathbb{Z}$ — the ring of Gaussian Integers.

PROPOSITION A.4.2. *Let $R \subset S$ be integral domains. The following statements are equivalent*

(1) *An element $s \in S$ is integral over R*
(2) *$R[s]$ is a finitely-generated R-module (see definition A.1.65 on page 370).*
(3) *$s \in T$ for some subring of S with $R \subseteq T \subseteq S$ and T is a finitely-generated R-module.*

REMARK. Note that being finitely generated as a module is very different from being finitely generated as a *ring or field.* For instance $R[X]$ is finitely generated as a *ring* over R but, as a *module,* it is

$$\bigoplus_{n=0}^{\infty} R \cdot X^n$$

PROOF. 1 $\implies$ 2. If s is integral over R, then

$$s^n + a_{n-1}s^{n-1} + \cdots + a_0 = 0$$

with the $a_i \in R$, so

$$s^n = -a_{n-1}s^{n-1} - \cdots - a_0$$

This means that $R[s]$ — the ring of polynomials in s will only have polynomials of degree $< n$, so $R[s]$ will be finitely generated as a *module* over R. Compare this argument to that used in proposition A.2.12 on page 390.

2 $\implies$ 3. Just set $T = R[s]$.

3 $\implies$ 1. Suppose that $t_1, \ldots, t_n \in T$ is a set of generators of T as an R-module. Then

$$st_i = \sum_{j=1}^{n} A_{i,j}t_j$$

for some $n \times n$ matrix A, so

$$\sum_{j=1}^{n} (\delta_{i,j}s - A_{i,j})t_j = 0$$

where

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Cramer's Rule implies that

$$\det(sI - A)t_i = C_i = 0$$

for all i, where C_i is the determinant of the matrix one gets by replacing the i[th]column by 0's. It follows that s is a root of the monic polynomial

$$\det(XI - A) = 0 \in R[X]$$

$\square$

DEFINITION A.4.3. If $R \subseteq S$ is an inclusion of integral domains and every element of S is integral over R, then S will be said to be *integral* over R.

REMARK. It is not hard to see that this property is preserved in *quotients.* If $\mathfrak{a} \subset S$ is an ideal then $S/\mathfrak{a}$ will be integral over $R/\mathfrak{a} \cap R$ because the monic polynomials satisfied by every element of S over R will map to monic polynomials in the quotients.

COROLLARY A.4.4. *Let $f: R \to S$ be an integral extension of integral domains. If $\mathfrak{a} \subset R$ is a proper ideal, then so is $\mathfrak{a} \cdot S \subset S$.*

PROOF. We will prove the contrapositive: If $\mathfrak{a} \cdot S = S$, then $\mathfrak{a} = R$. The statement that $\mathfrak{a} \cdot S = S$ and Nakayama's Lemma A.1.80 on page 378 imply that there exists $r \in R$ with $r \equiv 1 \pmod{\mathfrak{a}}$ with $r \cdot S = 0$. Since S is an integral domain, we must have $r = 0$ to $0 \equiv 1 \pmod{\mathfrak{a}}$ or $1 \in \mathfrak{a}$, so $\mathfrak{a} = R$. $\square$

PROPOSITION A.4.5. *Suppose $R \subseteq S$ are integral domains and let $s, t \in S$. Then:*

(1) *If s and t are integral over R, so are $t \pm s$ and st. Consequently integral elements over R form a ring.*

(2) *Let T be a commutative ring with $S \subseteq T$. If T is integral over S and S is integral over R, then T is integral over R.*

PROOF. If s and t are integral over R, let

$$\begin{aligned} R[s] &= Rs_1 + \cdots + Rs_k \\ R[t] &= Rt_1 + \cdots + Rt_\ell \end{aligned}$$

as R-modules. Then

(A.4.1)
$$\begin{aligned} R[s,t] &= Rs_1 + \cdots + Rs_k \\ &\quad + Rt_1 + \cdots + Rt_\ell \\ &\quad + \sum_{i=1,j=1}^{k,\ell} Rs_i t_j \end{aligned}$$

which contains $s \pm t$ and st and is still a finitely generated R-module. This proves the first statement.

To prove the second statement, suppose $t \in T$ satisfies the monic polynomial

$$t^k + s_{k-1}t^{k-1} + \cdots + s_0 = 0$$

with $s_i \in S$, and $S[t]$ is a finitely generated S-module. Since S is integral over R, $R[s_i]$ is a finitely-generated R-module, and so is

$$R' = R[s_0, \ldots, s_{k-1}]$$

— equation A.4.1 gives some ideal of how one could obtain a finite set of generators. The element t is also monic over R', so $R'[t]$ is a finitely-generated R'-module and

$$R[s_0, \ldots, s_{k-1}, t]$$

is a finitely-generated R-module. It follows that t is integral over R. $\square$

This immediately implies:

COROLLARY A.4.6. *If $R \subset S$ is an inclusion of integral domains and $\alpha \in S$ is integral over R, then $R \subset R[\alpha]$ is an integral extension of rings.*

Prime ideals behave in an interesting way in an integral extension:

PROPOSITION A.4.7 (Lying Over and Going Up). *Suppose $R \subset S$ is an integral extension of rings. Given a prime ideal $\mathfrak{p} \subset R$, there exists a prime ideal $\mathfrak{q} \subset S$ with*

$$R \cap \mathfrak{q} = \mathfrak{p}$$

The ideal $\mathfrak{q}$ may be chosen to contain an ideal $\mathfrak{q}_1$ that satisfies the condition $R \cap \mathfrak{q}_1 \subset \mathfrak{p}$.

REMARK. The first statement is called "lying over" since every prime $\mathfrak{p}$ in R has one in S "lying over it." The second is called "going up" because it gives a prime $\mathfrak{p}$ that is "up" from $\mathfrak{q}_1$.

PROOF. First, factor out $\mathfrak{q}_1$ and $R \cap \mathfrak{q}_1$ so that we may assume $\mathfrak{q}_1 = 0$. We need only prove that a prime $\mathfrak{q} \subset R$ exists with $R \cap \mathfrak{q} = \mathfrak{p}$. Let $M = R \setminus \mathfrak{p}$, a multiplicatively closed set. If we replace R by $R_\mathfrak{p} = R[M^{-1}]$ (see definition A.1.95 on page 385), we may assume R is local with a unique maximal ideal $\mathfrak{p}$. Since corollary A.4.4 on page 420 implies that $\mathfrak{p} \cdot S \neq S$, any maximal ideal of S containing $\mathfrak{p} \cdot S$ will have the desired properties. □

PROPOSITION A.4.8 (Incomparability Property). *Let $R \subset S$ be an integral extension of rings and let $\mathfrak{p} \subset \mathfrak{q}$ be two ideals of S. Then*

$$\mathfrak{p} \cap R = \mathfrak{q} \cap R$$

implies that $\mathfrak{p} = \mathfrak{q}$.

PROOF. We prove this by contradiction. Let $x \in \mathfrak{q} \setminus \mathfrak{p}$. It is not hard to see that $R/R \cap \mathfrak{p} \subset S/\mathfrak{p}$ is an integral extension (see the remark following definition A.4.3 on page 420). If $\bar{x}$ is the image of x in $S/\mathfrak{p}$ then $\bar{x}$ must satisfy a monic polynomial of minimal degree

$$\bar{x}^k + c_{k-1}\bar{x}^{k-1} + \cdots + c_0 = 0$$

with the $c_i \in R/\mathfrak{p}$.

Since $x \in \mathfrak{q}$, we conclude that $c_0 \in \mathfrak{q}/\mathfrak{p}$. Since $c_0 \in R/\mathfrak{p}$ and $\mathfrak{p} \cap R = \mathfrak{q} \cap R$, we conclude that $c_0 = 0$ and the polynomial has no constant term.

Since $R/\mathfrak{p}$ is an integral domain, we can factor off $\bar{x}$ to get a polynomial of *lower* degree that $\bar{x}$ satisfies, a contradiction. □

Integral extensions of rings have interesting properties where fields are concerned:

PROPOSITION A.4.9. *If R and S are integral domains, and S is an integral extension of R, then S is a field if and only if R is a field.*

PROOF. If R is a field, and $s \in S$ is a nonzero element, then $s \in R[s]$ is a finitely generated module over R — i.e., a vector space. Since S is an integral domain, multiplication by s induces a linear transformation of $R[s]$ whose kernel is 0. This means it is an isomorphism and has an inverse.

Conversely, if S is a field and $r \in R$. Then $r^{-1} \in S$ and it satisfies a monic polynomial over r:

$$r^{-n} + a_{n-1}r^{-(n-1)} + \cdots + a_0 = 0$$

with the $a_i \in R$. If we multiply this by r^{n-1}, we get

$$r^{-1} + a_{n-1} + \cdots + a_0 r^{n-1} = 0$$

□

DEFINITION A.4.10. If K is a field containing an integral domain, R, the ring of elements of K that are integral over R will be called the *integral closure* of R in K. If K is the field of fractions of R and its integral closure is equal to R itself, R will be called *integrally closed* or *normal*.

REMARK. Proposition A.4.5 on the previous page shows that the set of all integral elements over R form a ring — the integral closure of R.

PROPOSITION A.4.11. *Every unique factorization domain is integrally closed.*

REMARK. This shows that $\mathbb{Z}$ is integrally closed in $\mathbb{Q}$. It is possible for R to be normal but not integrally closed in a field *larger* than its field of fractions. For instance $\mathbb{Z}$ is integrally closed in $\mathbb{Q}$ but not in $\mathbb{Q}[\sqrt{2}]$.

PROOF. Let a/b be integral over A, with $a, b \in A$. If $a/b \notin A$ then there is an irreducible element p that divides b but not a. As a/b is integral,

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

Multiplying by b^n gives

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

Now p divides every term of this equation except the first. This is a contradiction! $\qquad\square$

A simple induction, using theorem A.3.7 on page 413 shows that

COROLLARY A.4.12. *For any $n > 0$, the rings $\mathbb{Z}[X_1, \ldots, X_n]$, $F[[X_1, \ldots, X_n]]$, and $F[X_1, \ldots, X_n]$, where F is any field, have unique factorization and are integrally closed (see definition A.4.10 on the preceding page) in their respective fields of fractions.*

REMARK. In most of these examples, the rings are not Euclidean.

Normality of a ring is a "local" property:

PROPOSITION A.4.13. *An integral domain, R, is normal if and only if its localizations, $R_{\mathfrak{p}}$, at all primes are normal.*

PROOF. If R is normal and $S \subset R$ is any multiplicative set, the solution to exercise 1 on page 427 implies that the integral closure of $S^{-1}R$ is $S^{-1}R$. The converse follows from the fact that

(A.4.2) $$R = \bigcap_{\text{all primes } \mathfrak{p} \subset R} R_{\mathfrak{p}} \subset F$$

$\qquad\square$

The following result gives a test for an element being integral over a ring

LEMMA A.4.14. *Let R be an integral domain with field of fractions, F, let $F \subset H$ be a finite extension of fields, and let $\alpha \in H$ be integral over R. Then*

(1) *all conjugates of α (in the algebraic closure of H) are integral over R,*
(2) *all coefficients of the characteristic polynomial, $\chi_\alpha(X) \in F[X]$, are integral over R,*
(3) *the norm $N_{H/F}(\alpha) \in F$ is integral over R.*
(4) *If $\alpha \in H$ is integral over R, then its minimal polynomial over F has coefficients in R.*

REMARK. If R is normal, this implies that $\chi_\alpha(X) \in R[X]$ and provides a necessary and sufficient condition for α to be integral.
For instance, $a + b\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$ is integral over $\mathbb{Z}$ if and only if

$$\chi_\alpha(X) = X^2 - 2aX + a^2 - 2b^2 \in \mathbb{Z}[X]$$

This implies that all elements $a + b\sqrt{2}$ with $a, b \in \mathbb{Z}$ are integral over $\mathbb{Z}$. Since $-2a \in \mathbb{Z}$, the only other possibility is for

$$a = \frac{2n+1}{2}$$

Plugging this into

$$\frac{(2n+1)^2}{4} - 2b^2 = m \in \mathbb{Z}$$

or

$$b^2 = \frac{(2n+1)^2 - 4m}{8}$$

which is never an integer much less a square, giving a contradiction.

PROOF. Let $p(X) \in R[X]$ be a monic polynomial such that $p(\alpha) = 0$. If α' is any conjugate of α, then the isomorphism

$$F[\alpha] \to F[\alpha']$$

that leaves F and $R \subset F$ fixed implies that $p(\alpha') = 0$ as well. The statement about the characteristic polynomial follows from the fact that its coefficients are elementary symmetric functions of the conjugates of α (see equation A.1.8 on page 364), and the fact that the set of integral elements form a ring (see proposition A.4.5 on page 421).

Statement 3, about the norm follows, from lemma A.2.34 on page 402.

The final statement follows from statement 1 and equation A.1.8 on page 364 which imply that all the coefficients of the minimal polynomial (in F) are integral over R — so they must be *in R*. $\square$

We conclude this section with a result on the behavior of integral closures under algebraic field extensions. To prove it, we will need the concept of *bilinear form*:

DEFINITION A.4.15. If V is a vector-space over a field, F, a *bilinear form* on V is a function

$$b \colon V \times V \to F$$

such that

(1) $b(c \cdot v_1, v_2) = b(v_1, c \cdot v_2) = c \cdot b(v_1, v_2)$ for all $v_1, v_2 \in V$ and $c \in F$.
(2) $b(v_1 + w, v_2) = b(v_1, v_2) + b(w, v_2)$ for all $v_1, v_2, w \in V$.
(3) $b(v_1, w + v_2) = b(v_1, w) + b(v_1, v_2)$ for all $v_1, v_2, w \in V$.

A bilinear form, $b(*, *)$, is called *symmetric* if $b(v_1, v_2) = b(v_2, v_1)$ for all $v_1, , v_2 \in V$. If $v = \{v_1, \ldots, v_n\}$ is a basis for V, then the *associated matrix* of b is M defined by

$$M_{i,j} = b(v_i, v_j)$$

A bilinear form, $b(*, *)$, is said to be *degenerate* if there exists a nonzero vector $v \in V$ such that $b(v, w) = 0$ for all $w \in W$.

REMARK. If M is the associated matrix of b, then we can write b as

(A.4.3) $$b(u, v) = u^T M v$$

where u and v are vectors expanded in the basis used to compute M, and u^T is the transpose.

PROPOSITION A.4.16. *Let V be a vector space over a field, F, equipped with a bilinear form*

$$b: V \times V \to F$$

Then b is nondegenerate if and only if its associated matrix is invertible.

PROOF. If M is invertible, then $u^T M \neq 0$ if $u \neq 0$ and we can define $v = (u^T M)^T = M^T u$ in which case

$$b(u, v) = \|u^T M\|^2 \neq 0$$

by equation A.4.3 on the preceding page. If M is not invertible, it sends some nonzero vector u to 0 and

$$b(u, v) = 0$$

for all $v \in V$.   □

We will be interested in nondegenerate bilinear forms because:

PROPOSITION A.4.17. *Let V be a vector-space over a field F with basis $\{u_1, \ldots, u_n\}$ and suppose that*

$$b: V \times V \to F$$

is a nondegenerate bilinear form. Then there exists a dual basis $\{v^1, \ldots v^n\}$ of V such that

$$b(u_i, v^j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

PROOF. If M is the associated matrix (with respect to the u-basis), simply define

$$v = M^{-1} u$$

The conclusion follows from equation A.4.3 on the facing page.   □

Now we introduce a special bilinear form significant in studying field extensions:

DEFINITION A.4.18. Let $F \subset H$ be a finite extension of fields. Then define the *trace form* of H over F via

$$b_{H/F}(h_1, h_2) = T_{H/F}(h_1 \cdot h_2)$$

(see section A.2.3 on page 396 for information about $T_{H/F}$).

REMARK. Lemma A.2.23 on page 396 implies that trace form is bilinear, and it is easy to see that it is also symmetric.

It is interesting to consider what happens if the trace form is degenerate. In this case, there exists $h \in H$ such that $b_{H/F}(h, h') = 0$ for all $h' \in H$, in particular, when $h' = h^{-1}$. It follows that $b_{H/F}(h, h^{-1}) = T_{H/F}(1) = 0$. But lemma A.2.34 on page 402 implies that

$$T_{H/F}(1) = [H:F] \cdot 1 = 0 \in F$$

The only way this can happen is if F has finite characteristic, p, and $p | [H:F]$. This happens when H is an inseparable extension of F (see definition A.2.19 on page 394).

LEMMA A.4.19. *If $F \subset H$ is a separable extension of fields, then the trace form, $b_{H/F}$, is nondegenerate.*

REMARK. Note that "separable" implies "finite."

PROOF. Since the extension is separable, theorem A.2.21 on page 394 implies that there exists a primitive element $\alpha \in H$ such that $H = F[\alpha]$. If $[H:F] = n$, then $\{1, \alpha, \ldots, \alpha^{n-1}\}$ are a basis for H over F.

We have $b_{H/F}(\alpha^i, \alpha^j) = T_{H/F}(\alpha^{i+j})$ and the associated matrix to $b_{H/F}$ is given by

$$M_{i,j} = T_{H/F}(\alpha^{i-1} \cdot \alpha^{j-1}) = T_{H/F}(\alpha^{i+j-2})$$

Let $\bar{H}$ be the algebraic closure of H and let $\alpha = \alpha_1, \ldots, \alpha_n$ be the conjugates of α in $\bar{H}$ (see definition A.2.33 on page 401). Lemma A.2.34 on page 402 implies that

$$T_{H/F}(\alpha^j) = \sum_{i=1}^{n} \alpha_i^j$$

Let V be the Vandermonde matrix $V(\alpha_1, \ldots, \alpha_n)$ — see exercise 34 on page 367. It is defined by

$$V_{i,j} = \alpha_i^{j-1}$$

Now, note that

$$\begin{aligned} M_{i,j} &= \sum_{\ell=1}^{n} \alpha_\ell^{i-1} \cdot \alpha_\ell^{j-1} \\ &= (V^T V)_{i,j} \end{aligned}$$

It follows that

$$\det M = (\det V)^2 = \prod_{1 \le i < j \le n} (\alpha_j - \alpha_i)^2$$

which is nonzero since the α_i are all distinct (because the field-extension was separable). □

Now we can prove our main result regarding integral extensions:

LEMMA A.4.20. *Suppose that A is integrally closed domain whose field of fractions is F. Let $F \subset H$ be a separable extension of fields of degree n, and let B be the integral closure of A in H. Then there exists a basis $\{v_1, \ldots, v_n\}$ for H over F such that*

$$B \subseteq \{v^1, \ldots, v^n\} \cdot A$$

If A is noetherian, this implies that B is a finitely generated module over A.

REMARK. Roughly speaking, this says that a finite extension of fields induces a finite extension of integrally closed rings.

PROOF. Let $\{u_1, \ldots, u_n\}$ be a basis for H over F. Each of the u_i satisfies an algebraic equation

$$a_n u_i^n + \cdots + a_0 = 0$$

and multiplying by a_n^{n-1} gives us a monic polynomial in $(a_n u_i)$ so it is integral over A. It follows that $a_n u_i \in B$ and — without loss of generality — we may assume that the basis elements $u_i \in B$.

This does not prove the result: we have only shown that every element of H can be expressed in terms of B and F.

Let $\{v^1, \ldots, v^n\}$ be the dual basis defined by the trace form, via proposition A.4.17 on page 425. This exists because the trace form is nondegenerate, by lemma A.4.19 on page 425.

If $x \in B$, let

(A.4.4)
$$x = \sum_{i=1}^{n} c_i v^i$$

where the $c_i \in F$. Note that $x \cdot u_i \in B$ since x and each of the u_i are in B. We claim that

$$b_{H/F}(x, u_i) = T_{H/F}(x \cdot u_i) \in A$$

This is because $x \cdot u_i$ satisfies a monic polynomial with coefficients in A — and $T_{H/F}(x \cdot u_i)$ is the negative of the coefficient of X^{n-1} (see definition A.2.22 on page 396 and the remark following it). We use the properties of the dual basis to conclude

$$
\begin{aligned}
T_{H/F}(x \cdot u_i) &= T_{H/F}\left(\left(\sum_{j=1}^{n} c_j v^j \right) \cdot u_i \right) \\
&= \sum_{j=1}^{n} c_j \cdot T_{H/F}(v^j \cdot u_i) \\
&= c_i
\end{aligned}
$$

So, in equation A.4.4, the c_i were elements of A all along and the conclusion follows. □

EXERCISES.

1. Let $R \subset T$ be an inclusion of rings and let $\bar{R}$ be its integral closure in T. Show, for any multiplicative set S, that $S^{-1}\bar{R}$ is the integral closure of $S^{-1}R$ in $S^{-1}T$.

2. Suppose R is an integral domain with field of fractions F and H is a finite extension of F. If $x \in H$ show that there exists an element $w \in R$ such that $r \cdot x$ is integral over R.

3. Let $R \subset T$ be an inclusion of rings with the property that T is a finitely-generated module over R. Now let $T \subset F$ where F is a field. Show that the integral closure of T in F is the same as the integral closure of R in F.

A.4.2. The Jacobson radical and Jacobson rings. In this section, we give a very brief treatment of a construct similar to the nilradical.

DEFINITION A.4.21. If R is a commutative ring, the *Jacobson radical*, $\mathfrak{J}(R)$, of R is defined by

$$\mathfrak{J}(R) = \bigcap_{\text{maximal } \mathfrak{m} \subset R} \mathfrak{m}$$

— the intersection of all of the maximal ideals of R.

REMARK. Since the nilradical is the intersection of all prime ideals and maximal ideals are prime, it is easy to see that

$$\mathfrak{N}(R) \subset \mathfrak{J}(R)$$

is always true.

DEFINITION A.4.22. A commutative ring, R, is called a *Jacobson ring* if for any ideal $\mathfrak{I} \subset R$

$$\sqrt{\mathfrak{I}} = \bigcap_{\mathfrak{I} \subset \mathfrak{m}} \mathfrak{m}$$

where the intersection is taken over all maximal ideals containing $\mathfrak{I}$.

REMARK. The term *Jacobson ring* was coined by Krull in [**99**] in honor of the notable American mathematician, Nathan Jacobson (1910–1999). Krull used Jacobson rings to generalize Hilbert's Nullstellensatz (theorem 2.2.5 on page 41). Because of their relation to the Nullstellensatz, they are sometimes called Hilbert rings or Jacobson-Hilbert rings.

Theorem A.1.47 on page 360 shows that $\sqrt{\mathfrak{I}}$ is the intersection of all prime ideals containing $\mathfrak{I}$. In a Jacobson ring, there are "enough" maximal ideals so the corresponding statement is true for the primes that are maximal.

We can characterize Jacobson rings by how prime ideals behave:

PROPOSITION A.4.23. *The following statements are equivalent*

(1) R is a Jacobson ring
(2) every prime ideal $\mathfrak{p} \subset R$ satisfies

(A.4.5) $$\mathfrak{p} = \bigcap_{\mathfrak{p} \subset \mathfrak{m}} \mathfrak{m}$$

where the intersections is taken over all maximal ideals containing $\mathfrak{p}$.
(3) $\mathfrak{J}(R') = 0$ for every quotient, R', of R that is an integral domain.

PROOF. If R is Jacobson, the statement is clearly true because prime ideals are radical, so 1 $\implies$ 2. Conversely, if $\mathfrak{I} \subset R$ is any ideal, theorem A.1.47 on page 360 implies that $\sqrt{\mathfrak{I}}$ is the intersection of all prime ideals that contain $\mathfrak{I}$ and equation A.4.5 implies that each of these is the intersection of all the maximal ideals that contain it. It follows that the condition in definition A.4.22 on page 428 is satisfied, so 2 $\implies$ 1. Statement 2 is equivalent to statement 2 because $R' = R/\mathfrak{p}$ for some prime ideal and lemma A.1.25 on page 352 implies that the maximal ideals of $R/\mathfrak{p}$ are in a one to one correspondence with the maximal ideals of R containing $\mathfrak{p}$.   □

This immediately implies:

COROLLARY A.4.24. *Every quotient of a Jacobson ring is Jacobson.*

It is not hard to find examples of Jacobson rings:

PROPOSITION A.4.25. *A principle ideal domain is a Jacobson ring if and only if it has an infinite number of prime ideals.*

REMARK. We immediately conclude that

(1) any field is a Jacobson ring,
(2) $\mathbb{Z}$ is a Jacobson ring,
(3) $k[X]$ is a Jacobson ring, where k is any field. An argument like that used in number theory implies that $k[X]$ has an infinite number of primes.

PROOF. We use the characterization of Jacobson rings in proposition A.4.23 on the facing page. Let R denote the ring in question — this is a unique factorization domain (see remark A.1.39 on page 357). *All* of the prime ideals of R are *maximal* except for (0). It follows that all prime ideals are equal to the intersection of maximal ideals that contain them, with the *possible* exception of (0).

If there are only a finite number of prime ideals, $(x_1), \ldots, (x_k)$ then
$$(x_1) \cap \cdots \cap (x_k) = (x_1 \cdots x_k) \neq (0)$$
so R fails to be Jacobson.

If there are an infinite number of prime ideals and $x \neq 0 \in R$ is an arbitrary element, then x factors as a finite product of primes. It follows that there exists a prime *not* in this factorization so that $x \notin \mathfrak{J}(R)$ — since nonzero prime ideals are maximal. It follows that the intersection of all maximal ideals that contain (0) *is* (0) and the ring is Jacobson. □

It is well-known that Jacobson rings are *polynomially-closed*: if J is a Jacobson ring, so is $J[X]$. To prove this, we need what is widely known as the Rabinowich Trick (which first appeared in [**141**]):

LEMMA A.4.26. *The following statements are equivalent:*

(1) the ring R is Jacobson
(2) if $\mathfrak{p} \subset R$ is any prime ideal and $S = R/\mathfrak{p}$ has an element $t \in S$ such that $S[t^{-1}]$ is a field, then S is a field.

PROOF. If R is Jacobson, so is S. The prime ideals of $S[t^{-1}]$ are those of S that do not contain t. Since $S[t^{-1}]$ is a field, it follows that t is contained in every nonzero prime ideal. If any nonzero prime ideals existed in S, t would be contained in them. Since R is Jacobson, so is S and $\mathfrak{J}(R) = 0$ (see proposition A.4.23 on page 428), so there cannot exist any nonzero prime ideals, and S must be a field.

Conversely, suppose the hypotheses are satisfied and $\mathfrak{p} \subset R$ is a prime ideal with
$$\mathfrak{p} \subsetneq \bigcap_{\mathfrak{p} \subset \mathfrak{m}} \mathfrak{m}$$
where the intersection is taken over all maximal ideals containing $\mathfrak{p}$. We will derive a contradiction.

If $t \in \bigcap_{\mathfrak{p} \subset \mathfrak{m}} \mathfrak{m} \setminus \mathfrak{p}$, the set of prime ideals, $\mathfrak{q}$, with $t \notin \mathfrak{q}$ has a maximal element (by Zorn's Lemma — A.1.27 on page 353), $\mathfrak{Q}$. This ideal is not maximal since t is contained in all maximal ideals, so $R/\mathfrak{Q}$ is not a field. On the other hand $\mathfrak{Q}$ generates a maximal ideal of $R[t^{-1}]$ so
$$R[t^{-1}]/\mathfrak{Q} \cdot R[t^{-1}] = (R/\mathfrak{Q})[t^{-1}]$$
(see lemma A.5.60 on page 469) is a field. The hypotheses imply that $R/\mathfrak{Q}$ is also a field — which is a contradiction. □

We need one more lemma to prove our main result:

LEMMA A.4.27. *Let R be a Jacobson domain and let S be an algebra over R generated by a single element, i.e. $S = R[\alpha]$ and an integral domain. If there exists an element $t \in S$ such that $S[t^{-1}]$ is a field, then R and S are both fields, and S is a finite extension of R.*

PROOF. Let F be the field of fractions of R. We have $S = R[X]/\mathfrak{p}$ where $\mathfrak{p} \subset R[X]$ is a prime ideal and X maps to α under projection to the quotient. We claim that $\mathfrak{p} \neq (0)$. Otherwise, there would exist an element $t \in R[X]$ that makes $R[X][t^{-1}]$ a field. Since $R[X][t^{-1}] = F[X][t^{-1}]$, the fact that $F[X]$ is known to be Jacobson (by proposition A.4.25 on page 428) and lemma A.4.26 on the preceding page imply that $F[X]$ is *also* a field, which is a contradiction.

Since $\mathfrak{p} \neq 0$, let $p(X) \in \mathfrak{p}$ be any nonzero polynomial

$$p_n X^n + \cdots + p_0$$

that vanishes in S. In $S[p_n^{-1}]$ we may divide by p_n to get a monic polynomial — showing that α is integral over $R[p_n^{-1}]$ (see definition A.4.1 on page 419) so corollary A.4.6 on page 421 implies that $S[p_n^{-1}]$ is integral over $R[p_n^{-1}]$.

Let

(A.4.6) $$c_n t^n + \cdots + c_0 = 0$$

be a polynomial that t satisfies in S (factor off copies of t to guarantee that $c_0 \neq 0$). Now, invert $p_n c_0$ in R and S, so we get $S[(p_n c_0)^{-1}]$ integral over $R[(p_n c_0)^{-1}]$. After doing this, we can divide equation A.4.6 by $c_0 t^n$ in $S[(c_0 p_n)^{-1}, t^{-1}]$ to get a monic polynomial for t^{-1}

$$t^{-n} + \left(\frac{c_1}{c_0}\right) t^{-(n-1)} + \cdots + \frac{c_n}{c_0} = 0$$

It follows that $S[(c_0 p_n)^{-1}, t^{-1}]$ is integral over $R[(p_n c_0)^{-1}]$. Since $S[t^{-1}]$ is a field, so is $S[(c_0 p_n)^{-1}, t^{-1}]$ (the same field) and proposition A.4.9 on page 422 implies that $R[(p_n c_0)^{-1}]$ is *also* a field. The fact that R is Jacobson, and lemma A.4.26 on the preceding page implies that R is also a field. So $R = R[p_n^{-1}]$ and $R[\alpha] = S$ is integral over R. Proposition A.4.9 on page 422 applied a second time implies that S is *also* a field and the conclusion follows. $\qquad \square$

We are now ready to prove the main result:

THEOREM A.4.28. *If R is a Jacobson ring, any finitely generated algebra over R is also a Jacobson ring.*

REMARK. This result provides a huge number of Jacobson rings:
- $\mathbb{Z}[X_1, \ldots, X_n]$
- all affine *k*-algebras (see definition 2.4.13 on page 67). Hilbert's Nullstellensatz (theorem 2.2.5 on page 41) already proved that affine *k*-algebras were Jacobson rings. In fact, theorem A.4.28 is often regarded as a *generalization* of the Nullstellensatz.

PROOF. We start with $S = R[\alpha]$. The general case follows by a simple induction. If $\mathfrak{p} \subset S$ is a prime ideal, then $S/\mathfrak{p}$ will be an integral domain and the image of R in $S/\mathfrak{p}$ will be a Jacobson domain. If there exists $t \in S/\mathfrak{p}$ such that $(S/\mathfrak{p})[t^{-1}]$ is a field, lemma A.4.27 implies that $S/\mathfrak{p}$ (and, for that matter $R/R \cap \mathfrak{p}$) is *also* a field — satisfying the conditions of lemma A.4.27 on the preceding page. It follows that S is Jacobson. □

It is also easy to find *non*-Jacobson rings:

EXAMPLE A.4.29. If $\mathfrak{t} = 2 \cdot \mathbb{Z} \subset \mathbb{Z}$, then $\mathbb{Z}_\mathfrak{t}$, is the ring[1] of rational numbers with odd denominators. This is a local ring with a unique maximal ideal, $2 \cdot \mathbb{Z}_{(2)}$ so $\mathfrak{J}(\mathbb{Z}_\mathfrak{t}) = 2 \cdot \mathbb{Z}_\mathfrak{t}$ but $\mathfrak{N}(\mathbb{Z}_\mathfrak{t}) = 0$, since it is an integral domain.

This example induces many more

EXAMPLE A.4.30. Let $R = \mathbb{Z}_\mathfrak{t}[X_1, \ldots, X_n]$ be a polynomial ring over $\mathbb{Z}_\mathfrak{t}$ from example A.4.29 above. The maximal ideals of $\mathbb{Q}[X_1, \ldots, X_n]$ are of the form $(X_1 - q_1, \ldots, X_n - q_n)$. If we restrict the q_i to be in $\mathbb{Z}_\mathfrak{t}$, we get ideals of R that are no longer maximal because the quotient of R by them is $\mathbb{Z}_\mathfrak{t}$, which is not a field. We can make these ideal maximal by adding one additional element. The ideals
$$\mathfrak{L}(q_1, \ldots, q_n) = (2, X_1 - q_1, \ldots, X_n - q_n)$$
are maximal because the quotient of R by them is $\mathbb{Z}_2$. The intersection of the ideals $\mathfrak{L}(q_1, \ldots, q_n)$ contains (at least) (2) or $2 \cdot R$. Since R is an integral domain, $\mathfrak{N}(R) = 0$ but $(2) \subset \mathfrak{J}(R)$. So R is not Jacobson, either.

A.4.3. Discrete valuation rings. In this section we define a class of rings that is important in algebraic geometry, especially in the theory of divisors in section 5.9 on page 284. Their main property is that they have an especially simple ideal-structure.

Krull introduced them with the concept of *valuation* in his work on algebraic number theory in [94]. A *valuation* on a field is a function that can be used to define a metric on this field. We have already seen an example of this in claim A.1.6 on page 343 — the function $v(x)$, there, is an example of a valuation.

DEFINITION A.4.31. Let F be a field and let $F^\times \subset F$ denote the subset of nonzero elements. A *discrete valuation* on F is a surjective function
$$v: F^\times \to \mathbb{Z}$$
with the properties:
  (1) $v(x \cdot y) = v(x) + v(y)$ for all $x, y \in F^\times$
  (2) $v(x + y) \geq \min(v(x), v(y))$

REMARK. Statement 2 implies that $v(1) = v(1 \cdot 1) = v(1) + v(1) = 0$. If $0 < \alpha < 1$ is some real number, it is not hard to see that
$$\alpha^{v(*)}: F^\times \to [0, 1]$$
defines a *metric* on F, where we define the metric of 0 to be 0.

The definition of valuation easily implies the following properties:

PROPOSITION A.4.32. *Let F be a field with valuation $v: F^\times \to \mathbb{Z}$. Then*

---

[1] We do not use the notation $\mathbb{Z}_{(2)}$ because that would conflict with the notation for 2-adic integers (see example A.5.30 on page 456).

(1) $v(x^{-1}) = -v(x)$ for all $x \in F^*$, since $v(x \cdot x^{-1}) = v(1) = 0 = v(x) + v(x^{-1})$.
(2) $v(-1) = 0$, because $v(1) = v((-1) \cdot (-1)) = v(-1) + v(-1)$ if the characteristic of F is $\neq 2$. If it is 2, then $-1 = 1$, so the statement still holds.
(3) $v(-x) = v(x)$, because $-x = x \cdot (-1)$.
(4) if $v(x) > v(y)$ then $v(x + y) = v(y)$. Certainly, it must be $\geq v(y)$ but, if we write $y = x + y - x$, we get $v(y) \geq \min(v(x + y), v(x))$.

It is easy to find examples of valuations:

EXAMPLE A.4.33. If $F = \mathbb{Q}$ and $p \in \mathbb{Z}$ is any prime then we can define the p-adic valuation, v_p, as follows:

For any $q \in \mathbb{Q}$ we have a unique representation

$$q = \prod p_i^{n_i}$$

where the p_i are primes and $n_i \in \mathbb{Z}$ are integers (which are negative if a prime only occurs in the denominator of q). If $p = p_j$, define

$$v_p(q) = n_j$$

It is well-known that the p-adic valuations constitute *all* of the discrete valuations on $\mathbb{Q}$ — see [**91**].

If a field, F, has a valuation

$$v \colon F^\times \to \mathbb{Z}$$

proposition A.4.32 on the preceding page implies that the set of elements $x \in F$ with $v(x) \geq 0$ form a ring, i.e., are closed under addition and multiplication.

Other interesting examples are provided by power-series rings and variants

EXAMPLE A.4.34. If k is a field, $R = k[[X]]$ is the ring power-series in X and $F = k((X))$ is the field of fractions of R, exercise 48 on page 386 implies that every element in F can be written uniquely in the form

$$f = X^\alpha \cdot r$$

with $r \in R$. It follows that F is a field with valuation given by $v(f) = \alpha$. The subring of elements with a valuation ≥ 0 is precisely $R \subset F$.

There are valuations that are *not* discrete in the sense above:

EXAMPLE A.4.35. We can also define the field of *Puiseux series*, discovered by Isaac Newton in 1676 ([**127**]) and rediscovered by Victor Puiseux ([**139**]) in 1850:

$$k\{\{X\}\} = \bigcup_{n=1}^{\infty} k((X^{1/n}))$$

An argument analogous to that used in the power-series case implies that every element of $k\{\{X\}\}$ can be uniquely written in the form

$$f = X^q \cdot \left(a_0 + a_1 X^{1/n} + a_2 X^{2/n} + \cdots \right)$$

for some $n \in \mathbb{Z}$, $n \geq 1$, some $q \in \mathbb{Q}$, and $a_0 \neq 0 \in k$. We can define the *valuation* of f to be q.

REMARK. If k is algebraically closed and of characteristic 0, it turns out that $k\{\{X\}\}$ is the *algebraic closure* of $k((X))$. Newton sketched a proof in a letter he wrote in 1676. See [**131**] for a short modern proof.

DEFINITION A.4.36. Let R be an integral domain with field of fractions F. Then R is a *discrete valuation ring* if there exists a valuation

$$v \colon F^\times \to \mathbb{Z}$$

such that
$$R = \{x \in F | v(x) \geq 0\}$$

This ring has an ideal
$$\mathfrak{m} = \{x \in F | v(x) > 0\}$$

The notation for a discrete valuation ring is $(R, \mathfrak{m})$.

REMARK. The properties of a valuation (in proposition A.4.32 on page 431) imply that $\mathfrak{m}$ is an ideal and that all $x \in R \setminus \mathfrak{m}$ are units, so $\mathfrak{m}$ is the unique maximal ideal, and R is a local ring.

For the p-adic valuation on $\mathbb{Q}$, the corresponding discrete valuation ring is $R_p \subset \mathbb{Q}$ of fractions whose denominator is relatively prime to p (when it is reduced to the lowest form). The maximal ideal is $p \cdot R_p$.

As mentioned above, discrete valuation rings have an extremely simple ideal-structure:

LEMMA A.4.37. Let $(R, \mathfrak{m})$ be a discrete valuation ring defined by a valuation
$$v: F^\times \to \mathbb{Z}$$
on the field of fractions, F, of R. Then there exists an element $r \in R$ such that $\mathfrak{m} = (r)$ and all ideals of R are of the form (r^n) for $n \in \mathbb{Z}^+$.

PROOF. Suppose $u \in R$ has $v(u) = 0$. Then $v(u^{-1}) = 0$ also, so u is a unit. If $\mathfrak{I} \subset R$ is an ideal, let $x \in \mathfrak{I}$ be the element with the smallest valuation. If $y \in \mathfrak{I}$, then $x^{-1}y \in F$ has a valuation $v(y) - v(x) \geq 0$ so $x^{-1}y \in R$ and $y = x \cdot x^{-1}y$ and $\mathfrak{I} = (x)$, and all ideals are principal. It follows that $\mathfrak{m} = (r)$ and $v(r) = 1$ (since the valuation-map is surjective).

Suppose $y \in R$ has the property that $v(y) = n$. Then $r^{-n}y$ has valuation 0 so it is a unit and $(y) = (r^n)$. $\square$

It is interesting to determine the properties a general ring must have to be a discrete valuation ring:

LEMMA A.4.38. Let R be a noetherian local domain with maximal ideal $\mathfrak{m} \in R$ and suppose that this is the only prime ideal (other than the trivial prime ideal, (0)). Then R is a discrete valuation ring if and only if it is integrally closed in its field of fractions.

PROOF. First, we show that a discrete valuation ring is integrally closed. If F is the field of fractions of R and $x/y \in F$ is integral over R, then
$$(x/y)^n + a_{n-1}(x/y)^{n-1} + \cdots + a_0 = 0$$
with the $a_i \in R$. If $v(x) < v(y)$ then
$$v((x/y)^n + a_{n-1}(x/y)^{n-1} + \cdots + (x/y)a_1) = v(-a_0) \geq 0$$

Proposition A.4.32 on page 431 implies that
$$v((x/y)^n + a_{n-1}(x/y)^{n-1} + \cdots + (x/y)a_1) = v((x/y)^n) < 0$$
which is a contradiction. It follows that $v(x) \geq v(y)$ and $x/y \in R$.

Now we work in the *other direction:* assume R satisfies the hypotheses (i.e., it is a noetherian local domain with a unique prime ideal) and show that it is a discrete valuation ring if it is integrally closed.

For every $u \in \mathfrak{m}$ and $v \in R \setminus (u)$ define[2]

(A.4.7) $$(u:v) = \{r \in R | rv \in (u)\}$$

---

[2]Compare this with the saturation of ideals defined in exercise 9 on page 71.

This is easily verified to be an ideal and nonempty (since it contains u at least). Let $(a:b)$ be the maximal such ideal (with respect to inclusion). We claim that it is a prime ideal. If $xy \in (a:b)$, then $xyb \in (a)$. Note that $(a:yb) \supseteq (a:b)$. If $x, y \notin (a:b)$, then $yb \notin (a)$ and $x \in (a:yb)$ so $(a:yb) \supsetneq (a:b)$, which contradicts the maximality of $(a:b)$.

Since $\mathfrak{m}$ is the only prime ideal of R, we have $\mathfrak{m} = (a:b)$. We claim that $\mathfrak{m} = (a/b)$ (so $b \,|\, a$). Equation A.4.7 on the previous page for $(a:b)$ implies that $(b/a) \cdot (a:b) = (b/a) \cdot \mathfrak{m} \subset R$.

If $(b/a) \cdot \mathfrak{m} \neq R$ then $(b/a) \cdot \mathfrak{m}$ must be an *ideal* of R, hence $(b/a) \cdot \mathfrak{m} \subset \mathfrak{m}$. Since R is noetherian, $\mathfrak{m}$ must be a finitely generated R-module. Since (b/a) maps a finitely-generated R-module to itself, proposition A.4.2 on page 419 implies that b/a is integral over R, hence *in R* (because R is integrally closed). This is a contradiction (by the condition above equation A.4.7 on the previous page), so we conclude that $(b/a)\mathfrak{m} = R$ and $\mathfrak{m} = R \cdot (a/b) = (a/b)$.

We claim that *all* ideals in R are principal. If not, there is a maximal non-principal ideal $\mathfrak{I}$ (because R is noetherian). We must have

$$\mathfrak{I} \subset \mathfrak{m} = (a/b)$$

Now consider

$$\mathfrak{I} \subset (b/a) \cdot \mathfrak{I} \subset (b/a) \cdot \mathfrak{m} = R$$

If $\mathfrak{I} = (b/a) \cdot \mathfrak{I}$, then by the reasoning above and proposition A.4.2 on page 419, we conclude that (b/a) is integral over R, hence in R. This is the same contradiction as before (with $\mathfrak{m}$) and we conclude that

$$\mathfrak{I} \subsetneq (b/a) \cdot \mathfrak{I}$$

which implies that the ideal $(b/a) \cdot \mathfrak{I}$ is principal, say $(b/a) \cdot \mathfrak{I} = (x)$. Then we get $\mathfrak{I} = (x \cdot a/b)$ which is a contradiction.

We conclude that all ideals are principal, and that R is a unique factorization domain by remark A.1.39 on page 357. The element $\pi = a/b$ that defines $\mathfrak{m}$ must be irreducible and a prime, so we can define a function

$$v: R \setminus \{0\} \to \mathbb{Z}$$

by setting $v(x)$ to the highest power of π that divides x. This extends to a valuation

$$\begin{aligned} v: F^\times &\to \mathbb{Z} \\ v(x/y) &= v(x) - v(y) \end{aligned}$$

and R is a discrete valuation ring. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

A.4.4. Graded rings and modules. A graded ring is a kind of ring subdivided into distinct direct summands. These appear in the study of projective varieties and sheaf-cohomology.

DEFINITION A.4.39. A ring, G, is called a *graded ring* over k if there exists a decomposition

$$G = G_0 \oplus G_1 \oplus \cdots$$

such that $G_i \cdot G_j \subset G_{i+j}$ for all $i, j \geq 0$. If $\mathfrak{I} \subset G$ is an ideal, we say that $\mathfrak{I}$ is a *graded ideal* if

$$\mathfrak{I} = \mathfrak{I}_0 \oplus \mathfrak{I}_1 \oplus \cdots$$

where $\mathfrak{I}_j = \mathfrak{I} \cap G_j$ for all $j \geq 0$. An ideal $\mathfrak{H} \subset G$ is *homogeneous* if all of its generators come from the same G_n for some n.

REMARK. Any ring, R, can be regarded as a graded ring, if we define $R_0 = R$, $R_i = 0$ for $i > 0$. A polynomial ring, $k[X_1, \ldots, X_t]$ is naturally graded with gradation

given by the total degree of a monomial (where we have specified the degree of *each* of the X_i):

$$k[X_0, \ldots, X_t] = k \oplus K_1 \oplus K_2 \oplus \cdots$$

where K_n is the vector space generated by all monomials of total degree n. For instance, let $G = k[X, Y]$ where X and Y are of degree 1 and let $H = k[X, Y]$ where X is of degree 1 and Y is of degree 2. Then:

$$
\begin{aligned}
G_0 &= k \\
G_1 &= k \cdot \{X, Y\} \\
G_2 &= k \cdot \{X^2, XY, Y^2\} \\
&\vdots
\end{aligned}
$$

and

$$
\begin{aligned}
H_0 &= k \\
H_1 &= k \cdot X \\
H_2 &= k \cdot Y \\
&\vdots
\end{aligned}
$$

so G and H are isomorphic as rings but not as *graded* rings.

Given a ring and ideal, one can define the associated graded ring

DEFINITION A.4.40. If R is a ring with ideal $\mathfrak{a} \subset R$, the *associated graded ring* is defined by

$$\mathbf{gr}(R, \mathfrak{a}) = \frac{R}{\mathfrak{a}} \oplus \frac{\mathfrak{a}}{\mathfrak{a}^2} \oplus \cdots \oplus \frac{\mathfrak{a}^n}{\mathfrak{a}^{n+1}} \oplus \cdots$$

with multiplication

$$\frac{\mathfrak{a}^n}{\mathfrak{a}^{n+1}} \otimes \frac{\mathfrak{a}^m}{\mathfrak{a}^{m+1}} \to \frac{\mathfrak{a}^{n+m}}{\mathfrak{a}^{n+m+1}}$$

REMARK. Note that each summand

$$\frac{\mathfrak{a}^n}{\mathfrak{a}^{n+1}}$$

is naturally a module over

$$\frac{R}{\mathfrak{a}}$$

Also note that the generators of $\mathfrak{a}$ generate $\mathbf{gr}(R, \mathfrak{a})$ as a ring., so the latter is finitely generated if $\mathfrak{a}$ is.

Given graded algebras over a ring R, we can define the graded tensor product

DEFINITION A.4.41. If A, B are graded algebras over a (non-graded) ring, R, the tensor product is the graded algebra $A \otimes_R B$ defined by

$$(A \otimes_R B)_n = \bigoplus_{i+j=n} A_i \otimes_R B_j$$

REMARK. This definition is consistent with the convention

$$(a \otimes b) \cdot (c \otimes d) = (ac \otimes bd)$$

It is not hard to see that:

PROPOSITION A.4.42. *If $\mathfrak{H} \subset G$ is a homogeneous ideal of a graded ring G, it is not hard to see that*

$$(A.4.8) \qquad \frac{G}{\mathfrak{H}} = \bigoplus_{i=0}^{\infty} \frac{G_i + \mathfrak{H}}{\mathfrak{H}}$$

is also a graded ring.

Here is a standard construction of a graded ring:

DEFINITION A.4.43. If $\mathfrak{a} \subset R$ is an ideal in a ring, we can define a graded ring

$$\Gamma(\mathfrak{a}) = R \oplus \mathfrak{a} \oplus \mathfrak{a}^2 \oplus \cdots = R[T \cdot \mathfrak{a}] \subset R[T]$$

by giving R a degree of 0, and $\mathfrak{a}$ a degree of 1 — or, equivalently, giving T a degree of 1. This is called the *Rees algebra* of $\mathfrak{a}$.

We begin with

DEFINITION A.4.44. If G is a graded ring (see definition A.4.39 on page 434), a module, M, over G is a *graded module* if

$$M = M_0 \oplus M_1 \oplus \cdots$$

with the property $G_i \cdot M_j \subset M_{i+j}$ for all $i, j \geq 0$.

If ℓ is an integer and M is a graded module over a graded ring, we define the *ℓ-twist* of M, denoted $M(\ell)$ is defined by

$$M(\ell)_i = M_{i+\ell}$$

REMARK. Although any k-algebra could be regarded as a graded algebra (put it all in G_0), some have a natural grading. For instance,

$$G = k[X_0, \ldots, X_n]$$

is naturally a graded ring by degrees of monomials, i.e., G_k consists of homogeneous polynomials of degree k. This grading is geometrically significant.

It is not hard to see

PROPOSITION A.4.45. *If $\mathfrak{I} \subset G$ is a graded ideal in a graded algebra, the quotient*

$$R = \frac{G}{\mathfrak{I}}$$

is naturally a graded algebra with

$$(A.4.9) \qquad R_j = \frac{G_j}{\mathfrak{I}_j}$$

for all j.

REMARK. Graded ideals are just graded submodules of G, regarding it as a graded module over itself. In general, all of the results of section A.1.6 on page 367 have versions for graded modules over graded algebras, given the following

LEMMA A.4.46. *Let $R = R_0 \oplus R_1 \oplus \cdots$ be a graded ring and let M be a graded module over R. If $m \in M$ and $\mathfrak{p} = \mathrm{ann}(m) \subset R$ is prime, then $\mathfrak{p}$ is homogeneous and $\mathfrak{p}$ is the annihilator of a homogeneous element of M.*

PROOF. If $r \in \mathfrak{p}$, we have a unique expression $r = \sum_{i=1}^{s} r_i$ where r_i is homogeneous of degree d_i, with $d_1 < d_2 < \cdots < d_s$. We will prove that $\mathfrak{p}$ is homogeneous by showing that $r \in \mathfrak{p}$ implies that all of the $r_i \in \mathfrak{p}$. By induction on s, it suffices to show that $r_1 \in \mathfrak{p}$.

Similarly, we have a unique expression $m = \sum_{j=1}^{t} m_j$ with m_j homogeneous of degree e_i with $e_1 < \cdots < e_t$. We claim that $r_1 \cdot m_1 = 0$, since this is the term in $r \cdot m = 0$ of

lowest degree. This proves the result in the case where $t = 1$. Now suppose it has been proved for all smaller values of t. The element

$$r_1 \cdot m = \sum_{j=2}^{t} r_1 \cdot m_j$$

is a sum of $< t$ homogeneous components. Let $\mathfrak{q} = \mathrm{ann}(r_1 \cdot m)$. By induction, we conclude that $\mathfrak{q}$ is homogeneous if it is prime, and $\mathfrak{p} \subseteq \mathfrak{q}$. If $\mathfrak{p} = \mathfrak{q}$, we are done. Otherwise, let $g \in \mathfrak{q} \setminus \mathfrak{p}$. Then $g \cdot r_1 \cdot m = 0$ so $gr_1 \in \mathrm{ann}(m) = \mathfrak{p}$. Since $\mathfrak{p}$ is prime and $g \notin \mathfrak{p}$, we conclude that $r_1 \in \mathfrak{p}$, and $\mathfrak{p}$ is homogeneous.

Now, since $\mathfrak{p}$ is homogeneous, $\mathfrak{p} \cdot m_j = 0$ for all j, so

$$\mathfrak{p} = \mathrm{ann}(m) \supset \bigcap_{j=1}^{t} \mathrm{ann}(m_j) \supset \mathfrak{p}$$

which implies that $\mathfrak{p} = \bigcap_{j=1}^{t} \mathrm{ann}(m_i) \supset \prod_{j=1}^{t} \mathrm{ann}(m_j)$. The fact that $\mathfrak{p}$ is prime implies that $\mathfrak{p} \supset \mathrm{ann}(m_j)$ (see exercise 11 on page 354) for some j, which means that $\mathfrak{p} = \mathrm{ann}(m_j)$. $\qquad\square$

With this in hand, we can easily generalize prime filtrations of modules to *graded* modules:

LEMMA A.4.47. *Let M be a graded module over a graded ring, R. Then there exists a finite ascending chain of graded-submodules*

$$0 = M_0 \subset M_1 \subset \cdots \subset M_t = M$$

with the property that for each i

$$\frac{M_{i+1}}{M_i} \cong \frac{R}{\mathfrak{p}_i}(\ell_i)$$

where $\mathfrak{p}_i \subset R$ is a homogeneous prime ideal and ℓ_i is an integer.

PROOF. The proof is exactly the same as that of theorem A.1.77 on page 375, except that we use lemma A.4.46 on the facing page to guarantee that the ideals $\{\mathfrak{p}_i\}$ are all homogeneous so that the quotients $R/\mathfrak{p}_i$ are now graded rings. The ℓ_i occur because the natural grading of $R/\mathfrak{p}_i$ may be shifted in forming iterated quotients. $\qquad\square$

We can also conclude something about other filtrations of modules:

DEFINITION A.4.48. If M is a module over a ring R and $\mathfrak{a} \subset R$ is an ideal, a filtration

$$\cdots \subset M_t \subset \cdots \subset M_0 = M$$

if called an $\mathfrak{a}$-*filtration* if $\mathfrak{a} \cdot M_n \subseteq M_{n+1}$ for all $n \geq 0$. It is called a *stable $\mathfrak{a}$-filtration* if $\mathfrak{a} \cdot M_n = M_{n+1}$ for all $n > n_0$.

REMARK. Note that these conditions only apply from some finite point on.

We can define a kind of module-analogue of the Rees algebra:

DEFINITION A.4.49. If M is a module over a ring, R, with a filtration

$$\cdots \subset M_t \subset \cdots \subset M_0 = M$$

define the Rees module of this to be

$$\Gamma(M) = M_0 \oplus M_1 \oplus \cdots$$

REMARK. If the filtration is an $\mathfrak{a}$-filtration for some ideal $\mathfrak{a} \subset R$ then $\Gamma(M)$ is naturally a graded-module over $\Gamma(\mathfrak{a})$, since $\mathfrak{a}^t \cdot M_n \subseteq M_{n+t}$.

One of our main results is:

LEMMA A.4.50. *If M is a finitely generated module over a ring R with an $\mathfrak{a}$-filtration*

$$\cdots \subset M_t \subset \cdots \subset M_0 = M$$

by finitely-generated submodules, for some ideal $\mathfrak{a} \subset R$, then $\Gamma(M)$ is finitely generated over $\Gamma(\mathfrak{a})$ if and only if this filtration is stable.

PROOF. If $\Gamma(M)$ is finitely generated over $\Gamma(\mathfrak{a})$, then for some n

$$M_0 \oplus \cdots \oplus M_n$$

generates all of $\Gamma(M)$. Since the filtration is an $\mathfrak{a}$-filtration, we have $\mathfrak{a}^i \cdot M_{n-i} \subseteq M_n$, so we can really say that M_n generates

$$M_n \oplus M_{n+1} \oplus \cdots$$

and considerations of grading imply that $M_{n+i} = \mathfrak{a}^i \cdot M_n$ so the filtration is stable from degree n on.

Conversely, if the filtration is stable from degree n on, then $M_{n+i} = \mathfrak{a}^i \cdot M_n$ so that $\Gamma(M)$ generated by $M_0 \oplus \cdots \oplus M_n$ over $\Gamma(\mathfrak{a})$.   □

The main result of this section is

THEOREM A.4.51 (Artin-Rees Theorem). *Suppose R is a noetherian ring with ideal $\mathfrak{a} \subset R$. If $N \subset M$ are finitely generated R-modules and M has a stable $\mathfrak{a}$-filtration*

$$\cdots \subset M_t \subset \cdots \subset M_0 = M$$

then the filtration

$$\cdots \subset M_t \cap N \subset \cdots \subset M_0 \cap N = N$$

is also a stable $\mathfrak{a}$-filtration. In other words there exists an integer n such that

$$\left(\mathfrak{a}^i \cdot M_n \right) \cap N = \mathfrak{a}^i \cdot (M_n \cap N)$$

PROOF. Since R is noetherian, $\mathfrak{a}$ is finitely generated and $\Gamma(\mathfrak{a})$ is a finitely-generated R-algebra, hence noetherian. Since the filtration on M is stable, $\Gamma(M)$ is finitely-generated over $\Gamma(\mathfrak{a})$. It is not hard to see that, computed with respect to the induced filtration, $\Gamma(N) \subset \Gamma(M)$, which means that it is *also* finitely generated (see lemma A.1.70 on page 372). The conclusion follows from lemma A.4.50).   □

A.5. A glimpse of category theory

A.5.1. Introduction. Category theory is a field as general as set theory that can be applied to many areas of mathematics. It is concerned with the patterns of mappings between mathematical structures and the types of conclusions one can draw from them.

Eilenberg and MacLane developed it with applications to algebraic topology in mind, see [**37**]. Today, it has applications to many other fields, including computer science — [**137**]

Once derided as "general nonsense," it has gained acceptance over time. Readers who want more than the "drive-by" offered here are invited to look at MacLane's classic, [**106**].

Here is an example of the kind of reasoning that category theory uses:

Suppose you want to define the product of two mathematical objects, *A* and *B*. One way to proceed is to say that $A \times B$ has the following *universal property:*

  (1) There exist maps from $A \times B$ to *A* and *B* (projections to the factors).

(2) Given *any* maps $f: Z \to A$ and $g: Z \to B$, there is a *unique* map

$$f \times g: Z \to A \times B$$

compatible with the maps from Z to A, B.

This is more succinctly stated with *commutative diagrams.* In a diagram like

(A.5.1)
$$
\begin{array}{ccc}
U & \xrightarrow{\ r\ } & V \\
{\scriptstyle t}\downarrow & & \downarrow{\scriptstyle s} \\
W & \xrightarrow{\ b\ } & X
\end{array}
$$

the arrows represent maps. We will say this diagram commutes if, whenever one can reach a node along different paths, the composite maps one encounters are equal. For instance, the statement that diagram A.5.1 commutes is equivalent to saying $s \circ r = b \circ t$.

DEFINITION A.5.1. We can define $A \times B$ by saying that,

(1) it has projection-maps $p_1: A \times B \to A$ and $p_2: A \times B \to B$
(2) whenever we have a diagram with solid arrows

(A.5.2)

where Z is an arbitrary "object" that maps to A and B — the dotted of arrow *exists*, is *unique*, and makes the whole diagram commute.

In other words, we define $A \times B$ by a general structural property that does not use the inner workings or A or B.

DEFINITION A.5.2. A *category*, $\mathscr{C}$, is a collection of *objects* and *morphisms*, which are maps between objects. These must satisfy the conditions:

(1) Given objects $x, y \in \mathscr{C}$, $\hom_{\mathscr{C}}(x, y)$ denotes the morphisms from x to y. This may be an empty set.
(2) Given objects $x, y, z \in \mathscr{C}$ and morphisms $f: x \to y$ and $g: y \to z$, the composition $g \circ f: x \to z$ is defined. In other words a dotted arrow exists in the diagram

making it commute.

(3) Given objects $x, y, z, w \in \mathscr{C}$ and morphisms $f\colon x \to y, g\colon y \to z$, $h\colon z \to w$, composition is associative, i.e., $h \circ (g \circ f) = (h \circ g) \circ f\colon x \to w$. This can be represented by a commutative diagram:
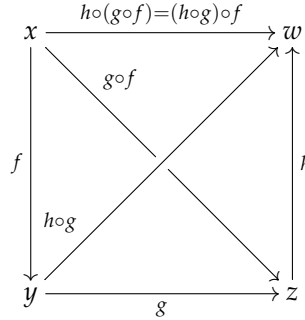
(4) Every object, $x \in \mathscr{C}$, has an identity map $1_x\colon x \to x$ such that, for any $f\colon x \to y$, $f \circ 1_x = 1_y \circ f = f\colon x \to y$. This is equivalent to saying that the diagram
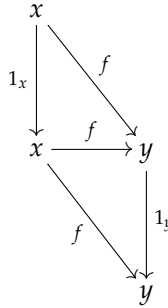
commutes.

After defining something so general, it is necessary to give many examples:

(1) The category, $\mathscr{V}$, of vector-spaces and linear transformations (when defining a category, one must specify the morphisms as well as the objects). Given two vector spaces, V and W, the set of morphisms, $\hom_{\mathscr{V}}(V, W)$, is *also* a vector-space.

(2) The category, $\mathscr{D}$, of vector-spaces where the only morphisms are *identity maps* from vector spaces to themselves. Categories in which the only morphisms are identity maps are called *discrete*. Discrete categories are essentially sets of objects.

(3) The category, $\mathscr{R}$, of rings where the morphisms are ring-homomorphisms.

(4) The category, $\mathscr{N}$, whose objects are positive integers and where the morphisms

$$m \to n$$

are all possible $n \times m$ matrices of real numbers. Composition of morphisms is just matrix-multiplication. This is an example of a category in which the morphisms aren't maps.

(5) The category, $\mathscr{S}$, of sets with the morphisms functions mapping one set to another.

(6) The category, $\mathscr{T}$, of topological spaces where the morphisms are continuous maps.
(7) The category, $\mathscr{M}_R$ of modules over a commutative ring, R. If $M, N \in \mathscr{M}_R$, the set of morphisms, $\hom_{\mathscr{M}_R}(M, N)$ is usually written $\hom_R(M, N)$.

DEFINITION A.5.3. A category, $\mathscr{C}$, is called *concrete* if

(1) its objects are sets (possibly with additional structure)
(2) morphisms that are equal as set-mappings are equal in $\mathscr{C}$.

REMARK. All of the examples given above except $\mathscr{N}$ are concrete.

We will also need the dual concept of *coproduct*. A coproduct of a set of objects is essentially their union. So why not just call it the union? Well the categorical definition below is essentially the same as that of the product, except that all of the *arrows* in definition A.5.1 on page 439 *are reversed* (hence, hardcore category-theorists insist that it is the *co*product):

DEFINITION A.5.4. An object Z in a category is a *coproduct* of A and B if

(1) there exist maps $i_1 : A \to Z$, $i_2 : B \to Z$, and
(2) *Any* maps $f_1 : A \to W$ $f_2 : B \to W$ induce a *unique* map $g : Z \to W$ making the diagram

commute. If this is true, we write

$$Z = A \coprod B$$

REMARK. Note that the symbol for a coproduct is an inverted product-symbol — which looks vaguely like a union-symbol. This is appropriate since coproducts have the structural properties of a union. As before, the universal property of coproducts imply that if they exist, they are unique.

Products *map to* their factors, and coproducts have *maps from* their factors. In some cases, products and coproducts are the same.

EXAMPLE A.5.5. Products and coproducts depend strongly on the category (and some categories do not even have these constructions):

(1) In the category of sets, the *union* is a coproduct. The Cartesian product is the product, so coproducts and products are very different.
(2) In the category of modules over a ring, the direct sum is the coproduct as well as the product.
(3) In the category of groups, the *free product* is the coproduct.

Category theory expresses the familiar concepts of monomorphism and epimorphism in "arrow-theoretic" terms:

DEFINITION A.5.6. A morphism $f : A \to B$ between objects of a category is:

(1) a *monomorphism* if, for any other object C and any two morphisms $g_1, g_2 \colon C \to A$

$$f \circ g_1 = f \circ g_2 \implies g_1 = g_2$$

(2) an *epimorphism* if, for any other object C and any two morphisms $g_1, g_2 \colon B \to C$

$$g_1 \circ f = g_1 \circ f \implies g_1 = g_2$$

EXERCISES.

1. If A and B are objects of a category $\mathscr{C}$, show that, for any object $W \in \mathscr{C}$

$$\hom_{\mathscr{C}}(W, A \times B) = \hom_{\mathscr{C}}(W, A) \times \hom_{\mathscr{C}}(W, B)$$

2. Prove the statement above that in the category of modules over a ring, the product and coproduct of two modules V and W is $V \oplus W$. Is the same thing true for infinite products and coproducts?

3. In a category $\mathscr{C}$, if

$$f \colon A \to B$$

is a monomorphism, show that

$$\hom_{\mathscr{C}}(C, A) \xrightarrow{\hom_{\mathscr{C}}(1, f)} \hom_{\mathscr{C}}(C, B)$$

is a monomorphism in the category of sets.

4. If $\mathscr{A}b$ is the category of abelian groups, show that a map is a monomorphism if and only if it is *injective* (in the usual sense) and is an epimorphism if and only if it is *surjective.*

A.5.2. Functors. A *functor* from one category to another is a kind of function of objects *and* morphisms.

DEFINITION A.5.7. Let $\mathscr{C}$ and $\mathscr{D}$ be categories. A *functor*

$$f \colon \mathscr{C} \to \mathscr{D}$$

is a function from the objects of $\mathscr{C}$ to those of $\mathscr{D}$ — i.e., if $x \in \mathscr{C}$ then $f(x) \in \mathscr{D}$ with the following additional property:

If $h \colon x \to y$ is a morphism in $\mathscr{C}$, then f defines, either

- a morphism $f(h) \colon f(x) \to f(y)$ in $\mathscr{D}$ — in which case f is called a *covariant functor* or just a *functor, or*
- a morphism $f(h) \colon f(y) \to f(x)$ in $\mathscr{D}$ — in which case f is called a *contravariant functor*.

In addition $f(1_x) = 1_{f(x)}$ and $f(j \circ h) = f(j) \circ f(h)$, if f is covariant or $f(j \circ h) = f(h) \circ f(j)$ is f is contravariant.

REMARK. Functors play an extremely important part in algebraic geometry, particularly contravariant ones.

Here are some examples:

(1) a functor $f: \mathscr{S} \to \mathscr{S}$ from the category of sets to itself. If $x \in \mathscr{S}$, $f(x) = 2^x$, the power-set or set of all subsets of x. If $d: x \to y$ is a set-mapping and $z \subset x$ is a subset, then $d|z: z \to y$ is a set-mapping whose images is a subset of y. It follows that d induces a natural map

$$2^d: 2^x \to 2^y$$

and this is what we define $f(d)$ to be.

(2) We can define $f: \mathscr{V} \to \mathscr{V}$ to send a real vector-space, $x \in \mathscr{V}$ to its *dual,* x^* — the vector-space of all linear transformations $\eta: x \to \mathbb{R}$.

If $m: x \to y$ is a linear transformation, and $\mu: y \to \mathbb{R}$ is an element of y^*, the composite $\eta \circ m: x \to \mathbb{R}$ is an element of x^*. We get a natural map $m^*: y^* \to x^*$ and we set $f(m) = m^*$. It follows that f is a *contravariant* functor.

(3) If $\mathscr{F}$ is the category of finite dimensional vector spaces, it is well-known that $x^{**} = x \in \mathscr{F}$, so the functor f defined in statement 2 above actually is a contravariant *isomorphism of categories*

$$f: \mathscr{F} \to \mathscr{F}$$

(4) If $\mathscr{G}$ is the category of groups and $\mathscr{R}$ is that of commutative rings, we can define a functor

$$g_n: \mathscr{R} \to \mathscr{G}$$

that sends a commutative ring $r \in \mathscr{R}$ to $GL_n(r)$, the group of $n \times n$ matrices whose determinant is a unit of r. Since homomorphisms of rings send units to units, it follows that any homomorphism of rings

$$h: r \to s$$

induces a natural homomorphism of groups $g_n(h): GL_n(r) \to GL_n(s)$.

We can classify functors in various ways:

DEFINITION A.5.8. A functor $f: \mathscr{C} \to \mathscr{D}$ is:

(1) an *isomorphism* if it is a bijection of objects and morphisms,
(2) *full* if it is "surjective on morphisms" — i.e., *every* morphism $g: f(c_1) \to f(c_2) \in \mathscr{D}$ is of the form $f(t)$ where t is a morphism $t: c_1 \to c_2$ (if f is covariant). In the contravariant case, reverse the arrows in $\mathscr{C}$ or $\mathscr{D}$ (but not both).
(3) *faithful* if it is "injective on morphisms" — i.e., given morphisms $m_1, m_2: c_1 \to c_2$ $f(m_1) = f(m_2)$ always implies that $m_1 = m_2$.

For instance, concrete categories are commonly defined as categories that have a faithful functor to the category of sets.

Isomorphism of categories is too stringent a condition in practice. Equivalence of categories is slightly weaker but very useful. To define it, we need:

DEFINITION A.5.9. Suppose $\mathscr{C}$ is a category and $f: \mathscr{C} \to \mathscr{C}$ is a functor such that $f(x)$ is isomorphic to x for all $x \in \mathscr{C}$. A *natural isomorphism*

$$j_x: x \to f(x)$$

is an isomorphism defined for all objects $x \in \mathscr{C}$ with the property that, for any morphism $g: x \to y$ the diagram

$$
\begin{array}{ccc}
x & \xrightarrow{\ j_x\ } & f(x) \\
{\scriptstyle g}\Big\downarrow & & \Big\downarrow{\scriptstyle f(g)} \\
y & \xrightarrow[\ j_y\ ]{} & f(y)
\end{array}
$$

commutes.

REMARK. The thing that makes an isomorphism *natural* is that it is defined for all objects and in a way compatible with all maps between them. Prior to the introduction of category theory, it was common to call certain maps natural without giving any precise definition.

This is a special case of a *natural transformation* of functors:

DEFINITION A.5.10. If $\mathscr{C}$, $\mathscr{D}$ are categories and $f, g: \mathscr{C} \to \mathscr{D}$ are functors, a *natural transformation*

$$t: f \to g$$

is a morphism

$$t(x): f(x) \to g(x)$$

defined for all $x \in \mathscr{C}$, such that, for any morphism $m: x \to y$ the diagram

$$
\begin{array}{ccc}
f(x) & \xrightarrow{\ t(x)\ } & g(x) \\
{\scriptstyle f(m)}\Big\downarrow & & \Big\downarrow{\scriptstyle g(m)} \\
f(y) & \xrightarrow[\ t(y)\ ]{} & g(y)
\end{array}
$$

commutes.

REMARK. In the notation of definition A.5.9 on the preceding page, a natural *isomorphism* is a natural transformation from the identity functor to f.

It is possible to form a category out of all of the functors between two categories. Natural transformations are the *morphisms* in this "category of functors."

Here's an example of a natural isomorphism:

EXAMPLE A.5.11. If V is a vector-space, there is a morphism of vector-spaces

$$V \to V^{**}$$

that sends $v \in V$ to the linear function, $t \in V^{**}$, on V^* with $t(r) = r(v)$ for $r \in V^*$. It clearly commutes with all maps of vector-spaces. This is well-known to be an isomorphism if V is finite-dimensional.

And here is one of a natural transformation:

EXAMPLE A.5.12. If V is a vector-space, define

$$\begin{aligned} f(V) &= V \oplus V \\ g(V) &= V \end{aligned}$$

Now, for every vector-space, V, define

$$t(V): f(V) \to g(V)$$

to be the homomorphism that sends $(v_1, v_2) \in V \oplus V$ to $v_1 + v_2 \in V$. This is easily verified to be a natural transformation.

In considering when two categories are "equivalent," it turns out that requiring them to be isomorphic is usually too restrictive. Instead, we require them to be equivalent in the following sense:

DEFINITION A.5.13. Given categories $\mathscr{C}$ and $\mathscr{D}$, a pair of functors

$$\begin{aligned} f: \mathscr{C} &\to \mathscr{D} \\ g: \mathscr{D} &\to \mathscr{C} \end{aligned}$$

define an *equivalence of categories* if there exist natural isomorphisms

$$j_x: x \to g \circ f(x)$$

for all $x \in \mathscr{C}$ and

$$k_y: y \to f \circ g(x)$$

for all $y \in \mathscr{D}$.

EXERCISES.

5. Let f be the functor defined in statement 2 on page 443 above and suppose we have a morphism

$$m: V_1 \to V_2$$

between vector spaces that is represented by a matrix, A. Describe the matrix-representation of

$$f(m): V_2^* \to V_1^*$$

6. If $\mathscr{F}$ is the category of finite-dimensional vector-spaces, show that the functor f defined in statement 2 on page 443 above is an equivalence of categories

$$f: \mathscr{F} \to \mathscr{F}$$

Why isn't it an isomorphism?

A.5.3. Adjoint functors. *Adjoint functors* are ones that *complement* each other in a certain sense. They occur naturally in many settings — Daniel Kan was the first to recognize these patterns (see [**84**]) and develop a general concept.

As often happens in category theory, the definition is very cryptic without several examples:

DEFINITION A.5.14. Given two categories, $\mathscr{A}$ and $\mathscr{B}$, functors

$$
\begin{aligned}
f : \mathscr{A} &\to \mathscr{B} \\
g : \mathscr{B} &\to \mathscr{A}
\end{aligned}
$$

are said to be *adjoint* if there exists a natural isomorphism

(A.5.3) $$ \hom_{\mathscr{A}}(x, g(y)) = \hom_{\mathscr{B}}(f(x), y) $$

for all $x \in \mathscr{A}$ and $y \in \mathscr{B}$. In this situation, f is called a *left-adjoint* to g and g is called a *right-adjoint* to f. The collection, $(f, g, \mathscr{A}, \mathscr{B})$ is called an *adjunction*.

REMARK. Note that (with rare exceptions) f and g are not *inverses* of each other.

Our terminology was taken from Hilbert space theory: U_1 and U_2 are adjoint *operators* if

$$ \langle U_1 x, y \rangle = \langle x, U_2 y \rangle $$

in the Hilbert space, where $\langle *, * \rangle$ is the inner product[3]. Kan was inspired by this equation's similarity (in appearance, not function!) to equation A.5.3 to name his constructs "adjoints". Hilbert-space adjoints are not adjoints in *our* sense except in certain odd settings (see [**11**]).

Here is an example of a common pattern — where one of the functors forgets extra structure an object has and regards it as something more primitive (these are called *forgetful functors*):

EXAMPLE A.5.15. Let $\mathscr{V}_k$ be the category of vector-spaces over a field, k, and let $\mathscr{S}$ be the category of sets. The functor

$$ g : \mathscr{V}_k \to \mathscr{S} $$

simply maps a vector space onto the set of its nonzero elements — it *forgets* the extra structure a vector-space has. The functor

$$ f : \mathscr{S} \to \mathscr{V}_k $$

maps a set $x \in \mathscr{S}$ to

$$ f(x) = \bigoplus_{y \in x} k \cdot x $$

— the vector-space with basis x. Any set-mapping $t : x \to f(V)$ extends *uniquely* to a vector-space homomorphism

$$ f(t) : f(x) \to V $$

---

[3] In *finite dimensions*, U_2 is the conjugate-transpose of U_2.

since a homomorphism of vector-spaces is determined by its effect on basis-elements. On the other hand, any homomorphism of vector-spaces *is* a *unique* map of their nonzero elements (regarded as *sets*) so we get a natural equality

$$\hom_{\mathscr{S}}(x, g(y)) = \hom_{\mathscr{V}_k}(f(x), y)$$

for all $y \in \mathscr{V}_k$ and $x \in \mathscr{S}$.

Here's another example of adjoint functors where forgetful functors are *not* involved:

EXAMPLE A.5.16. Suppose $\mathscr{C}$ is some category and assume that the categorical product (defined in A.5.1 on page 439) in $\mathscr{C}$ exists. Strictly speaking, it is a functor

$$\prod : \mathscr{C} \times \mathscr{C} \to \mathscr{C}$$

where $\mathscr{C} \times \mathscr{C}$ is the category of pairs (x, y) for $x, y \in \mathscr{C}$ and morphisms are defined in a similar way. Now consider the diagonal functor

(A.5.4) $$\Delta : \mathscr{C} \to \mathscr{C} \times \mathscr{C}$$

that sends every $x \in \mathscr{C}$ to $(x, x) \in \mathscr{C} \times \mathscr{C}$. The definition of product, and diagram A.5.2 on page 439 implies that every pair of morphisms

$$\begin{array}{ccc} x & \to & y \\ x & \to & z \end{array}$$

induces a unique morphism $x \to y \prod z$. Such pairs of morphisms are really morphisms

$$\Delta x \to (y, z) \in \mathscr{C} \times \mathscr{C}$$

so we get an equivalence

$$\hom_{\mathscr{C} \times \mathscr{C}}(\Delta x, (y, z)) = \hom_{\mathscr{C}}(x, y \prod z)$$

which implies that $\prod$ is a right-adjoint to Δ. In this case, the adjunction involves a functor of two variables.

EXERCISES.

7. Example A.5.16 shows that the diagonal functor

$$\Delta : \mathscr{C} \to \mathscr{C} \times \mathscr{C}$$

in equation A.5.4 is a left-adjoint to the product functor

$$\prod : \mathscr{C} \times \mathscr{C} \to \mathscr{C}$$

Show that it is a right-adjoint to the coproduct functor (definition A.5.4 on page 441), showing that a functor can be a left-adjoint to one functor and a right-adjoint to another.

A.5.4. Limits. Limits in category theory are universal constructions somewhat like the union construction in the introduction. We will look at something similar but more complex:

DEFINITION A.5.17. Suppose $\mathscr{C}$ is a category and I is a partially ordered set of indices. Suppose $\{X_\alpha\}$, for $\alpha \in I$, is a sequence of objects of $\mathscr{C}$. Whenever $\alpha \leq \beta$ suppose there is a morphism

$$f_{\alpha,\beta}\colon X_\alpha \to X_\beta$$

and whenever $\alpha \leq \beta \leq \gamma$ the diagram

$$
\begin{array}{ccc}
X_\alpha & \xrightarrow{\;f_{\alpha,\beta}\;} & X_\beta \\
 & {\scriptstyle f_{\alpha,\gamma}} \searrow & \downarrow {\scriptstyle f_{\beta,\gamma}} \\
 & & X_\gamma
\end{array}
$$

commutes. Then the *direct limit*, $\varinjlim X_\alpha$ has

(1) morphisms $\phi_\alpha\colon X_\alpha \to \varinjlim X_\alpha$ that make the diagrams

(A.5.5)
$$
\begin{array}{ccc}
X_\alpha & \xrightarrow{\;f_{\alpha,\beta}\;} & X_\beta \\
 & {\scriptstyle \phi_\alpha} \searrow & \downarrow {\scriptstyle \phi_\beta} \\
 & & \varinjlim X_\alpha
\end{array}
$$

commute for all $\alpha, \beta \in I$ with $\alpha \leq \beta$.

(2) the *universal property* that whenever there is an object $Z \in \mathscr{C}$ and morphisms $h_\alpha\colon X_\alpha \to Z$ for all $\alpha \in I$ that make the diagrams

$$
\begin{array}{ccc}
X_\alpha & \xrightarrow{\;f_{\alpha,\beta}\;} & X_\beta \\
 & {\scriptstyle h_\alpha} \searrow & \downarrow {\scriptstyle h_\beta} \\
 & & Z
\end{array}
$$

commute for all $\alpha \leq \beta$, then there exists a *unique morphism* $u\colon \varinjlim X_\alpha \to Z$ that makes the diagrams

$$
\begin{array}{ccc}
X_\alpha & \xrightarrow{\;\phi_\alpha\;} & \varinjlim X_\alpha \\
 & {\scriptstyle h_\alpha} \searrow & \downarrow {\scriptstyle u} \\
 & & Z
\end{array}
$$

commute for all $\alpha \in I$.

REMARK. To roughly summarize: whenever the X's map to some object, Z, in a way compatible with the $f_{\alpha,\beta}$'s, the direct limit *also* maps to Z.

Some authors require I to be a *directed set*, i.e., for any $\alpha, \beta \in I$ there exists some $\gamma \in I$ with $\alpha \leq \gamma$ and $\beta \leq \gamma$.

Suppose we have two objects K_1 and K_2 that satisfy all of the conditions listed above. Then statement 2 on the preceding page above implies the existence of *unique* maps

$$K_1 \xrightarrow{f} K_2$$
$$K_2 \xrightarrow{g} K_1$$

The composites

$$K_1 \xrightarrow{g \circ f} K_1$$
$$K_2 \xrightarrow{f \circ g} K_2$$

are also unique maps satisfying all of the conditions in definition A.5.17 on the facing page. But the respective *identity map*s satisfy these conditions, so we must have

$$g \circ f = 1_{K_1}$$
$$f \circ g = 1_{K_2}$$

Note that the word "unique" is crucial to this discussion. Also note that we have not promised that direct limits exist — only that *if* they exist, they are unique (up to isomorphism). Whether they exist depends on the category.

Hardcore category-theorists prefer the term "filtered colimit" for direct limit. It is also sometimes called the inductive limit. The term "direct limit" seems to be favored by algebraists.

In the case where $\mathscr{C}$ is a concrete category (see definition A.5.3 on page 441) we can explicitly construct the direct limit.

PROPOSITION A.5.18. *Let $\mathscr{C}$ be a concrete category (see definition A.5.3 on page 441) and assume the notation of definition A.5.17 on the preceding page. Then*

$$(A.5.6) \qquad \varinjlim X_\alpha = \coprod_{\alpha \in I} X_\alpha / \sim$$

the coproduct (see definition A.5.4 on page 441) or union modulo an equivalence relation, $\sim$, defined by

$$x_\alpha \sim x_\beta$$

for $x_\alpha \in X_\alpha$, $x_\beta \in X_\beta$ if and only if there exists a $\gamma \in I$ with $\alpha \leq \gamma$ and $\beta \leq \gamma$ and

$$f_{\alpha,\gamma}(x_\alpha) = f_{\beta,\gamma}(x_\beta)$$

The maps $\phi_\alpha \colon X_\alpha \to \varinjlim X_\alpha$ are the composites

$$(A.5.7) \qquad X_\alpha \to \coprod_{\alpha \in I} X_\alpha \to \coprod_{\alpha \in I} X_\alpha / \sim$$

REMARK. So the maps $f_{\alpha,\beta}$ "glue together" the pieces, X_α, in the union. Elements of the X_α are equivalent if they eventually get glued together.

Concrete categories include the category of rings, vector spaces, and sets. Coproducts of *sets* are just their union. Coproducts of *vector spaces* are their direct sum (see exercise 2 on page 442). Coproducts of *rings* are more complicated (see proposition 2.7.3 on page 95) and so is the corresponding definition of direct limit.

PROOF. Checking the commutativity of diagrams A.5.5 on page 448 is straightforward.

If we have morphisms $h_\alpha\colon X_\alpha \to Z$ for all $\alpha \in I$, the disjoint union also maps to Z:

$$\coprod h_\alpha\colon \coprod_\alpha X_\alpha \to Z$$

and in a *unique* way compatible with the inclusions $X_\alpha \hookrightarrow \coprod_\alpha X_\alpha$. The commutativity of the diagrams

$$X_\alpha \xrightarrow{\ f_{\alpha,\beta}\ } X_\beta$$

$$h_\alpha \searrow \quad \downarrow h_\beta$$

$$Z$$

implies that *equivalent* elements under $\sim$ will map to the *same* element of Z via $\bigsqcup h_\alpha$, so that we get a well-defined map

$$u = \coprod h_\alpha/\!\sim\colon \coprod_\alpha X_\alpha/\!\sim\; \to Z$$

that is *unique* (because $\bigsqcup h_\alpha$ was). Since our construction has the same universal property as the direct limit, it must be isomorphic to it (in a unique way). $\square$

The reader may still find the concept of direct limit hard to grasp. We claim that direct limits are a kind of "generalized union," something implied the following:

PROPOSITION A.5.19. *Assuming the notation of definition A.5.17 on page 448 and that $\mathscr{C}$ is a concrete category, we have*

(A.5.8) $$\varinjlim X_\alpha = \bigcup_{\alpha \in I} \phi_\alpha(X_\alpha)$$

If all of the maps $f_{\alpha,\beta}$ are injective, then the maps $\phi_\alpha\colon X_\alpha \to \varinjlim X_\alpha$ are also injective.

REMARK. If the $f_{\alpha,\beta}$ are injective, the direct limit is *literally* a union of the X_α.

If they are *not* injective, and $\mathscr{C}$ is a category of groups, rings, or vector-spaces, the direct limit essentially divides out by the kernels of the $f_{\alpha,\beta}$ — "forcing them" to be injective — and *then* takes the union.

PROOF. Equation A.5.8 follows immediately from equations A.5.6 and A.5.7 on the previous page.

If all of the $f_{\alpha,\beta}$ are injective, then the only way two elements $x_1, x_2 \in X_\alpha$ can become equivalent is for $x_1 = x_2 \in X_\alpha$. $\square$

EXAMPLE A.5.20. Suppose I is the set of positive integers and $i \leq j$ is $i|j$. Let

$$R_n = \mathbb{Z}\left[\frac{1}{n}\right]$$

Then

$$f_{n,m}\colon R_n \to R_m$$

when $n|m$, is defined to send

$$\frac{1}{n} \mapsto \frac{k}{m}$$

where $k = m/n$. We claim that $\varinjlim R_n = \mathbb{Q}$. The maps $f_{n,m}$ are all injective and each $R_n \subset \mathbb{Q}$ so

$$\varinjlim R_n = \bigcup_{n=1}^{\infty} R_n \subset \mathbb{Q}$$

Since every possible denominator occurs in some R_n this inclusion must actually be an equality.

DEFINITION A.5.21. In the notation of definition A.5.17 on page 448, a subset $I' \subset I$ is said to be *cofinal*, if for every $\alpha \in I$, there exists a $\beta \in I'$ such that $\alpha \leq \beta$.

REMARK. Cofinal subsets are important because they *determine* colimits and limits:

PROPOSITION A.5.22. *In the notation of definition A.5.17 on page 448, if $I' \subset I$ is a cofinal subset then*

$$\varinjlim X_\beta = \varinjlim X_\alpha$$

where α runs over I and β runs over I'.

REMARK. This is significant because direct limits are sometimes easier to compute with cofinal subsets.

PROOF. This follows immediately from the universal properties: since all X_β map to $\varinjlim X_\alpha$, we get a unique map

$$\varinjlim X_\beta \to \varinjlim X_\alpha$$

Since every $\alpha \in I$ is $\leq \beta$ for some $\beta(\alpha) \in I'$, we get unique maps from all of the $X_\alpha \to X_{\beta(\alpha)}$ inducing a unique map to

$$\varinjlim X_\alpha \to \varinjlim X_\beta$$

$\square$

Recall the concept of rings of fractions in definition A.1.89 on page 383. We can define this in terms of a universal property:

PROPOSITION A.5.23. *In the category, $\mathscr{R}$, of commutative rings the pair $(S^{-1}A, \iota)$ has the universal property: every element of S maps to a unit in $S^{-1}A$, and any other homomorphism $f \colon A \to B$ with this property factors uniquely through ι:*

$$A \xrightarrow{\quad \iota \quad} S^{-1}A$$
$$f \searrow \qquad \downarrow \beta$$
$$B$$

PROOF. If β exists

$$s\frac{a}{s} = a \implies \beta(s)\beta\left(\frac{a}{s}\right) = \beta(a) = f(a)$$

so just *define*

$$\beta\left(\frac{a}{s}\right) = f(a)f(s)^{-1}$$

Now

$$\frac{a}{s} = \frac{b}{t} \implies z(at - bs) = 0 \text{ for some } z \in S$$

and this implies

$$f(a)f(t) - f(b)f(s) = 0$$

since $f(z)$ is a unit.   □

Modules of fractions also have a universal property:

PROPOSITION A.5.24. *If M is a module over a ring A with multiplicative set $S \subset A$ and N is a module over $S^{-1}A$ then N is also a module over A via the standard inclusion $\iota\colon A \to S^{-1}A$. Any homomorphism*

$$f\colon M \to N$$

over A extends uniquely to a homomorphism of $S^{-1}A$ modules

$$\bar{f}\colon S^{-1}M \to N$$

that makes the diagram

$$
\begin{array}{ccc}
S^{-1}M & \xrightarrow{\ \bar{f}\ } & N \\
\uparrow & & \uparrow \\
M & \xrightarrow{\ f\ } & N
\end{array}
$$

where the vertical maps are homomorphisms of modules covering the map $\iota\colon A \to S^{-1}A$.

PROOF. Left as an exercise to the reader.   □

One bonus of this approach is the following (this is very similar to example A.5.20 on page 450):

COROLLARY A.5.25. *Suppose A is a commutative ring with a multiplicative set $S \subset A$. Define an order on the elements of S via:*

$s_1 \le s_2$ *if there exists an element $x \in R$, such that $s_2 = x \cdot s_1$.*

Define maps $f_{s,t}\colon A_s \to A_t$ for $s, t \in S$ with $t = x \cdot s$ by

$$\frac{a}{s} \mapsto \frac{a \cdot x}{t}$$

Then

$$S^{-1}A = \varinjlim A_s$$

REMARK. Recall the notation A_h in definition A.1.91 on page 384.

The proof below almost seems like "cheating" — we ignore algebraic subtleties and give an "arrow-theoretic" argument. This was one of the early complaints against category theory (and [**37**]).

The philosophy of category theory is that if one can prove something *merely* by analyzing patterns of mappings, one should do so.

PROOF. The ring of fractions, $S^{-1}A$, and the direct limit, $\varinjlim A_s$, have the same universal property. $\square$

If we *reverse* all of the arrows that occur in the diagrams of definition A.5.17 on page 448, we get another important construction — the *inverse limit:*

DEFINITION A.5.26. Suppose $\mathscr{C}$ is a category and I is a partially ordered set of indices. Suppose $\{X_\alpha\}$, for $\alpha \in I$, is a sequence of objects of $\mathscr{C}$. Whenever $\alpha \leq \beta$ suppose there is a morphism

$$f_{\alpha,\beta} \colon X_\alpha \leftarrow X_\beta$$

and whenever $\alpha \leq \beta \leq \gamma$ the diagram

commutes. Then the *inverse limit,* $\varprojlim X_\alpha$ has

(1) morphisms $\pi_\alpha \colon X_\alpha \leftarrow \varprojlim X_\alpha$ that make the diagrams

(A.5.9)

commute for all $\alpha, \beta \in I$ with $\alpha \leq \beta$.

(2) the *universal property* that whenever there is an object $Z \in \mathscr{C}$ and morphisms $h_\alpha \colon X_\alpha \leftarrow Z$ for all $\alpha \in I$ that make the diagrams

commute for all $\alpha \leq \beta$, then there exists a *unique morphism* u: $\varprojlim X_\alpha \leftarrow Z$ that makes the diagrams

$$X_\alpha \xleftarrow{\quad \pi_\alpha \quad} \varinjlim X_\alpha$$

commute for all $\alpha \in I$.

REMARK. So anything that maps to all of the X's in a way compatible with the maps $f_{\alpha,\beta}$ also maps to the inverse limit.

Since the inverse limit has a universal property, it is unique up to isomorphism (if it exists at all!). Hardcore category-theorists prefer the term "limit" for the inverse limit.

As with the direct limit, we have an explicit construction of the inverse limit in categories of groups, rings, and vector-spaces:

PROPOSITION A.5.27. *Let $\mathscr{C}$ be a category of groups, rings, or vector-spaces and assume the notation of definition A.5.26 on the preceding page. Then*

$$\varprojlim X_\alpha \subset \prod_{\alpha \in I} X_\alpha$$

is the subset of (possibly infinite) sequences

$$(\dots, x_\alpha, \dots)$$

where $x_\alpha \in X_\alpha$ for all $\alpha \in I$, and with the property that, whenever $\alpha \leq \beta$, $f_{\alpha,\beta}(x_\beta) = x_\alpha$.

The maps π_β: $X_\beta \leftarrow \varinjlim X_\alpha$ are the composites

$$\varprojlim X_\alpha \hookrightarrow \prod_{\alpha \in I} X_\alpha \to X_\beta$$

where $\prod_{\alpha \in I} X_\alpha \to X_\beta$ is just projection to a factor.

REMARK. Whereas the direct limit *glues together* the X's via the $f_{\alpha,\beta}$, the inverse limit *selects* infinite sequences compatible with the $f_{\alpha,\beta}$. If $\mathscr{C}$ is a category of groups, rings, or vector-spaces, then $f_{\alpha,\beta}$ will preserve this structure and the inverse limit will also have it.

PROOF. We only have to verify that this construction has the same universal property as the inverse limit. If $Z \in \mathscr{C}$ and has maps h_β: $Z \to X_\beta$ for all $\beta \in I$, then we get a *unique* map

$$\prod h_\alpha \colon Z \to \prod_{\alpha \in I} X_\alpha$$

—see the definition of product in diagram A.5.2 on page 439 and extend it to an arbitrary number of factors. The commutativity of the diagrams

$$
\begin{array}{ccc}
X_\alpha & \xleftarrow{\;f_{\alpha,\beta}\;} & X_\beta \\
& \nwarrow\; h_\alpha \qquad \uparrow h_\beta & \\
& Z &
\end{array}
$$

implies that the image of $\prod h_\alpha$ will *actually* lie within $\varprojlim X_\alpha \subset \prod X_\alpha$. This verifies the universal property. □

As we noted earlier, direct limits are "generalized unions". Under some circumstances, inverse limits are like "generalized intersections:"

PROPOSITION A.5.28. *Under the assumptions of proposition A.5.27, suppose the $f_{\alpha,\beta}\colon X_\beta \to X_\alpha$ are injective for all $\alpha, \beta \in I$. Then so is*

$$
\pi_\beta \colon \varprojlim X_\alpha \to X_\beta
$$

for all $\beta \in I$.
If there exists $X \in \mathscr{C}$ such that

$$
X_\alpha \subset X
$$

for all $\alpha \in I$ and $\alpha \leq \beta$ if and only if $X_\beta \subset X_\alpha$ where $f_{\alpha,\beta}\colon X_\beta \to X_\alpha$ is the inclusion, then

$$
\varprojlim X_\alpha = \bigcap_{\alpha \in I} X_\alpha
$$

PROOF. If all of the $f_{\alpha,\beta}$ are injective, then a sequence

$$
(\dots, x_\alpha, \dots)
$$

is *uniquely* determined by its αth member: the βth element, x_β, to the right of x_α will be $f_{\alpha,\beta}^{-1}(x_\alpha)$, and this is unique. It follows that the projection map

$$
\varprojlim X_\beta \to X_\alpha
$$

is injective. Its image is the set of all $x_\alpha \in X_\alpha$ of the form $f_{\alpha,\beta}(x_\beta)$ for *all β* with $\alpha \leq \beta$.
Moving on to the second statement, we have proved that

$$
\varprojlim X_\alpha \subset \bigcap_{\alpha \in I} X_\alpha
$$

Equality follows from both objects having the same universal property. □

From the proof of proposition A.5.28, it is clear that $x_\beta = f_{\alpha,\beta}^{-1}(x_\alpha)$ (if it exists). If $f_{\alpha,\beta}$ is *not* injective, *all* elements of $f_{\alpha,\beta}^{-1}(x_\alpha)$ give rise to new sequences from the β-position on. For instance:

PROPOSITION A.5.29. *Under the assumptions of proposition A.5.27, suppose that the set of indices, I, is the disjoint union of $\{I_j\}$ — i.e. no $\alpha \in I_j$ is comparable with any $\beta \in I_k$ with $j \neq k$. Then*

$$\underbrace{\varprojlim X_\alpha}_{\alpha \in I} = \prod_j \underbrace{\varprojlim X_\beta}_{\beta \in I_j}$$

PROOF. Since the I_j are disjoint they have no influence over each other — all sequences from I_j are paired with all sequences from I_k, $j \neq k$.   □

It follows that $\varprojlim X_\alpha$ can be very large indeed:

EXAMPLE A.5.30. Let I be positive integers ordered in the usual way, let $p \in \mathbb{Z}$ be a prime, and let $X_n = \mathbb{Z}_{p^n}$ for all n. The maps $f_{n,m} \colon \mathbb{Z}_{p^m} \to \mathbb{Z}_{p^n}$ are reduction modulo p^n (where $n \leq m$).
Then

$$\mathbb{Z}_{(p)} = \varprojlim X_n$$

is called the *p-adic integers* and its field of fractions is called the *p-adic numbers*, $\mathbb{Q}_{(p)}$. Reduction modulo p^n (for all n) defines an injection

$$\mathbb{Z} \hookrightarrow \mathbb{Z}_{(p)}$$

and, like $\mathbb{R}$, $\mathbb{Z}_{(p)}$ is *uncountable* for all p. These rings were first described by Kurt Hensel in 1897 (see [**75**]), with a definition *wildly* different from ours. Hensel showed that one could define infinite series in $\mathbb{Q}_{(p)}$ like that for e^x with many number-theoretic applications.
Technically, elements of $\mathbb{Z}_{(p)}$ are "infinite series"

$$n_0 + n_1 \cdot p + n_2 \cdot p^2 + \cdots$$

such that $0 \leq n_i < p$ for all i. The image $\mathbb{Z} \subset \mathbb{Z}_{(p)}$ consists of the "series" that terminate after a finite number of terms. Two such "series" are equal if all corresponding n_i's are equal. Define a metric on $\mathbb{Z}$ via the *p-adic valuation* defined in A.4.33 on page 432

$$d(m_1, m_2) = \left(\frac{1}{2}\right)^{v_p(m_1 - m_2)}$$

so $v_p(m_1 - m_2)$ is the highest power of p such that

$$p^k | (m_1 - m_2)$$

Then $\mathbb{Z}_{(p)}$ is the *completion* of $\mathbb{Z}$ in this metric, and two elements, P, Q of $\mathbb{Z}_{(p)}$ are equal if and only if

$$\lim_{i \to \infty} d(P_i, Q_i) = 0$$

where P_i and Q_i are, respectively, the ith partial sums of P and Q.

Here's another interesting example:

EXAMPLE A.5.31. Let A be a ring and let $\mathfrak{m} = (X) \subset A[X]$ be an ideal. Then

$$A[[X]] = \varprojlim A[X]/\mathfrak{m}^n$$

On the other hand, if there is a top index in I, the inverse limit is well-behaved:

PROPOSITION A.5.32. *Under the assumptions of proposition A.5.27 on page 454, suppose there exists $\gamma \in I$ such that $\alpha \leq \gamma$ for all $\alpha \in I$. Then*

$$\varprojlim X_\alpha = X_\gamma$$

PROOF. We could do an algebraic analysis of this statement, but it is easier to "cheat," so our proof is: they both have the same universal property. □

EXERCISES.

8. Let $\mathscr{C}$ be a category and let $\mathscr{C}_\infty$ be the category of infinite sequences

$$\cdots \to x_2 \to x_1$$

of morphisms of objects of $\mathscr{C}$. Then

$$\varprojlim *: \mathscr{C}_\infty \to \mathscr{C}$$

is a functor. Show that this is an adjoint of the functor

$$\Delta_\infty: \mathscr{C} \quad \to \quad \mathscr{C}_\infty$$
$$x \quad \mapsto \quad \cdots \xrightarrow{1} x \xrightarrow{1} x$$

9. Suppose $\{X_\alpha\}$, for $\alpha \in I$, is a sequence of objects of a concrete category, $\mathscr{C}$. Whenever $\alpha \leq \beta$ suppose there is a morphism

$$f_{\alpha,\beta}: X_\alpha \to X_\beta$$

and whenever $\alpha \leq \beta \leq \gamma$ the diagram

commutes. If $x, y \in X_\alpha$ map to the same element of $\varinjlim X_\alpha$, show that there exists a $\beta \geq \alpha$ such that $f_{\alpha,\beta}(x) = f_{\alpha,\beta}(y)$.

A.5.5. Abelian categories. An abelian category is essentially one in which morphisms of objects have kernels and cokernels. The standard example is the category of modules over a commutative ring. The official definition is:

DEFINITION A.5.33. A category $\mathscr{A}$ is *abelian* if:
  (1) it has *products* and *coproducts* of all pairs of objects,
  (2) it has a *zero object* (which behaves like an identity for products and coproducts),

(3) all morphisms have a kernel and cokernel:

(a) if $A \xrightarrow{f} B$ is a morphism, there exists a monomorphism $K \xrightarrow{m} A$ such that $f \circ m = 0$, and if $C \xrightarrow{g} A$ is any morphism with $f \circ g = 0$, there exists a *unique* morphism $v \colon C \to K$ such that

$$
\begin{array}{ccc}
C & \xrightarrow{\ g\ } & A \\
& \searrow{\scriptstyle v} & \big\uparrow{\scriptstyle m} \\
& & K
\end{array}
$$

commutes.

(b) if $A \xrightarrow{f} B$ is a morphism, there exists an epimorphism $B \xrightarrow{e} E$ such that $e \circ f = 0$, and if $g \colon B \to D$ is any morphism with $g \circ f = 0$, then there exists a *unique* morphism $v \colon E \to D$ such that

$$
\begin{array}{ccc}
B & \xrightarrow{\ g\ } & D \\
& \searrow{\scriptstyle e} & \big\uparrow{\scriptstyle v} \\
& & E
\end{array}
$$

(4) the set of morphisms between two objects, $\hom_{\mathscr{A}}(A, B)$, has the structure of an abelian group for which composition is distributive over sums.

If $F \colon \mathscr{A} \to \mathscr{A}'$ is a functor between abelian categories, F is said to be *additive* if, whenever we have morphisms $g_1, g_2 \colon M \to N$ in $\mathscr{A}$,

$$
F(g_1 + g_2) = F(g_1) + F(g_2) \colon F(M) \to F(N)
$$

REMARK. Since kernels and cokernels are defined by universal properties, they are unique up to isomorphism. Also note that the kernel and cokernel defined *here* are "arrow-theoretic" versions of the more familiar algebraic concepts — i.e., morphisms.

Examples of *additive* functors include $M \otimes_{\mathbb{Z}} *$ and $\hom_{\mathscr{A}}(M, *)$. The functor $F \colon \mathscr{A}b \to \mathscr{A}b$ that sends an abelian group G to $G \otimes_{\mathbb{Z}} G$ is an example of a functor that is *not* additive (it is "quadratic").

The concept of a projective module is well-defined for an abelian category

DEFINITION A.5.34. If $A, B \in \mathscr{A}$ are objects of an abelian category with $f \colon A \to B$ an epimorphism (see definition A.5.6 on page 441), an object P is *projective* if, for any morphism $g \colon P \to B$, there exists a morphism $\ell \colon P \to A$ that fits into a commutative diagram

$$
\begin{array}{ccc}
& & A \\
& \nearrow{\scriptstyle \ell} & \big\downarrow{\scriptstyle f} \\
P & \xrightarrow{\ g\ } & B
\end{array}
$$

The category $\mathscr{A}$ will be said to have *enough projectives* if, for any object A there exists a projective object P and an epimorphism $P \to A$.

REMARK. For instance, the category of modules over a ring *always* has enough projectives because every module is the surjective image of a free module.

If we reverse all of the arrows in A.5.34 on the facing page, we get a definition of *injective* objects:

DEFINITION A.5.35. If $A, B \in \mathscr{A}$ are objects of an abelian category with $f: B \to A$ a monomorphism (see definition A.5.6 on page 441), an object I is *injective* if, any morphism $g: B \to I$, there exists a morphism $e: A \to I$ that fits into a commutative diagram

$$
\begin{array}{ccc}
B & \xrightarrow{\ f\ } & A \\
{\scriptstyle g}\downarrow & \swarrow{\scriptstyle e} & \\
I & &
\end{array}
$$

The category $\mathscr{A}$ will be said to have *enough injectives* if, for any object A there exists an injective object I and a monomorphism $A \to I$.

REMARK. Homomorphisms into injective objects *extend* to other objects containing them.

The categorical property of a product in definition A.5.1 on page 439 implies that arbitrary *products* of injective objects are injective.

Over the category of modules, we have a criterion for injectivity:

PROPOSITION A.5.36 (Baer's Criterion). *If R is a commutative ring, an R-module, I is injective if and only if every homomorphism $\mathfrak{J} \to I$ from an ideal $\mathfrak{J} \subset R$ extends to a homomorphism $R \to I$.*

REMARK. In other words, in the category of modules over a ring, injectivity only has to be verified for *ideals* of the ring.

PROOF. The only-if part follows from the definition of injective modules.

Conversely, suppose $A \subset B$ are R-modules and $f: A \to I$ is a homomorphism and we consider extensions to submodules B' with

$$A \subset B' \subseteq B$$

These extensions are partially ordered by inclusion. Zorn's lemma (A.1.27 on page 353) implies that there is a maximal one, B' say. If $B' \neq B$, we will get a contradiction. If $b \in B \setminus B'$ then $\mathfrak{J}(b) = \{r \in R | r \cdot b \in B'\}$ is an ideal of R and $\mathfrak{J}(b) \xrightarrow{\cdot b} B' \xrightarrow{f} I$ defines a homomorphism into I. The hypotheses imply that this extends to all of R, so $b \in B'$. □

Since all abelian groups are modules over $\mathbb{Z}$ and all ideals of $\mathbb{Z}$ are of the form (m) for $m \in \mathbb{Z}$, Baer's Criterion implies that

PROPOSITION A.5.37. *An abelian group, G, is injective if and only if it is divisible — i.e. for any $g \in G$ and any integer n there exists an element $g' \in G$ such that $g = n \cdot g'$.*

Since quotients of divisible groups are divisible, we conclude:

PROPOSITION A.5.38. *Any quotient of an injective object in $\mathscr{A}b$ is injective in $\mathscr{A}b$.*

EXAMPLE A.5.39. In the category of abelian groups, $\mathscr{A}b$, $\mathbb{Q}$ and $\mathbb{Q}/\mathbb{Z}$ are injective.

This allow us to conclude that the category of abelian groups has enough injectives (see [**10**]):

PROPOSITION A.5.40. *If A is an abelian group and*

$$A = \frac{F}{K}$$

where F is a free abelian group, then $F \otimes_{\mathbb{Z}} \mathbb{Q}$ is injective and

$$A \hookrightarrow \frac{F \otimes_{\mathbb{Z}} \mathbb{Q}}{K}$$

is an injective abelian group containing A so the category of abelian groups, $\mathscr{A}b$, has enough injectives.

It is interesting that this result immediately extends to the category of modules over an arbitrary ring (see [**36**]):

PROPOSITION A.5.41. *If R is a ring and M is an R-module and I is an injective abelian group, then*

$$I(R) = \hom_{\mathscr{A}b}(R, I)$$

is an injective R-module — with R acting on the first factor via

$$(r' \cdot \varphi)(r) = \varphi(r' \cdot r)$$

for $\varphi \in \hom_{\mathscr{A}b}(R, I)$. In addition, there exists an injective R-module N and an inclusion

$$M \hookrightarrow N$$

so that the category of modules over R, $\mathscr{M}_R$, has enough injectives.

PROOF. Suppose $A \subset B$ is an inclusion of R-modules and $g \colon A \to \hom_{\mathscr{A}b}(R, I)$ is a homomorphism. We will show that this extends to B. Define a natural map

$$\iota \colon I(R) \quad \to \quad I$$
$$f \quad \mapsto \quad f(1)$$

The composite $\iota \circ g \colon A \to I$, regarded as a map of abelian groups, extends to $\bar{g} \colon B \to I$ and we define

$$G \colon B \quad \to \quad \hom_{\mathscr{A}b}(R, I)$$
$$b \quad \mapsto \quad (r \mapsto \bar{g}(r \cdot b))$$

— a homomorphism of R-modules and the desired extension.

To prove the second statement, note the existence of a monomorphism

$$f \colon M \quad \to \quad \hom_{\mathscr{A}b}(R, M)$$
$$m \quad \mapsto \quad (r \mapsto r \cdot m)$$

of R-modules. If we "forget" the module structure of M and regard it only as an abelian group, there exists an injective abelian group and a morphism

$$g\colon M \to I$$

The composite

$$M \xrightarrow{f} \hom_{\mathscr{A}b}(R, M) \xrightarrow{\hom_{\mathscr{A}b}(1, g)} \hom_{\mathscr{A}b}(R, I)$$

is a monomorphism (see exercise 12). $\qquad\square$

EXERCISES.

10. If A is an abelian group, show that $\hom(A, \mathbb{Q}/\mathbb{Z}) = 0$ if and only if $A = 0$.

11. Show that if we have a monomorphism

$$f\colon A \to B$$

where A is injective, there exists a map

$$g\colon B \to A$$

such that $g \circ f = 1\colon A \to A$.

12. If $\mathscr{A}$ is an abelian category and

$$0 \to A \xrightarrow{r} B \xrightarrow{s} C \to 0$$

is an exact sequence — i.e., $r = \ker s$ and $s = \operatorname{coker} r$, show that

$$0 \to \hom_{\mathscr{A}}(D, A) \xrightarrow{\hom_{\mathscr{A}}(1, r)} \hom_{\mathscr{A}}(D, B) \xrightarrow{\hom_{\mathscr{A}}(1, s)} \hom_{\mathscr{A}}(D, C)$$

is exact.

A.5.6. Tensor products. The standard example of an abelian category is $\mathscr{M}_R$ — the category of modules over a commutative ring, R.

There are various operations with modules that are most easily understood in terms of category theory. The simplest is the direct sum

DEFINITION A.5.42. If $M_1, M_2 \in \mathscr{M}_R$, the *direct sum*, $M_1 \oplus M_2 \in \mathscr{M}_R$, is the module of pairs

$$(m_1, m_2) \in M_1 \oplus M_2$$

with R acting via

$$r \cdot (m_1, m_2) = (r \cdot m_1, r \cdot m_2)$$

for all $r \in R$, $m_1 \in M_1$ and $m_2 \in M_2$.

REMARK. This is just a straightforward generalization of the concept of direct sum of vector-spaces — and the direct sum is a product and coproduct in the category of R-modules.

For instance, the free module R^n is a direct sum

$$R^n = \underbrace{R \oplus \cdots \oplus R}_{n \text{ factors}}$$

The direct sum is a functor of two variables:

PROPOSITION A.5.43. *If $f_1 \colon M_1 \to N_1$ and $f_2 \colon M_2 \to N_2$ are morphisms in $\mathscr{M}_R$, then there is an induced morphism*

$$
\begin{aligned}
f_1 \oplus f_2 \colon M_1 \oplus M_2 &\to N_1 \oplus N_2 \\
(m_1, m_2) &\mapsto (f_1(m_1), f_2(m_2))
\end{aligned}
$$

for all $m_1 \in M_1$ and $m_2 \in M_2$. In addition, $\ker(f_1 \oplus f_2) = \ker f_1 \oplus \ker f_2$.

PROOF. The only thing that needs to be proved is the statement about the kernels. Clearly $\ker f_1 \oplus \ker f_2 \subset \ker(f_1 \oplus f_2)$. If (m_1, m_2) maps to 0 in $N_1 \oplus N_2$, we must have $f_1(m_1) = f_2(m_2) = 0$ so this proves $\ker(f_1 \oplus f_2) \subset \ker f_1 \oplus \ker f_2$. $\qquad\square$

The following concept also originated with linear algebra, but is more complex than the direct sum. It is another functor of two variables:

DEFINITION A.5.44. If $M_1, M_2 \in \mathscr{M}_R$, then define the *tensor product of M_1 and M_2 over R*

$$M_1 \otimes_R M_2$$

to be the R-module that is a quotient of the free abelian group generated by symbols $\{m_1 \otimes m_2\}$ with $m_1 \in M_1$, $m_2 \in M_2$ subject to the identities

    (1) $(r \cdot m_1) \otimes m_2 = m_1 \otimes (r \cdot m_2) = r \cdot (m_1 \otimes m_2)$ (defines the R-action on $M_1 \otimes_R M_2$) for all $r \in R$, $m_1 \in M_1$, $m_2 \in M_2$,
    (2) $(m_1 + m_1') \otimes m_2 = m_1 \otimes m_2 + m_1' \otimes m_2$ for all $m_1, m_1' \in M_1$, $m_2 \in M_2$,
    (3) $m_1 \otimes (m_2 + m_2') = m_1 \otimes m_2 + m_1 \otimes m_2'$ for all $m_1 \in M_1$, $m_2, m_2' \in M_2$.

REMARK. Rule 1 implies that $0 \otimes m_2 = m_1 \otimes 0 = 0$. Here is another way to define the tensor product:

Form the free abelian group $\mathbb{Z}[M_1 \times M_2]$. Its elements are formal linear combinations of symbols $[m \times n]$ for all $m \in M_1$ and $n \in M_2$. Then $M_1 \otimes_R M_2 = \mathbb{Z}[M_1 \times M_2]/W$, where $W \subset \mathbb{Z}[M_1 \times M_2]$ is the subgroup generated by

    (1) $[r \cdot m_1 \times m_2] - [m_1 \times (r \cdot m_2)]$, $[e \cdot m_1 \times m_2] - e \cdot [m_1 \times m_2]$, for all $e \in \mathbb{Z}$, $r \in R$, $m_1 \in M_1$, and $m_2 \in M_2$
    (2) $[(m_1 + m_1') \times m_2] - [m_1 \times m_2] - [m_1' \times m_2]$, for all $m_1, m_1' \in M_1$, $m_2 \in M_2$,
    (3) $[m_1 \times (m_2 + m_2')] - [m_1 \otimes m_2] - [m_1 \otimes m_2']$ for all $m_1 \in M_1$, $m_2, m_2' \in M_2$.

The R-module structure is defined by setting $r \cdot [m_1 \times m_2] = [r \cdot m_1 \times m_2]$ for all $r \in R$, $m_1 \in M_1$, $m_2 \in M_2$

EXAMPLE A.5.45. If $M \in \mathscr{M}_R$, then

$$
\begin{aligned}
R \otimes_R M &\xrightarrow{\ \cong\ } M \\
r \otimes m &\mapsto r \cdot m
\end{aligned}
$$

Clearly this map is surjective. If $r \otimes m$ is in the kernel, then $r \cdot m = 0 \in M$. In this case, rule 1 on the facing page in definition A.5.44 on the preceding page implies that

$$r \cdot 1 \otimes m \sim 1 \otimes r \cdot m = 0$$

In the category-theoretic sense, $M \otimes_R N$ is neither a product nor a coproduct. It does have a universal property, though:

PROPOSITION A.5.46. *Let M, N, and T be modules over a commutative ring, R, and let*

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

be a mapping with the property that

(1) $f|m \times N \to T$ is an R-module-homomorphism for any $m \in M$
(2) $f|M \times n \to T$ is an R-module homomorphism for any $n \in N$
(3) $f(r \cdot m, n) = f(m, r \cdot n)$ for all $m \in M, n \in N$, and $r \in R$

Then there exists a unique map

$$g: M \otimes_R N \to T$$

that makes the diagram

$$M \times N \xrightarrow{\ c\ } M \otimes_R N$$

commute, where $c(m,n) = m \otimes n$, for all $m \in M$ and $n \in N$.

REMARK. Here $M \times N$ is simply a Cartesian product of *sets.* A map satisfying statements 1 and 2 above is said to be *bilinear.*

The canonical map

$$c: M \times N \to M \otimes_R N$$

is not surjective in general since $M \otimes_R N$ consists of *formal linear combinations* of symbols $m \otimes n$. Elements of $M \otimes_R N$ in the image of c are called *decomposable tensors* or elementary tensors. The paper **[61]** gives criteria for elements of $M \otimes_R N$ to be decomposable when R is a field.

This result implies that an decomposable tensor $m \otimes n$ *vanishes* if and only if every bilinear map

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

sends $m \times n$ to 0.

PROOF. The *only* map $\mathbb{Z}[M \times N] \to T$ compatible with f is

$$\begin{aligned} \mathbb{Z}[M \times N] &\to T \\ m \times n &\mapsto f(m,n) \end{aligned}$$

for all $m \in M$ and $n \in N$. The defining relations for $M \otimes_R N$ and the conditions on the map f imply that this gives a well-defined map

$$\begin{aligned} M \otimes N &\to T \\ m \otimes n &\mapsto f(m,n) \end{aligned}$$

for all $m \in M$ and $n \in N$. Since any such map must lift to a map $\mathbb{Z}[M \times N] \to T$, this must be *unique.* $\square$

Tensor-products are functors of two variables:

PROPOSITION A.5.47. *Let $f\colon V_1 \to V_2$ and $g\colon W_1 \to W_2$ be homomorphisms of vector-spaces. Then there is a natural map*

$$
\begin{aligned}
f \otimes g\colon V_1 \otimes_k W_1 \;&\to\; V_2 \otimes_k W_2 \\
(f \otimes g)(v \otimes w) \;&=\; f(v) \otimes g(w)
\end{aligned}
$$

REMARK. Exercise 16 on page 473 gives some idea of what the homomorphism $f \otimes g$ looks like.

Tensor products are distributive over direct sums, a property that allows us to do many computations:

PROPOSITION A.5.48. *Let M, N, T be modules over the commutative ring R. Then there are standard isomorphisms*

$$
M \otimes_R (N \oplus T) = M \otimes_R N \oplus M \otimes_R T
$$

and

$$
(M \oplus N) \otimes_R T = M \otimes_R T \oplus N \otimes_R T
$$

PROOF. We will prove the first case: the second is similar. We could use a detailed algebraic argument, but it is easier to "cheat" and use the universal property of a tensor product.

We will show that, given any bilinear map $z\colon M \times (N \oplus T) \to Z$, where Z is an R-module, there exists a *unique* homomorphism $d\colon M \otimes_R N \oplus M \otimes_R T \to Z$ making the diagram

(A.5.10)

$$
\begin{array}{ccc}
M \times (N \oplus T) & \xrightarrow{\ b\ } & M \otimes_R N \oplus M \otimes_R T \\
& {\scriptstyle z}\searrow & \;\;\downarrow{\scriptstyle d} \\
& & Z
\end{array}
$$

commute. Here, b is a bilinear map taking the place of the c map in A.5.46 on the previous page. This will show that $M \otimes_R N \oplus M \otimes_R T$ has the same universal property as $M \otimes_R (N \oplus T)$ so it must be isomorphic to it.

We begin by constructing a bilinear map $b\colon M \times (N \oplus T) \to M \otimes_R N \oplus M \otimes_R T$ via $b(m, (n, t)) = (m \otimes n, m \otimes t)$ for all $m \in M$, $n \in N$, and $t \in T$. This is easily verified to be bilinear:

(1) for any fixed $m_0 \in M$, $\ell(n, t) = b(m_0, (n, t)) = (m_0 \otimes n, m_0 \otimes t)$ for all $n \in N$ and $t \in T$, defines an R-module homomorphism

$$
\ell\colon N \oplus T \to M \otimes_R N \oplus M \otimes_R T
$$

since the composites

$$
\begin{aligned}
N \;&\to\; m_0 \otimes N \subset M \otimes_R N \\
T \;&\to\; m_0 \otimes T \subset M \otimes_R T
\end{aligned}
$$

are module-homomorphisms.

(2) a similar argument shows that for any fixed $n_0 \in N$ and $t_0 \in T$, the map $\ell(m) = b(m, (n_0, t_0))$, for all $m \in M$ defines a module homomorphism

$$
\ell\colon M \to M \otimes_R N \oplus M \otimes_R T
$$

Now, suppose we have a bilinear map

$$z: M \times (N \oplus T) \to Z$$

We will show that there exists a unique map

$$d: M \otimes_R N \oplus M \otimes_R T \to Z$$

that makes diagram A.5.10 on the facing page commute.

We define d on the direct summands of $M \otimes_R N \oplus M \otimes_R T$:

(1) $d_1: M \otimes_R N \to Z$ *must* send $m \otimes n$ to $z(m, (n, 0)) \in Z$ for all $m \in M$ and $n \in N$ so we define $d_1(m \otimes n) = z(m, (n, 0))$. The bilinearity of z implies that $d_1 | M \otimes n_0: M \to Z$ is a module homomorphism for any fixed $n_0 \in N$ and $d_1 | m_0 \otimes N: N \to Z$ is also a module homomorphism. It follows that d_1 is a module homomorphism.

(2) We define $d_2: M \otimes_R T \to Z$ by $d_2(m \otimes t) = z(m, (0, t))$. This is the only definition compatible with z and an argument like that used above shows that it is a module-homomorphism.

We set

$$d = d_1 + d_2: M \otimes_R N \oplus M \otimes_R T \to Z$$

This is a module-homomorphism that makes diagram A.5.10 on the preceding page commute. It is *unique* because it is uniquely determined on the two summands. □

COROLLARY A.5.49. *If $M \in \mathscr{M}_R$, R, then*

$$M \otimes_R R^n = \underbrace{M \oplus \cdots \oplus M}_{n \text{ times}}$$

and

$$R^n \otimes_R R^m = R^{n \cdot m}$$

PROOF. This follows from example A.5.45 on page 462, proposition A.5.48 on the preceding page and induction on n. □

If R is an algebra over another ring S, we can define the structure of an R-module on $A \otimes_S B$ by $f \cdot (a \otimes b) = (r \cdot a \otimes r \cdot b)$, for $r \in R$. We can also define an R-action on groups of homomorphisms:

DEFINITION A.5.50. *If M and N are R-modules, where R is an S-algebra, then*

$$\hom_R(M, N)$$

denotes morphisms that are R-linear (i.e. morphisms of $\mathscr{M}_R$) and

$$\hom_S(M, N)$$

are morphisms of $\mathscr{M}_S$, i.e. morphisms that are S-linear. Then we can equip $\hom_S(M, N)$ with the structure of an R-module via the rule

If $f \in \hom_S(M, N)$ is such that $f(m) = n$, then $(r \cdot f)(m) = f(r \cdot m)$.

We have important relations between $\hom_S$ and $\hom_R$:

PROPOSITION A.5.51. *If $A, B, C \in \mathscr{M}_R$, where R is an S-algebra, then there exists a unique isomorphism*

$$s \colon \hom_R(A, \hom_S(B, C)) \to \hom_R(A \otimes_S B, C)$$

REMARK. This is clearly natural with respect to all homomorphisms of A, B, or C.

PROOF. We define the map by $s(\varphi)(a \otimes b) = \varphi(a)(b)$. If $s(\varphi) = 0$, if is the 0-map for all b or a, so it vanishes in $\hom_R(A, \hom_S(B, C))$. It follows that s is injective. If $f \in \hom_R(A \otimes_S B, C)$ then $f(a, *)$ for a fixed $a \in A$ defines a function $B \to C$ which is S-linear. This implies that s is surjective. $\square$

Suppose $\mathfrak{a} \subset R$ is an ideal and M is a module over R. Then it is easy to see that $\mathfrak{a} \cdot M \subset M$ is a submodule and we have

PROPOSITION A.5.52. *If $M \in \mathscr{M}_R$ and $\mathfrak{a} \subset R$ is an ideal, there exists a natural isomorphism:*

$$
\begin{aligned}
q \colon M \otimes_R \left(\frac{R}{\mathfrak{a}}\right) &\to \frac{M}{\mathfrak{a} \cdot M} \\
m \otimes r &\mapsto R \cdot m \pmod{\mathfrak{a} \cdot M}
\end{aligned}
$$

PROOF. It is not hard to see that q is surjective. Consider the composite

$$M = M \otimes_R R \xrightarrow{1 \otimes p} M \otimes_R \left(\frac{R}{\mathfrak{a}}\right)$$

where $p \colon R \to R/\mathfrak{a}$ is the projection. The surjectivity of q implies that $\ker 1 \otimes p \subset \mathfrak{a} \cdot M$. On the other hand, if $x \in \mathfrak{a}$, $x \cdot m \otimes 1 \sim m \otimes x \cdot 1 = 0 \in M \otimes_R (R/\mathfrak{a})$, by rule 1 in definition A.5.44 on page 462. This shows that $\mathfrak{a} \cdot M \subset \ker 1 \otimes p$ and that q is also *injective.* $\square$

We can use tensor-products to convert modules over one ring into modules over another:

PROPOSITION A.5.53. *Let M be a module over a commutative ring R and let $f \colon R \to S$ be a homomorphism of rings. Then S is a module over R and*

$$M \otimes_R S$$

is a module over S with S-action given by

$$t \cdot m \otimes s = m \otimes st$$

for all $s, t \in S$ and $m \in M$.

REMARK. This operation is called a change of base. If $R \hookrightarrow S$ is an inclusion, it is called extension of scalars, the idea being that the action of R on M is "extended" to the larger ring, S.

Recall the concept of a module of fractions, defined in section A.1.8 on page 383.

PROPOSITION A.5.54. *If $M \in \mathscr{M}_R$ and $S \subset R$ is a multiplicative set, then*

$$S^{-1}M \cong M \otimes_R (S^{-1}R)$$

is a module over $S^{-1}R$. If $\mathfrak{p} \subset R$ is a prime ideal and $S = R \setminus \mathfrak{p}$, then $S^{-1}R = R_{\mathfrak{p}}$ and

$$M_{\mathfrak{p}} \cong M \otimes_R R_{\mathfrak{p}}$$

and is a module over $R_{\mathfrak{p}}$.

REMARK. As defined in definition A.1.89 on page 383, $S^{-1}M$ is a module over R. Proposition A.5.53 on the preceding page shows that $S^{-1}M$ is *also* a module over $S^{-1}R$. If $S^{-1}R$ is the field of fractions, then $S^{-1}M$ is a *vector space* over that field.

PROOF. The map

$$f: M \otimes_R (S^{-1}R) \to S^{-1}M$$

is defined by $f(m \otimes s^{-1}r) = r \cdot m/s$ for all $m \in M$, $s \in S$, and $r \in R$. If $r_1/s_1 \equiv r_2/s_2 \in S^{-1}R$, then $u \cdot (s_2 r_1 - s_1 r_2) = 0$ for some $u \in S$, and

$$u \cdot (s_2 r_1 \cdot m - s_1 r_2 \cdot m) = u \cdot (s_2 r_1 - s_1 r_2) \cdot m = 0$$

so f is well-defined. The inverse map

$$g: S^{-1}M \to M \otimes_R (S^{-1}R)$$

is defined by $g(m/s) = m \otimes s^{-1}$. If $m_1/s_1 \equiv m_2/s_2 \in S^{-1}M$ then

$$u \cdot (s_2 \cdot m_1 - s_1 \cdot m_2) = 0$$

for some $u \in S$, or $us_2 \cdot m_1 = us_1 \cdot m_2$, so

$$us_2 \cdot m_1 \otimes u^{-1}s_1^{-1}s_2^{-1} = us_1 \cdot m_2 \otimes u^{-1}s_1^{-1}s_2^{-1}$$

By rule 1 of definition A.5.44 on page 462, both sides of this equation are equal to

$$us_2 \cdot m_1 \otimes u^{-1}s_1^{-1}s_2^{-1} = m_1 \otimes s_1^{-1}$$
$$us_1 \cdot m_2 \otimes u^{-1}s_1^{-1}s_2^{-1} = m_2 \otimes s_2^{-1}$$

It follows that $g(m_1/s_1) = g(m_2/s_2)$, so g is well-defined and clearly the inverse of f. $\square$

It is easy to verify that tensor products preserve *surjectivity* of maps:

PROPOSITION A.5.55. *If $M \in \mathscr{M}_R$ and $f: N \to T$ is a surjective morphism in $\mathscr{M}_R$, then*

$$1 \otimes f: M \otimes_R N \to M \otimes_R T$$

is also surjective.

REMARK. If

$$0 \to A \to B \to C \to 0$$

is an exact sequence of modules and we take the tensor product of this with M, the resulting sequence is exact on the *right*

$$M \otimes_R A \to M \otimes_R B \to M \otimes_R C \to 0$$

and we say that the *functor $M \otimes_R *$ is *right-exact*. This sequence might not be exact on the left — $M \otimes_R A \to M \otimes_R B$ might not be an inclusion. For instance if

$$f = \times 2: \mathbb{Z} \to \mathbb{Z}$$

and $M = \mathbb{Z}_2$, then

$$f \otimes 1 = 0 \colon \mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}_2 = \mathbb{Z}_2 \to \mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}_2 = \mathbb{Z}_2$$

PROOF. If $\sum m_i \otimes t_i \in M \otimes_R T$ then it is the image of $\sum m_i \otimes n_i$, where $f(n_i) = t_i$. $\square$

This leads to another consequence of Nakayama's Lemma:

COROLLARY A.5.56. *Let R be a noetherian local ring with maximal ideal $\mathfrak{m}$, let M be a finitely generated R-module, and let*

$$f \colon M \to \frac{M}{\mathfrak{m} \cdot M} = M \otimes_R \left(\frac{R}{\mathfrak{m}} \right)$$

be the projection to the quotient. If $\{m_1, \ldots m_t\} \in M$ are elements with the property that $\{f(m_1), \ldots, f(m_t)\}$ generate $M/\mathfrak{m} \cdot M$, then $\{m_1, \ldots m_t\}$ generate M.

REMARK. Note that $R/\mathfrak{m}$ is a field so that $M/\mathfrak{m} \cdot M$ is a vector space.

PROOF. Let $M' \subset M$ be the submodule generated by $\{m_1, \ldots m_t\}$. Then M/M' is a finitely generated R module with the property that

$$\left(\frac{M}{M'} \right) \otimes_R \left(\frac{R}{\mathfrak{m}} \right) = 0$$

which implies that

$$\mathfrak{m} \cdot \left(\frac{M}{M'} \right) = \left(\frac{M}{M'} \right)$$

Corollary A.1.82 on page 379 implies that $M/M' = 0$. $\square$

We also get the interesting result

COROLLARY A.5.57. *If R is a noetherian local ring with maximal ideal $\mathfrak{m}$, then finitely generated projective modules over R are free.*

PROOF. Let P be a projective module over R and let $p_1, \ldots, p_n \in P$ be a set of elements with the property that their image in

$$P \otimes_R \left(\frac{R}{\mathfrak{m}} \right) = \frac{P}{\mathfrak{m} \cdot P} = V$$

generate the vector-space V. Corollary A.5.56 implies that $p_1, \ldots, p_n$ generate P. If R^n is a free module on generators $x_1, \ldots, x_n$, the homomorphism

$$f \colon R^n \to P$$

that sends x_i to p_i for $i = 1, \ldots, n$ is surjective. If $K = \ker f$, we get a short exact sequence

$$0 \to K \to R^n \xrightarrow{f} P \to 0$$

Since P is projective, this is split and we get an isomorphism

$$R^n \cong K \oplus P$$

(see exercise 42 on page 380). Now take the tensor product with $R/\mathfrak{m}$ to get

$$R^n \otimes_R \left(\frac{R}{\mathfrak{m}} \right) \cong P \otimes_R \left(\frac{R}{\mathfrak{m}} \right) \oplus K \otimes_R \left(\frac{R}{\mathfrak{m}} \right)$$

Since $R^n \otimes_R \left(\frac{R}{\mathfrak{m}}\right)$ and $P \otimes_R \left(\frac{R}{\mathfrak{m}}\right)$ are both n-dimensional vector-spaces over $R/\mathfrak{m}$, it follows that

$$K \otimes_R \left(\frac{R}{\mathfrak{m}}\right) = \frac{K}{\mathfrak{m} \cdot K} = 0$$

This implies that $K = \mathfrak{m} \cdot K$ and corollary A.1.82 on page 379 implies that $K = 0$ and $P = R^n$. $\qquad\square$

DEFINITION A.5.58. A module $M \in \mathscr{M}_R$ will be called *flat* if the functor $M \otimes_R *$ preserves *injections* as well as surjections. In other words, M is flat if, whenever

$$N \to T$$

is an injective homomorphism of R-modules, so is

$$M \otimes_R N \to M \otimes T$$

REMARK. For instance, R is a flat module over itself, as example A.5.45 on page 462 shows. In general, every free module, R^n, is flat over R, by proposition A.5.48 on page 464.

The term *flat module* first appeared in Serre's paper, [**152**].

Flat modules are very useful because:

PROPOSITION A.5.59. *Let R be a commutative ring and let A be a flat R-module. If*

$$\cdots \xrightarrow{f_{n+1}} M_{n+1} \xrightarrow{f_n} M_n \xrightarrow{f_{n-1}} M_{n-1} \to \cdots$$

is an exact sequence in $\mathscr{M}_R$, then so is

$$\cdots \xrightarrow{1 \otimes f_{n+1}} A \otimes_R M_{n+1} \xrightarrow{1 \otimes f_n} A \otimes_R M_n \xrightarrow{1 \otimes f_{n-1}} A \otimes_R M_{n-1} \to \cdots$$

PROOF. The exactness of the (long) sequence above is equivalent to saying that the short sequences

$$0 \to \operatorname{im} f_n \to M_n \to \operatorname{im} f_{n-1} \to 0$$

are exact (see definition A.1.64 on page 369) for all n. Since tensor products preserve surjections (proposition A.5.55 on page 467), we know that $\operatorname{im}(1 \otimes f_n) = A \otimes_R (\operatorname{im} f_n)$ for all n (and this is for *any* module, A, not just a flat one). The conclusion follows by the fact that flat modules preserve *injections* as well as surjections (definition A.5.58). $\qquad\square$

The following result describes a very important class of flat modules:

LEMMA A.5.60. *Let R be a commutative ring and let S be a multiplicative set. Then $S^{-1}R$ is a flat module over R.*

PROOF. If $f: N \to T$ is an injective homomorphism of R-modules, we will show that

$$S^{-1}R \otimes_R N \to S^{-1}R \otimes_R T$$

is also injective. We replace these tensor products by modules of fractions, using proposition A.5.54 on page 466, to get the equivalent map

$$S^{-1}N \to S^{-1}T$$

Extend this to

$$N \to S^{-1}N \to S^{-1}T$$

An element $n \in N$ maps to zero in $S^{-1}T$ if and only if $s \cdot n = 0$ for some $s \in S$ (see definition A.1.89 on page 383). If this happens, n also maps to 0 in $S^{-1}N$ so the map is injective. $\qquad\square$

We know that free modules are flat by proposition A.5.48 on page 464. It turns out that

PROPOSITION A.5.61. *Projective modules are flat.*

REMARK. Projective modules over $\mathbb{Z}$ have elements that are *not* multiples of 2. On the other hand, lemma A.5.60 on the previous page shows that $\mathbb{Q}$ is a flat module over $\mathbb{Z}$ that *cannot* be projective since *all* of its elements are divisible by 2.

PROOF. Let P be a projective module over a ring R and let Q be another (projective) module such that $P \oplus Q = R^n$. If

$$f \colon M \to N$$

is an injective homomorphism, we know that

$$f \otimes 1 \colon M \otimes_R R^n \to N \otimes_R R^n$$

is also injective, and is equal to

$$f \otimes 1 \colon M \otimes_R (P \oplus Q) \to N \otimes_R (P \oplus Q)$$

which is equal to

$$(f \otimes 1_P) \oplus (f \otimes 1_Q) \colon M \otimes_R P \oplus M \otimes_R Q \to N \otimes_R P \oplus N \otimes_R Q$$

where 1_P and 1_Q are the identity maps of P and Q, respectively. Since $(f \otimes 1_P) \oplus (f \otimes 1_Q)$ is injective, proposition A.5.43 on page 462 implies that $f \otimes 1_P$ and $f \otimes 1_Q$ must be injective too. $\qquad\square$

It is interesting to see what forming modules of fractions does to prime filtrations:

COROLLARY A.5.62. *Let M be a module over a ring R and let $S \subset R$ be a multiplicative set. Let*

$$0 = M_0 \subset M_1 \subset \cdots \subset M_n = M$$

with

$$\frac{M_{i+1}}{M_i} \cong \frac{R}{\mathfrak{p}_i}$$

for prime ideals $\mathfrak{p}_i \subset R$, be the prime filtration of M. Then the prime filtration of $S^{-1}R \otimes_R M$ is

$$0 = S^{-1}R \otimes_R M_{j_0} \subset S^{-1}R \otimes_R M_{j_1} \subset \cdots \subset S^{-1}R \otimes_R M_{j_t} = S^{-1}R \otimes_R M$$

where

$$\frac{S^{-1}R \otimes_R M_{j_{i+1}}}{S^{-1}R \otimes_R M_{j_i}} \cong \frac{S^{-1}R}{\mathfrak{p}_{j_i} \cdot S^{-1}R}$$

where $\{\mathfrak{p}_{j_0}, \ldots, \mathfrak{p}_{j_t}\} \subseteq \{\mathfrak{p}_0, \ldots, \mathfrak{p}_n\}$ is the subset of prime ideals that do not contain any elements of S.

PROOF. Since $S^{-1}R$ is flat over R, every short exact sequence

$$0 \to \frac{R}{\mathfrak{p}_i} \to M_i \to M_{i+1} \to 0$$

gives rise to

$$0 \to S^{-1}R \otimes_R \left(\frac{R}{\mathfrak{p}_i}\right) \to S^{-1}R \otimes_R M_i \to S^{-1}R \otimes_R M_{i+1} \to 0$$

and the short exact sequence

$$0 \to S^{-1}R \otimes_R \mathfrak{p}_i \to S^{-1}R \otimes_R R \to S^{-1}R \otimes_R \left(\frac{R}{\mathfrak{p}_i}\right) \to 0$$

where $S^{-1}R \otimes_R \mathfrak{p}_i = \mathfrak{p}_i \cdot S^{-1}R$ and $S^{-1}R \otimes_R R = S^{-1}R$. If $\mathfrak{p}_i$ contains an element of S, $\mathfrak{p}_i \cdot S^{-1}R = S^{-1}R$ and the quotient

$$S^{-1}R \otimes_R \left(\frac{R}{\mathfrak{p}_i}\right)$$

will be the trivial ring. It follows that those primes do not participate in the prime filtration of $S^{-1}R \otimes_R M$. □

We conclude this section with a converse to lemma A.5.60 on page 469:

LEMMA A.5.63. *Let R be a noetherian ring and let A be a finitely-generated R-module. Then $A_{\mathfrak{m}}$ is a free $R_{\mathfrak{m}}$-module for all maximal ideals $\mathfrak{m} \subset R$ if and only if A is projective.*

REMARK. In other words, locally free modules are *projective.*

PROOF. Since A is finitely-generated, there exists a finitely-generated free module, F, and a surjective homomorphism

$$f: F \to A$$

inducing surjective homomorphisms

$$f_{\mathfrak{m}} = 1 \otimes f: F_{\mathfrak{m}} \to A_{\mathfrak{m}}$$

Since $A_{\mathfrak{m}}$ is free, there exist splitting maps

$$g_{\mathfrak{m}}: A_{\mathfrak{m}} \to F_{\mathfrak{m}}$$

with $f_{\mathfrak{m}} \circ g_{\mathfrak{m}} = 1: A_{\mathfrak{m}} \to A_{\mathfrak{m}}$ for all maximal ideals $\mathfrak{m} \subset R$. Since A is finitely-generated, there exists an element $s_{\mathfrak{m}} \in R \setminus \mathfrak{m}$ for each maximal ideal such that

$$s_{\mathfrak{m}} \cdot f_{\mathfrak{m}} \circ g_{\mathfrak{m}}(A) \subset A \subset A_{\mathfrak{m}}$$

i.e., $s_{\mathfrak{m}}$ "clears the denominators" of $f_{\mathfrak{m}} \circ g_{\mathfrak{m}}(A)$. Let $\mathfrak{S}$ denote the ideal generated by all of the $s_{\mathfrak{m}}$. Since R is noetherian, $\mathfrak{S}$ is generated by some finite set of the $s_{\mathfrak{m}}$

$$\mathfrak{S} = (s_{\mathfrak{m}_1}, \ldots, s_{\mathfrak{m}_t})$$

If $\mathfrak{S} \subsetneq R$, then it is contained in some maximal ideal, which contradicts the fact that it contains an element not in every maximal ideal. We conclude that $\mathfrak{S} = R$ and that there exist elements $\{x_1, \ldots, x_t\}$ such that

$$\sum_{i=1}^{t} x_i \cdot s_{\mathfrak{m}_i} = 1$$

If we define

$$g = \sum_{i=1}^{t} x_i \cdot f_{\mathfrak{m}_i} \circ g_{\mathfrak{m}_i} : A \to F$$

and $f \circ g = 1 : A \to A$, so A is a direct summand of F and is projective.

The *only if* part comes from corollary A.5.57 on page 468.   □

We conclude this section with a generalization of the *dual* of a module:

DEFINITION A.5.64. If M is a module over a ring, R, define $M^* = \hom_R(M, R)$ — the dual of M. It is also a module over R (just let R act on it by multiplying the values of homomorphisms).

REMARK. Clearly, $R^* = R$ since a homomorphism $R \to R$ is determined by its effect on $1 \in R$. It is also not hard to see that the dual of a finitely generated free module is free of the same rank. If F is a free module, the isomorphism between F and F^* is not natural.

There *is* a natural isomorphism

$$F \to F^{**}$$

where we map $x \in F$ to the homomorphism $F^* \to R$ given by $f(x)$ for $f \in F^*$.

This and the way hom behaves with direct sums implies that:

COROLLARY A.5.65. *Let P be a finitely generated projective module over a commutative ring R. Then P^* is also a finitely generated projective module and*

$$P = P^{**}$$

EXERCISES.

13. Suppose M is a module over a ring, R, and $N \subset M$ is a submodule. If A is a flat module over R, show that $A \otimes_R N$ is a submodule of $A \otimes_R M$ and

$$\frac{A \otimes_R M}{A \otimes_R N} = A \otimes_R \left(\frac{M}{N} \right)$$

14. If M, N, T are modules over a ring, R, show that there are natural isomorphisms

$$\begin{aligned} M \otimes_R (N \otimes_R T) &\cong (M \otimes_R N) \otimes_R T \\ M \otimes_R N &\cong N \otimes_R M \end{aligned}$$

15. Let V be a vector space over a field, k, with basis $\{e_1, \ldots, e_n\}$ and let W be a vector space with basis $\{f_1, \ldots, f_n\}$. Show that

$$V \otimes_k W$$

is $n \cdot m$ dimensional, with basis

$$\{e_i \otimes f_j\}$$

$i = 1, \ldots, n$ and $j = 1, \ldots, m$.

Show that $\{e_i \otimes f_j\}$ are a basis for $V \otimes_k W$ even if they are infinite dimensional.

16. Suppose k is a field and $f: k^n \to k^m$ and $g: k^s \to k^t$ are given by $m \times n$ and $t \times s$ matrices A and B, respectively. What is the matrix representing $A \otimes B$?

17. If $\mathfrak{a}, \mathfrak{b} \subset R$ are two ideals in a commutative ring, show that

$$\frac{R}{\mathfrak{a}} \otimes_R \frac{R}{\mathfrak{b}} = \frac{R}{\mathfrak{a} + \mathfrak{b}}$$

This implies that

$$\mathbb{Z}_n \otimes_{\mathbb{Z}} \mathbb{Z}_m = \mathbb{Z}_{\gcd(n,m)}$$

18. If R is a ring and M is a flat R-module. Show that

$$\mathfrak{a} \otimes_R M = \mathfrak{a} \cdot M$$

for all ideals $\mathfrak{a} \subset R$.

19. Show that tensor products commute with direct limits, i.e. if

$$M_0 \xrightarrow{f_0} \cdots \xrightarrow{f_{n-1}} M_n \xrightarrow{f_n} \cdots$$

is a direct system of modules over a ring R and N is an R-modules, show that

$$\left(\varinjlim M_j \right) \otimes_R N = \varinjlim \left(M_j \otimes_R N \right)$$

20. If $\{A_i, a_i\}$, $\{B_i, b_i\}$, and $\{C_i, c_i\}$ are three systems of homomorphisms of modules such that the diagram

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & A_{i+1} & \xrightarrow{f_{i+1}} & B_{i+1} & \xrightarrow{g_{i+1}} & C_{i+1} & \longrightarrow & 0 \\
& & \Big\uparrow a_i & & \Big\uparrow b_i & & \Big\uparrow c_i & & \\
0 & \longrightarrow & A_i & \xrightarrow{f_i} & B_i & \xrightarrow{g_i} & C_i & \longrightarrow & 0
\end{array}
$$

commutes for all i and each row is exact, show that

$$0 \to \varinjlim A_i \xrightarrow{\varinjlim f_i} \varinjlim B_i \xrightarrow{\varinjlim g_i} \varinjlim C_i \to 0$$

is an exact sequence.

21. Suppose $f: R \to S$ is a homomorphism of rings with the property that S is a flat module over R. If $\alpha \in R$ is a non-zero-divisor, show that $f(\alpha) \in S$ is a non-zero-divisor.

22. Suppose $f: R \to S$ is a homomorphism of rings and M is a flat module over R. Show that $M \otimes_R S$ is a flat module over S.

23. Let $\mathscr{M}_{\mathbb{Z}}$ be the category of modules over $\mathbb{Z}$ (otherwise known as abelian groups, $\mathscr{A}b$), the set, $\hom_{\mathscr{M}_{\mathbb{Z}}}(A, B)$, is naturally a module over $\mathbb{Z}$. For any $A, B, C \in \mathscr{M}_{\mathbb{Z}}$, show that there exists a natural isomorphism

$$\hom_{\mathscr{M}_{\mathbb{Z}}}(A \otimes_{\mathbb{Z}} B, C) \cong \hom_{\mathscr{M}_{\mathbb{Z}}}(A, \hom_{\mathscr{M}_{\mathbb{Z}}}(B, C))$$

so that the functors $* \otimes_{\mathbb{Z}} B$ and $\hom_{\mathscr{M}_{\mathbb{Z}}}(B, *)$ are adjoints.

24. Let $M, N \in \mathscr{M}_R$ and let $S \subset R$ be a multiplicative set. Show that

$$S^{-1}R \otimes_R (M \otimes_R N) = (S^{-1}R \otimes_R M) \otimes_{S^{-1}R} (S^{-1}R \otimes_R N)$$

25. If M is a finitely generated projective module, show that M^* is also a finitely generated projective module.

A.6. Tensor Algebras and variants

In this section, we will discuss several algebras one can construct from modules over a ring. The most general case is the tensor algebra, with the symmetric and exterior algebras being quotients.

Historically, the first of these to appear were *exterior algebras*, described in [**60**] by Hermann Grassmann. Grassmann developed exterior algebras in the context of vector spaces — and many linear algebra constructs (like determinants) have elegant formulations in terms of exterior algebras[4].

Tensor algebras appeared later, in the context of category theory and are more general than exterior algebras.

DEFINITION A.6.1. If R is a commutative ring and M is an R-module, define:
$$M^{\otimes n} = \underbrace{M \otimes_R \cdots \otimes_R M}_{n \text{ times}}$$
with $M^{\otimes 0} = R$ and $M^{\otimes 1} = M$. Given this definition, we define the *tensor algebra* over M:
$$T(M) = R \oplus M \oplus M^{\otimes 2} \oplus M^{\otimes 3} \oplus \cdots$$

This is a (noncommutative) algebra over R by defining
$$(m_1 \otimes \cdots \otimes m_s) \cdot (n_1 \otimes \cdots \otimes n_t) = m_1 \otimes \cdots \otimes m_s \otimes n_1 \otimes \cdots \otimes n_t$$
and extending this to all of $T(M)$ R-linearly.

REMARK. Tensor algebras are often called free algebras. Any module homomorphism
$$f: M \to N$$
induces a unique algebra-homomorphism
$$T(f): T(M) \to T(N)$$
Furthermore, if A is any algebra over R and $g: M \to A$ is a homomorphism of R-modules, there exists a unique homomorphism of R-algebras
$$T(M) \to A$$
whose restriction to $M = M^{\otimes 1}$ is g. If $\mathscr{A}_R$ is the category of R-algebras and $\mathscr{M}_R$ that of R-modules, let
$$F: \mathscr{A}_R \to \mathscr{M}_R$$
be the forgetful functor that maps an R-algebra to its underlying R-module (forgetting that we can multiply elements of this module), we get a natural isomorphism

(A.6.1) $$\hom_{\mathscr{A}_R}(T(M), A) \cong \hom_R(M, FA)$$

---

[4]Often called Grassmann algebras

making $T(*)$ and F adjoints (compare with example A.5.15 on page 446).

The tensor algebra is an example of a *graded ring* (see definition A.4.39 on page 434) with

$$T(M)_n = M^{\otimes n}$$

Corollary A.5.49 on page 465 immediately implies that

PROPOSITION A.6.2. *If M is a free module of rank t (see example A.1.58 on page 368) over a ring R, then $T_n(M)$ is a free module of rank t^n.*

We also have:

PROPOSITION A.6.3. *If M is any module over a commutative ring, R, and $S \subset R$ is any multiplicative set, then*

$$T(S^{-1}R \otimes_R M) = S^{-1}R \otimes_R T(M)$$

PROOF. This follows immediately from the solution to exercise 24 on page 473. □

There are two important variants on tensor algebras that we need:

DEFINITION A.6.4. Let M be a module over a commutative ring, R, and let $\mathfrak{s} \subset T(M)$ be the (two-sided) ideal generated by elements

$$x \otimes y - y \otimes x$$

for all $x, y \in M$. The quotient, $\mathcal{S}(M) = T(M)/\mathfrak{s}$, is called the *symmetric algebra* on M.

REMARK. This is clearly a commutative ring. Since $T(M)$ is not commutative, the ideal $\mathfrak{s}$ must be two-sided — it is the sum

$$\sum_{x,y \in M} T(M) \cdot (x \otimes y - y \otimes x) \cdot T(M)$$

Symmetric algebras also have a defining universal property:

PROPOSITION A.6.5. *Let $\mathscr{C}_R$ denote the category of commutative algebras over a (commutative) ring R and let $\mathscr{M}_R$ denote the category of R-modules. There is a forgetful functor*

$$f \colon \mathscr{C}_R \to \mathscr{M}_R$$

that "forgets" the multiplication operation in an R-algebra (so it becomes a mere module). The symmetric algebra is an adjoint to f in the sense that

$$\hom_R(M, f(A) = \hom_{\mathscr{C}_R}(\mathcal{S}M, A)$$

PROOF. We already know that

(A.6.2) $$\hom_{\mathscr{A}_R}(T(M), A) \cong \hom_R(M, FA)$$

If A is a *commutative* algebra, then the map

$$T(M) \to A$$

factors through $\mathcal{S}M$:

$$T(M) \to \mathcal{S}M \to A$$

□

It is not hard to see that

PROPOSITION A.6.6. *If M is a free module of rank t over a commutative ring, R, then*

$$\mathcal{S}(M) = R[X_1, \ldots, X_t]$$

PROOF. Suppose $\{e_1, \ldots, e_t\}$ is a free basis for M. It is straightforward to see that

$$e_{j_1}^{\otimes n_1} \otimes \cdot \otimes e_{j_\ell}^{\otimes n_\ell}$$

with $\sum n_i = n$ and $j_1 < \cdots < j_\ell$ is a free basis for $\mathcal{S}_n(M)$ — and these are in a 1-1 correspondence with monomials in the X_i of total degree n.   □

The second variant of tensor algebras is called *exterior algebras or* Grassmann algebras in honor of Hermann Grassmann (since he first described them in [**60**]). For our purposes, they are more interesting than symmetric algebras and have more applications. Although Grassman originally defined them for vector-spaces over fields, this definition can easily be extended to modules over a commutative ring:

DEFINITION A.6.7. If M is a module over a commutative ring, R, the *exterior algebra* over M is defined to be

$$\wedge M = T(M)/\mathfrak{a}$$

where $\mathfrak{a}$ is the two-sided ideal generated by elements $\{x \otimes x\}$ for all $x \in M$. This is a graded ring with

$$\Lambda^n M = M^{\otimes n}/M^{\otimes n} \cap \mathfrak{a}$$

The product-operation is written $x \wedge y$ for $x, y \in \bigwedge M$.

REMARK. If $x, y \in M$, then

$$(x + y) \wedge (x + y) = 0$$

because of how the ideal $\mathfrak{a}$ is defined. The distributive laws implies that

$$
\begin{aligned}
(x + y) \wedge (x + y) &= x \wedge x + x \wedge y + y \wedge x + y \wedge y \\
&= x \wedge y + y \wedge x
\end{aligned}
$$

so $x \wedge y = -y \wedge x$ for elements of M. The level $\Lambda^n M$ is generated by all expressions of the form

$$x_1 \wedge \cdots \wedge x_n$$

for $x_1, \ldots, x_n \in M$.

Exterior algebras have applications to fields as varied as differential geometry (see [**157**]), partial differential equations (see [**20**]) and physics (see [**136**]) — besides algebraic geometry. Grassmann's original definition was *axiomatic,* using axioms based on linearity, associativity, and anti-commutativity — and only for vector-spaces.

We have some direct-sum relations:

PROPOSITION A.6.8. *If M and N are modules over R then*

(1) $T(M \oplus N) = T(M) \otimes_R T(N) \otimes_R T(M) \otimes_R \cdots$ *— as graded algebras (see definition A.4.41 on page 435), i.e.,*

$$T(M \oplus N)_m = \bigoplus_{\sum_{j=1}^{\infty} (i_j + n_j) = m} T(M)_{i_1} \otimes_R T(N)_{n_1} \otimes_R \cdots$$

(2) $\mathcal{S}(M \oplus N) = \mathcal{S}(M) \otimes_R \mathcal{S}(N)$ — *as graded algebras, so*

$$\mathcal{S}(M \oplus N)_m = \bigoplus_{i+j=m} \mathcal{S}(M)_i \otimes_R \mathcal{S}(N)_j$$

(3) $\Lambda(M \oplus N) \cong \Lambda(M) \otimes_R \Lambda(N)$ — *as graded algebras, so*

$$\Lambda^m(M \oplus N) \cong \bigoplus_{i+j=m} \Lambda^i(M) \otimes_R \Lambda^j(N)$$

REMARK. Note that, in line 1 all but a finite number of the i_j, n_j must be 0.

PROOF. The first statement follows from the general properties of the tensor product.

The second statement follows from the first and the fact that the commutativity relations between $T(M)$ and $T(N)$ reduces the "infinite tensor product" to $T(M) \otimes_R T(N)$. Imposing the commutativity relations within $T(M)$ and $T(N)$ gives $\mathcal{S}(M) \otimes_R \mathcal{S}(N)$.

The third statement follows by a similar argument except that we may have to permute factors in an expression like $n_1 \wedge m_2 \wedge \dots, \wedge m_i$ so that all of the m-factors occur to the left of the n-factors. This multiplies by ± 1, so we get an isomorphism. $\square$

Here's an example of computations in an exterior algebra:

EXAMPLE. Let M be a free module over R on a free basis $\{e_1, e_2, e_3\}$ and let $v = 2e_1 + e_2 - e_3$ and $w = e_1 - 3e_2 + e_3$. Then

$$
\begin{aligned}
v \wedge w &= (2e_1 + e_2 - e_3) \wedge (e_1 - 3e_2 + e_3) \\
&= 2e_1 \wedge (e_1 - 3e_2 + e_3) + e_2 \wedge (e_1 - 3e_2 + e_3) \\
&\quad - e_3 \wedge (e_1 - 3e_2 + e_3) \\
&= 2e_1 \wedge e_1 - 2e_1 \wedge 3e_2 + 2e_1 \wedge e_3 \\
&\quad + e_2 \wedge e_1 - 3e_2 \wedge e_2 + e_2 \wedge e_3 \\
&\quad - e_3 \wedge e_1 + 3e_3 \wedge e_2 - e_3 \wedge e_3
\end{aligned}
$$

Here, we have used the distributive rule several times. After applying the annihilation and linearity conditions, we get

$$
\begin{aligned}
v \wedge w &= -6e_1 \wedge e_2 + 2e_1 \wedge e_3 + e_2 \wedge e_1 + e_2 \wedge e_3 \\
&\quad - e_3 \wedge e_1 + 3e_3 \wedge e_2
\end{aligned}
$$

And after "standardizing" by replacing any $e_j \wedge e_i$ by $-e_i \wedge e_j$ whenever $j > i$, we get

$$v \wedge w = -7e_1 \wedge e_2 + 3e_1 \wedge e_3 - 2e_2 \wedge e_3$$

Clearly, the set $\{e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3\}$ forms a free basis for $\Lambda^2 V$ (any relation between them would imply a relation between basis elements of $T_2(M)$).

In general, we have:

PROPOSITION A.6.9. *Let M be a free module over R with free basis $\{e_1, \dots, e_n\}$. Then $\Lambda^k M$ has a free basis consisting of symbols*

$$\{e_{i_1} \wedge \dots \wedge e_{i_k}\}$$

for all sequences $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. *Consequently, the rank of* $\Lambda^k V$ *is* $\binom{n}{k}$, *and* $\bigwedge^k V = 0$ *whenever* $k > n$.

PROOF. By definition, $\Lambda^k V$ consists of all sequences $v_1 \wedge \cdots \wedge v_k$ and, using the linearity and distributivity properties, we can write these as linear combinations of all length-k sequences of basis elements

$$\{e_{j_1} \wedge \cdots \wedge e_{j_k}\}$$

The annihilation property implies that any such sequence with two *equal* indices will vanish. It also implies that we can arrange these indices in ascending order (multiplying terms by -1 if necessary). □

Proposition A.6.3 on page 475 and the fact that $S^{-1}R$ is *flat* over R (see lemma A.5.60 on page 469) imply that

PROPOSITION A.6.10. *Let M be a module over a commutative ring, R, and let $S \subset R$ be a multiplicative set. Then*

$$\bigwedge(S^{-1}R \otimes_R M) = S^{-1}R \otimes_R \bigwedge M$$
$$\mathcal{S}(S^{-1}R \otimes_R M) = S^{-1}R \otimes_R \mathcal{S}(M)$$

PROOF. The fact that $S^{-1}R$ is flat over R implies that

$$S^{-1}R \otimes_R \left(\frac{T(M)}{\mathfrak{a}} \right) = \frac{S^{-1}R \otimes_R T(M)}{S^{-1}R \otimes_R \mathfrak{a}} = \frac{T(S^{-1}R \otimes_R M)}{\mathfrak{a}'}$$

where $\mathfrak{a}'$ is the form of $\mathfrak{a}$ in $T(S^{-1}R \otimes_R M)$. It follows that $\bigwedge(S^{-1}R \otimes_R M) = S^{-1}R \otimes_R \bigwedge M$. The proof for the symmetric algebra is entirely analogous. □

We will often be interested in certain elements of $\bigwedge M$ with an especially simple structure (particularly when we study Grassmannians):

DEFINITION A.6.11. If M is a module over a commutative ring, elements of $\Lambda^k M$ of the form

$$m_1 \wedge \cdots \wedge m_k$$

for $m_i \in M$ will be said to be *decomposable*.

REMARK. An exterior algebra consists of formal linear combinations of decomposable elements. If $x \in \Lambda^k M$ is decomposable then

$$\begin{aligned} x \wedge x &= (m_1 \wedge \cdots \wedge m_k) \wedge (m_1 \wedge \cdots \wedge m_k) \\ &= 0 \end{aligned}$$

because of the annihilation condition. Suppose M is a free module on the free basis $\{e_1, e_2, e_3, e_4\}$ and

$$x = e_1 \wedge e_2 + e_3 \wedge e_4$$

Then

$$x \wedge x = 2e_1 \wedge e_2 \wedge e_3 \wedge e_4 \neq 0$$

so this x is *not* decomposable.

For the rest of this section, we will assume that R is a *field* so that modules over R are vector-spaces. The following result is key to understanding the geometric meaning of $\Lambda^k V$:

LEMMA A.6.12. *Let $v_1, \ldots, v_k \in V$ be vectors in a vector space. Then, in $\Lambda^k V$,*

$$v_1 \wedge \cdots \wedge v_k = 0$$

if and only if the set $\{v_1, \ldots, v_k\}$ is linearly dependent.

PROOF. If they are linearly *independent*, they are part of a basis for V and proposition A.6.9 on page 477 implies that their wedge-product is part of a basis for $\Lambda^k V$, hence nonzero.

Suppose they are linearly dependent and, without loss of generality, suppose

$$v_1 = \sum_{j=2}^{k} a_j v_j$$

Then

$$
\begin{aligned}
v_1 \wedge \cdots \wedge v_k &= \sum_{j=2}^{k} a_j v_j \wedge v_2 \wedge \cdots \wedge v_k \\
&= 0
\end{aligned}
$$

since each term in the sum on the right will have v_j equal to one of the vectors in $v_2 \wedge \cdots \wedge v_k$. $\square$

COROLLARY A.6.13. *Let $W \subset V$ be a k-dimensional subspace with basis $\{w_1, \ldots, w_k\}$. Then the element*

$$\bar{w} = w_1 \wedge \cdots \wedge w_k \in \Lambda^k V$$

determines W uniquely. In fact the kernel of the linear map

$$\bar{w} \wedge *: V \to \Lambda^{k+1} V$$

is precisely W.

REMARK. This gives a kind of geometric interpretation of a wedge-product like $w_1 \wedge \cdots \wedge w_k$: it represents a *k*-dimensional subspace of V, and $\Lambda^k V$ is "all formal linear combinations" of such subspaces.

In three dimensions, the *cross-product* is really a wedge-product in disguise, i.e. $v \times w$ is the wedge-product, $v \wedge w$, that represents the plane spanned by v and w. It "looks like" a vector because in $\mathbb{R}^3$ there is a 1-1 correspondence between planes and normal vectors to those planes. This is a special case of something called *Hodge duality*: if V is n-dimensional, a fixed element $\alpha \neq 0 \in \Lambda^n V$ defines an isomorphism

$$\Lambda^k V^* \to \Lambda^{n-k} V$$

where V^* is the dual of V (see 2 on page 443) — also n-dimensional. See exercise 1 on page 234.

PROOF. Lemma A.6.12 implies that, for any $v \in V$, $\bar{w} \wedge v = 0$ if and only if the set of vectors $\{w_1, \ldots, w_k, v\}$ is linearly dependent. Since the set $\{w_1, \ldots, w_k\}$ is linearly independent, it follows that $\bar{w} \wedge v = 0$ if and only if $v \in W$. $\square$

We get a cool way to compute determinants:

LEMMA A.6.14. *Suppose V is a vector space with basis $\{e_1, \ldots, e_n\}$ and A is an $n \times n$ matrix. If the columns of A are vectors $\{v_1, \ldots, v_n\}$ then*

$$v_1 \wedge \cdots \wedge v_n = \det A \cdot e_1 \wedge \cdots \wedge e_n$$

PROOF. We do induction on n. If $n = 1$, there is nothing to prove. Suppose the result is true for $(n-1) \times (n-1)$ matrices and $n-1$-dimensional vector spaces, and we are computing

$$v_1 \wedge \cdots \wedge v_n$$

Let $v = \sum_{i=1}^n a_i \cdot e_i$ and plug this into the formula. We get

$$v_1 \wedge \cdots \wedge v_n = \sum_{i=1}^n a_i \cdot e_i \wedge v_2 \wedge \cdots \wedge v_n$$

Consider the ith term of this, $a_i \cdot e_i \wedge v_2 \wedge \cdots \wedge v_n$. The vectors in $v_2 \wedge \cdots \wedge v_n$ will also be linear combinations of the e_j but the presence of e_i in the wedge product will annihilate all of their terms containing e_i, i.e.

$$a_i \cdot e_i \wedge v_2 \wedge \cdots \wedge v_n = a_i e_i \wedge v_2' \wedge \cdots \wedge v_n'$$

where $v_j' = v_j - $ (its ith component). In other words, v_j' will be a vector in an $(n-1)$-dimensional vector space that is the result of deleting e_i from V. By induction, we get

$$v_2' \wedge \cdots \wedge v_n' = \det A_{(i,1)} \cdot e_1 \wedge \cdots \wedge e_{i-1} \wedge e_{i+1} \wedge \cdots \wedge e_n$$

where $A_{(i,1)}$ is the sub-matrix of A one gets by deleting the first column and the ith row — i.e., the ith *minor* (expanding the determinant using minors from the first column). We get

$$a_i \cdot e_i \wedge v_2 \wedge \cdots \wedge v_n = a_i \det A_{(i,1)} e_i \wedge e_1 \wedge \cdots \wedge e_{i-1} \wedge e_{i+1} \wedge \cdots \wedge e_n$$

Shifting e_i into its proper place multiplies this by $(-1)^{i+1}$ so we get

$$a_i \cdot e_i \wedge v_2 \wedge \cdots \wedge v_n = (-1)^{i+1} a_i \det A_{(i,1)} \cdot e_1 \wedge \cdots \wedge e_n$$

and

$$\begin{aligned} v_1 \wedge \cdots \wedge v_n &= \left(\sum_{i=1}^n (-1)^{i+1} a_i \det A_{(i,1)} \right) \cdot e_1 \wedge \cdots \wedge e_n \\ &= \det A \cdot e_1 \wedge \cdots \wedge e_n \end{aligned}$$

$\square$

COROLLARY A.6.15. *Let V be an n-dimensional vector space with k-dimensional subspace W, and suppose*

$$\{b_1, \ldots, b_k\}$$

is a basis for W. If

$$A \colon W \to W$$

is a change of basis, to a basis

$$\{c_1, \ldots, c_k\}$$

then

$$c_1 \wedge \cdots \wedge c_k = \det A \cdot b_1 \wedge \cdots \wedge b_k$$

PROOF. Extend the bases for W to bases for all of V, i.e.

$$\{b_1, \ldots, b_k, e_{k+1}, \ldots, e_n\}$$

and

$$\{c_1, \ldots, c_k, e_{k+1}, \ldots, e_n\}$$

The change of basis can be represented by an $n \times n$ matrix that is A extended by the identity matrix, i. e.,

$$A' = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}$$

Lemma A.6.14 on the preceding page implies that

$$c_1 \wedge \cdots \wedge c_k \wedge e_{k+1} \wedge \cdots \wedge e_n$$
$$= \det A' \cdot b_1 \wedge \cdots \wedge b_k \wedge e_{k+1} \wedge \cdots \wedge e_n$$
$$= \det A \cdot b_1 \wedge \cdots \wedge b_k \wedge e_{k+1} \wedge \cdots \wedge e_n$$

so

$$(c_1 \wedge \cdots \wedge c_k - \det A \cdot b_1 \wedge \cdots \wedge b_k) \wedge e_{k+1} \wedge \cdots \wedge e_n = 0$$

The conclusion follows from lemma A.6.12 on page 479 since

$$x = c_1 \wedge \cdots \wedge c_k - \det A \cdot b_1 \wedge \cdots \wedge b_k$$

is not in the span of $z = e_{k+1} \wedge \cdots \wedge e_n$ so that $x \wedge z = 0$ implies $x = 0$.   □

EXERCISES.

1. If

$$0 \to U \xrightarrow{f} V \xrightarrow{g} W \to 0$$

is an exact sequence of k-vector-spaces of dimensions, respectively, u, v, w, show that

$$\Lambda^v V \cong \Lambda^u U \otimes_k \Lambda^w W$$

and if the diagram

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & U_1 & \xrightarrow{f_1} & V_1 & \xrightarrow{g_1} & W_1 & \longrightarrow & 0 \\
& & \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} & & \\
0 & \longrightarrow & U_2 & \xrightarrow[f_2]{} & V_2 & \xrightarrow[g_2]{} & W_2 & \longrightarrow & 0
\end{array}
$$

commutes and columns that are isomorphisms, then the diagram

$$
\begin{array}{ccc}
\Lambda^v V_1 & \xrightarrow{\cong} & \Lambda^u U_1 \otimes_k \Lambda^w W_1 \\
{\scriptstyle \Lambda^w b} \downarrow & & \downarrow {\scriptstyle \Lambda^u a \otimes_k \Lambda^w c} \\
\Lambda^v V_2 & \xrightarrow[\cong]{} & \Lambda^u U_2 \otimes_k \Lambda^w W_2
\end{array}
$$

also commutes (so it is natural with respect to isomorphisms of exact sequences).

2. If V is 3-dimensional with basis $\{e_1, e_2, e_3\}$, compute
$$(2e_1 + 3e_2 - e_3) \wedge (e_1 - e_2 + e_3)$$

3. Compute the determinant of
$$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ 2 & 0 & 0 & -1 \end{bmatrix}$$

using exterior products.

A.7. The module of Kähler differentials

Given an algebra, A, over a field (or even over a ring) we can define the module of Kähler differentials — a module over A. In a manner of speaking, this module describes how one could define differentials over the algebra. It has important geometric applications (see section 3.5 on page 154) and gives criteria for elements of a field to be algebraically independent.

We begin with a definition:

DEFINITION A.7.1. Let R be a ring, and let M be an R-module. Then we define the *trivial extension of R by M*, denoted $R \star M$, by

(1) as abelian groups, $R \star M = R \oplus M$,
(2) with a ring-operation
$$(r_1, m_1)(r_2, m_2) = (r_1 r_2, r_2 m_1 + r_1 m_2)$$
for all $r_1, r_2 \in R$ and $m_1, m_2 \in M$.

We clearly have a canonical inclusion $i: M \to R \star M$ sending $m \in M$ to $(0, m) \in R \star M$, and a projection $p: R \star M \to R$ that sends (r, m) to r.

REMARK. Note that $i(M) \subset R \star M$ is an ideal with the property that $i(M)^2 = 0$.

begin with:

DEFINITION A.7.2. Let R be a k-algebra, for some ring k and let M be an R-module. A map
$$f: R \to M$$
is called a *derivation over k* if

(1) it is a k-linear in the sense that $f(x \cdot y) = x \cdot f(y)$ for all $x \in k, y \in M$.
(2) it satisfies the condition
$$f(r_1 \cdot r_2) = r_1 \cdot f(r_2) + r_2 \cdot f(r_1)$$
for all $r_1, r_2 \in R$.

The set of all derivations over k from R to M is denoted $\text{Der}_k(R, M)$.

REMARK A.7.3. Note that, for any derivation $D: R \to M$, that $D^{-1}(0)$ is a subring of R. In particular, the fact that $1^2 = 1$ implies that $D(1) = 0$. If R is a field, then so is $D^{-1}(0)$. It also follows that $D(c) = 0$ for any $c \in k$.

LEMMA A.7.4. *Let A and B be rings and let $\mathfrak{N} \subset B$ be an ideal such that $\mathfrak{N}^2 = 0$. Let $p: B \to B/\mathfrak{N}$ be the natural map, and let $u, v: A \to B$ be homomorphisms such that $p \circ u = p \circ v$ so that $\mathrm{im}\,(u - v) \subset \mathfrak{N}$.*

Then $D = u - v$ is a derivation, $D: A \to \mathfrak{N}$.

Conversely, if $f: A \to B$ is any homomorphism and $D: A \to \mathfrak{N}$ is any derivation, then $f + D: A \to B$ is also a homomorphism.

If M is an A-module and $f, g: R \to A \star M$ are two homomorphisms whose composite with the canonical projection

$$p: A \star M \to A$$

are the same, then

$$f - g: A \to M$$

is a derivation.

PROOF. The homomorphisms u and v induce the same A-module structure on $\mathfrak{N}$ because

$$(u - v)(x) \cdot y \in \mathfrak{N}^2 = 0$$

for all $x \in A, y \in \mathfrak{N}$. Set

$$
\begin{aligned}
u(xy) = u(x)u(y) &= (v(x) + D(x))\,(v(y) + D(y)) \\
&= v(xy) + v(x)D(y) + v(y)D(x) \\
&= v(xy) + x \cdot D(y) + y \cdot D(x)
\end{aligned}
$$

for the A-module structure on $\mathfrak{N}$ induced any homomorphism like u or v. The condition that $\mathfrak{N}^2 = 0$ implies that $D(x)D(y) = 0$. $\square$

Next, we need the concept of Kähler differentials:

DEFINITION A.7.5. *If A is a k-algebra, the module of Kähler differentials, $\Omega_{A/k}$, is the A-module generated by symbols da for $a \in A$ and subject to the relations*

(1) $d(c \cdot a) = c \cdot da$ for all $a \in A$ and $c \in k$,
(2) $d(a_1 + a_2) = da_1 + da_2$ for all $a_1, a_2 \in A$
(3) $d(a_1 \cdot a_2) = a_1 \cdot da_2 + a_2 \cdot da_1$

REMARK. The module of *Kähler differentials* allows us to describe *all* of the derivations that can exist for a ring.

In algebraic varieties, A will be an conventional algebra (i.e., a vector-space) over a *field, k*. When discussing general schemes, k and A are arbitrary commutative rings.

> Erich Kähler (1906–2001) was a German mathematician who contributed to many different fields, including mathematical physics (the n-body problem), algebraic and differential geometry, and the theory of differential equations.

DEFINITION A.7.6. *If A is a k-algebra, then the map*

$$u: A \quad \to \quad \Omega_{A/k}$$

defined by $u(a) = da$ for all $a \in A$, is called the universal derivation of A, and may be denoted $d_{A/k}$.

This derivation is universal because:

PROPOSITION A.7.7. *If $f: A \to M$ is a derivation, then there exists a unique homomorphism of A-modules*

$$g: \Omega_{A/k} \to M$$

that makes the diagram

(A.7.1)
$$A \xrightarrow{\quad u \quad} \Omega_{A/k}$$

commute.

REMARK. This result immediately implies that $\mathrm{Der}_k(A, M) \cong \hom_A(\Omega_{A/k}, M)$. This correspondence is functorial and even an *isomorphism of functors*

(A.7.2) $\mathrm{Der}_k(A, *) \cong \hom_A(\Omega_{A/k}, *)$

PROOF. Since $\Omega_{A/k}$ is generated by the symbols da, it suffices to define g on them. The *only* way to do this that is compatible with f is to set $g(da) = f(a)$. We must verify that the relations in $\Omega_{A/k}$ are satisfied in the image of g. This follows from condition 2 in definition A.7.2 on page 482.   □

EXAMPLE A.7.8. If k is a ring and $R = k[X_1, \dots, X_n]$, we claim that

$$\Omega_{R/k} = R \cdot dX_1 \oplus \cdots \oplus R \cdot dX_n$$

— the free module of rank n generated by the $\{dX_i\}$. The way to see this is to note that each of the partial derivatives $\{\partial/\partial X_i\}$ is a derivation, $\partial_i \colon R \to R$ — so it induces a homomorphism

$$d_i \colon \Omega_{R/k} \to R$$

We get a homomorphism of R-modules

(A.7.3)
$$\begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} \colon \Omega_{R/k} \to R^n$$

This is clearly surjective and its kernel consists of *constants* (i.e., polynomials whose derivatives are all 0). If $c \in k$ and $r \in R$, we claim $dc = 0$ since

$$\begin{aligned} d(c \cdot r) &= c \cdot dr & \text{because of rule 1 in definition A.7.5} \\ &= c \cdot dr + r \cdot dc & \text{because of rule 3 in definition A.7.5} \end{aligned}$$

so $r \cdot dc = 0$ for all $r \in R$ including 1. It follows that the kernel of the map in equation A.7.3 is 0.

If $f \in k[X_1, \dots, X_n]$ then

(A.7.4) $df = \dfrac{\partial f}{\partial X_1} \cdot dX_1 + \cdots + \dfrac{\partial f}{\partial X_n} \cdot dX_n$

The module of Kähler differentials determines a great deal of information about a ring. If V is an irreducible affine variety, corollary A.7.17 on page 491 shows that $\Omega_{k[V]/k}$ determines the *dimension* of a V, and corollary A.7.19 on page 492 shows that it determines the *tangent space* (technically, the cotangent space) at each point of V. Finally, theorem 3.5.12 on page 158 proves that it defines the cotangent *bundle,* under the Serre correspondence (theorem 3.5.9 on page 157).

In order to prove these statements, we need to develop some properties of Kähler differentials. We begin by showing that they are well-behaved with respect to localization:

PROPOSITION A.7.9. *Given a commutative square*

$$\begin{array}{ccc} R & \longrightarrow & R' \\ \uparrow & & \uparrow \\ k & \longrightarrow & k' \end{array}$$

of rings and homomorphisms, there is a natural homomorphism of R-modules

$$\Omega_{R/k} \to \Omega_{R'/k'}$$

hence also a natural homomorphism of R'-modules

$$h: \Omega_{R/k} \otimes_R R' \to \Omega_{R'/k'}$$

If $R' = R \otimes_k k'$ then h is an isomorphism, so

$$\Omega_{R'/k'} = \Omega_{R/k} \otimes_k k' = \Omega_{R/k} \otimes_R R'$$

If $k \subset R$ is an inclusion of rings and $S \subset R$ is a multiplicative set, then there exists a natural isomorphism

(A.7.5) $$e: \Omega_{S^{-1}R/k} \to S^{-1}\Omega_{R/k} = S^{-1}R \otimes_R \Omega_{R/k}$$

defined by

$$e\left(d\left(\frac{r}{s}\right)\right) = \frac{s \cdot dr - r \cdot ds}{s^2}$$

REMARK. See definition A.5.44 on page 462. If $K = k(X_1, \ldots, X_n)$, this and example A.7.8 on the facing page imply that $\Omega_{K/k}$ is a vector-space of dimension n with basis $dX_1, \ldots, dX_n$ — just regard K as a localization of the polynomial ring.

PROOF. The composite

$$R \to R' \xrightarrow{u'} \Omega_{R'/k'}$$

is a derivation (where u' is the universal derivation for $\Omega_{R'/k'}$), so it induces a unique homomorphism

$$g: \Omega_{R/k} \to \Omega_{R'/k'}$$

of R-modules. Since $\Omega_{R'/k'}$ is an R'-module with action of R induced by the homomorphism $R \to R'$, we get a homomorphism of R'-modules

$$h: \Omega_{R/k} \otimes_R R' \to \Omega_{R'/k'}$$

If $R' = R \otimes_k k'$, then $\Omega_{R/k} \otimes_R R' = \Omega_{R/k} \otimes_R R \otimes_k k' = \Omega_{R/k} \otimes_k k'$ so h is a homomorphism

$$h: \Omega_{R/k} \otimes_k k' \to \Omega_{R'/k'}$$

and the universal derivation

$$u: R \to \Omega_{R/k}$$

induces a derivation

$$R' = R \otimes_k k' \to \Omega_{R/k} \otimes_k k' = \Omega_{R/k} \otimes_R R'$$

which induces a unique homomorphism of R'-modules

$$\Omega_{R'/k'} \to \Omega_{R/k} \otimes_k k'$$

which is the inverse of h above. It follows that h is an isomorphism. □

Lemma A.7.4 on page 483 suggests a way to compute modules of Kähler differentials:

THEOREM A.7.10. *Let k be a ring, let R be a k-module, let*

$$\mu\colon R \otimes_k R \to R$$

be the multiplication homomorphism, and let

$$\mathfrak{M} = \ker \mu$$

Let

$$\lambda\colon R \to R \otimes_k R$$
$$r \mapsto r \otimes 1$$

for all $r \in R$. This defines an R-module structure on $R \otimes_k R$. Then there exists a natural R-module-isomorphism

$$\alpha\colon \Omega_{R/k} \to \mathfrak{M}/\mathfrak{M}^2$$
$$dx \mapsto (1 \otimes x - x \otimes 1) \pmod{\mathfrak{M}^2}$$

for all $x \in R$.

PROOF. We claim that

$$(\mu, 1 - \lambda \circ \mu)\colon R \otimes_k R \to R \oplus \mathfrak{M}$$

$$\lambda + 1_{\mathfrak{M}}\colon R \oplus \mathfrak{M} \to R \otimes_k R$$

are inverse isomorphisms of R-modules. We factor out $\mathfrak{M}^2$ to get

(A.7.6) $$R \otimes_k R \xrightarrow{\nu} \frac{R \otimes_k R}{\mathfrak{M}^2} \xrightarrow{\cong} R \oplus \frac{\mathfrak{M}}{\mathfrak{M}^2} = R \star \frac{\mathfrak{M}}{\mathfrak{M}^2} \xrightarrow{p} R$$

Define

$$\bar{\lambda}\colon R \to R \otimes_k R$$
$$r \mapsto 1 \otimes r$$

$$\delta = \bar{\lambda} - \lambda\colon R \to \mathfrak{M}$$
$$r \mapsto 1 \otimes r - r \otimes 1$$

then

$$x \otimes y = xy \otimes 1 + x \cdot (1 \otimes y - y \otimes 1) = \mu(x \otimes y) \cdot 1 + x \cdot \delta(y)$$

for all $x, y \in R \otimes_k R$, so that $\sum_i x_i \otimes y_i \in \mathfrak{M}$ satisfies

$$\sum_i x_i \otimes y_i = \sum_i x_i \cdot \delta(y_i)$$

For all $y \in R$, set $\nu\delta y = dy$. It follows that $\mathfrak{M}/\mathfrak{M}^2$ is generated as an R-module by symbols $\{dy\}$ for all $y \in R$.

We claim that the homomorphism of k-modules

$$d\colon R \to \mathfrak{M}/\mathfrak{M}^2$$
$$y \mapsto dy$$

is a derivation. This follows from lemma A.7.4 on page 483 because it is the difference of two homomorphisms $\nu \circ \bar{\lambda}$ and $\nu \circ \lambda$ whose composite with p in equation A.7.6 are the *same*.

If M is an R-module and

$$D\colon R \to M$$

is any derivation, define a homomorphism of R-algebras

$$\varphi\colon R \otimes_k R \to R \star M$$

by $\varphi(x \otimes y) = (xy, x \cdot D(y))$. The ideal $\mathfrak{M}$ maps to $(0, M) \subset R \star M$ and since $M^2 = 0$, the map

$$\bar{\varphi} \colon \frac{R \otimes_k R}{\mathfrak{M}^2} \xrightarrow{\cong} R \oplus \frac{\mathfrak{M}}{\mathfrak{M}^2} = R \star \frac{\mathfrak{M}}{\mathfrak{M}^2} \to R \star M$$

is well-defined and gives rise to an R-module homomorphism

$$f \colon \frac{\mathfrak{M}}{\mathfrak{M}^2} \to M$$

with $D = f \circ d$. It follows that $\mathfrak{M}/\mathfrak{M}^2$ has the same universal property as $\Omega_{R/k}$, so equation A.7.2 on page 484 implies that, for *every* R-module, M

$$\hom_R(\Omega_{R/k}, M) = \hom_R(\mathfrak{M}/\mathfrak{M}^2, M)$$

It we set $M = \Omega_{R/k}$, then $1 \colon \Omega_{R/k} \to \Omega_{R/k}$ maps to a homomorphism $\mathfrak{M}/\mathfrak{M}^2 \to \Omega_{R/k}$. If we set $M = \mathfrak{M}/\mathfrak{M}^2$, then $1 \colon \mathfrak{M}/\mathfrak{M}^2 \to \mathfrak{M}/\mathfrak{M}^2$ gives rise to the *inverse* homomorphism $\Omega_{R/k} \to \mathfrak{M}/\mathfrak{M}^2$. $\square$

To compute modules of Kähler differentials in general, we use two exact sequences:

THEOREM A.7.11 (The first fundamental exact sequence). *Let k, R, S be rings with homomorphisms*

$$k \xrightarrow{\varphi} R \xrightarrow{\psi} S$$

Then

(1) *the sequence*

$$\Omega_{R/k} \otimes_R S \xrightarrow{v} \Omega_{S/k} \xrightarrow{u} \Omega_{S/R} \to 0$$

is exact.

(2) *the map v has a left-inverse (or what amounts to the same thing, v is injective and im v is a direct summand of $\Omega_{S/k}$) if and only if every derivation of R over k into any S-module M can be extended to a derivation $S \to M$.*

PROOF. The maps v is defined by $v\left(d_{R/k}(x) \otimes y\right) = y \cdot d_{S/k}\psi(x)$, and the map u is defined by

$$u\left(y \cdot d_{S/k}(y')\right) = y \cdot d_{S/R}(y')$$

for all $x \in R$ and $y, y' \in S$. It is clear that u is surjective. Since

$$d_{S/R}\psi(x) = 0$$

because $\psi(x) \in R$ counts as a constant (see remark A.7.3 on page 482), we have $u \circ v = 0$.

It remains to show that $\ker u = \operatorname{im} v$. It is enough to show

(A.7.7) $\qquad \hom_S(\Omega_{R/k} \otimes_R S, T) \xleftarrow{\hom(v,1)} \hom_S(\Omega_{S/k}, T) \xleftarrow{\hom(u,1)} \hom_S(\Omega_{S/R}, T)$

is exact for all S-modules, T (for instance $T = \operatorname{coker}(v)$). We begin with

$$\hom_S(\Omega_{R/k} \otimes_R S, T) = \hom_R(\Omega_{R/k}, \hom_R(S, T)) = \hom_R(\Omega_{R/k}, T)$$
$$= \operatorname{Der}_k(R, T)$$

— see proposition A.5.51 on page 466. The exact sequence in equation A.7.7 becomes

$$\operatorname{Der}_k(R, T) \xleftarrow{D \mapsto D \circ \psi} \operatorname{Der}_k(S, T) \leftarrow \operatorname{Der}_R(S, T)$$

Suppose the map $D \mapsto D \circ \psi$, sends D to 0. Then, for any element $r \in R, s \in S$,

$$D(\psi(r)s) = D(r \cdot s) = s \cdot D(\psi(r)) + r \cdot D(s)$$
$$= 0 + r \cdot D(s)$$

so D is R-linear, and in the image of $\operatorname{Der}_R(S, T)$. $\square$

The second fundamental exact sequence is also called *conormal exact sequence*:

LEMMA A.7.12 (Conormal exact sequence). *If $f\colon R \to S$ is a surjective homomorphism of algebras over a ring k with kernel $\mathfrak{A}$, then the sequence*

(A.7.8)
$$\frac{\mathfrak{A}}{\mathfrak{A}^2} \xrightarrow{d} S \otimes_R \Omega_{R/k} \xrightarrow{t} \Omega_{S/k} \to 0$$

is exact. Here d sends $x \in \mathfrak{A}$ to $1 \otimes u(x) \in S \otimes_R \Omega_{R/k}$ (where $u\colon R \to \Omega_{R/k}$ is the universal derivation) and t sends $s \otimes da \in S \otimes_R \Omega_{R/k}$ to $s \cdot d(f(a)) \in \Omega_{S/k}$.

REMARK. Note that
$$\frac{\mathfrak{A}}{\mathfrak{A}^2} = \frac{R}{\mathfrak{A}} \otimes_R \mathfrak{A} = S \otimes_R \mathfrak{A}$$

in the definition of d.

PROOF. Since $\Omega_{S/k}$ is generated by elements $d(f(a))$, the map t is surjective and is a homomorphism of S-modules when we equip $S \otimes_R \Omega_{R/k}$ with the action of S coming from the tensor product.

The map d is defined on $\mathfrak{A}$ but is a map to
$$S \otimes_R \Omega_{R/k} = \left(\frac{R}{\mathfrak{A}}\right) \otimes_R \Omega_{R/k} = \frac{\Omega_{R/k}}{\mathfrak{A} \cdot \Omega_{R/k}}$$

so its kernel contains $\mathfrak{A}^2$ and it really defines a map from $\mathfrak{A}/\mathfrak{A}^2$. The composite $t \circ d$ is clearly 0. We show that
$$\frac{S \otimes_R \Omega_{R/k}}{\operatorname{im} d} \cong \Omega_{S/k}$$

by showing that they both have the same universal property (see proposition A.7.7 on page 483).

We define a derivation
$$z\colon S \to \frac{S \otimes_R \Omega_{R/k}}{\operatorname{im} d}$$

by sending $s \in S$ to $1 \otimes dr$ where $r \in R$ is any element with $f(r) = s$. If $f(r_1) = f(r_2) = s$, then
$$1 \otimes dr_1 \equiv 1 \otimes dr_2 \quad (\operatorname{mod} \operatorname{im} d)$$

so this is well-defined. This map is a derivation because of the identities in $\Omega_{R/k}$, and proposition A.7.7 on page 483 implies the existence of a *unique* homomorphism
$$\Omega_{S/k} \to \frac{S \otimes_R \Omega_{R/k}}{\operatorname{im} d}$$

and its composite with t is the identity map (since it is unique). On the other hand, any derivation
$$\alpha\colon S \to M$$

(where M is an S-module) induces a derivation
$$\beta = \alpha \circ f\colon R \to M$$

where the R-action on M is given by composition with f, and (by proposition A.7.7 on page 483) a *unique* map
$$\Omega_{R/k} \to M$$

and a unique map
$$S \otimes_R \Omega_{R/k} \to S \otimes_R M = M$$

since M is an S-module. This clearly vanishes on $\operatorname{im} d$ so it defines a *unique* homomorphism
$$\frac{S \otimes_R \Omega_{R/k}}{\operatorname{im} d} \to M$$

making the appropriate version of diagram A.7.1 on page 484 commute. The conclusion follows.   □

Lemma A.7.12 on the facing page also gives us a way to compute $\Omega_{S/k}$ and explore its properties:

COROLLARY A.7.13. *If* $S = k[X_1, \ldots, X_n]/\mathfrak{A}$ *and* $\mathfrak{A} = (f_1, \ldots, f_m)$, *with projection* $g: k[X_1, \ldots, X_n] \to S$, *define the Jacobian matrix by*

$$\mathfrak{J}_{i,j} = g\left(\frac{\partial f_i}{\partial X_j}\right)$$

for $i = 1, \ldots, m$ *and* $j = 1, \ldots, n$ *(compare this with equation 3.3.2 on page 120). This defines a module-homomorphism*

$$\mathfrak{J}: S^m \to S^n$$

that fits into an exact sequence

$$S^m \xrightarrow{\mathfrak{J}} S^n \to \Omega_{S/k} \to 0$$

REMARK. In other words, $\Omega_{S/k}$ is the cokernel of the Jacobian map.

PROOF. This follows immediately from lemma A.7.12 on the facing page with $R = k[X_1, \ldots, X_n]$, $\Omega_{R/k} = R^n$ (by example A.7.8 on page 484) so $S \otimes_R \Omega_{R/k} = S^n$. The definition of the Jacobian map follows from the definition of the map, d, in equation A.7.8 on the facing page and equation A.7.4 on page 484. $\square$

EXAMPLE. We return to example 3.3.12 on page 125. Let V be the curve defined by $Y^2 = X^3$, with $k[V] = k[X, Y]/(Y^2 - X^3)$, and assume the characteristic of k is neither 2 nor 3.

As before, let x and y be the images of X and Y, respectively, under the projection

$$k[X, Y] \to k[V]$$

Note that x generates an isomorphic copy of $k[X]$ in $k[V]$ — i.e., it doesn't satisfy any relations. We can write all elements of $k[V]$ uniquely in the form $a(x) + yb(x)$, where $a(x)$ and $b(x)$ are polynomials in x — with the multiplication rule

$$(a(x) + yb(x))(c(x) + yd(x))$$
$$= a(x)c(x) + x^3 b(x)d(x) + y(a(x)d(x) + b(x)c(x))$$

We begin by computing $\Omega_{k[V]/k}$ using corollary A.7.13. In this case, the Jacobian is

$$f = \begin{bmatrix} -3x^2 & 2y \end{bmatrix}: k[V] \to k[V] \oplus k[V]$$

Applying this to $a(x) + yb(x)$ gives us

$$\left(-3x^2 a(x) - 3yx^2 b(x), 2x^3 b(x) + 2ya(x)\right)$$

— elements of $k[V] \oplus k[V]$ equated to *zero* in the quotient, $\Omega_{k[V]/k}$.

We claim that $\Omega_{k[V]/k}$ has $k[V] \oplus k^3$ as its underlying abelian group. Suppose

(A.7.9) $$(u(x) + yv(x), w(x) + yz(x)) \in k[V] \oplus k[V]$$

Let $w_3(x)$ be the portion of $w(x)$ of degree 3 and higher — so that $w_3(x)/x^3$ is well-defined. Then we add

$$f\left(-z(x)/2 - yw_3(x)/2x^3\right)$$
$$= \left(-3x^2 z(x)/2 - 3yx^2 w_3(x)/2x^3, -w_3(x) - yz(x)\right)$$

to equation A.7.9 and get

$$\left(u(x) - 3x^2 z(x)/2 + y\left(v(x) - 3x^2 w_3(x)/2x^3\right), w(x) - w_3(x) + 0\right)$$

so every element of $\Omega_{k[V]/k}$ can be *uniquely* written as

$$\left(u(x) + yv(x), c_0 + c_1 x + c_2 x^2\right)$$

Now we consider the action of $k[V]$ on this. It will suffice to describe the action of x and y.

$$x \cdot (u(x) + yv(x), c_0 + c_1 x + c_2 x^2)$$
$$= \left(xu(x) + yxv(x), c_0 x + c_1 x^2 + c_2 x^3\right)$$
$$= \left(xu(x) + y\left(xv(x) - 3c_2 x^2/2\right), c_0 x + c_1 x^2\right)$$

and

$$y \cdot (u(x) + yv(x), c_0 + c_1 x + c_2 x^2)$$
$$= \left(x^3 v(x) + yu(x), y\left(c_0 + c_1 x + c_2 x^2\right)\right)$$
$$= \left(x^3 v(x) - 3x^2\left(c_0 + c_1 x + c_2 x^2\right)/2 + yu(x), 0\right)$$

We will be concerned with the question of when the leftmost map in the conormal sequence (equation A.7.8 on page 488) is *injective* — and even *split* injective. In this case, the conormal sequence will effectively compute $\Omega_{R/k}$.

It is often said that Kähler differentials "linearize" rings, converting tensor products into *direct sums:*

COROLLARY A.7.14. *If A and B are be finitely generated k-algebras with $R = A \otimes_k B$, then*

$$\Omega_{R/k} = R \otimes_A \Omega_{A/k} \oplus R \otimes_B \Omega_{B/k}$$

PROOF. First note that the result is true for $A = k[X_1, \ldots, X_n]$ and $B = k[Y_1, \ldots, Y_m]$ by example A.7.8 on page 484. Furthermore, if $A = k[X_1, \ldots, X_n]/\mathfrak{A}$ and $B = k[Y_1, \ldots, Y_m]/\mathfrak{B}$, then

$$A \otimes_k B = k[X_1, \ldots, X_n, Y_1, \ldots, Y_m]/(\mathfrak{A} + \mathfrak{B})$$

(see corollary 2.7.2 on page 94). The conclusion follows from the fact that the Jacobians satisfy

$$\mathfrak{J}_{\mathfrak{A}+\mathfrak{B}} = \left[\begin{array}{cc} \mathfrak{J}_\mathfrak{A} & 0 \\ 0 & \mathfrak{J}_\mathfrak{B} \end{array}\right]$$

since $\partial X_i/\partial Y_j = \partial Y_k/\partial X_\ell = 0$. $\qquad\square$

We get an immediate corollary:

COROLLARY A.7.15. *If A is a finitely generated k algebra, and $R = A[X]$, then*

$$\Omega_{R/k} = R \otimes_A \Omega_{R/k} \oplus R \cdot dX$$

PROOF. This follows immediately from corollary A.7.14 and $R = A \otimes_k k[X]$. $\qquad\square$

The corollary above is used to show how Kähler differentials behave with respect to finite extensions of fields:

COROLLARY A.7.16. *Let F be a finitely generated field over k and let $F \subset G$ be a finite separable extension. Then*

$$\Omega_{G/k} = G \otimes_F \Omega_{F/k}$$

REMARK. If F is a finite extension of k, this implies that $\Omega_{F/k} = 0$ (since $\Omega_{k/k} = 0$).

Separability is actually necessary here: Suppose $k = \mathbb{Z}_p(X)$ and $F = k(X^{1/p}) = k[Y]/(Y^p - X)$, a finite inseparable extension. The conormal exact sequence is

$$(Y^p - X) \xrightarrow{\partial/\partial Y} F \otimes_{k[Y]} \Omega_{k[Y]/k} \to \Omega_{F/k} \to 0$$

The derivative of $Y^p - X$ vanishes identically and $\Omega_{k[Y]/k} = k[Y] \cdot dY$, so that

$$\Omega_{F/k} = F \otimes_{k[Y]} \Omega_{k[Y]/k} = F \cdot dY \neq 0$$

PROOF. Since G is a finite extension of F, it is algebraic and there exists an element $x \in G$ such that $G = F[x]$. Let $q(X) \in F[X]$ be the minimal polynomial of x. Then $\Omega_{F[X]/k} = F[X] \otimes_F \Omega_{F/k} \oplus F[X] \cdot dX$ and

$$G \otimes_{F[X]} \Omega_{F[X]/k} = G \otimes_F \Omega_{F/k} \oplus G \cdot dX$$

so we get an exact sequence

$$(q)/(q^2) \to G \otimes_F \Omega_{F/k} \oplus G \cdot dX \to \Omega_{G/k} \to 0$$

where the image of $q(X)$ is $q'(X) \cdot dX$ (via the chain-rule, etc). Since the extension is separable, $q'(X) \neq 0$ (see lemma A.2.16 on page 392) hence its image in G is invertible. It follows that the image of dX in $\Omega_{G/k}$ is 0.   □

Now we can prove that the module of Kähler differentials determines the dimension of an affine k-algebra:

COROLLARY A.7.17. *Let k be a field of characteristic 0 and let A be an affine k-algebra that is an integral domain with field of fractions F. Then*

$$\dim_F F \otimes_A \Omega_{A/k} = \dim A$$

REMARK. Of course, $\dim A$ denotes the Krull dimension.

PROOF. If $n = \dim A$, the Noether Normalization Theorem (2.2.2 on page 40) implies that A is an integral extension of a sub-ring $B = k[X_1, \ldots, X_n]$. Since A is a finitely generated module over B, its field of fractions is one over $G = k(X_1, \ldots, X_n)$, i.e. a finite extension. This is also separable[5], by lemma A.2.17 on page 393. Example A.7.8 on page 484 and proposition A.7.9 on page 485 imply that $\Omega_{G/k} = G^n$ and corollary A.7.16 on the preceding page implies that

$$\Omega_{F/k} \cong F \otimes_G \Omega_{G/k} = F^n$$

But proposition A.7.9 on page 485 implies that $\Omega_{F/k} \cong F \otimes_A \Omega_{A/k}$ and the conclusion follows.   □

Lemma A.7.4 on page 483 allows us to prove a sharper version of the conormal exact sequence:

THEOREM A.7.18. *If $f\colon A \to B$ is a surjective homomorphism of k-algebras with kernel $\mathfrak{K}$, then the map d in the conormal sequence*

(A.7.10) $$\frac{\mathfrak{K}}{\mathfrak{K}^2} \xrightarrow{d} B \otimes_A \Omega_{A/k} \xrightarrow{1 \otimes Df} \Omega_{B/k} \to 0$$

is a split injection if and only if there exists a homomorphism

$$g\colon B \to \frac{A}{\mathfrak{K}^2}$$

splitting the projection

$$\frac{A}{\mathfrak{K}^2} \to \frac{A}{\mathfrak{K}} = B$$

_____

[5]This is the only place where we use the hypothesis that the characteristic of k is 0.

REMARK. When A is a polynomial algebra, $\Omega_{A/k}$ is a free module and this result guarantees that $\Omega_{B/k}$ is a direct summand of a free module — i.e., a *projective* module. If $B = k[V]$ and $\Omega_{B/k}$ is projective, then $\Omega_{B/k}$ represents a vector-bundle on V under the Serre correspondence (theorem 3.5.9 on page 157).

PROOF. We begin by reducing to the case where $\mathfrak{K}^2 = 0$. The conormal sequence for $A \to A/\mathfrak{K}^2$ is

$$\mathfrak{K}^2 \xrightarrow{d} A/\mathfrak{K}^2 \otimes_A \Omega_{A/k} \to \Omega_{(A/\mathfrak{K}^2)/k} \to 0$$

so

$$\Omega_{(A/\mathfrak{K}^2)/k} \cong \frac{\Omega_{A/k}}{\mathfrak{K}^2 \cdot \Omega_{A/k} + \operatorname{im} d(\mathfrak{K}^2)}$$

Since d is a derivation, $d(a_1 a_2) = a_1 d(a_2) + a_2 d(a_1)$, so $d(\mathfrak{K}^2) \subset \mathfrak{K} \cdot \Omega_{A/k}$. It follows that

$$B \otimes_A \Omega_{A/k}$$
$$= \frac{\Omega_{A/k}}{\mathfrak{K} \cdot \Omega_{A/k}} = \frac{\Omega_{A/k}}{\mathfrak{K}^2 \cdot \Omega_{A/k} + \operatorname{im} d(\mathfrak{K}^2) + \mathfrak{K} \cdot \Omega_{A/k}}$$
$$= B \otimes_A \Omega_{(A/\mathfrak{K}^2)/k}$$

so, without loss of generality, we can assume $\mathfrak{K}^2 = 0$.

Now suppose d is split by a map $\sigma: B \otimes_A \Omega_{A/k} \to \mathfrak{K}$. If $\gamma = f \otimes 1: \Omega_{A/k} = A \otimes_A \Omega_{A/k} \to B \otimes_A \Omega_{A/k}$, then d is the restriction of $\gamma \circ u$ to $\mathfrak{K}$ (recall that $u: A \to \Omega_{A/k}$ is the universal derivation) and we can define

$$\delta = \sigma \circ \gamma \circ u: A \to \mathfrak{K} \subset A$$

— a k-linear derivation. Lemma A.7.4 on page 483 implies that

$$(1 - \delta): A \to A$$

is a homomorphism of k-algebras. If $x \in \mathfrak{K}$ then $\delta(x) = \sigma \circ \gamma \circ u(x) = x$, since σ *splits* $d = \gamma \circ u$. It follows that $(1 - \delta)(x) = 0$ and $1 - \delta$ induces a map

$$\xi: B \to A$$

inducing a map $B \to A$. Now $f \circ \xi = f \circ (1 - \delta) = f$, since $f \circ \delta = 0$ because the image of δ lies in $\mathfrak{K}$. It follows that ξ splits f and that there exists a map $D\xi$ splitting Df in equation A.7.10 on the preceding page.

Conversely, suppose that $\tau: B \to A$ is a map splitting $f: A \to B$. If $\delta = 1 - \tau \circ f: A \to A$, then $\delta(A) \subset \mathfrak{K}$ and lemma A.7.4 on page 483 (and the fact that $\mathfrak{K}^2 = 0$) shows that δ is a derivation from $A \to \mathfrak{K}$. The universal properties of $\Omega_{A/k}$ implies that this corresponds to a homomorphism of A-modules

$$\delta': \Omega_{A/k} \to \mathfrak{K}$$

such that $\delta' \circ u = \delta$. The A-linearity of δ' implies that $\delta'(\mathfrak{K} \cdot \Omega_{A/k}) = 0$ so that δ' induces a homomorphism

$$\delta'': B \otimes_A \Omega_{A/k} \to \mathfrak{K}$$

We claim that this splits d in equation A.7.10 on the previous page. If $x \in \mathfrak{K}$, then

$$\delta'' \circ d(x) = \delta' \circ u(x) = \delta(x) = (1 - \tau \circ f)(x) = x$$

The conclusion follows. □

The following result is interesting when $R = k[V]_{\mathfrak{m}} = \mathcal{O}_{V,p}$:

COROLLARY A.7.19. *If $(R, \mathfrak{m})$ is a local algebra over a field k with $R/\mathfrak{m} = k$, then the natural map*

$$\frac{\mathfrak{m}}{\mathfrak{m}^2} \to k \otimes_R \Omega_{R/k}$$

is an isomorphism.

REMARK. Suppose $A = k[V]$ for an irreducible affine variety, V. This result, combined with corollary A.7.9 on page 485 implies that

$$k \otimes_{A_{\mathfrak{m}}} (A_{\mathfrak{m}} \otimes_A \Omega_{A/k}) = k \otimes_A \Omega_{A/k} = \frac{\mathfrak{m}}{\mathfrak{m}^2}$$

for all maximal ideals $\mathfrak{m} \subset A$. In the rightmost tensor product, k is an A-module via the projection $A \to A/\mathfrak{m} = k$. It follows that $\Omega_{A/k}$ *determines* the cotangent spaces at all points of V. This supports the notion that $\Omega_{A/k}$ corresponds to the cotangent *bundl*e of V (if one exists).

PROOF. Just apply lemma A.7.12 on page 488 to the exact sequence

$$0 \to \mathfrak{m} \to R \to k \to 0$$

and use the fact that $\Omega_{k/k} = 0$. This implies that the map is surjective. Theorem A.7.18 on page 491 implies that this map is also *injective* (the fact that R is a k-algebra implies the existence of a map $k \to R$ that splits the projection to the quotient). $\square$

LEMMA A.7.20. *Let $(R, \mathfrak{m})$ be a noetherian local domain with field of fractions F, and let $k = R/\mathfrak{m}$. If M is a finitely-generated R-module then*

(A.7.11) $$\dim F \otimes_R M = \dim k \otimes_R M$$

if and only if M is free.

PROOF. The dimensions mentioned above are clearly equal if M is free. To see the converse, let

$$R^n \to M$$

be a surjective map with kernel K so we have a short exact sequence

$$0 \to K \to R^n \to M \to 0$$

and the sequence

$$0 \to K \otimes_R F \to F^n \to M \otimes_R F \to 0$$

is also exact since F is flat over R (see lemma A.5.60 on page 469). In addition

(A.7.12) $$K \otimes_R k \to k^n \to M \otimes_R k \to 0$$

is also exact. Equation A.7.11 and counting dimensions imply that this second exact sequence is also *short* exact (so that $K \otimes_R k \to k^n$ is injective). Since this is true for *any* surjection $R^n \to M$, choose $n = \dim k \otimes_R M$. Then $f: R^n \to M$ has the property that $f \otimes 1: R^n \otimes_R k \to M \otimes_R k$ is surjective so

$$\left(\frac{M}{\operatorname{im} f} \right) \otimes_R k = 0$$

Nakayama's lemma (corollary A.1.82 on page 379) implies that $M/\operatorname{im} f = 0$, i.e., f is surjective. The fact that equation A.7.12 is a short exact sequence implies that $K \otimes_R k = 0$ and Nakayama's lemma tells us that $K = 0$. $\square$

EXERCISES.

1. If $k \to K$ is an extension of fields of characteristic 0 and $f_1, \ldots, f_m \in K$, show that $\{f_1, \ldots, f_m\}$ are *algebraically* independent if and only if $\{df_1, \ldots, df_m\} \in \Omega_{K/k}$ are *linearly* independent.

2. If F is the field of fractions of $\mathbb{Q}[[T]]$, show that $\dim \Omega_{F/\mathbb{Q}}$ is uncountable, so $\Omega_{\mathbb{Q}[[T]]/\mathbb{Q}}$ is uncountably generated over $\mathbb{Q}$.

Sheaves and ringed spaces

"Geometry is one and eternal shining in the mind of God. That share in it accorded to men is one of the reasons that Man is the image of God."
—Johannes Kepler, *Conversation with the Sidereal Messenger (an open letter to Galileo Galilei)*, [**144**].

B.1. Sheaves

General algebraic varieties are the result of gluing together affine algebraic varieties along regular maps.

One problem that arises in the general case is that the coordinate ring no longer contains all the significant geometric information. For instance, the only regular functions on projective spaces are *constants* — see example 4.4.8 on page 186. On the other hand, a projective space is a union of affine subvarieties which have coordinate rings that express geometric properties.

We require more mathematical machinery, particularly a construct called a *sheaf*. Sheaves are used in many areas of mathematics including differential topology and complex function theory and topology.

Sheaf theory was used implicitly for years before the formal definition appeared. In 1936, Eduard Čech introduced a topological construction called *nerves* that are essentially special cases of sheaves. Others, like Steenrod and Leray also used concepts similar to sheaves, usually in algebraic topology. The first "official" definition appeared in [**42**].

The first person to apply sheaves to algebraic geometry (in [**150**]) was Serre. Hirzebruch extended this to topological constructions in [**80**].

> Born 1926, Jean-Pierre Serre is a French mathematician who has made remarkable contributions to algebraic topology, algebraic geometry, group theory and other fields.

We will use sheaves to associate "coordinate rings" to the open affines that are glued together to form a variety.

DEFINITION B.1.1. Let V be a topological space and let $\mathscr{C}$ be a category of sets with additional structure. Suppose that, for every open subset $U \subset V$ we have an object $\mathcal{O}_V(U) \in \mathscr{C}$. Then $\mathcal{O}_V$ is called a *sheaf of objects of $\mathscr{C}$* if:

(1) If U' is an open subset of U then there is a natural morphism (called the *restriction map*) $\mathcal{O}_V(U) \to \mathcal{O}_V(U')$. If $x \in \mathcal{O}_V(U)$, we write the restriction of x to U' as $x|U'$.
(2) The restriction map $\mathcal{O}_V(U) \to \mathcal{O}_V(U)$ is the identity map.

(3) If $U' \subset U_1 \cap U_2$ are open sets, then the diagram of restriction maps

$$
\begin{array}{ccc}
 & \mathcal{O}_V(U_1) & \\
 \nearrow & & \searrow \\
\mathcal{O}_V(U) & & \mathcal{O}_V(U') \\
 \searrow & & \nearrow \\
 & \mathcal{O}_V(U)_2 &
\end{array}
$$

commutes.

(4) if $U \subset V$ is an open set with an open cover $U = \bigcup U_i$ and $x, y \in \mathcal{O}_V(U)$ have the property that $x|U_i = y|U_i$, then $x = y$.

(5) If $x_i \in \mathcal{O}_V(U_i)$ for all U_i in some open cover of U, and

$$
x_i|U_i \cap U_j = x_j|U_i \cap U_j
$$

for all i, j then there is an element $x \in \mathcal{O}_V(U)$ such that $x_i = x|U_i$.

If $x \in V$, the *stalk* of $\mathcal{O}_V$ at x, denoted $\mathcal{O}_{V,x}$ is the direct limit $\varinjlim \mathcal{O}_V(U)$ (see definition A.5.17 on page 448) as U runs over all open sets with $x \in U$. The direct limit of restriction-maps induces a canonical map

$$
r_x\colon \mathcal{O}_V(U) \to \mathcal{O}_{V,x}
$$

for all $x \in U$.

REMARK. The term "category of sets with additional structure" is vague. We will actually want to work with sheaves of rings or k-algebras, in which case the "additional structure" is operations like addition and multiplication.

We will often be interested in sheaves of functions on V with restriction maps that are restrictions of functions. In this case, conditions 1, 2, and 3 are automatically satisfied and the definition becomes simpler.

Conditions 1 and 5 require that a $x \in \mathcal{O}_V(U)$ if and only if each point $p \in U$ has a neighborhood U_p such that $x|U_{Pp} \in \mathcal{O}_V(U_p)$. In other words, the condition for $x \in \mathcal{O}_V(U)$ is *local*.

Condition 5 implies that natural *extensions* of elements of $\mathcal{O}_V(U_i)$ are also in $\mathcal{O}_V(\bigcup_i U_i)$. If $\mathcal{O}_V$ doesn't satisfy conditions 4 and 5, it is a *presheaf*. If a presheaf satisfies condition 4, it is a *separated presheaf*.

We also have:

EXAMPLE. Recall that a function $f\colon U \to \mathbb{C}$, where U is open in $\mathbb{C}^n$ is *analytic* if it is described by a convergent power series in a neighborhood of each point $P \in U$. Let V be an open subset of $\mathbb{C}^n$, and for each open subset $U \subset V$, let $\mathcal{A}_V(U)$ be the set of all analytic functions on U.

Then $\mathcal{A}_V$ is a sheaf of $\mathbb{C}$-algebras.

Here's a very simple example:

DEFINITION B.1.2. Let V be a topological space and let A be an abelian group. The *constant sheaf* $\underline{A}$ has the property that $\underline{A}_V(U)$ is the group of continuous maps $U \to A$ for all open sets U where A has the *discrete topology* (every element of A is a *distinct component*).

REMARK. If U is *connected*, $\underline{A}(U) = A$, and if $U' \subset U$ is an inclusion of open sets with U' connected, the induced map

$$
\underline{A}_V(U) \to \underline{A}_V(U')
$$

is the identity map. If U is an open set that is the union of n disjoint components then $\underline{A}(U) = A^n$.

It is interesting to contrast definition B.1.2 on the preceding page with:

EXAMPLE B.1.3. Let V be a topological space and, for each open subset $U \subset V$, let $\mathcal{O}_V(U)$ be the set of all constant functions on U. Then $\mathcal{O}_V$ is a *presheaf* that is not a *sheaf* unless V is *irreducible*.

Condition 5 states that *different* constants on *different* irreducible components must patch together to give a function in $\mathcal{O}_V(V)$. Such a function would not be a constant over all of V.

DEFINITION B.1.4. Given two sheaves $\mathcal{O}_1$ and $\mathcal{O}_2$ on a space X, a *morphism of sheaves* is a system of morphisms

$$r(U) \colon \mathcal{O}_1(U) \to \mathcal{O}_2(U)$$

for all open sets, $U \subset V$, that commute with all of the structure-morphisms in definition B.1.1 on page 495.

Note: the reader may have seen material on sheaves that uses the notation

$$\Gamma(U, \mathcal{O}_X)$$

when $U \subset X$ and $\mathcal{O}_X$ is a sheaf on a space, X. This consists of *sections* of $\mathcal{O}_X$ — maps

$$s \colon U \to \mathcal{O}_X$$

with the property that $\pi \circ s = 1 \colon U \to U$, where $\pi \colon \mathcal{O}_X \to X$ is the map that sends $\mathcal{O}_{X,x}$ to x for all $x \in X$. In section 3 of [**150**], Serre proves that the sections of a sheaf of abelian groups forms a sheaf that is *canonically isomorphic* to the original sheaf. In other words

$$\Gamma(U, \mathcal{O}_X) = \mathcal{O}_X(U)$$

so we will rarely use the $\Gamma(*, *)$-notation. Sometimes this notation is useful when we want to regard evaluation as a functor of a sheaf.

We will also need the concept of the *support* of a sheaf — basically where it is nonvanishing:

DEFINITION B.1.5. If $\mathcal{O}_X$ is a sheaf of abelian groups on a space X, the *support* of $\mathcal{O}_X$, denoted $\mathrm{Supp}\, \mathcal{O}_X$, is the set $X \setminus \bar{U}$, where

$$\bar{U} = \bigcup_{\mathcal{O}_X(U)=0} U$$

and the sets U are open.

We also have the related concept of the support of a *section* of a sheaf:

DEFINITION B.1.6. If $\mathcal{O}_X$ is a sheaf of abelian groups on a space X and $f \in \mathcal{O}_X(U)$ is a section over an open set $U \subset X$, the *support* of f, denoted $\mathrm{Supp}\, f$, is the set of points $x \in U$ with the property that

$$r_x(f) \neq 0 \in \mathcal{O}_{X,x}$$

— where r_x is the restriction-map defined in B.1.1 on page 495.

Here is an extreme example of a sheaf — a kind of sheaf over a *point*

EXAMPLE B.1.7. Given a space X, a point $x \in X$, and a module M, the *skyscraper sheaf* at x, denoted $i_x(M)$ is defined by

$$i_x(U) = \begin{cases} M & \text{if } x \in U \\ 0 & \text{otherwise} \end{cases}$$

All restriction maps are the identity map or the zero map.

REMARK. The only nonvanishing stalk of a skyscraper sheaf is the one over the point of definition, x. The *support* of this sheaf is the single point, x.

EXERCISES.

1. If $\mathscr{F}$ is a sheaf over a space X and U is any open set, show that the map

$$\prod_{x \in U} r_x \colon \mathscr{F}(U) \to \prod_{x \in U} \mathscr{F}_x$$

is injective.

2. Show that a morphism of sheaves is determined by its behavior on stalks.

3. Suppose $f\colon X \to Y$ is a surjective continuous map of topological spaces and let $\mathcal{O}_X$ be a sheaf on X. Show that we can define a sheaf $f_*\mathcal{O}_X$ on Y by setting

$$f_*\mathcal{O}_X(U) = \mathcal{O}_X(f^{-1}(U))$$

for every open set $U \subset Y$. This is called the *direct image sheaf*.

B.2. Presheaves verses sheaves

As mentioned above, a presheaf satisfies all of the conditions in definition B.1.1 on page 495 except for condition 5.

Here's an example of a presheaf that is not a sheaf. It also illustrates how a homomorphic image of a sheaf may fail to be a sheaf:

EXAMPLE B.2.1. Consider the sequence of morphisms of sheaves over $\mathbb{C}$:

$$\underline{2\pi i \mathbb{Z}} \to \mathcal{A}_{\mathbb{C}} \xrightarrow{\exp} \mathcal{N}_{\mathbb{C}^*}$$

where:

(1) $\underline{2\pi i \mathbb{Z}}$ is the constant sheaf equal to the abelian group $2\pi i \mathbb{Z} \subset \mathbb{C}$.
(2) $\mathcal{A}_{\mathbb{C}}$ is the sheaf that associates to each open set $U \subset \mathbb{C}$ the $\mathbb{C}$-algebra, $\mathcal{A}_{\mathbb{C}}(U)$, of complex-analytic functions defined on it. This is an algebra under addition.
(3) $\mathcal{N}_{\mathbb{C}^*}$ is the sheaf of complex-analytic *nowhere-zero* functions on $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, which form an algebra over $\mathbb{C}^*$ under *multiplication*.
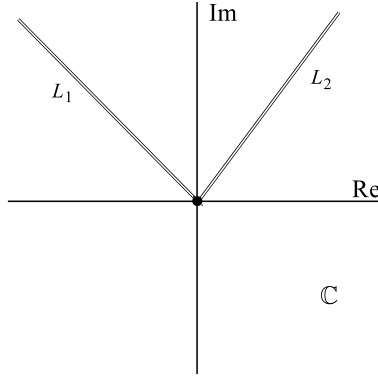
FIGURE B.2.1.   Branch-cuts

(4) exp is the exponential function, which defines a map of sheaves be-
cause it sends sums into products.

It is well-known that $e^{2\pi ni} = 1$ so the image of the constant sheaf $\underline{2\pi i \mathbb{Z}}$ under
exp is the constant function $1 \in \mathbb{C}^*$.

The image of exp is a presheaf, $\mathcal{E}_{\mathbb{C}^*}$, that associates to each open set $U \subset \mathbb{C}$
the set of functions $e^{f(z)}$ where $f(z)$ is complex-analytic on U. We claim that
this is a presheaf that is *not* a sheaf.

In figure B.2.1, consider the open sets $U_1 = \mathbb{C} \setminus L_1$ and $U_2 = \mathbb{C} \setminus L_2$, where

- The line-segments, L_1 and L_2, are rays from the origin.
- On each of these sets, the image of exp $= \mathcal{E}_{\mathbb{C}^*}$ has a function equal to
  z.
- On the other hand, $\mathcal{E}_{\mathbb{C}^*}(\mathbb{C})$ does *not* have a *global* function equal to z
  even though $\mathbb{C}^* = U_1 \cup U_2$.

This is because ln is *only* well-defined after one removes a branch-cut from $\mathbb{C}^*$.

There is a standard procedure for completing a presheaf to make it a sheaf:
just add functions to the global set that restrict to compatible functions on each
open set. We must also add restrictions of these *new* global functions to all of
the open sets.

DEFINITION B.2.2. Let $\mathcal{O}_V$ be a presheaf on a space V. If $U \subset V$ is open
and has an open covering, $U = \bigcup_i U_i$, define a *compatible family* $\mathcal{C}(\{U_i\})$ to be a
sequence

$$x_i \in \mathcal{O}_V(U_i)$$

with

$$x_i|U_i \cap U_j = x_j|U_i \cap U_j$$

for all i, j. A compatible family forms a set in a natural way and operations on
this set are preserved by restriction maps they will be well-defined on compat-
ible families.

If $U = \bigcup_j W_i$ is a another open covering that is a refinement of $\{U_i\}$, there is a natural restriction homomorphism

$$\mathcal{C}(\{U_i\}) \to \mathcal{C}(\{W_j\})$$

LEMMA B.2.3. *Let $\mathcal{O}_V$ be a presheaf of groups on a topological space V. Define a new presheaf $\overline{\mathcal{O}_V}$ on V as follows:*

$$\overline{\mathcal{O}_V}(U) = \varinjlim \mathcal{C}(\{U_i\})$$

where the direct limit (see proposition A.5.18 on page 449) is taken over all open coverings of U. Then $\overline{\mathcal{O}_V}$ is called the completion of $\mathcal{O}_V$ and is a sheaf.

There is a canonical homomorphism of presheaves

$$f\colon \mathcal{O}_V \to \overline{\mathcal{O}_V}$$

REMARK. Applying this construction to the presheaf $\mathcal{E}_{\mathbb{C}^*}$ in the example above will result in the sheaf $\mathcal{N}_{\mathbb{C}^*}$. For instance, the open sets U_1 and U_2 will have functions equal to z that will form a compatible family that is globally equal to z.

According to Mumford ([**122**]):

> F' is "the best possible sheaf you can get from F". It is easy to imagine how to get it: first identify things which have the same restrictions, and then add in all the things which can be patched together.

Note that the direct limit identifies elements whose restriction of all open sets are equal — guaranteeing that the result is separated.

PROOF. It is easy to see that $\overline{\mathcal{O}_V}$ will still satisfy conditions 1, 2 and 3 of definition B.1.1 on page 495. Each ring $\overline{\mathcal{O}_V}(U)$ will contain all of the elements of $\mathcal{O}_V(U)$ plus any extras that arise from compatible families on open covers. It follows that the result will be a sheaf. $\qquad\square$

This process of converting a presheaf into a sheaf preserves stalks:

PROPOSITION B.2.4. *Let $\mathscr{F}$ be a presheaf on a space X and let*

$$f\colon \mathscr{F} \to \mathscr{F}'$$

be the canonical map to the completion. Then f induces isomorphisms

$$f_x\colon \mathscr{F}_x \to \mathscr{F}'_x$$

for all $x \in X$.

REMARK. One way to think of this is "the direct limit used to construct $\mathscr{F}'$ commutes with the direct limits used to construct the stalks."

PROOF. If $s \in \mathscr{F}_x$ is in the kernel of f_x, then there exists an open set U and an element $t \in \mathscr{F}(U)$ with $t_x = s$. Since $f_x(s) = 0$ it follows that there exists an open set $U' \subset U$ such that $f(yt)|U' = 0$. Since $\mathscr{F}'$ is a direct limit over open covers, there must exist an open set $U'' \subset U$ with $t|U'' = 0$. But this implies that $s = 0 \in \mathscr{F}_x$.

To prove surjectivity, let $s \in \mathscr{F}'_x$ and note that there exists an open set U and element $t \in \mathscr{F}'(U)$ that gives rise to s in the inverse limit. Since $\mathscr{F}'(U)$ is a direct limit over open covers, there exists an open set $U' \subset U$ such that $t \in \mathscr{F}(U')$, from which we conclude that $s \in \operatorname{im} f_x$. $\qquad\square$

Since the homomorphic image of a sheaf is generally only a presheaf, we make

DEFINITION B.2.5. In the category of sheaves, $\mathscr{S}$, the image of a morphism is the completion of its image as a presheaf.

We can define subsheaves and quotient sheaves:

DEFINITION B.2.6. Let $f\colon \mathcal{O}_1 \to \mathcal{O}_2$ be an injective morphism of sheaves of modules over a space X. We can define a presheaf over X by setting

$$\mathcal{O}_q(U) = \frac{\mathcal{O}_2(U)}{\mathcal{O}_1(U)}$$

for open sets $U \subset V$. The completion of this presheaf is called the *quotient sheaf* and denoted $\mathcal{O}_2/\mathcal{O}_1$.

REMARK. Example B.2.1 on page 498 shows why we must take the completion. On open sets, $U \subset \mathbb{C}^*$, that do *not* circle the origin

$$\mathcal{N}_{\mathbb{C}^*}(U) = \frac{\mathcal{A}_{\mathbb{C}}(U)}{2\pi i \mathbb{Z}(U)} = \frac{\mathcal{A}_{\mathbb{C}}(U)}{2\pi i \mathbb{Z}}$$

but the quotient, $\mathcal{E}_{\mathbb{C}^*}$, is not a sheaf since it does not include the function z. Its *completion* is $\mathcal{N}_{\mathbb{C}^*}$.

In one case, we do not *need* to complete a presheaf

DEFINITION B.2.7. A sheaf $\mathscr{X}$ over a space X is *flasque* if every restriction map

$$p_U^{U'}\colon \mathscr{X}(U') \to \mathscr{X}(U)$$

for $U \subset U'$, is *surjective*.

REMARK. Most of the sheaves we have dealt with before are not flasque — indeed, the restriction maps have usually been injective.
Flasque sheaves are also called *flabby* sheaves.

Surjectivity of restriction maps implies that compatible families *always* extend to the entire space so:

PROPOSITION B.2.8. *A flasque separated presheaf is a sheaf.*

REMARK. This means we can define quotients of flasque sheaves in the straightforward way — without considering whether they are sheaves.

EXERCISES.

1. If

$$0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$$

is an exact sequence of sheaves where $\mathscr{F}$ and $\mathscr{G}$ are flasque, show that $\mathscr{H}$ is also flasque, so that

$$0 \to \mathscr{F}(U) \to \mathscr{G}(U) \to \mathscr{H}(U) \to 0$$

is exact for all open sets U.

2. If $\mathscr{F}$ is a flasque sheaf on a space Y that is a closed subspace of a space X with inclusion map
$$i\colon Y \to X$$
show that $i_*\mathscr{F}$ is flasque on X.

3. Show that the category of presheaves is abelian (see definition A.5.33 on page 457) but the category of sheaves is not.

4. Show that a homomorphism of sheaves of abelian groups $f\colon \mathscr{F} \to \mathscr{G}$ is injective if and only if it induces an injection on all stalks.

5. Show that a homomorphism of sheaves of abelian groups $f\colon \mathscr{F} \to \mathscr{G}$ is an isomorphism if and only if it induces an isomorphism on all stalks.

6. Show that a homomorphism of sheaves of abelian groups $f\colon \mathscr{F} \to \mathscr{G}$ on a space X is surjective if and only if it induces a surjection on all stalks.

7. Show that a sequence of sheaves of abelian groups
$$0 \to \mathscr{F} \to \mathscr{G} \to \mathscr{H} \to 0$$
over a space X is exact if and only if the induced sequence of stalks
$$0 \to \mathscr{F}_x \to \mathscr{G}_x \to \mathscr{H}_x \to 0$$
is exact, for all $x \in X$.

8. If $f\colon X \to Y$ is a map of topological spaces and $\mathscr{F}$ is a sheaf on Y, show that
$$f^{-1}(\mathscr{F})(U) = \varinjlim_{V \supset f(U)} \mathscr{F}(V)$$
for open sets $V \subset Y$, defines a presheaf on X. The sheaf associated to this is called the inverse image sheaf of $\mathscr{F}$ under f. If f is an *open* mapping (i.e., if it sends open sets to open sets), then $f^{-1}(\mathscr{F})(U) = \mathscr{F}(f(U))$).

9. If A is an abelian group, why isn't the constant sheaf $\underline{A}$ flasque?

B.3. Ringed spaces

DEFINITION B.3.1. If V is a topological space and $\mathcal{O}_V$ is a sheaf over V, the pair $(V, \mathcal{O}_V)$ is called a *ringed space*.

A *morphism of ringed spaces* $(f, f^\sharp)\colon V \to W$ is
  (1) a continuous map $f\colon V \to W$ of topological spaces
  (2) for every open set $U \subset W$ there is a ring-homomorphism $f^\sharp_U\colon \mathcal{O}_W(U) \to \mathcal{O}_V(f^{-1}U)$ compatible with restriction maps. In other words, given open sets $U_1 \subset U_2 \subset W$, the diagram

$$
\begin{array}{ccc}
\mathcal{O}_W(U_2) & \xrightarrow{\ f^\sharp_{U_2}\ } & \mathcal{O}_V(f^{-1}(U_2)) \\
\downarrow & & \downarrow \\
\mathcal{O}_W(U_1) & \xrightarrow{\ f^\sharp_{U_1}\ } & \mathcal{O}_V(f^{-1}(U_1))
\end{array}
$$

commutes.

The set of morphisms $f\colon V \to W$ is denoted hom (V, W).

A ringed space with the property that the stalk of each point (see definition B.1.1 on page 495) is a local ring (see definition A.1.24 on page 351) is called a *locally ringed space*.

REMARK. For historical reasons, we often write $\Gamma(U, \mathcal{O}_V)$ for $\mathcal{O}_V(U)$ and call its elements *sections of $\mathcal{O}_V$ over U*. Unfortunately both notations are in widespread use.

It is not hard to see that the restriction of a sheaf to an open set is also a sheaf on that open set.

EXAMPLE. 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 neighborhood U for which $(U, \mathcal{O}_V|U)$ is isomorphic to the ringed space of continuous functions on an open subset of $\mathbb{R}^n$. This is slicker (or shorter) than the usual definition found in [**104**].

A *differential 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 neighborhood U for which $(U, \mathcal{O}_V|U)$ is isomorphic to the ringed space of smooth functions on an open subset of $\mathbb{R}^n$.

We will often be interested in *local* properties of sheaves:

DEFINITION B.3.2. Let $(V, \mathcal{O}_V)$ be a ringed space and let $p \in V$ be a point. Consider pairs (f, U) consisting of an open neighborhood U of p and an $f \in \mathcal{O}_V(U)$.

Write $(f, U) \frown (f', U')$ if there exists an open set $U'' \subset U \cap U'$ such that $f|U'' = f'|U''$.

This is an equivalence relation and an equivalence class of pairs is called a *germ* of a function at p (relative to $\mathcal{O}_V$).

The germs of functions at p will be denoted $\mathcal{O}_{V,p}$ or just $\mathcal{O}_p$ — the *stalk* of $\mathcal{O}_V$ over p.

If $\mathcal{O}_{V,p}$ is a local ring for all $p \in V$, then $(V, \mathcal{O}_V)$ is called a *locally ringed space*.

REMARK. It is not hard to see that

$$\mathcal{O}_p = \varinjlim \mathcal{O}_V(U)$$

— the direct limit (see proposition A.5.18 on page 449) over all open sets containing p. In all interesting cases, this is a local ring with maximal ideal consisting of functions that vanish at p.

EXAMPLE. Let $\mathcal{O}_V$ be the sheaf of complex analytic functions on $V = \mathbb{C}$. These are functions that are equal to power series expansions at their nonsingular points (points where they don't "blow up" as $1/z$ does when $z = 0$).

A power series $\sum_{n \geq 0} a_n (z - c)^n$ is called *convergent* if it has a nonzero (but arbitrarily small) radius of convergence. The set of all such power series is a $\mathbb{C}$-algebra.

This $\mathbb{C}$-algebra (of convergent power series) is canonically isomorphic to the $\mathbb{C}$-algebra of germs of functions.

DEFINITION B.3.3. If $(X, \mathcal{O}_X)$ is a ringed space, a *sheaf of modules*, $\mathscr{F}$ over $\mathcal{O}_X$, or a module over $\mathcal{O}_X$ is a sheaf with the property that, for every open set $U \subset X$, $\mathscr{F}(U)$ is a module over $\mathcal{O}_X(U)$ such that the restriction map for $V \subset U$,

$$p_V^U: \mathscr{F}(U) \to \mathscr{F}(V)$$

is compatible with module structures via the ring homomorphism $\mathcal{O}_X(U) \to \mathcal{O}_X(V)$.

REMARK. Having defined our objects of study, we must define morphisms:

DEFINITION B.3.4. A morphism $g: \mathscr{F} \to \mathscr{G}$ of $\mathcal{O}_X$-modules is a morphism of sheaves such that $g(U): \mathscr{F}(U) \to \mathscr{G}(U)$ is a homomorphism of $\mathcal{O}_X(U)$-modules. The group of homomorphisms is denoted $\hom_{\mathcal{O}_X}(\mathscr{F}, \mathscr{G})$ or $\hom_X(\mathscr{F}, \mathscr{G})$.

If $U \subset X$ is an open set, then $\mathscr{F}|U$ is a module over $\mathcal{O}_X|U$ and we can define the sheaf-hom functor, $\mathscr{H}om_X(\mathscr{F}, \mathscr{G})$ by

$$\mathscr{H}om_X(\mathscr{F}, \mathscr{G})(U) = \hom_U(\mathscr{F}|U, \mathscr{G}|U)$$

REMARK. Although $\mathscr{H}om_X(\mathscr{F}, \mathscr{G})$ is only defined as a presheaf, the result is actually a sheaf: If $\{U_\alpha\}$ is an open cover of X and $f_\alpha: \mathscr{F}|U_\alpha \to \mathscr{G}|U_\alpha$ is a set of morphisms that agree on intersections, they define a morphism $f: \mathscr{F} \to \mathscr{G}$ because $\mathscr{F}$ and $\mathscr{G}$ are sheaves.

We can define tensor products of modules over $\mathcal{O}_X$:

DEFINITION B.3.5. If $(X, \mathcal{O}_X)$ is a ringed space and $\mathscr{F}$ and $\mathscr{G}$ are modules over $\mathcal{O}_X$, the we define

$$\mathscr{F} \otimes_{\mathcal{O}_X} \mathscr{G}$$

to be the *completion* (see lemma B.2.3 on page 500) of the presheaf, P, defined by

$$P(U) = \mathscr{F}(U) \otimes_{\mathcal{O}_X(U)} \mathscr{G}(U)$$

for every open set $U \subset X$.

REMARK. Example 5.3.4 on page 236 shows that the completion-step is actually necessary.

DEFINITION B.3.6. If $f(X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ is a morphism of ringed spaces (as in definition B.3.1 on page 502) and $\mathscr{F}$ is a $\mathcal{O}_X$-module, then $f_* \mathscr{F}$ is naturally a $\mathcal{O}_X$-module too. The maps $f^\sharp(U): \mathcal{O}_Y(U) \to \mathcal{O}_X(f^{-1}(U))$ equips $f_* \mathscr{F}$ with a $\mathcal{O}_Y$-module structure.

EXERCISES.

1. Suppose we want to define a sheaf of finitely generated k-algebras whose stalks are local rings on a topological space V. In this case, show that definition B.1.1 on page 495 is equivalent to the following shorter definition (valid for affine varieties):

DEFINITION B.3.7. If k is algebraically closed, and V is a topological space, a sheaf of finitely generated k-algebras whose stalks are local rings on V assigns a set $\mathcal{O}_V(U)$ of functions $U \to k$ to each open set $U \subset V$ such that:

a. $\mathcal{O}_V(U)$ contains the constant functions and is closed under addition and multiplication,
b. if $U' \subset U$ is an open subset and $f \in \mathcal{O}_V(U)$ then $f|U' \in \mathcal{O}_V(U')$,
c. a function $f: U \to k$ is in $\mathcal{O}_V(U)$ if $f|U_i \in \mathcal{O}_V(U_i)$ for some open covering $\{U_i\}$ of U.

2. Let V be a topological space and let $\mathscr{X}(V)$ be the category of open sets of V where the only morphisms allowed are inclusions of open sets. Show that we can give the very short but barely comprehensible definition of a presheaf:

DEFINITION. A presheaf on V is a contravariant functor (see definition A.5.7 on page 442)

$$\mathscr{X}(V) \to \mathscr{C}$$

3. If $(X, \mathcal{O}_X)$ is a ringed space and $\mathscr{F}$ is a module over $\mathcal{O}_X$, show that there exists a natural homomorphism of sheaves

$$\mathscr{F} \otimes_{\mathcal{O}_X} \mathscr{H}om(\mathscr{F}, \mathcal{O}_X) \to \mathcal{O}_X$$

4. Under the conditions of exercise 3 above, assume that $\mathscr{F}$ is a locally-free sheaf of rank 1. Show that the homomorphism

$$\mathscr{F} \otimes_{\mathcal{O}_X} \mathscr{H}om(\mathscr{F}, \mathcal{O}_X) \to \mathcal{O}_X$$

is an *isomorphism* in this case.

Vector bundles

> "...this miracle of analysis, this marvel of the world of ideas, an almost amphibian object between Being and Non-being that we call the imaginary number..."
> —Gottfried Wilhelm Leibniz

C.1. Introduction

In this section, we will define an important topological concept closely related to sheaves. A *vector bundle* attaches a vector-space to every point of a ringed space, V, in a way that looks like a *product* in a neighborhood of every point.

Throughout, k will denote a fixed field and

(1) all vector-spaces will be over k
(2) all sheaves will have values in k

We begin with a definition:

DEFINITION C.1.1. If $(X, \mathcal{O}_X)$ is a ringed space, $U \subset X$ is an open set, then a map $f: U \to \mathbb{A}^n$, is said to be *compatible with* $\mathcal{O}_X$ if the composite

$$f: U \to \mathbb{A}^n \xrightarrow{p_i} k$$

is in $\mathcal{O}_X(U)$ for all $1 \leq i \leq n$, where $p_i: \mathbb{A}^n \to \mathbb{A}^1 = k$ is projection to the ith copy of $\mathbb{A}^1 = k$.

REMARK. For instance, if $\mathcal{O}_X$ is the sheaf of continuous functions, then f must be continuous to be compatible with it. If $\mathcal{O}_X$ is the constant sheaf, then the image of f must be a single point.

A map $f: U \to \hom(\mathbb{A}^n, \mathbb{A}^m) = \mathbb{A}^{n \cdot m}$ will be regarded as compatible with $\mathcal{O}_X$ if the corresponding map $f: U \to \mathbb{A}^{n \cdot m}$ is. Here, $\hom(\mathbb{A}^n, \mathbb{A}^m)$ denotes the set of $m \times n$ matrices with coefficients in k. The fact that $\mathcal{O}_X(U)$ is a *ring* implies that:

PROPOSITION C.1.2. *If $(X, \mathcal{O}_X)$ is a ringed space, $U \subset X$ is an open set, and*

- *$f: U \to \mathbb{A}^n$ and*
- *$g: U \to \hom(\mathbb{A}^m, \mathbb{A}^m)$*

are both compatible with $\mathcal{O}_X$, then so is the composite $g \circ f: U \to \mathbb{A}^m$.

Now we can define the main object of this section:

DEFINITION C.1.3. Let $(X, \mathcal{O}_X)$ be a locally ringed space (see definition B.3.2 on page 503). A *vector bundle*, ξ, *over* X is a topological space W and a continuous map

$$\xi: W \to X$$

such that

    (1) if $x \in X$ is any point, $\xi^{-1}(x) = \mathbb{A}^n$ for some n.

    (2) there exists a covering

(C.1.1)
$$X = \bigcup_{\alpha} U_{\alpha}$$

    by open sets $\{U_{\alpha}\}$ (called a *trivializing cover* of ξ) and homeomorphisms (called *charts*)

$$\iota_{\alpha}: \xi^{-1}(U_{\alpha}) \to U_{\alpha} \times \mathbb{A}^n$$

    that make the diagram

$$
\begin{array}{ccc}
\xi^{-1}(U_{\alpha}) & \xrightarrow{\iota_{\alpha}} & U_{\alpha} \times \mathbb{A}^n \\
\downarrow{\scriptstyle p} & & \downarrow \\
U_{\alpha} & =\!=\!=\!=\!= & U_{\alpha}
\end{array}
$$

    commute, and such that $\iota_{\alpha}|\xi^{-1}(u): \xi^{-1}(u) \to u \times \mathbb{A}^n$ is a linear isomorphism of vector spaces for all $u \in U_{\alpha}$. If U_{α} and U_{β} are two open sets in equation C.1.1, the map $\iota_{\beta} \circ \iota_{\alpha}^{-1}|U_{\alpha} \cap U_{\beta}$ has the form

(C.1.2) $\iota_{\beta} \circ \iota_{\alpha}^{-1}|U_{\alpha} \cap U_{\beta} = (1 \times \varphi_{\beta,\alpha}): U_{\alpha} \cap U_{\beta} \times \mathbb{A}^n \to U_{\beta} \cap U_{\alpha} \times \mathbb{A}^n$

    where $\varphi_{\beta,\alpha}: U_{\alpha} \cap U_{\beta} \to GL(k,n)$ is compatible with $\mathcal{O}_X(U_{\alpha} \cap U_{\beta})$ (see definition C.1.1 on the previous page), called a *transition function*. It is not hard to see that transition functions must satisfy the consistency conditions

(C.1.3)
$$
\begin{aligned}
\varphi_{\gamma,\beta} \circ \varphi_{\beta,\alpha} &= \varphi_{\gamma,\alpha} \\
\varphi_{\beta,\alpha} &= \varphi_{\alpha,\beta}^{-1}
\end{aligned}
$$

    on $U_{\alpha} \cap U_{\beta} \cap U_{\gamma}$.

The number n is called the *rank* of the vector bundle, and the map p is called its *projection*.

    A vector-bundle of rank 1 (i.e. $n = 1$) is called a *line bundle*. The space, W, is called the *total-space* of the vector-bundle and X is called the *base-space*.

    REMARK. The standard definition of vector-bundle makes no reference to sheaves and simply requires maps $\varphi_{\alpha,\beta}$ and $\psi_{\alpha,\beta}$ to be continuous, smooth, or algebraic. All of these special cases are covered by using the appropriate sheaf.

    Note the notation $(1 \times \varphi_{\beta,\alpha}): U_{\alpha} \cap U_{\beta} \times \mathbb{A}^n \to U_{\beta} \cap U_{\alpha} \times \mathbb{A}^n$ where "$U_{\alpha} \cap U_{\beta}$" means $U_{\alpha} \cap U_{\beta} \subset U_{\alpha}$ — i.e. the intersection, *regarded* as part of the U_{α} chart — and $U_{\beta} \cap U_{\alpha}$ represents the *same* intersection regarded as part of the U_{β} chart.

    The following is straightforward:

PROPOSITION C.1.4. *If $(X, \mathcal{O}_X)$ is a locally ringed space with an open cover*

$$X = \bigcup_{\alpha} U_{\alpha}$$
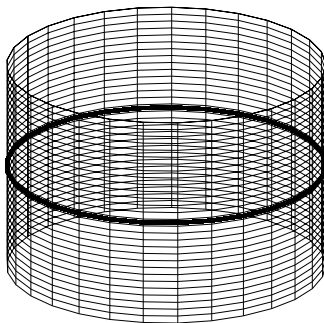
FIGURE C.1.1.   A trivial line-bundle over a circle

and

$$\varphi_{\alpha,\beta} \colon U_\alpha \cap U_\beta \to GL(n,k)$$

are a set of functions satisfying the consistency conditions (equations C.1.3 on the facing page) and compatible with $\mathcal{O}_X(U_\alpha \cap U_\beta)$, we can construct a vector bundle by forming the union and taking equivalence classes

$$\bigcup_\alpha U_\alpha \times \mathbb{A}^n \, / \sim$$

where $\sim$ is the equivalence relation defined on $U_\alpha \cap U_\beta$ that identifies $(u,v) \sim (u, \varphi_{\beta,\alpha}(v))$, where $u \in U_\alpha \cap U_\beta$, $v \in \mathbb{A}^n$ and (u,v) and $(u, \varphi_{\beta,\alpha}(v))$ are regarded as elements of $U_\alpha \times \mathbb{A}^n$ and $U_\beta \times \mathbb{A}^n$, respectively.

EXAMPLE C.1.5. The simplest example of a vector-bundle is a product

$$V \times \mathbb{A}^n$$

This is called a *trivial bundle* — see figure C.1.1 for an example of a trivial bundle over a circle. Its transition functions

$$\varphi_{\alpha,\beta} \colon U_\alpha \cap U_\beta \to GL(k,n)$$

can be regarded as sending every point $u \in U_\alpha \cap U_\beta$ to the *identity* matrix.

Condition 2 in definition C.1.3 on page 507 means that every vector-bundle is *locally trivial*.

EXAMPLE C.1.6. The vector-bundle (a Möbius strip!) in figure C.1.3 on the next page can be constructed by decomposing the circle into a union of two line-segments

$$S^1 = U_1 \cup U_2$$

such that $U_1 \cap U_2 = W_1 \cup W_2$, as in figure C.1.2 on the following page and defining $\varphi_{1,2}$ by

(C.1.4) $$\varphi_{1,2}(u) = \begin{cases} \times 1 & \text{if } u \in W_1 \\ \times -1 & \text{if } u \in W_2 \end{cases}$$
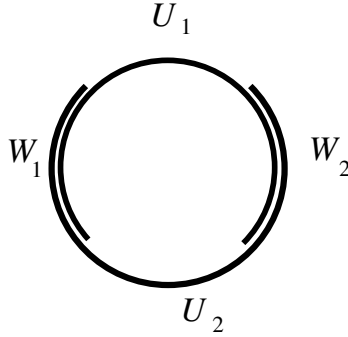
FIGURE C.1.2.   Constructing a vector bundle by patching
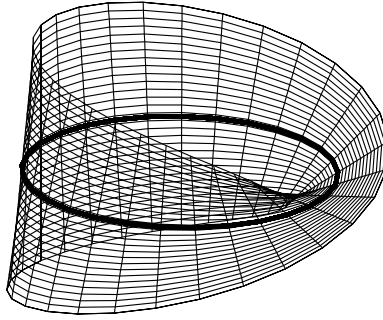
FIGURE C.1.3.   A nontrivial line-bundle over a circle

We can also consider the set of vector-bundles over a fixed base space. They form a category with morphisms defined by

DEFINITION C.1.7. Let $\mathscr{B}(X, \mathcal{O}_X)$ denote the category of vector-bundles over the locally ringed space, $(X, \mathcal{O}_X)$, where a morphism $\xi_1 \to \xi_2$ is a map F

$$F: W_1 \quad \to \quad W_2$$

that:

(1) makes the diagram

$$
\begin{array}{ccc}
W_1 & \xrightarrow{\ F\ } & W_2 \\
{\scriptstyle \xi_1}\downarrow & & \downarrow{\scriptstyle \xi_2} \\
X & =\!=\!=\!= & X
\end{array}
$$

commute,

(2) with $F|\xi_1^{-1}(v): \xi_1^{-1}(v) \to \xi_2^{-1}(v)$ is a linear map of vector-spaces for every $v \in X$, and

(3) if $\{U_\alpha\}$ is a trivializing cover of ξ_1 with charts $\iota_\alpha: \xi_1^{-1}(U_\alpha) \to U_\alpha \times \mathbb{A}^n$, $\{V_\beta\}$ is one for ξ_2 with charts $\iota'_\beta: \xi_2(V_\beta) \to V_\beta \times \mathbb{A}^m$ then

$$\iota'_\beta \circ F \circ \iota_\alpha^{-1}: U_\alpha \cap V_\beta \times \mathbb{A}^n \to U_\alpha \cap V_\beta \times \mathbb{A}^m$$

is of the form $\iota'_\beta \circ F \circ \iota_\alpha^{-1} \quad = \quad (1 \times \psi_{\alpha,\beta})$ and $\psi_{\alpha,\beta} \colon U_\alpha \cap V_\beta \quad \to \quad \hom(\mathbb{A}^m, \mathbb{A}^m)$ is compatible with $\mathcal{O}_V(U_\alpha \cap V_\beta)$.

REMARK. Figures C.1.1 on page 509 and C.1.3 on the facing page are two line-bundles over the circle that turn out to not be isomorphic.

Up to isomorphism, vector-bundles are determined by their transition-functions:

LEMMA C.1.8. *Let ξ be a vector-bundle over $(X, \mathcal{O}_X)$ with trivializing cover*

$$X = \bigcup_\alpha U_\alpha$$

charts

$$\iota_\alpha \colon \xi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{A}^n$$

and transition functions

$$\varphi_{\alpha,\beta} \colon U_\alpha \cap U_\beta \to GL(k, n)$$

If ξ' is the vector-bundle created by proposition C.1.4 on page 508 using these transition-functions, there is an isomorphism

$$F \colon \xi \to \xi'$$

PROOF. Simply define

$$F|\xi^{-1}(U_\alpha) = \iota_\alpha \colon \xi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{A}^n$$

$\square$

PROPOSITION C.1.9. *Let $\xi \colon W \to V$ be a vector-bundle over a space, V and let $U \subset V$ be a subspace. Then the fiber of f over U, $\xi^{-1}(U)$, has the structure of a vector-bundle over U. This is called the restriction of W to U and is denoted $\xi|U$.*

PROOF. Clearly, the fiber of $E|U$ over each point of U will still be a vector space and of the same dimension as the rank of E. Local triviality (condition 2 in definition C.1.3 on page 507 is also easy to verify. $\square$

Having attached a vector-space to each point of a variety, we are interested in defining *vector-fields* over the variety.

DEFINITION C.1.10. Let $\xi \colon W \to X$ be a rank-n vector-bundle over a ringed space $(X, \mathcal{O}_X)$ over a field k with trivializing open cover

$$X = \bigcup_\alpha U_\alpha$$

Then a *global section* of f is a morphism $s \colon X \to W$ such that $\xi \circ s = 1 \colon X \to X$ and $\iota_\alpha \circ s = 1 \times \sigma_\alpha |\xi^{-1}(U_\alpha) \colon \xi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{A}^n$ where

$$\iota_\alpha \colon \xi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{A}^n$$

are the trivializing morphisms, and σ_α is compatible with $\mathcal{O}_X$ for all α.

REMARK. All vector-bundles have at least one section: the 0-section, $s_0 \colon V \to W$, where $s(v) = 0 \in \xi^{-1}(v)$ for all points $v \in V$.

Since $\xi^{-1}(v)$ is a *vector space* over k for any point $v \in V$, the set of sections forms a vector space over k:
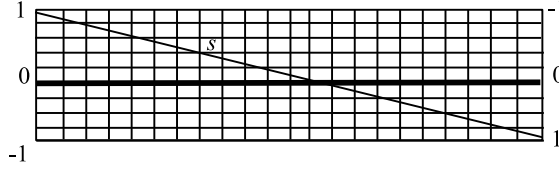
FIGURE C.1.4.   A nontrivial bundle "unrolled"

(1) If $s_1, s_2: V \to W$ are sections of f, $(s_1 + s_2): V \to W$ is the section whose value at a point $v \in V$ is just $s_1(v) + s_2(v) \in \xi^{-1}(v)$,
(2) If $s: V \to W$ is a section of ξ and $x \in k$, $(x \cdot s): V \to W$ is the section with $(x \cdot gs)(v) = x \cdot s(v)$ for all points $v \in V$.

The vector-space of sections of a vector-bundle, ξ, over V is denoted $\Gamma(V, \xi)$. An isomorphism of vector-bundles induces an isomorphism of their vector-spaces of sections.

The vector-bundle in figure C.1.1 on page 509 has a section that is nonzero everywhere — in fact, we can define $s(v) = 1 \in \xi^{-1}(v) = \mathbb{R}^1$. The vector-bundle in figure C.1.3 on page 510 does not, as becomes clear if we cut and "unroll" it as in figure C.1.4 on page 512. Every section of this bundle must vanish somewhere, which shows that the bundles in figures C.1.1 on page 509 and C.1.3 are *not* isomorphic.

PROPOSITION C.1.11. *Let ξ be a vector-bundle of rank n on a ringed space X and let $\{U_\alpha\}$ be a trivializing cover with trivializing functions*

$$\iota_\alpha: \xi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{A}^n$$

and transition functions $\varphi_{\alpha,\beta}: U_\alpha \cap U_\beta \to U_\beta \cap U_\alpha$. Then functions

$$s_\alpha: U_\alpha \to \mathbb{A}^n$$

define sections $\iota_\alpha^{-1} \circ (1 \times s_\alpha): \xi^{-1}(U_\alpha) \to U_\alpha$ that form a global section of ξ if and only if $s_\beta = \varphi_{\beta,\alpha} \circ s_\alpha$ for all α and β.

PROOF. This follows immediately from the definitions of the various terms.   □

If $U \subset V$ is a subvariety, restriction of maps defines a homomorphism

(C.1.5) $\Gamma(V, \xi) \to \Gamma(U, \xi|U)$

of vector-spaces.

Essentially any functorial operations we can perform on vector *spaces* extend to vector-*bundles:*

PROPOSITION C.1.12. *The functors $\oplus$, $\otimes$, and $()^*$ on vector-spaces extend to functors of trivial vector bundles*

$$\xi_1 = X \times \mathbb{A}^n$$
$$\xi_2 = X \times \mathbb{A}^m$$

by defining $\xi_1 \oplus \xi_2 = X \times (\mathbb{A}^n \oplus \mathbb{A}^m)$, $\xi_1 \otimes \xi_2 = X \times (\mathbb{A}^n \otimes \mathbb{A}^m)$, $\xi_1^* = X \times (\mathbb{A}^n)^*$. These extend uniquely to arbitrary vector-bundles.

PROOF. Suppose ξ_1 is an arbitrary vector bundle over X of rank n and ξ_2 is one of rank m. Then:

(1) Find a trivializing cover for both ξ_1 and ξ_2

$$X = \bigcup_\alpha U_\alpha$$

— for instance take the intersection of one for each.

(2) For $\xi_1 \oplus \xi_2$, use proposition C.1.4 on page 508 to glue together

$$U_\alpha \times \mathbb{A}^{n+m}$$

by transition functions

$$\varphi_{\alpha,\beta} \oplus \varphi'_{\alpha,\beta} = \left[\begin{array}{cc} \varphi_{\alpha,\beta} & 0 \\ 0 & \varphi'_{\alpha,\beta} \end{array} \right]$$

— where $\varphi_{\alpha,\beta}$ are the transition functions for ξ_1 and $\varphi'_{\alpha,\beta}$ are those for ξ_2.

(3) For $\xi_1 \otimes \xi_2$ do the same thing with copies of

$$U_\alpha \times \mathbb{A}^{nm}$$

with transition functions $\varphi_{\alpha,\beta} \otimes \varphi'_{\alpha,\beta}$ — the Kronecker product (see exercise 16 on page 473) of the matrices representing $\varphi_{\alpha,\beta}$ and $\varphi'_{\alpha,\beta}$.

(4) For ξ_1^*, use transition functions $\varphi^*_{\alpha,\beta} = \varphi^{\mathrm{tr}}_{\beta,\alpha}$, where $\varphi^{\mathrm{tr}}_{\beta,\alpha}$ is the transpose of the matrix representing $\varphi_{\beta,\alpha}$. The subscripts α and β are interchanged because $()^*$ is a *contravariant functor* (see definition A.5.7 on page 442 and example 2 on page 443): A homomorphism of vector spaces $f\colon V_1 \to V_2$ induces a homomorphism $f^*\colon V_2^* \to V_1^*$.

It is left as an exercise to the reader to show that these define vector bundles. $\square$

The following construction

DEFINITION C.1.13. Let X be a ringed space and let k be a field. The *Picard group* of X, denoted $\mathrm{Pic}(X)$ is the group of isomorphism classes of line-bundles on X with the operation $\otimes$. The trivial bundle is the identity element.

REMARK. The line-bundles form a group under this operation because $\otimes$ is associative on vector-spaces. If ξ is a vector-bundle over X with transition-functions $\varphi_{\alpha,\beta}$, we can construct another vector bundle, η, with transition functions $\varphi^{-1}_{\alpha,\beta}$ and $\xi \otimes \eta$ will be trivial.

Note that the Picard group depends on the field k as well as X.

EXERCISES.

1. Show that a vector-bundle, ξ, over $(X, \mathcal{O}_X)$ of rank n is trivial if and only if it has n linearly independent sections — i.e. n sections $\{s_1, \ldots, s_n\}$ such that, at every point $p \in X$, $\{s_1(p), \ldots, s_n(p)\}$ are linearly independent vectors.

2. If $\xi \in \text{Pic}(X)$ is a line bundle on a space, X, show that ξ^{-1} is the line bundle ξ^*, as in proposition C.1.12 on page 512. This gives a nice functorial description of ξ^{-1}.

3. If ξ is the line-bundle in example C.1.6 on page 509, show that $\xi \otimes \xi$ is a trivial line bundle. If S^1 is a circle and the field of definition for vector-bundles is $\mathbb{R}$, show that $\text{Pic}(S^1) = \mathbb{Z}_2$.

4. Show that $\oplus$ is a coproduct (see definition A.5.4 on page 441) in the category of vector-bundles over $(X, \mathcal{O}_X)$.

5. Suppose

$$0 \to E \to F \to G \to 0$$

is a short exact sequence of vector-bundles of rank e, f, g, respectively. Show that there exists an isomorphism of line-bundles

$$\Lambda^f F \cong \Lambda^e E \otimes \Lambda^g G$$

C.2. Vector-bundles and sheaves

This section makes extensive use of the material in appendix B on page 495. The connection between vector-bundles and sheaves is given by:

DEFINITION C.2.1. Let $\xi \colon W \to X$ be a vector-bundle over a ringed space, $(X, \mathcal{O}_X)$. The *section-sheaf*, $\mathscr{S}_\xi$, is defined by

$$\mathscr{S}_\xi(U) = \Gamma(U, \xi)$$

for all open sets, $U \subset X$ with restriction-maps defined by restriction as in equation C.1.5 on page 512.

REMARK C.2.2. If $U \subset X$ is an open set with the property that $\xi|U$ is trivial, and $p \in U$ is any point, then

$$\iota_U(\mathscr{S}_{\xi,p}) = p \times \mathbb{A}^n$$

where n is the rank of ξ and $\iota_U \colon \xi^{-1}(U) \to U \times \mathbb{A}^n$ is a chart — for instance, the *constant* sections give us this. In this way, we can recover the *vector-spaces* from the sheaf of *sections*.

Sheaves also have associated section-sheaves: $\Gamma(U, \mathcal{O}_V)$. The sections of a sheaf, $\mathcal{O}_X$, is a sheaf, $\Gamma(*, \mathcal{O}_X)$, that is *canonically isomorphic* to $\mathcal{O}_X$ (proposition 3 in section 3 of [**150**]), rendering the notation $\Gamma(*, \mathcal{O}_X)$ superfluous. It is still in wide use, so $\Gamma(U, \mathcal{O}_X)$ should be regarded as *completely synonymous* with $\mathcal{O}_X(U)$. Occasionally, when discussing sheaf cohomology, there will be an advantage to using the section-functor, $\Gamma(U, \mathcal{O}_X) = \mathcal{O}_X(U)$.

If the vector-bundle, ξ, is of rank one — i.e., a *line-bundle* — then the sheaf $\mathscr{S}_\xi$ is a sheaf of invertible functions and is called an *invertible sheaf*.

In order to understand this connection better, we need:

DEFINITION C.2.3. Let $\mathcal{O}_1, \mathcal{O}_2$ be two sheaves over a space, V, where $\mathcal{O}_1$ is a sheaf of rings and $\mathcal{O}_2$ is a sheaf of abelian groups. The sheaf, $\mathcal{O}_2$, is a *module over* $\mathcal{O}_1$, if there exist homomorphisms

$$m_U: \mathcal{O}_1(U) \otimes \mathcal{O}_2(U) \to \mathcal{O}_2(U)$$

of abelian groups for all open sets $U \subset V$ defining $\mathcal{O}_2(U)$ as a module over $\mathcal{O}_1(U)$. In addition, these module structure-maps must be compatible with restriction-homomorphisms, i.e. for every inclusion of open sets $U' \subset U$, the diagram

$$
\begin{array}{ccc}
\mathcal{O}_1(U) \otimes \mathcal{O}_2(U) & \xrightarrow{\;\;m_U\;\;} & \mathcal{O}_2(U) \\
{\scriptstyle r_{U,U'} \otimes s_{U,U'}}\downarrow & & \downarrow{\scriptstyle s_{U,U'}} \\
\mathcal{O}_1(U') \otimes \mathcal{O}_2(U') & \xrightarrow[\;\;m_{U'}\;\;]{} & \mathcal{O}_2(U')
\end{array}
$$

commutes, where $r_{U,U'}: \mathcal{O}_1(U) \to \mathcal{O}_1(U')$ and $s_{U,U'}: \mathcal{O}_2(U) \to \mathcal{O}_2(U')$ are the restriction-homomorphisms of $\mathcal{O}_1$ and $\mathcal{O}_2$. respectively.

A *homomorphism*

$$g: \mathcal{O}_2 \to \mathcal{O}_3$$

of modules over $\mathcal{O}_1$ is a homomorphism of modules $g(U): \mathcal{O}_2(U) \to \mathcal{O}_3(U)$ for all open sets U that is compatible with restriction-homomorphisms.

The sheaf, $\mathcal{O}_2$, is a *free module over* $\mathcal{O}_1$, if

$$\mathcal{O}_2 \cong \bigoplus_{i=1}^{n} \mathcal{O}_1$$

for some integer $n > 0$, which is called the rank of the free module.

The sheaf, $\mathcal{O}_2$, is a *locally free module over* $\mathcal{O}_1$ if there exists an open cover

$$V = \bigcup_\alpha U_\alpha$$

such that $\mathcal{O}_2(U_\alpha)$ is a free module over $\mathcal{O}_1(U_\alpha)$ for all α.

REMARK. If $\mathcal{O}_2$ is locally free over $\mathcal{O}_1$ and V is connected, it is not hard to see that the rank of $\mathcal{O}_2(U_\alpha)$ over $\mathcal{O}_1(U_\alpha)$ is the same for all U_α.

We can clearly extend operations on modules like $\oplus$ and $\otimes$ to modules over a sheaf. The direct sum, $\oplus$, is clearly a coproduct in the category of modules over a sheaf.

PROPOSITION C.2.4. *If $\xi: W \to X$ is a rank-n vector-bundle over a ringed space, $(X, \mathcal{O}_X)$, then the sheaf of sections, $\mathscr{S}_\xi$ module over $\mathcal{O}_X$ that is free (of rank n) if ξ is trivial. In general $\Gamma(\xi)$ is locally free of rank n.*

Any morphism of vector-bundles

$$F: \xi_1 \to \xi_2$$

induces a homomorphism

$$\mathscr{S}_F: \mathscr{S}_{\xi_1} \to \mathscr{S}_{\xi_2}$$

of modules over $\mathcal{O}_X$.

PROOF. It is not hard to see that $\Gamma(\xi)$ is naturally a module over $\mathcal{O}_X$. If ξ is trivial, then

$$W = X \times \mathbb{A}^n$$

and any section is of the form

$$s = (1 \times \sigma): X \to X \times \mathbb{A}^n$$

where $\sigma: X \to \mathbb{A}^n$ is compatible with $\mathcal{O}_X$. The composite of σ with the n projections $\mathbb{A}^n \to \mathbb{A}^1 = k$ give rise to n elements of $\mathcal{O}_X$ so

$$\mathscr{S}_\xi = \mathcal{O}_X^n$$

In general, X has a trivializing cover

$$X = \bigcup_\alpha U_\alpha$$

with $\xi|U_\alpha$ a trivial bundle so that $\Gamma(\xi)(U_\alpha) = \mathcal{O}_X(U_\alpha)^n$, making $\Gamma(\xi)$ *locally free*. $\square$

It is interesting that the converse is also true:

PROPOSITION C.2.5. *Let ξ_1, ξ_2 be two vector-bundles over a ringed space $(X, \mathcal{O}_X)$ of ranks n and m, respectively. Then any homomorphism of modules*

$$g: \mathscr{S}_{\xi_1} \to \mathscr{S}_{\xi_2}$$

over $\mathcal{O}_X$ induces a morphism of vector-bundles

$$\bar{g}: \xi_1 \to \xi_2$$

such that $g = \mathscr{S}_{\bar{g}}$ (in the notation of proposition C.2.4 on the preceding page).

PROOF. Let

$$X = \bigcup_\alpha U_\alpha$$

be a trivializing open cover (see equation C.1.1 on page 508) for both ξ_1 and ξ_2 (i.e., intersect trivializing covers for the two vector-bundles) and with sets of transition-functions $\{\varphi_{\alpha,\beta}\}$ and $\{\varphi'_{\alpha,\beta}\}$, respectively. Then the map

$$\iota'_\alpha \circ g(U_\alpha) \circ \iota_\alpha^{-1}|p: \iota_\alpha(\mathscr{S}_{\xi_1,p}) = \mathbb{A}^n \to \iota'_\alpha(\mathscr{S}_{\xi_2,p}) = \mathbb{A}^m$$

is a linear homomorphism of vector spaces for any point $p \in U_\alpha$ — see remark C.2.2 on page 514. It follows that $g(U_\alpha)$ can be regarded as defining a map

(C.2.1) $$\iota'_\alpha \circ g(U_\alpha) \circ \iota_\alpha^{-1}: U_\alpha \to \hom(\mathbb{A}^n, \mathbb{A}^m)$$

compatible with $\mathcal{O}_X$. We can define

$$\bar{g}_\alpha = 1 \times (\iota'_\alpha \circ g(U_\alpha) \circ \iota_\alpha^{-1}): U_\alpha \times \mathbb{A}^n \to U_\alpha \times \mathbb{A}^m$$

We claim that the maps $\{\bar{g}_\alpha\}$ are compatible with the transition functions on $U_\alpha \cap U_\beta$ in the sense that the diagram

$$
\begin{array}{ccc}
U_\alpha \cap U_\beta \times \mathbb{A}^n & \xrightarrow{\bar{g}_\alpha} & U_\alpha \cap U_\beta \times \mathbb{A}^m \\
{\scriptstyle 1 \times \varphi_{\alpha,\beta}} \downarrow & & \downarrow {\scriptstyle 1 \times \varphi'_{\alpha,\beta}} \\
U_\alpha \cap U_\beta \times \mathbb{A}^n & \xrightarrow[\bar{g}_\beta]{} & U_\alpha \cap U_\beta \times \mathbb{A}^m
\end{array}
$$

commutes. This follows from the fact that $1 \times \varphi_{\alpha,\beta} = \iota_\beta \circ \iota_\alpha^{-1}$ and $1 \times \varphi'_{\alpha,\beta} = \iota'_\beta \circ (\iota'_\alpha)^{-1}$ (see equation C.1.2 in definition C.1.3 on page 507), so

$$
\begin{aligned}
(1 \times \varphi'_{\alpha,\beta}) \circ \bar{g}_\alpha &= (1 \times \varphi'_{\alpha,\beta}) \circ (1 \times \iota'_\alpha \circ g(U_\alpha) \circ \iota_\alpha^{-1}) \\
&= 1 \times \left(\iota'_\beta \circ (\iota'_\alpha)^{-1} \circ \iota'_\alpha \circ g(U_\alpha) \circ \iota_\alpha^{-1} \right) \\
&= 1 \times \left(\iota'_\beta \circ g(U_\alpha) \circ \iota_\alpha^{-1} \right) \\
&= 1 \times \left(\iota'_\beta \circ g(U_\alpha) \circ \iota_\beta^{-1} \circ \iota_\beta \circ \iota_\alpha^{-1} \right) \\
&= (1 \times \iota'_\beta \circ g(U_\alpha) \circ \iota_\beta^{-1}) \circ (1 \times \varphi_{\alpha,\beta}) \\
&= (1 \times \iota'_\beta \circ g(U_\beta) \circ \iota_\beta^{-1}) \circ (1 \times \varphi_{\alpha,\beta}) \quad \text{since } g(U_\alpha) = g(U_\beta) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{on } U_\alpha \cap U_\beta \\
&= \bar{g}_\beta \circ (1 \times \varphi_{\alpha,\beta})
\end{aligned}
$$

It follows that the $\{\bar{g}_\alpha\}$ patch together to define a morphism $\bar{g} \colon \xi_1 \to \xi_2$.

It is also not hard to see that the map of section-sheaves will coincide with g: on U_α form the composite $(\iota'_\alpha)^{-1} \circ \bar{g}_\alpha \circ \iota_\alpha$ and we recover $g(U_\alpha)$. $\qquad\square$

The previous result implies that there is an equivalence of categories between the category of vector bundles over $(X, \mathcal{O}_X)$ and that of modules over $\mathcal{O}_X$ that are section-sheaves of vector bundles. Our next step will be to identify these:

PROPOSITION C.2.6. *Let $(X, \mathcal{O}_X)$ be a connected, locally ringed space with $\mathfrak{m}_p \subset \mathcal{O}_{X,p}$ the unique maximal ideal and such that $\mathcal{O}_{X,p}/\mathfrak{m}_p = k$ for all points $p \in X$. If M is a locally free module over $\mathcal{O}_X$ of finite rank, then M is isomorphic to the section-sheaf of a vector bundle over $(X, \mathcal{O}_X)$.*

PROOF. Since M is locally free, there is an open covering

$$
X = \bigcup_\alpha U_\alpha
$$

such that $M(U_\alpha)$ is free over $\mathcal{O}_X(U_\alpha)$. Since X is connected, the rank of $M(U_\alpha)$ over $\mathcal{O}_X(U_\alpha)$ is the same for all α, say $n > 0$, and

$$
\frac{M_p}{\mathfrak{m}_p \cdot M_p} = \mathbb{A}^n
$$

for all $p \in X$. The restriction-maps

$$
\begin{aligned}
p_{U_\alpha, U_\alpha \cap U_\beta} \colon M(U_\alpha) = \mathcal{O}_X(U_\alpha)^n &\to M(U_\alpha \cap U_\beta) = \mathcal{O}_X(U_\alpha \cap U_\beta)^n \\
p_{U_\beta, U_\alpha \cap U_\beta} \colon M(U_\beta) = \mathcal{O}_X(U_\beta)^n &\to M(U_\alpha \cap U_\beta) = \mathcal{O}_X(U_\alpha \cap U_\beta)^n
\end{aligned}
$$

induce isomorphisms

$$
\begin{aligned}
f_\alpha(p) \colon \frac{M(U_\alpha)_p}{\mathfrak{m}_p \cdot M(U_\alpha)_p} = \mathbb{A}^n &\to \frac{M(U_\alpha \cap U_\beta)_p}{\mathfrak{m}_p \cdot M(U_\alpha \cap U_\beta)_p} = \mathbb{A}^n \\
f_\beta(p) \colon \frac{M(U_\beta)_p}{\mathfrak{m}_p \cdot M(U_\alpha)_p} = \mathbb{A}^n &\to \frac{M(U_\alpha \cap U_\beta)_p}{\mathfrak{m}_p \cdot M(U_\alpha \cap U_\beta)_p} = \mathbb{A}^n
\end{aligned}
$$

for any point $p \in U_\alpha \cap U_\beta$. If we define $\varphi_{\alpha,\beta}(p) = f_\beta^{-1}(p) \circ f_\alpha(p)$, we get maps

$$\varphi_{\alpha,\beta} \colon U_\alpha \cap U_\beta \to GL(k,n)$$

compatible with $\mathcal{O}_X$ (since they are induced by restriction-maps of $\mathcal{O}_X$). It is left as an exercise to the reader to verify that these satisfy all of the compatibility conditions of transition functions (see equation C.1.3 on page 508). It follows that we can glue the $U_\alpha \times \mathbb{A}^n$ together as in proposition C.1.4 on page 508 to get a vector-bundle, ξ, on X.

It is not hard to see that every element $e \in M$ defines a section of ξ — its value at a point p is its image in $\mathbb{A}^n = M_p/\mathfrak{m}_p \cdot M_p$. Conversely, any section of ξ defines sections of $\xi|U_\alpha = U_\alpha \times \mathbb{A}^n$ for all α that are compatible with restriction-maps of M, so it defines an element of M. $\square$

We summarize these results by:

THEOREM C.2.7. *Let $(X, \mathcal{O}_X)$ be a connected, locally ringed space with $\mathfrak{m}_p \subset \mathcal{O}_{X,p}$ the unique maximal ideal and such that $\mathcal{O}_{X,p}/\mathfrak{m}_p = k$ for all points $p \in X$. Then the section-sheaf functor defines an equivalence of categories between the category of vector-bundles over $(X, \mathcal{O}_X)$ and that of locally-free modules over $\mathcal{O}_X$ of finite rank.*

EXERCISES.

1. If ξ_1 and ξ_2 are vector-bundles over $(X, \mathcal{O}_X)$ show that

$$\mathscr{S}_{\xi_1 \oplus \xi_2} = \mathscr{S}_{\xi_1} \oplus \mathscr{S}_{\xi_2}$$

2. Suppose $(X, \mathcal{O}_X)$ is a ringed space and

$$0 \to \mathscr{E} \to \mathscr{F} \to \mathscr{G} \to 0$$

is a short exact sequence of locally free modules over $\mathcal{O}_X$ of ranks e, f, g, respectively. Show that there exists an isomorphism of invertible sheaves over $\mathcal{O}_X$

$$\Lambda^f \mathscr{F} \cong \Lambda^e \mathscr{E} \otimes_{\mathcal{O}_X} \Lambda^g \mathscr{G}$$

Cohomology

"God exists since mathematics is consistent, and the Devil exists since
we cannot prove it."
—André Weil, *as quoted in* [**46**].

D.1. Chain complexes and cohomology

Homology theory is one of the pillars of algebraic topology and a variant
called *sheaf cohomology* is widely used in algebraic geometry. The first step to
developing this theory involves defining cochain complexes — a purely alge-
braic construct that will be coupled to geometry later.

We will assume all objects here are in a fixed abelian category (see sec-
tion A.5.5 on page 457), $\mathscr{A}$. For instance, they could be abelian groups or mod-
ules over any ring, or even certain types of sheaves.

We begin with the most basic construct:

DEFINITION D.1.1. A *chain complex* (C_i, ∂_i) is a sequence of objects of $\mathscr{A}$ and
homomorphisms

$$\cdots \to C_{i+1} \xrightarrow{\partial_{i+1}} C_i \xrightarrow{\partial_i} C_{i-1} \to \cdots$$

where, for all i, $\partial_i \circ \partial_{i+1} = 0$. A *morphism* of cochain complexes $\{f_i\}: (C_i, \partial_i) \to$
(D_i, ∂_i') (or *chain-map*) is a sequence of homomorphisms

$$f_i: C_i \to D_i$$

such that the diagrams

(D.1.1)
$$\begin{CD} C_i @>f_i>> D_i \\ @V\partial_i VV @VV\partial_i' V \\ C_{i-1} @>>f_{i-1}> D_{i-1} \end{CD}$$

commute for all i. The maps, ∂_i, are called the *boundary maps* or *differentials*
of the chain-complex. The category of chain-complexes with chain-maps as
morphisms is denoted $\mathscr{C}h$.

REMARK. The condition $\partial_{i-1} \circ \partial_i = 0$ implies that $\operatorname{im} \partial_i \subseteq \ker \partial_{i-1}$. In alge-
braic topology, chain-complexes are geometrically defined objects that contain
a great deal of topological information.

Now we define a dual concept that is very similar:

DEFINITION D.1.2. A *cochain complex* (C^i, δ_i) is a sequence of objects of $\mathscr{A}$ and homomorphisms

$$\cdots \to C^{i-1} \xrightarrow{\delta_{i-1}} C^i \xrightarrow{\delta_i} C^{i+1} \to \cdots$$

where, for all i, $\delta_{i+1} \circ \delta_i = 0$. A *morphism* of cochain complexes $\{f_i\}: (C^i, \delta_i) \to (D^i, \delta_i')$ (or *chain-map*) is a sequence of homomorphisms

$$f_i: C^i \to D^i$$

such that the diagrams

(D.1.2)
$$\begin{array}{ccc} C^{i+1} & \xrightarrow{f_{i+1}} & D^{i+1} \\ \delta_i \uparrow & & \uparrow \delta_i' \\ C^i & \xrightarrow{f_i} & D^i \end{array}$$

commute for all i. The maps, δ_i, are called the *coboundary maps* or *codifferentials* of the cochain-complex. The category of cochain-complexes with chain-maps as morphisms is denoted $\mathscr{C}o$.

REMARK. The superscripts are not exponents! At this point, the reader may wonder what essential difference exists between chain complexes and cochain complexes. The answer is "none!" We can define cochain-complexes as chain-complexes with *negative subscripts:*

$$C^i = C_{-i}$$

(or equivalently, defining chain-complexes as cochain-complexes with negative superscripts). Anything we can prove for one is valid for the other under this equivalence.

Historically, chain-complexes appeared first and were geometrically defined. The generators of the C_i were i-dimensional building blocks for a topological space and the ∂_i mapped one of these to its boundary. Cochain complexes appeared later as sets of functions one could define on these building blocks.

In actual applications (in the next section), this symmetry will break down to some extent and they both will express complementary information. We will give greater emphasis to *cochain* complexes because they are the ones that are most significant in algebraic geometry.

The condition $\delta_{i+1} \circ \delta_i = 0$ implies that $\operatorname{im} \delta_i \subseteq \ker \delta_{i+1}$ for all i. With this in mind, we can define

DEFINITION D.1.3. Given:

- a chain-complex, (C_i, ∂_i), we can define its *homology groups*, $H_i(C)$ via

$$H_i(C) = \frac{\ker \partial_i}{\operatorname{im} \partial_{i+1}}$$

- a cochain complex (C^i, δ_i), we can define its associated *cohomology groups*, $H^i(C)$, via

$$H^i(C) = \frac{\ker \delta_i}{\operatorname{im} \delta_{i-1}}$$

REMARK. These will also be objects in the category $\mathscr{A}$. If $H_i(C) = 0$, then the original chain complex was an exact sequence. Such chain-complexes are said to be *exact* or *acyclic*. A similar definition exists for cochain complexes.

Historically, $H_i(C)$ measured the number of i-dimensional "holes" a topological space had (so an n-sphere has $H_n = \mathbb{Z}$ and $H_i = 0$ for $0 < i < n$).

Note that the diagrams D.1.2 on the preceding page imply that chain maps preserve images and kernels of the boundary or coboundary homomorphisms. This implies that

PROPOSITION D.1.4. *A chain map or morphism:*
- *of chain complexes* $\{f_i\}\colon (C_i, \partial_i) \to (D_i, \partial_i')$ *induces homomorphisms of homology*

$$f_*^i\colon H_i(C) \to H_i(D)$$

or
- *of cochain-complexes* $\{f_i\}\colon (C^i, \delta_i) \to (D^i, \delta_i')$ *induces homomorphisms of cohomology groups*

$$f_i^*\, H^i(C) \to H^i(D)$$

Next, we consider a property of chain maps:

DEFINITION D.1.5. Two
- chain maps $f, g\colon (C, \partial_C) \to (D, \partial_D)$ of chain-complexes are said to be *chain-homotopic* if there exists a set of homomorphisms

$$\Phi_i\colon C^i \to D^{i+1}$$

for all $i > 0$ called a *homotopy*, such that

$$f_i - g_i = \Phi_{i-1} \circ \partial_C + \partial_D \circ \Phi_i$$

- chain maps $f, g\colon (C, \delta_C) \to (D, \delta_D)$ of cochain-complexes are said to be chain-homotopic if there exists a set of homomorphisms

$$\Phi_i\colon C^i \to D^{i-1}$$

called a *cohomotopy* for all $i > 0$, such that

$$f_i - g_i = \Phi_{i+1} \circ \delta_C + \delta_D \circ \Phi_i$$

REMARK. Chain-homotopy clearly defines an equivalence relation on chain-maps. Although the definition seems odd, the maps Φ arise naturally in certain topological settings.

The main significance of chain-homotopy is that:

PROPOSITION D.1.6. *If* $f, g\colon (C, \delta_C) \to (D, \delta_D)$ *are chain-homotopic chain-maps of cochain complexes then*

$$f^* = g^*\colon H^i(C) \to H^i(D)$$

REMARK. A corresponding result exists for chain-complexes, by "reversing all the arrows." The symmetry between chain and cochain complexes persists.

PROOF. If $x \in H^i(C)$, then there exists an element $y \in \ker(\delta_C)_i \subset C^i$ such that $x \equiv y \pmod{\operatorname{im}(\delta_C)_{i-1}}$. If we evaluate $(f - g)(y)$, we get

$$
\begin{aligned}
(f - g)(y) &= (\Phi \circ \delta_C + \delta_D \circ \Phi)(y) \\
&= \delta_D \circ \Phi(y) \qquad\qquad\text{because } y \in \ker(\delta_C)_i
\end{aligned}
$$

It follows that $f(y) \equiv g(y) \pmod{\operatorname{im}(\delta_D)_i}$ and $f^*(x) = g^*(x) \in H^i(D)$. $\square$

We can also define an equivalence relation on cochain-complexes:

DEFINITION D.1.7. Two cochain-complexes (C, δ_C) and (D, δ_D) are *chain-homotopy equivalent* if there exist chain maps

$$
\begin{aligned}
f &: (C, \delta_C) &\to& \quad (D, \delta_D) \\
g &: (D, \delta_D) &\to& \quad (C, \delta_C)
\end{aligned}
$$

such that $f \circ g : (D, \delta_D) \to (D, \delta_D)$ and $g \circ f : (C, \delta_C) \to (C, \delta_C)$ are both chain-homotopic to their respective identity maps.

REMARK. Clearly, homotopy equivalent cochain-complexes have isomorphic cohomology groups

$$
H^i(C) \cong H^i(D)
$$

for all i. Chain-homotopy equivalence is a much sharper relationship than simply having isomorphic cohomology groups. In a certain sense, (C, δ_C) and (D, δ_D) may be regarded as equivalent in every important respect.

Our final topic in the basic algebra of chain-complexes is:

DEFINITION D.1.8. If (C, δ_C), (D, δ_D), and (E, δ_E) are cochain complexes, an *exact sequence*

$$
0 \to (C, \delta_C) \xrightarrow{f} (D, \delta_D) \xrightarrow{g} (E, \delta_E) \to 0
$$

are chain-maps f, g such that

$$
0 \to C^i \xrightarrow{f_i} D^i \xrightarrow{g_i} E^i \to 0
$$

are exact for all i.

REMARK. Exact sequences of cochain complexes arise in many natural settings and can be used to compute cohomology because of the next result:

PROPOSITION D.1.9. *An exact sequence*

$$
0 \to (C, \delta_C) \xrightarrow{f} (D, \delta_D) \xrightarrow{g} (E, \delta_E) \to 0
$$

of cochain-complexes induces a homomorphism

$$
c \colon H^i(E) \to H^{i+1}(C)
$$

for all i, called the connecting map, that fits into a long exact sequence in cohomology:

$$
\cdots \to H^i(C) \xrightarrow{f^*} H^i(D) \xrightarrow{g^*} H^i(E) \xrightarrow{c} H^{i+1}(C) \to \cdots
$$

Here f^* and g^* are the induced maps and $c \colon H^i(E) \to H^{i+1}(C)$, called the connecting map is defined as

$$
c = f^{-1} \circ \delta_D \circ g^{-1}
$$

or in more detail by

(1) *If $x \in H^i(E)$, then there exists $y \in \ker(\delta_C)_i$ such that $x \equiv y$ (mod im $(\delta_C)_{i-1}$).*

(2) *Since $g: D^i \to E^i$ is surjective, there exists $z \in D^i$ with $g(z) = y$.*

(3) *Now take $(\delta_D)_i(z) = w \in D^{i+1}$. Since $y \in \ker(\delta_C)_i$ and chain-maps commute with coboundaries, $w \in \ker g$.*

(4) *Since the sequence is exact, this w is in the image of (the injective map) f so we may regard $w \in C^{i+1}$.*

(5) *This $w \in \ker(\delta_C)_{i+1}$ because it is in the image of δ_C and its image in D is in $\ker(\delta_D)_{i+1}$ since $(\delta_D)_{i+1} \circ (\delta_D)_i = 0$.*

REMARK. This will turn out to be very useful for *computing* cohomology groups.

PROOF. The proof follows by analyzing the commutative diagram

in a visual process affectionately (or angrily!) called a "diagram chase".

We show that c is well-defined: any two distinct lifts of y to D^i will differ by an element of C^i. Since the right square commutes, the final result will differ by an element of $(\delta_C)^i$, hence define the same element of $H^{i+1}(C)$. If y, y' both represent the same x, they will differ by an element of $(\delta_C)_{i-1}$ and their lifts to D^i will differ by an element of $(\delta_D)_{i-1}$, which will be annihilated when we plug it into $(\delta_D)_i$.

The proof of the remaining assertions (about the sequence being exact) follows by similar arguments and is left to the reader. $\square$

If we define $C_i = C^{-i}$ and $H_i(C) = H^{-i}(C)$, we *immediately* get the corresponding result for exact sequences of *chain*-complexes:

PROPOSITION D.1.10. *Given a short exact sequence of chain-complexes*

$$0 \to (C, \partial_C) \xrightarrow{f} (D, \partial_D) \xrightarrow{g} (E, \partial_E) \to 0$$

of chain-complexes, there exists a homomorphism

$$c: H_i(E) \to H_{i-1}(C)$$

for all i, called the connecting map, that fits into a long exact sequence in cohomology:

$$\cdots \to H_i(C) \xrightarrow{f^*} H_i(D) \xrightarrow{g^*} H_i(E) \xrightarrow{c} H_{i-1}(C) \to \cdots$$

We also need two more basic concepts:

DEFINITION D.1.11. Given a chain-map of cochain complexes

$$f \colon (C, \delta_C) \to (D, \delta_D)$$

the *algebraic mapping cone* of f is a cochain complex defined by

$$\mathbf{A}(f)^n = C^{n+1} \oplus D^n$$

with a differential

$$\delta_A^n = \begin{bmatrix} -\delta_C^{n+1} & 0 \\ f^{n+1} & \delta_D^n \end{bmatrix} \colon \begin{bmatrix} C^{n+1} \\ D^n \end{bmatrix} = \mathbf{A}(f)^n \to \begin{bmatrix} C^{n+2} \\ D^{n+1} \end{bmatrix} = \mathbf{A}(f)^{n+1}$$

and giving a short exact sequence of cochain complexes

(D.1.3) $$0 \to D \to \mathbf{A}(f) \to C[+1] \to 0$$

where $C[+1]$ is C shifted upwards by one degree so $C[+1]^n = C^{n+1}$ with $\delta_{C[+1]}^n = -\delta_C^{n+1}$.

REMARK. It is left as an exercise to the reader to verify that $\delta_A^2 = 0$. As the name hints, this was originally an algebraic version of a geometric construction.

The short exact sequence in D.1.3 induces a long exact sequence in cohomology (proposition D.1.9 on page 522):

$$\cdots \to H^i(D) \to H^i(\mathbf{A}(f)) \to H^i(C[+1]) \to H^{i+1}(D) \to \cdots$$

with $H^i(C[+1]) = H^{i+1}(C)$. Analysis of the connecting map $H^i(C[+1]) \to H^{i+1}(D)$ shows that it is identical to the map in cohomology induced by f so we can rewrite the long exact sequence as

(D.1.4) $$\cdots \to H^i(C) \xrightarrow{f^*} H^i(D) \to H^i(\mathbf{A}(f))$$

$$\to H^{i+1}(C) \xrightarrow{f^*} H^{i+1}(D) \to \cdots$$

D.1.1. "Topological" homology and cohomology. In this section, we will give a crude and (very!) non-rigorous overview of how homology and cohomology were originally developed and what they "mean" — see [**73**] for rigor and more details.

One way of studying topological spaces involved breaking them up into a union of discrete pieces of every dimension, called *simplices*[1], and a chain-complex was constructed from these simplices. The boundary operator actually represented taking the boundary of n-dimensional simplices and expressing them as $n-1$-dimensional simplices. Although the chain-complex one gets from this construction is not unique (far from it!), it can be proved that its homology is.

In dimension 0, $H_0(X; \mathbb{C}) = \mathbb{C}^k$ where k is the number of components of X. Higher dimensional homology encodes the number of d-dimensional "holes" a space has. For instance, if S^n is an n-sphere, then $H_i(S^n; \mathbb{C}) = 0$ for $0 < i < n$ and $H_n(S^n; \mathbb{C}) = \mathbb{C}$.

---

[1]Simplices are essentially polyhedral pieces of Euclidean space. Singular homology and cohomology involves *mappings* of simplices into a space.

Cohomology originally studied the behavior of functions on a topological space — C^i was the set of functions on i-dimensional simplices and the coboundary operator $\delta^i \colon C^i \to C^{i+1}$ determined a function on $i+1$-dimensional simplices by taking its value on the boundary. For instance $H^0(X;\mathbb{C})$ is the set of *locally-constant* functions on X. If X has k components, this is $\mathbb{C}^k$.

In higher dimensions, $H^i(X;\mathbb{C})$ measures the extent to which certain functions on simplices are determined by their behavior on the on the boundaries of those simplices. Roughly speaking, $H^1(\mathbb{R}^2;\mathbb{C}) = 0$ is equivalent to Green's Theorem in multivariate calculus, and $H^2(\mathbb{R}^3;\mathbb{C}) = 0$ is equivalent to the Divergence Theorem.

EXERCISES.

1. If
$$0 \to (C,\delta_C) \xrightarrow{f} (D,\delta_D) \xrightarrow{g} (E,\delta_E) \to 0$$
is an exact sequence of cochain complexes, and two out of the three complexes are acyclic, show that the third must be acyclic also.

2. If
$$0 \to (C,\delta_C) \xrightarrow{f} (D,\delta_D) \xrightarrow{g} (E,\delta_E) \to 0$$
is an exact sequence of cochain-complexes and D is acyclic, show that
$$H^i(E) \cong H^{i+1}(C)$$
for all i.

3. Show that, if two chain-maps $f,g \colon (C,\delta_C) \to (D,\delta_D)$ are chain-homotopic and $F \colon \mathscr{A} \to \mathscr{A}'$ is any additive functor (for instance, $\hom_{\mathscr{A}}(M,*)$ for any $M \in \mathscr{A}$), then the induced chain-maps
$$F(f), F(G) \colon (F(C^i), F(\delta_C)) \to (F(D^i), F(\delta_D))$$
are also chain-homotopic. It follows that, if (C,δ_C) and (D,δ_D) are chain-homotopy *equivalent*, then $(F(C^i), F(\delta_C))$ and $(F(D^i), F(\delta_D))$ also are.

4. Given a commutative diagram

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & A & \xrightarrow{\ r\ } & B & \xrightarrow{\ s\ } & C & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle u} & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle w} & & \\
0 & \longrightarrow & A' & \xrightarrow[\ r'\ ]{} & B' & \xrightarrow[\ s'\ ]{} & C' & \longrightarrow & 0
\end{array}
$$

with exact rows, do a diagram-chase to show that, if u and w are isomorphisms, then so is v. This is a watered-down version of what is called the 5-Lemma.

D.1.2. Resolutions and Derived functors. Now we will consider special types of chain- and cochain-complexes called *resolutions.* We will assume that our abelian category, $\mathscr{A}$, has enough projectives and injectives (see definition A.5.35 on page 459). This is true for abelian groups and modules over any ring, for instance.

Resolutions are used to compute constructs called *derived functors. Roughly speaking,* given a functor $\mathscr{F}\colon\mathscr{A} \to \mathscr{B}$ between abelian categories, the *first* derived functor measures the extent to which $\mathscr{F}$ fails to be *exact* (i.e. map exact sequences to exact sequence) — if it *vanishes,* then the functor is exact. The *second* derived functor (again, roughly speaking) measures the extent to which the *first* derived functor fails to be exact, and so on. See corollary D.1.20 on page 529 for a more precise statement.

In algebraic geometry, we will want to compute derived functors of the *global-sections* functor of sheaves like the sheaf of regular functions (see section D.3 on page 545 and the remarks following definition D.3.2 on page 546).

DEFINITION D.1.12. If $M \in \mathscr{A}$ is an object, a *right resolution*, I^*, of M is a cochain-complex

$$I_0 \xrightarrow{\delta_0} I_1 \xrightarrow{\delta_1} \cdots$$

where there exists a monomorphism $M \to I_0$ that makes the complex

$$0 \to M \to I_0 \xrightarrow{\delta_0} I_1 \xrightarrow{\delta_1} \cdots$$

exact or acyclic. If all of the I_j are injective, this is called an *injective resolution.*

The *injective dimension* of an object, $M \in \mathscr{A}$, denoted inj-dim M, is the largest subscript of the shortest possible injective resolution of M — if M has a finite injective resolution — or ∞.

REMARK. The definition immediately implies that

$$H^i(I^*) = \begin{cases} M & \text{if } i = 0 \\ 0 & \text{otherwise} \end{cases}$$

Since $\mathscr{A}$ has enough injectives, every object has *some* injective resolution:
   (1) set I_0 to some injective containing M and I_1 to an injective object containing I_0/M.
   (2) set I_{j+1} to an injective object containing $I_j/\delta_{j-1}(I_{j-1})$.

We also have projective resolutions — usually not that interesting in the category of *sheaves* because it doesn't have enough projectives:

DEFINITION D.1.13. If $M \in \mathscr{A}$ is an object, a *(left) resolution*, P_*, of M is a chain-complex

$$\cdots \to P_1 \to P_0 \to 0$$

where there exists a epimorphism $P_0 \to M$ that makes the chain complex

$$\cdots \to P_1 \to P_0 \to M \to 0$$

exact or acyclic. A resolution is called a projective resolution if all of the P_i are projective objects. The *projective dimension* of an object, $M \in \mathscr{A}$, denoted proj-dim M is the largest subscript that occurs in a minimal projective resolution — if M has a finite projective resolution — or ∞.

Injective resolutions are by no means unique although they have an interesting property:

PROPOSITION D.1.14. *Suppose $M, N \in \mathscr{A}$ are two objects with right resolutions I^* and J^*, respectively. If J^* is an injective resolution, then any morphism*

$$f \colon M \to N$$

induces a chain-map

$$\hat{f} \colon I^* \to J^*$$

Although $\hat{f}$ is not unique, any two such induced chain-maps are chain-homotopic.

REMARK. This implies that injective resolutions are unique up to chain-homotopy type.

A similar statement can be proved for projective resolutions (reverse all the arrows!).

PROOF. We make extensive use of the property of injective modules described in exercise 11 on page 461. In the diagram

$$
\begin{array}{ccc}
I_0 & \dashrightarrow & J_0 \\
\uparrow & & \uparrow \\
M & \xrightarrow{\;f\;} & N
\end{array}
$$

it is clear that a portion of I_0 maps to J_0 (namely the portion in the image of M). The injective property of J_0 implies that this extends to all of I_0. In a similar fashion, we inductively construct the chain-map $\hat{f}$ in all higher dimensions.

Suppose g_1 and g_2 are two chain-maps $g_1, g_2 \colon (I^*, \delta) \to (J^*, \sigma)$ that cover the same map $f \colon M \to N$. It follows that $g = g_1 - g_2$ is chain-map that covers the *zero* map. We will show that it is homotopic to zero, i.e. there exists a map

$$\Phi_i \colon I_i \to J_{i-1}$$

such that

(D.1.5) $$g_i = \Phi_{i+1} \circ \delta_i + \sigma_{i-1} \circ \Phi_i$$

Since M maps to 0, we have that $g_0 \colon I_0 \to J_0$ maps the kernel of δ_0 to 0, which means that it maps $\operatorname{im} \delta_0 \subset I_1$ to J_0. The injective property of J_1 implies that this extends to *all* of I_1, giving

$$\Phi_1 \colon I_1 \to J_0$$

with $g_0 = \Phi_1 \circ \delta_0$. Suppose equation D.1.5 is true for degrees $< t$. In degree t, consider

$$g_t - \sigma_{t-1} \circ \Phi_t \colon I_t \to J_t$$

$$
\begin{aligned}
(g_t - \sigma_{t-1} \circ \Phi_t) \circ \delta_{t-1} &= g_t \circ \delta_{t-1} - \sigma_{t-1} \circ \Phi_t \circ \delta_{t-1} \\
&= \sigma_{t-1} \circ g_{t-1} - \sigma_{t-1} \circ \Phi_t \circ \delta_{t-1} \\
&= \sigma_{t-1} \circ \Phi_t \circ \delta_{t-1} + \sigma_{t-1} \circ \sigma_{i-2} \circ \Phi_{i-1} \\
&\quad - \sigma_{t-1} \circ \Phi_t \circ \delta_{t-1} \\
&= \sigma_{t-1} \circ \Phi_t \circ \delta_{t-1} - \sigma_{t-1} \circ \Phi_t \circ \delta_{t-1} \\
&= 0
\end{aligned}
$$

So $(g_t - \sigma_{t-1} \circ \Phi_t)|\operatorname{im}\delta_{t-1} = 0$ which means that $(g_t - \sigma_{t-1} \circ \Phi_t)|\ker\delta_t = 0$. The argument used above implies that $g_t - \sigma_{t-1} \circ \Phi_t$ defines a map $\Phi_{t+1}\colon I_{t+1} \to J_t$ such that

$$g_t - \sigma_{t-1} \circ \Phi_t = \Phi_{t+1} \circ \delta_t$$

The conclusion follows.   □

Reversing the arrows proves the chain-complex version

PROPOSITION D.1.15. *Suppose $M, N \in \mathscr{A}$ are two objects with left resolutions P_* and Q_*, respectively. If P_* is a projective resolution, then any morphism*

$$f\colon M \to N$$

induces a chain-map

$$\hat{f}\colon P_* \to Q_*$$

Although $\hat{f}$ is not unique, any two such induced chain-maps are chain-homotopic.

REMARK. This implies that projective resolutions are unique up to chain-homotopy type.

We will be interested in functors $F\colon \mathscr{A} \to \mathscr{A}'$ to other abelian categories:

DEFINITION D.1.16. A functor $F\colon \mathscr{A} \to \mathscr{A}'$ is *left-exact* if an exact sequence

$$0 \to A \xrightarrow{r} B \xrightarrow{s} C \to 0$$

in $\mathscr{A}$ implies that the sequence

$$0 \to F(A) \xrightarrow{F(r)} F(B) \xrightarrow{F(s)} F(C)$$

is exact.

REMARK. Exercise 12 on page 461 shows that $\hom_{\mathscr{A}}(A, *)$ is left-exact.
The inclusion of the category of sheaves into that of presheaves is left exact. Example B.2.1 on page 498 shows why it is not also right-exact.

Since injective resolutions are *unique* up to chain-homotopy type, the solution to exercise 3 on page 525 that the following constructs will be well-defined:

DEFINITION D.1.17. If $F\colon \mathscr{A} \to \mathscr{A}'$ is a left-exact functor and $C \in \mathscr{A}$ has an injective resolution $I_0 \to \cdots$, then the *right derived functors* of F are

$$R^i F(C) = H^i(F(I^*)) \in \mathscr{A}'$$

for $i \geq 0$.

REMARK. It is not hard to see that $R^0 FC = C$. The $\{R^i F(C)\}$ for $i > 0$ essentially measure how much F fails to be right-exact.

DEFINITION D.1.18. If $M, N \in \mathscr{A}$ and I^* is an injective resolution of N, the cohomology *groups* (in $\mathscr{A}b$)

$$\operatorname{Ext}^i_R(M, N) = H^i(\hom_{\mathscr{A}}(M, I))$$

depend only on M and N and are functorial.

To analyze the behavior of derived functors, we need the following result

LEMMA D.1.19 (Injective Horseshoe Lemma). *Suppose*

$$0 \to A \xrightarrow{r} B \xrightarrow{s} C \to 0$$

is a short exact sequence in $\mathscr{A}$ and I^* and J^* are injective resolutions of A and C, respectively. Then there exists an injective resolution W^* of B fitting into a commutative diagram

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & A & \xrightarrow{\ r\ } & B & \xrightarrow{\ s\ } & C & \longrightarrow & 0 \\
 & & \downarrow{\epsilon_A} & & \downarrow{\epsilon_B} & & \downarrow{\epsilon_C} & & \\
0 & \longrightarrow & I^* & \xrightarrow{\ u\ } & W^* & \xrightarrow{\ v\ } & J^* & \longrightarrow & 0
\end{array}
$$

where the bottom row is a short exact sequence of chain-complexes.

PROOF. Clearly, $W^n = I^n \oplus J^n$ for all n. The map $\epsilon_A \colon A \to I^0$ shows that a *sub*-object of B maps to I^0. Injectivity implies that this map extends to *all* of B, so we get $\iota \colon B \to I^0$ and can define

$$\epsilon_B = \iota \oplus \epsilon_C \circ s \colon B \to I^0 \oplus J^0 = W^0$$

We claim that this is injective: if $b \in B$ maps to zero, it must map to zero in J^0 so that $s(b) = 0$. This means that $b \in \operatorname{im} A$ which maps to I^0 via the injective map, ϵ_A.

Suppose this exact sequence of resolutions has been constructed up to degree n so we have

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & I^n/\operatorname{im}\delta_A & \xrightarrow{\ r\ } & W^n/\operatorname{im}\delta_B & \xrightarrow{\ s\ } & J^n/\operatorname{im}\delta_C & \longrightarrow & 0 \\
 & & \downarrow{\delta_A^n} & & \downarrow{f} & & \downarrow{\delta_C^n} & & \\
0 & \longrightarrow & I^{n+1} & \xrightarrow{\ u\ } & W^{n+1} & \xrightarrow{\ v\ } & J^{n+1} & \longrightarrow & 0
\end{array}
$$

where the vertical maps are inclusions. Now construct f *exactly* the way ϵ_B was constructed above. $\qquad\square$

This immediately implies that

COROLLARY D.1.20. *If*

$$0 \to A \xrightarrow{r} B \xrightarrow{s} C \to 0$$

is a short exact sequence in $\mathscr{A}$, and $F \colon \mathscr{A} \to \mathscr{A}'$ is a left-exact additive functor, there exists a natural long exact sequence

$$0 \to F(A) \to F(B) \to F(C) \to R^1 F(A) \to$$
$$\cdots \to R^i F(A) \to R^i F(B) \to R^i F(C) \to R^{i+1} F(A) \to \cdots$$

REMARK. This long exact sequence is often useful for computing the $R^i F(A)$. For instance, if $R^1 F(A) = 0$ then the sequence

$$0 \to F(A) \to F(B) \to F(C) \to 0$$

is exact. If $R^1 F(*)$ is *always* 0, then F is an exact functor.

Here is an application:

DEFINITION D.1.21. If $F\colon \mathscr{A} \to \mathscr{A}'$ is a left-exact additive functor, an object $M \in \mathscr{A}$ is called *F-acyclic* if

$$R^i F(M) = 0$$

for $i > 0$.

The long exact sequence in corollary D.1.20 on the previous page implies that

COROLLARY D.1.22. *Let $F\colon \mathscr{A} \to \mathscr{A}'$ be a left-exact additive functor, and let M be injective or F-acyclic. Then*

(1) If

$$0 \to A \to M \to B \to 0$$

is a short exact sequence in $\mathscr{A}$, then

$$R^i F(A) \cong R^{i-1} F(B)$$

for $i > 1$ and $R^1 F(A)$ is coker $F(M) \to R(B)$.

(2) if

$$0 \to A \to M_0 \to \cdots \to M_{n-1} \to B \to 0$$

with the M_i injective or F-acyclic then

$$R^i F(A) \cong R^{i-n} R(B)$$

and $R^n F(A)$ is coker $F(M_{n-1}) \to F(B)$.

(3) If

$$0 \to A \to M_0 \to \cdots$$

is a resolution by F-acyclic objects, then $R^i F(A) = H^i(F(M))$.

PROOF. To prove the first statement, note that the long exact sequence in corollary D.1.20 on the preceding page reduces to

$$0 \to F(A) \to F(M) \to F(B) \to R^1 F(A) \to 0$$

and

$$0 \to R^n F(B) \xrightarrow{\delta} R^{n+1} F(A) \to 0$$

for all $n > 0$. The second statement follows from the first applied to short exact sequences

$$0 \to A \to M_0 \to K_1 \to 0$$
$$0 \to K_i \to M_i \to K_{i+1} \to 0$$
$$0 \to K_{n-1} \to M_{n-1} \to B \to 0$$

and induction on n.

To prove the third statement, note that we can truncate the resolution by *F*-acyclic objects at any point to get

$$0 \to A \to M_0 \to \cdots \to M_{n-1} \to \ker \delta^n \to 0$$

and

$$R^n F(A) = \operatorname{coker} F(\delta^{n-1})\colon F(M_{n-1}) \to F(\ker \delta^n) = \ker F(\delta^n) = H^n(F(M))$$

where $F(\ker \delta^n) = \ker F(\delta^n)$ is due to the left-exactness of F. $\square$

EXERCISES.

5. Show that $\text{Ext}_R^i(A \oplus B, C) = \text{Ext}_R^i(A, C) \oplus \text{Ext}_R^i(B, C)$

6. If N is an injective object of $\mathscr{A}$, show that

$$\text{Ext}_R^i(M, N) = 0$$

for $i > 0$ and *any* object $M \in \mathscr{A}$. Conclude that $\hom(*, N)$ is an exact functor.

7. If $\mathscr{A} = \mathscr{M}_R$, the category of modules over a commutative ring, R, show that

$$\text{Ext}_R^i(R, M) = 0$$

for $i > 0$ and any R-module, M.

8. If $\mathscr{A} = \mathscr{M}_R$, the category of modules over a commutative ring, R, and P is any projective module over M, show that

$$\text{Ext}_R^i(P, M) = 0$$

for $i > 0$ and any R-module, M, so that $\hom_R(P, *)$ is an exact functor.

9. Suppose $\mathscr{A} = \mathscr{M}_R$, M and N are R-modules, and

$$\cdots \to P_1 \to P_0 \to M \to 0$$

is a projective resolution (see definition D.1.13 on page 526). Show that

$$\text{Ext}_R^i(M, N) = H^i(\hom_R(P_*, N))$$

so *projective* resolutions could be used to compute the Ext^i-groups.

10. Find an injective resolution for $\mathbb{Z}$ in $\mathscr{A}b$.

11. If $\mathscr{A} = \mathscr{A}b$ show that every abelian group, A, has an injective resolution that ends in degree 1 — i.e. is of the form

$$I_0 \to I_1$$

so that $\text{Ext}_{(}^i A, B) = 0$ for $A, B \in \mathscr{A}b$ and $i > 1$.

12. If $A \in \mathscr{A}b$ is a finite abelian group, show that there is an (unnatural!) isomorphism

$$A \cong \hom_{\mathscr{A}b}(A, \mathbb{Q}/\mathbb{Z})$$

and that $\text{Ext}_{\mathbb{Z}}^1(A, \mathbb{Z}) = A$.

13. Suppose $\mathscr{A} = \mathscr{M}_R$, the category of modules over a ring, R. If A and B are R-modules, an *extension* of A by B is a short exact sequence

$$0 \to B \to E \to A \to 0$$

where E is some module. Two such extensions are considered equivalent if they fit into a commutative diagram

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & B & \overset{r}{\longrightarrow} & E_1 & \overset{s}{\longrightarrow} & A & \longrightarrow & 0 \\
& & \| & & \downarrow{\scriptstyle v} & & \| & & \\
0 & \longrightarrow & B & \underset{r'}{\longrightarrow} & E_2 & \underset{s'}{\longrightarrow} & A & \longrightarrow & 0
\end{array}
$$

Exercise 4 on page 525 implies that v is an isomorphism. Regard an extension as equivalent to 0 if it is split, i.e. of the form

$$0 \to B \to B \oplus A \to A \to 0$$

Show that there is a 1-1 correspondence between equivalence-classes of extensions of A by B and the elements of $\operatorname{Ext}_R^1(A, B)$. This is one reason for the name "Ext^*" for the derived functors of $\hom_{\mathscr{A}}(A, *)$. An extension (no pun intended!) of this argument shows that $\operatorname{Ext}_R^n(A, B)$ can be regarded as equivalence-classes of n-fold extensions.

D.1.3. δ-functors. In the course of analyzing cohomology groups, proofs involving certain diagrams and exact sequences (like those in the proofs of propositions D.1.25 on page 533 and D.1.26 on the next page) appear with monotonous regularity. Since the proofs are essentially the same, we abstract out the structural features that they have in common:

DEFINITION D.1.23. If $\mathscr{A}$ and $\mathscr{B}$ are abelian categories, a *covariant δ-functor* from $\mathscr{A}$ to $\mathscr{B}$ is a sequence of additive functors

$$F^n \colon \mathscr{A} \to \mathscr{B}$$

such that for every short exact sequence

$$0 \to A \to B \to C \to 0$$

in $\mathscr{A}$, there exists a natural morphism

$$\delta^n \colon F^n C \to F^{n+1} A$$

for all $n \geq 0$ that fits into a natural long exact sequence

$$0 \to F^0 A \to F^0 B \to F^0 C \xrightarrow{\delta^0} F^1 A \to$$

$$\cdots \to F^n A \to F^n B \to F^n C \xrightarrow{\delta^n} F^{n+1} A \to \cdots$$

A contravariant δ-functor is like a covariant one with all the arrows reversed

REMARK. Note the similarity to the behavior of $H^*(X, *)$. This clearly defines a cohomological δ-functor.

DEFINITION D.1.24. A δ-functor, F, is *universal* if, for any other δ-functor, G (with the same variance!), any natural transformation (see definition A.5.10 on page 444)

$$t \colon F^0 A \to G^0 A$$

for all $A \in \mathscr{A}$, induces *unique* natural transformations $t^n \colon F^n A \to G^n A$ for all $n > 0$.

REMARK. It is not hard to see that two *distinct* universal δ-functors that are naturally isomorphic in degree 0 (i.e. $F^0 A \cong G^0 A$) must be isomorphic in *all* degrees. This is the same "universal property" nonsense that served us so well in section A.5 on page 438.

Universal δ-functors are not hard to find:

PROPOSITION D.1.25. *If $F\colon \mathscr{A} \to \mathscr{B}$ is a covariant δ-functor that has the property that, for any $A \in \mathscr{A}$, there exists a monomorphism $g\colon A \to B$ such that $F^n g = 0$, $n > 0$ (g may depend on n and A), then F is universal.*

PROOF. Given $A \in \mathscr{A}$, let

$$0 \to A \xrightarrow{g} B \to C \to 0$$

be a short exact sequence with $F^1 g = 0$. If G is any other δ-functor, we can compare the long exact sequences

$$
\begin{array}{ccccccccc}
F^0 A & \longrightarrow & F^0 B & \longrightarrow & F^0 C & \xrightarrow{\delta_F} & F^1 A & \xrightarrow{\ 0\ } & F^1 B \\
\ \downarrow{\scriptstyle t} & & {\scriptstyle t}\downarrow & & \ \downarrow{\scriptstyle t} & & \ \vdots{\scriptstyle t_1} & & \\
G^0 A & \longrightarrow & G^0 B & \longrightarrow & G^0 C & \xrightarrow{\delta_G} & G^1 A & \longrightarrow & G^1 B
\end{array}
$$

and we can fill in the dotted arrow with the *unique* morphism that makes the diagram commute: Since the upper row is exact δ_F is the cokernel of $F^0 B \to F^0 C$, following the definition of cokernel in definition A.5.33. Since the diagram commutes, the composite $F^0 B \to F^0 C \xrightarrow{t} G^1 C \to G^1 A$ vanishes, so there exists a *unique* morphism

$$F^1 A \to G^1 A$$

that makes the diagram commute. We prove the higher cases by induction on n, where we choose a morphism $g_{n+1}\colon A \to B_{n+1}$ so that $F(g_{n+1})$ is 0 in degree $n+1$

$$
\begin{array}{ccccccccc}
F^n A & \xrightarrow{F(g_{n+1})} & F^n B_{n+1} & \longrightarrow & F^n C & \xrightarrow{\delta_F} & F^{n+1} A & \xrightarrow{\ 0\ } & F^{n+1} B_{n+1} \\
\ \downarrow{\scriptstyle t_n} & & {\scriptstyle t_n}\downarrow & & \ \downarrow{\scriptstyle t_n} & & \ \vdots{\scriptstyle t_{n+1}} & & \\
G^n A & \longrightarrow & G^n B_{n+1} & \longrightarrow & G^n C & \xrightarrow{\delta_G} & G^{n+1} A & \longrightarrow & G^{n+1} B_{n+1}
\end{array}
$$

Naturality of the t_i (in the sense of definition A.5.10 on page 444) involves a large diagram chase and is left to the reader. $\square$

PROPOSITION D.1.26. *If $F\colon \mathscr{A} \to \mathscr{B}$ is a contravariant δ-functor that has the property that, for any $A \in \mathscr{A}$, there exists an epimorphism $g\colon P \to A$ such that $F^n g = 0$, $n > 0$ (g may depend on n and A), then F is universal.*

PROOF. The proof of this case is literally the same as the previous one except that the induced maps run in the opposite direction. Let $P \to A$ be an epimorphism with kernel $K \to P$. Then we get

$$
\begin{array}{ccccccccc}
F^0 A & \longrightarrow & F^0 P & \longrightarrow & F^0 K & \xrightarrow{\delta_F} & F^1 A & \xrightarrow{\ 0\ } & F^1 P \\
\ \downarrow{\scriptstyle t} & & {\scriptstyle t}\downarrow & & \ \downarrow{\scriptstyle t} & & \ \vdots{\scriptstyle t_1} & & \\
G^0 A & \longrightarrow & G^0 P & \longrightarrow & G^0 K & \xrightarrow{\delta_G} & G^1 A & \longrightarrow & G^1 P
\end{array}
$$

The rest of the proof is similar. $\square$

D.2. Rings and modules

D.2.1. Ext* and Tor$_*$. Since the category of modules over a ring has enough projectives (just map a suitable free module to a module) and enough injectives (proposition A.5.41 on page 460), we can define homology and cohomology-based functors.

We have seen the Ext^i-functors — derived functors of the hom$_R$-functor (definition D.1.18 on page 528). We can also define derived functors of the $\otimes$-functor — these are *homology*-type objects:

DEFINITION D.2.1. If $M, N \in \mathscr{M}_R$ are modules over a ring R and

$$P_* \to M \to 0$$

is a projective resolution of M, then

$$\mathrm{Tor}_i^R(M, N) = H_i(P_* \otimes_R N)$$

REMARK. Although $\mathrm{Ext}_R^i(M, N)$ can be computed using an *injective* reso-lution of N or a *projective* resolution (exercise 9 on page 531) of M, $\mathrm{Tor}_i^R(M, N)$ requires a projective resolution of M[2]. It follows that this is undefined in a category that does not have enough projectives (like that of sheaves).

Since Tor_*^R-is a homology-type functor we get

LEMMA D.2.2. *If R is a ring and*

$$0 \to A \xrightarrow{u} B \xrightarrow{v} C \to 0$$

is a short exact sequence of R-modules, then there is an induced long exact sequence

$$\cdots \to \mathrm{Tor}_i^R(A, D) \to \mathrm{Tor}_i^R(B, D) \to \mathrm{Tor}_i^R(C, D) \to \mathrm{Tor}_{i-1}^R(A, D) \to \cdots$$
$$\cdots \to \mathrm{Tor}_1^R(C, D) \to A \otimes_R D \to B \otimes_R D \to C \otimes_R D \to 0$$

for any R-module D.

PROOF. A projective version of lemma D.1.19 on page 529 (just reverse all of the arrows!) implies that we can get a commutative diagram with exact rows

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & P_* & \xrightarrow{r} & Q_* & \xrightarrow{s} & W_* & \longrightarrow & 0 \\
& & \downarrow{\epsilon_A} & & \downarrow{\epsilon_B} & & \downarrow{\epsilon_C} & & \\
0 & \longrightarrow & A & \xrightarrow{u} & B & \xrightarrow{v} & C & \longrightarrow & 0
\end{array}
$$

where the upper row is projective resolutions of the modules in the lower row. The conclusion follows by taking the tensor product with D and invoking proposition D.1.10 on page 523. $\qquad\square$

Just as Ext^i is sensitive to projectivity or injectivity, Tor$_i$ is sensitive to *flat-ness* (see definition A.5.58 on page 469) of a module:

PROPOSITION D.2.3. *If R is a ring, an R-module, M, is flat if and only if*

$$\mathrm{Tor}_1^R(M, N) = 0$$

for all R-modules, N.

---

[2]Or of N — both produce the same result.

REMARK. We know that all projective modules are flat (proposition A.5.61 on page 470), but the converse is *not* true: $\mathbb{Q}$ is a flat $\mathbb{Z}$-module that is *not* projective.

We also have a version of corollary D.1.22 on page 530

COROLLARY D.2.4. *If R is a ring and*

$$0 \to K \to F \to A \to 0$$

is a short exact sequence of R-modules with F flat, then

$$\mathrm{Tor}_i^R(A, D) = \mathrm{Tor}_{i-1}^R(K, D)$$

for any R-module D. If

$$0 \to K_m \to F_m \to \cdots \to F_0 \to A \to 0$$

is an exact sequence with the F_i flat, then $\mathrm{Tor}_i^R(A, D) \cong \mathrm{Tor}_{i-m-1}^R(K_m, D)$ for $i \geq m + 2$ and any R-module D.

PROOF. Proposition D.2.3 on the preceding page implies that $\mathrm{Tor}_i^R(F_j, D) = 0$, so the F_j are F-acyclic for the Tor_*^R-functor. At this point, the proof is the same as that of corollary D.1.22 on page 530 except that we substitute the exact sequence in lemma D.2.2 on the preceding page for the one used in that proof. $\square$

Along these lines, we also have

PROPOSITION D.2.5. *If $f\colon R \to S$ is a flat morphism of rings (i.e., S is a flat module over R) and M, N are R-modules, then*

$$\mathrm{Tor}_i^R(M, N) \otimes_R S = \mathrm{Tor}_i^S(M \otimes_R S, N \otimes_R S)$$

If R is noetherian and M is finitely generated then

$$\mathrm{Ext}_R^i(M, N) \otimes_R S = \mathrm{Ext}_S^i(M \otimes_R S, N \otimes_R S)$$

PROOF. If

$$\cdots \to P_1 \to P_0 \to M \to 0$$

is a projective resolution of M, the flatness of S implies that

$$\cdots \to P_1 \otimes_R S \to P_0 \otimes_R S \to M \otimes_R S \to 0$$

is also exact, and is a projective resolution of $M \otimes_R S$ over S. Now

$$\mathrm{Tor}_i^R(M, N) = H_i(P_* \otimes_R S)$$

and

$$
\begin{aligned}
\mathrm{Tor}_i^S(M \otimes_R S, N \otimes_R S) &= H_i(P_* \otimes_R S \otimes_S (N \otimes_R S)) \\
&= H_i(P_* \otimes_R N \otimes_R S) \\
&= H_i(P_* \otimes_R N) \otimes_R S
\end{aligned}
$$

— the last equality being due to the exactness of $* \otimes_R S$, so it commutes with homology.

To prove the statement about Ext_R^i note that the hypotheses imply that the P_i can all be finitely generated *free R-modules*. If $P_i = R^n$, then

$$
\begin{aligned}
\hom_S(P_i \otimes_R S, N \otimes_R S) &= \bigoplus_{j=1}^{n} \hom_S(S, N \otimes_R S) \\
&= \bigoplus_{j=1}^{n} N \otimes_R S \\
&= \bigoplus_{j=1}^{n} \hom_R(R, N) \otimes_R S \\
&= \hom_R(P_i, N) \otimes_R S
\end{aligned}
$$

so the conclusion follows by the same reasoning used above. $\qquad\square$

PROPOSITION D.2.6. *If R is a ring and let M is an R-module, the following two statements are equivalent:*

(1) proj-dim $M \le n$
(2) $\text{Ext}_R^{n+1}(M, N) = 0$ *for all R-modules, N.*

PROOF. It is clear that $1 \implies 2$. Suppose statement 2 is true and let

$$
\cdots \to P_{n+1} \xrightarrow{f} P_n \to \cdots \to P_0 \to M \to 0
$$

is a projective resolution of M. Statement 2 implies that

$$
\cdots \to \hom_R(P_n, N) \to \hom_R(P_{n+1}, N) \to \hom_R(P_{n+2}, N) \to \cdots
$$

is exact for all N. If we set $N = \text{im } f$, then $f \in \hom_R(P_{n+1}, N)$ and its image in $\hom_R(P_{n+2}, N)$ is 0 because the exactness of the original resolution. It follows that $f \in \hom_R(P_{n+1}, N)$ is in the image of some $\alpha \in \hom_R(P_n, N)$ so that there exists a homomorphism

$$
\alpha \colon P_n \to \text{im } P_{n+1}
$$

such that $f \circ \alpha = f$. The projective property of P_n implies that α lifts to a map $P_n \to P_{n+1}$ and $P_n = \text{im } f \oplus K$ where $\text{im } f$ and K must both be projective modules (where $K = \ker \alpha$). It follows that the resolution of M can be shortened to

$$
0 \to K \to \cdots \to P_0 \to M \to 0
$$

$\square$

PROPOSITION D.2.7. *If R is a ring, an R-module, M, is injective if and only if*

$$
\text{Ext}_R^1(R/\mathfrak{a}, M) = 0
$$

for all ideals $\mathfrak{a} \subset R$.

PROOF. If M is injective, $\hom_R(*, M)$ is exact and $\text{Ext}_R^1(N, M) = 0$ for all modules N. Conversely, if $\text{Ext}_R^1 R/\mathfrak{a}, M) = 0$, then the sequence

$$
0 \to \hom_R(\mathfrak{a}, M) \to \hom_R(R, M) \to \hom_R(R/\mathfrak{a}, M) \to 0
$$

is exact for all ideals $\mathfrak{a} \subset R$ and the conclusion follows from Baer's Criterion (A.5.36). $\qquad\square$

Now we are ready to prove a sequence of results regarding rings and modules:

LEMMA D.2.8. *If R is a ring and $n \geq 0$ is an integer, then the following statements are equivalent:*

(1) proj-dim $M \leq n$ *for all R-modules M*
(2) proj-dim $M \leq n$ *for all finitely generated R-modules M*
(3) inj-dim $M \leq n$ *for all R-modules M*
(4) $\text{Ext}_R^{n+1}(M, N) = 0$ *for all R-modules M and N.*

PROOF. It is clear that $1 \implies 2$. To see that $2 \implies 3$ form an exact sequence

$$0 \to M \to I_0 \to \cdots \to I_{n-1} \to U \to 0$$

with the I_j injective. The long exact sequence in corollary D.1.20 on page 529 and the fact that $\text{Ext}_R^i(*, I_j) = 0$ implies that

$$\text{Ext}_R^1(N, U) = \text{Ext}_R^{n+1}(N, M)$$

for any R-modules, N (use an induction like that in corollary D.1.22 on page 530, statement 2). Now set $N = R/\mathfrak{a}$ — a finitely generated R-module. Then statement 2 above implies that

$$\text{Ext}_R^{n+1}(R/\mathfrak{a}, M) = 0 = \text{Ext}_R^1(R/\mathfrak{a}, U)$$

for all ideals $\mathfrak{a} \subset R$. Proposition D.2.7 on the facing page implies that U is injective so the injective dimension of M is $\leq n$.

Clearly, $3 \implies 4$ and proposition D.2.6 on the preceding page shows that $4 \implies 1$. $\qquad\square$

It follows that

$$\sup_{M \in \mathscr{M}_R} \text{proj-dim } M = \sup_{M \in \mathscr{M}_R} \text{inj-dim } M$$

We give this common value a name:

DEFINITION D.2.9. If R is a ring, the common value the global dimension of R:

$$\text{gl-dim } R = \sup_{M \in \mathscr{M}_R} \text{proj-dim } M = \sup_{M \in \mathscr{M}_R} \text{inj-dim } M$$

REMARK. Note that this dimension might be ∞.

It might seem that we have a bewildering number of ways of defining the "dimension" of a ring: Krull dimension and now a kind of homological dimension. These different concepts will turn out to agree with each other in the cases that concern us (even though they are different in general).

In some cases, Tor_i^R can be used to compute dimension

PROPOSITION D.2.10. *If $(R, \mathfrak{m})$ is a noetherian local ring (with maximal ideal $\mathfrak{m}$) and M is a finitely generated R-module then*

$$\text{proj-dim } M \leq n \iff \text{Tor}_{n+1}^R(M, k) = 0$$

where $k = R/\mathfrak{m}$.

PROOF. The implication $\implies$ is clear. For the reverse implication, corollary D.2.4 on page 535 allows us to reduce to the case where $n = 0$. So, if $\operatorname{Tor}_1^R(M, k) = 0$ we will show that M is projective. If

$$f: M \to M \otimes_R k = M/\mathfrak{m} \cdot M$$

is the projection, corollary A.5.56 on page 468 implies that if $S = \{m_1, \ldots, m_t\}$ have the property that $\{f(m_1), \ldots, f(m_t)\}$ generate $M \otimes_R k$ then the set S generate M.

Let

$$0 \to K \to F \xrightarrow{p} M \to 0$$

be a short exact sequence with F a free module of rank t with p sending the jth basis element to m_j in the set of generators of M. The map p is surjective and, by construction,

$$p \otimes 1: F \otimes_R k \to M \otimes_R k$$

is an isomorphism. Since $\operatorname{Tor}_1^R(M, k) = 0$, the kernel of $p \otimes 1$ is $K \otimes_R k$, which must be 0. So K is an R-module with $K/\mathfrak{m} \cdot K = 0$ or $K = \mathfrak{m} \cdot K$. Corollary A.1.82 on page 379 implies that $K = 0$, so that M is a free module. $\square$

LEMMA D.2.11. *If R is a noetherian ring with a finitely generated R-module M, then*

(1) $\operatorname{proj-dim} M = \sup_{\mathfrak{m}} \operatorname{proj-dim} M_{\mathfrak{m}}$, *where $\mathfrak{m}$ runs over all the maximal ideals of R.*
(2) $\operatorname{proj-dim} M \leq n$ *if and only if $\operatorname{Tor}_{n+1}^R(M, R/\mathfrak{m}) = 0$ for all maximal ideals $\mathfrak{m} \subset R$.*

PROOF. Statement 1 follows from the second statement in proposition D.2.5 on page 535, lemma D.2.8 on the preceding page, and exercise 47 on page 386. $\square$

LEMMA D.2.12. *If R is a noetherian ring, the following statements are equivalent:*

(1) $\operatorname{gl-dim} R \leq n$
(2) $\operatorname{proj-dim} M \leq n$ *for all finitely generated R-modules M.*
(3) $\operatorname{inj-dim} M \leq n$ *for all finitely generated R-modules M.*
(4) $\operatorname{Ext}_R^{n+1}(M, N) = 0$ *for all finitely generated R-modules M and N*
(5) $\operatorname{Tor}_{n+1}^R(M, N) = 0$ *for all finitely generated R-modules M and N*

PROOF. Lemma D.2.8 on the preceding page implies that statements 1 and 2 are equivalent and imply statements 3 and 5. Statement 3 implies statement 4. Lemma D.2.11 shows that statement 5 implies statement 2 and proposition D.2.6 on page 536 shows that statement 4 implies statement 2. $\square$

In the case of a noetherian local ring, it is easy to characterize the global dimension:

COROLLARY D.2.13. *If $(R, \mathfrak{m})$ is a noetherian local ring with $k = R/\mathfrak{m}$, then*

$$\operatorname{gl-dim} R \leq n \iff \operatorname{Tor}_{n+1}^R(k, k)$$

so that $\operatorname{gl-dim} R = \operatorname{proj-dim} k$

PROOF. The right implication is clear. If we assume that $\text{Tor}^R_{n+1}(k,k) = 0$ then proj-dim $k \leq n$, by proposition D.2.10 on page 537. This implies that $\text{Tor}^R_{n+1}(M,k) = 0$ for *all* R-modules M — in particular, *finitely generated* modules M. If M is finitely generated, then proposition D.2.10 on page 537 shows that

$$\text{Tor}^R_{n+1}(M,k) = 0 \implies \text{proj-dim } M \leq n$$

But, by lemma D.2.12 on the preceding page, this means that gl-dim $R \leq n$. $\square$

We need the following definition

DEFINITION D.2.14. If M is a module over a ring R, an element $x \in R$ will be called *M-regular* if

$$M \xrightarrow{x\cdot} M$$

is injective and $M/(x) \cdot M \neq 0$. An *M-regular sequence* $\{x_1, \ldots, x_t\} \subset R$ is a sequence with the property that

(1) x_1 is M-regular in M, and
(2) for all i, x_i is M-regular in $M/(x_1, \ldots, x_{i-1}) \cdot M$.

The maximum length of an M-regular sequence in a module is called its *depth*, denoted depth M.

The depth of a module is related to its projective dimension:

LEMMA D.2.15. *If $(R, \mathfrak{m})$ is a noetherian local ring with $k = R/\mathfrak{m}$, M is a finitely generated R-module of projective dimension t, and $x \in \mathfrak{m}$ is M-regular then*

$$\text{proj-dim } (M/x \cdot M) = t + 1$$

REMARK. Note that an M-regular element of R *must* come from the maximal ideal $\mathfrak{m}$: if $x \in R \setminus \mathfrak{m}$, then the image of x in k is $\neq 0$. It follows that $(M/x \cdot M) \otimes_R k = 0$, and Nakayama's Lemma (A.1.82 on page 379) implies that $M/x \cdot M = 0$.

PROOF. The sequence

$$0 \to M \xrightarrow{x\cdot} M \to M/x \cdot M \to 0$$

is exact, by the hypotheses. The long exact sequence of Tor-functors (D.2.2 on page 534) gives

$$\text{Tor}^R_n(M/x \cdot M, k) = 0$$

if $n > t + 1$. In degree $t + 1$, we get

$$0 = \text{Tor}^R_{t+1}(M,k) \to \text{Tor}^R_{t+1}(M/x \cdot M, k) \to \text{Tor}^R_t(M,k) \xrightarrow{x\cdot} \text{Tor}^R_t(M,k)$$

The map $x\cdot$ on the right is zero because $x \in \mathfrak{m}$ so that it annihilates k. It follows that

$$\text{Tor}^R_{t+1}(M/x \cdot M, k) \cong \text{Tor}^R_t(M,k) \neq 0$$

and lemma D.2.12 on the facing page implies the conclusion. $\square$

We get the following interesting result that explains how a module can have a projective dimension less than the global dimension of its ring:

COROLLARY D.2.16. *If $(R, \mathfrak{m})$ is a noetherian local ring with $k = R/\mathfrak{m}$, and M is a finitely generated module, then*

$$
\begin{aligned}
\text{depth } M + \text{proj-dim } M &\leq \text{gl-dim } R \\
\text{depth } R &\leq \text{gl-dim } R
\end{aligned}
$$

REMARK. It will turn out that these inequalities are actually *equations* in cases that interest us.

Cohen-Macaulay rings are particularly important in algebraic geometry:

DEFINITION D.2.17. A local noetherian ring, R, is called a local Cohen-Macaulay ring if its *depth* (as a module over itself) is equal to its *Krull dimension* (see 2.8.7 on page 100). A noetherian ring is called Cohen-Macaulay if all of its localizations are local Cohen-Macaulay.

REMARK. If $p \in V$ is a simple point of an algebraic variety, exercise 8 on page 139 shows that $\mathcal{O}_{V,p}$ is a local Cohen-Macaulay ring. Its local parameters form an M-sequence.

D.2.2. Koszul complexes. We will define a particular type of cochain complex that will have important geometric applications.

DEFINITION D.2.18. If R is a ring, $\mathfrak{A} = (x_1, \dots, x_n) \subset R$ is an ideal, the *Koszul cochain complex* $C = \mathbf{K}(x_1, \dots, x_n)$ is the cochain complex with

$$
C^i = \Lambda^i F
$$

for $i = 0, \dots, n$, where $F = \bigoplus_{i=1}^n R \cdot e_i$ — the free R-module on a basis $\{e_i\}$ and the differential

$$
\delta \colon C^i \to C^{i+1}
$$

is exterior multiplication $(\sum_{i=1}^n x_i \cdot e_i) \wedge *$.

REMARK. The self-annihilation properties of $\wedge$-products immediately imply that $\delta^2 = 0$. As usual, we follow the convention that $\Lambda^0 F = R$.

This construction is clearly functorial with respect to homomorphisms that preserve the sequence $\{x_i\}$: if $f \colon F \to G$ is a homomorphism of free modules that sends $(x_1, \dots, x_n) \in R^n = F$ to $(y_1, \dots, y_n) \in R^n = G$, it induces a homomorphism

$$
\mathbf{K}(f) \colon \mathbf{K}(x_1, \dots, x_n) = \mathbf{K}(y_1, \dots, y_n)
$$

We will frequently want to consider cochain complexes

$$
M \otimes_R \mathbf{K}(x_1, \dots, x_n)
$$

where M is some R-module.

The Koszul complex $\mathbf{K}(x_1, \dots, x_n)$ is sensitive to the extent to which the x_i are *independent* of each other:

PROPOSITION D.2.19. *If $\{x_1, \dots, x_n\} \subset R$ is a set of elements, $\{y_1, \dots, y_m\} \subset (x_1, \dots, x_m)$ — the ideal generated by the x_i, and M is any R-module then*

$$
M \otimes_R \mathbf{K}(x_1, \dots, x_n, y_1, \dots, y_m) = M \otimes_R \mathbf{K}(x_1, \dots, x_n) \otimes_R \Lambda \, R^m
$$

so that

(D.2.1) $H^n(M \otimes_R \mathbf{K}(x_1, \ldots, x_n, y_1, \ldots, y_m)) =$

$$\bigoplus_{i+j=n} H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_n)) \otimes_R \Lambda^j R^m$$

for all $n \geq 0$. It follows that

$$H^n(M \otimes_R \mathbf{K}(x_1, \ldots, x_n, y_1, \ldots, y_m)) = 0$$

if and only if

$$H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_n)) = 0$$

for all $n - m \leq i \leq n$.

PROOF. There is an isomorphism of free R-modules

$$\alpha \colon R^{n+m} \quad \rightarrow \quad R^{n+m}$$
$$(x_1, \ldots, x_n, y_1, \ldots, y_m) \quad \mapsto \quad (x_1, \ldots, x_n, 0, \ldots, 0)$$

— if $y_i = \sum_{j=1}^n s_{i,j} x_j$ then

$$\left[\begin{array}{c|c} I & 0 \\ \hline -S & I \end{array} \right]$$

where the upper left block is an $n \times n$ identity matrix, the lower right block is an $m \times m$ identity matrix and the lower left block is $-S = -\left[s_{i,j} \right]$.

The isomorphism α induces an isomorphism of Koszul complexes

$$M \otimes_R \mathbf{K}(x_1, \ldots, x_n, y_1, \ldots, y_m) = M \otimes_R \mathbf{K}(x_1, \ldots, x_n, 0, \ldots, 0)$$

As cochain *modules* (see proposition A.6.8 on page 476):

$$M \otimes_R \mathbf{K}(x_1, \ldots, x_n, 0, \ldots, 0) = M \otimes_R \mathbf{K}(x_1, \ldots, x_n) \otimes_R \Lambda \, R^m$$

and, since the coboundaries on the right factor are 0, this is true as cochain-*complexes*. It follows that

$$M \otimes_R \mathbf{K}(x_1, \ldots, x_n, 0, \ldots, 0)^n = \bigoplus_{i+j=n} M \otimes_R \mathbf{K}(x_1, \ldots, x_n)^i \otimes_R \Lambda^j R^m$$

and, since taking the tensor product with a free module is exact, we get equation D.2.1.   □

It turns out that the Koszul complex has a simple algebraic interpretation — it is essentially an *iterated algebraic mapping cone* (see definition D.1.11 on page 524:

PROPOSITION D.2.20. *If R is a ring, $\{x_1, \ldots, x_t\} \subset R$ is a set of elements, $x_{t+1} \in R$ is one additional element, and M is any R-module, then there exists a short exact sequence of cochain-complexes*

$$0 \to \mathbf{K}(x_1, \ldots, x_t)[-1] \xrightarrow{\wedge x_{t+1} e_{t+1}} \mathbf{K}(x_1, \ldots, x_{t+1}) \to \mathbf{K}(x_1, \ldots, x_t) \to 0$$

It follows that we get a long exact sequence in cohomology

$$\cdots \to H^{i-1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_t)) \to H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t+1}))$$

$$\to H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_t)) \xrightarrow{x_{t+1} \cdot} H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_t))$$

$$\to H^{i+1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t+1})) \to \cdots$$

where $x_{t+1} \cdot$ represents multiplication by x_{t+1}.

PROOF. It is not hard to see that

$$\mathbf{K}(x_1, \ldots, x_{t+1})^n = \mathbf{K}(x_1, \ldots, x_t)^n \oplus \mathbf{K}(x_1, \ldots, x_t)^{n-1} \wedge e_{t+1}$$

where the left summand is the set of terms that do *not* include a factor of e_{t+1} and the right summand is the set of those that *do.*

Let $\delta = \left(\sum_{j=1}^{t} x_j \cdot e_j \right) \wedge *$ denote the coboundary of $\mathbf{K}(x_1, \ldots, x_t)$ and $\delta' = \left(\sum_{j=1}^{t+1} x_j \cdot e_j \right) \wedge *$ that of $\mathbf{K}(x_1, \ldots, x_{t+1})$.

If $v \in \mathbf{K}(x_1, \ldots, x_t)^{n-1}$, the fact that $e_{t+1} \wedge e_{t+1} = 0$ implies that

$$\left(\sum_{j=1}^{t+1} x_j \cdot e_j \right) \wedge (v \wedge e_{t+1}) = (\delta v) \wedge e_{t+1}$$

If $w \in \mathbf{K}(x_1, \ldots, x_t)^n$

$$\left(\sum_{j=1}^{t+1} x_j \cdot e_j \right) \wedge w = \delta w + (-1)^n x_{t+1} \cdot (w \wedge e_{t+1})$$

where the $(-1)^n$ comes from our need to permute the factor of e_{t+1} with the n factors in w. We get

$$
\begin{aligned}
\delta'|\mathbf{K}(x_1, \ldots, x_t)^{n-1} \wedge e_{t+1} &= \delta \wedge 1 \\
\delta'|\mathbf{K}(x_1, \ldots, x_t)^n &= \delta + (-1)^n x_{t+1} \cdot e_{t+1}
\end{aligned}
$$

It follows that we get an inclusion of cochain complexes

$$\mathbf{K}(x_1, \ldots, x_t)[-1] \xrightarrow{\wedge e_{t+1}} \mathbf{K}(x_1, \ldots, x_{t+1})$$

where $\mathbf{K}(x_1, \ldots, x_t)[-1]^n = \mathbf{K}(x_1, \ldots, x_t)^{n-1}$ that extends to a short exact sequence

(D.2.2) $\quad 0 \to \mathbf{K}(x_1, \ldots, x_t)[-1] \xrightarrow{\wedge e_{t+1}} \mathbf{K}(x_1, \ldots, x_{t+1}) \to \mathbf{K}(x_1, \ldots, x_t) \to 0$

and this induces a long exact sequence in cohomology

$$\cdots \to H^i(\mathbf{K}(x_1, \ldots, x_t)[-1]) \to H^i(\mathbf{K}(x_1, \ldots, x_{t+1}))$$
$$\to H^i(\mathbf{K}(x_1, \ldots, x_t)) \xrightarrow{c} H^{i+1}(\mathbf{K}(x_1, \ldots, x_t)[-1]) \to \cdots$$

or (using the fact that $H^i(\mathbf{K}(x_1, \ldots, x_t)[-1]) = H^{i-1}(\mathbf{K}(x_1, \ldots, x_t))$)

$$\cdots \to H^{i-1}(\mathbf{K}(x_1, \ldots, x_t)) \to H^i(\mathbf{K}(x_1, \ldots, x_{t+1}))$$
$$\to H^i(\mathbf{K}(x_1, \ldots, x_t)) \xrightarrow{c} H^i(\mathbf{K}(x_1, \ldots, x_t)) \to \cdots$$

where the connecting map c, is given by

$$c = (-1)^i x_{t+1} \cdot : H^i(\mathbf{K}(x_1, \ldots, x_t)) \to H^i(\mathbf{K}(x_1, \ldots, x_t))$$

— see proposition D.1.9 on page 522. We claim that the sequence

$$0 \to M \otimes_R \mathbf{K}(x_1, \ldots, x_t)[-1]$$
$$\xrightarrow{\wedge e_{t=1}} M \otimes_R \mathbf{K}(x_1, \ldots, x_{t+1}) \to M \otimes_R \mathbf{K}(x_1, \ldots, x_t) \to 0$$

is *also* short exact. This is because the exact sequence in D.2.2 on page 542 is a *split* exact sequence of *free* modules (ignoring coboundaries). □

The Koszul complex is a powerful tool for studying the *depth* of a module (see definition D.2.14 on page 539):

COROLLARY D.2.21. *If R is a noetherian ring and M is a finitely generated R-module with an M-sequence $\{x_1, \ldots, x_t\} \subset R$ then*

$$H^i(M \otimes \mathbf{K}(x_1, \ldots, x_t)) = \begin{cases} 0 & \text{if } i \neq t \\ M/(x_1, \ldots, x_t)M \neq 0 & \text{if } i = t \end{cases}$$

If $\mathfrak{I} = (x_1, \ldots, x_n) \subset R$ is an ideal, $\{x_1, \ldots, x_t\}$ is a maximal M-sequence of elements in $\mathfrak{I}$ and $M \neq \mathfrak{I} \cdot M$ then

$$H^i(M \otimes \mathbf{K}(x_1, \ldots, x_n)) = \begin{cases} 0 & \text{if } i < t \\ \neq 0 & \text{if } i = t \end{cases}$$

PROOF. We do induction on t. If $t = 1$, then $M \otimes_R \mathbf{K}(x_1)$ is

$$0 \to M \xrightarrow{x_1 \cdot} M \to 0$$

so $H^0(M \otimes_R \mathbf{K}(x_1)) = 0$ and $H^1(M \otimes_R \mathbf{K}(x_1)) = M/x_1 \cdot M$. If it has been proved up to $t - 1$, we use proposition D.2.20 on page 541 to get a long exact sequence

$$\cdots \to H^{i-1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t-1})) \to H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_t))$$
$$\to H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t-1})) \xrightarrow{x_t \cdot} H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t-1}))$$
$$\to H^{i+1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_t)) \to \cdots$$

If $i < t - 1$ all of the $H^i(M \otimes \mathbf{K}(x_1, \ldots, x_{t-1}))$-terms vanish, so the $H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_t))$-terms also vanish.

When $i = t - 1$, we get

$$0 \to H^{t-1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_t))$$
$$\to H^{t-1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t-1})) \xrightarrow{x_t \cdot} H^{t-1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t-1}))$$
$$\to H^t(M \otimes_R \mathbf{K}(x_1, \ldots, x_t)) \to H^t(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t-1})) = 0$$

and $H^{t-1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_t))$ must vanish because

$$H^{t-1}(M \otimes_R \mathbf{K}(x_1, \ldots, x_{t-1})) = M/(x_1, \ldots, x_{t-1})M$$

and

$$x_t \cdot M/(x_1, \ldots, x_{t-1})M \xrightarrow{x_t \cdot} M/(x_1, \ldots, x_{t-1})M$$

must be *injective* (see definition D.2.14 on page 539).

The second statement follows by continuing this induction. As before, we conclude that $H^i(M \otimes \mathbf{K}(x_1, \ldots, x_n)) = 0$ for $i < t$ and the significant thing we must prove is that

$$H^t(M \otimes \mathbf{K}(x_1, \ldots, x_n)) \neq 0$$

If t is the length of a maximal M-sequence, then *all* of the elements of $\mathfrak{I}$ annihilate some elements of $N = M/(x_1, \ldots, x_t)M$ (so they cannot be used to extend

the M-sequence). If Z is the set of elements of R that annihilate elements of N — so $\mathfrak{I} \subset Z$ then corollary A.1.76 on page 375 implies that

$$\mathfrak{I} \subset Z = \bigcup_{\mathfrak{p} \in \mathrm{Assoc}(N)} \mathfrak{p}$$

(see definition A.1.73 on page 374), a *finite* union by the remark following theorem A.1.77 on page 375. Prime avoidance (exercise 6 on page 71) implies that $\mathfrak{I} \subset \mathfrak{p}_\gamma$ for some γ. This prime ideal annihilates a nonzero element, $m \in M$, so that, in the long exact sequence (with $j \geq t$)

$$\cdots \to H^t(M \otimes_R \mathbf{K}(x_1, \ldots, x_j)) \to H^t(M \otimes_R \mathbf{K}(x_1, \ldots, x_{j-1}))$$
$$\xrightarrow{x_j} H^t(M \otimes_R \mathbf{K}(x_1, \ldots, x_{j-1})) \to \cdots$$

the element m must be in the kernel of $\xrightarrow{x_j}$, so $H^t(M \otimes_R \mathbf{K}(x_1, \ldots, x_j)) \neq 0$. $\quad\square$

It follows that Koszul complexes can be used to determine the *depth* of a module:

THEOREM D.2.22. *Let R be a noetherian ring with an ideal $\mathfrak{I} = (x_1, \ldots, x_n) \subset R$ and with a finitely generated R-module M. If $t \geq 0$ has the property that*

$$H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_n)) = 0$$

if $i < t$ and

$$H^t(M \otimes_R \mathbf{K}(x_1, \ldots, x_n)) \neq 0$$

then every a maximal M-sequence of M taken from elements of the ideal $\mathfrak{I}$ has length t.

PROOF. Suppose that $\{y_1, \ldots, y_s\} \subset \mathfrak{I}$ is a maximal M-sequence taken from $\mathfrak{I}$. Then

$$H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_n, y_1, \ldots, y_s)) = \bigoplus_{u+v=i} H^u(M \otimes_R \mathbf{K}(x_1, \ldots, x_n)) \otimes_R \Lambda^v R^s$$

so

$$H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_n, y_1, \ldots, y_s)) = 0$$

if $i < s$ and

$$H^s(M \otimes_R \mathbf{K}(x_1, \ldots, x_n, y_1, \ldots, y_s)) \neq 0$$

and corollary D.2.21 on the preceding page implies that $r = t$. $\quad\square$

Now we have some interesting results regarding Cohen-Macaulay rings (see definition D.2.17 on page 540):

PROPOSITION D.2.23. *If $(R, \mathfrak{m})$ is a local Cohen-Macaulay ring, then its Krull dimension is equal to its global dimension. If M is an R-module*

(D.2.3) $$\mathrm{Tor}_i^R(M, k) = H^{n-i}(M \otimes_R \mathbf{K}(x_1, \ldots, x_n))$$

for all i.

PROOF. Let $R/\mathfrak{m} = k$ and let $\mathfrak{m} = (x_1, \ldots, x_n)$ — where $\{x_1, \ldots, x_n\}$ are a regular sequence for R. Then corollary D.2.21 on the preceding page implies that

$$H^i(\mathbf{K}(x_1, \ldots, x_n)) = 0$$

for $i \neq n$ and $H^n(\mathbf{K}(x_1, \ldots, x_n)) = k$. Since the $\mathbf{K}(x_1, \ldots, x_n)^i$ are all projective modules (actually free modules), we can flip it "upside down" by defining

$$(\text{D.2.4}) \qquad\qquad C_i = \mathbf{K}(x_1, \ldots, x_n)^{n-i}$$

and

$$(\text{D.2.5})$$

$$
\begin{array}{ccc}
C_i & \xrightarrow{\ \partial_i\ } & C_{i-1} \\
\| & & \| \\
\mathbf{K}(x_1, \ldots, x_n)^{n-i} & \xrightarrow[\delta^{n-i}]{} & \mathbf{K}(x_1, \ldots, x_n)^{n-i+1}
\end{array}
$$

and the new chain-complex, C is a projective resolution of k — which implies the statement about $\operatorname{Tor}_i^R(M, k)$. It also follows that the gl-dim $R \leq n$ and corollary D.2.16 on page 540 implies that it is equal to n. □

The following result is due to Auslander and Buchsbaum (see [**7**]):

COROLLARY D.2.24 (Auslander-Buchsbaum Formula). *If $(R, \mathfrak{m})$ is a local Cohen-Macaulay ring and M is an R-module, then*

$$\operatorname{depth} M + \operatorname{proj-dim} M = \operatorname{gl-dim} R$$

PROOF. The depth of M is the lowest value of d for which $H^i(M \otimes_R \mathbf{K}(x_1, \ldots, x_n)) \neq 0$. But this means that

$$\operatorname{Ext}_R^{n-d}(M, k) = H_{n-d}(C) \neq 0$$

so that proj-dim $M \geq n - d$, and

$$\operatorname{depth} M + \operatorname{proj-dim} M \geq \operatorname{gl-dim} R$$

so corollary D.2.16 on page 540 implies the conclusion. □

D.3. Cohomology of sheaves

D.3.1. Basic construction. Given a sheaf on a ringed space, we can define cochain-complexes from it whose cohomology gives vital information about the sheaf. There are several ways to define cohomology of sheaves.

Most of the material in this section comes from Jean-Pierre Serre's remarkable paper [**150**], which invented this field and proved most of its interesting results.

It will turn out that we can use flasque modules to compute cohomology of sheaves — which is a good thing since the injectives in the category of sheaves tend to be huge.

PROPOSITION D.3.1. *If $(X, \mathcal{O}_X)$ is a ringed space, and $\mathscr{M}$ is a module over $\mathcal{O}_X$ then there exists an injective, flasque module over $\mathcal{O}_X$ and an inclusion*

$$(\text{D.3.1}) \qquad\qquad \mathscr{M} \hookrightarrow \mathscr{F}$$

It follows that the category of modules over $\mathcal{O}_X$ has enough injectives.

PROOF. If $\mathscr{M}$ is a module over $\mathcal{O}_X$, each stalk $\mathscr{M}_x$ is a module over $\mathcal{O}_{X,x}$ and we have

$$\mathscr{M}_x \subset J_x$$

where I_x is an *injective* module over $\mathcal{O}_{X,x}$ (see proposition A.5.41 on page 460). Now define

$$\mathcal{F}(U) = \prod_{x \in U} J_x$$

for all open sets $U \subset X$. This is a product of the skyscraper sheaves $i_x(J_x)$ for all $x \in X$ — see example B.1.7 on page 498. The $\mathcal{F}(U)$ are all injective modules because products of injectives are injective. This *huge* sheaf is clearly flasque (by construction) and $\mathcal{F}_x = J_x$ for all $x \in X$.

If $\mathcal{G}$ is any module over $\mathcal{O}_X$, we have

$$(D.3.2) \qquad \hom_{\mathcal{O}_X}(\mathcal{G}, \mathcal{F}) = \prod_{x \in X} \hom_{\mathcal{O}_X}(\mathcal{G}, i_x(J_x)) = \prod_{x \in X} \hom_{\mathcal{O}_{X,x}}(\mathcal{G}_x, J_x)$$

so that the injections $\mathcal{M}_x \subset J_x$ induce an injection of sheaves as in D.3.1 on the previous page.

We claim that equation D.3.2 implies that $\mathcal{F}$ is an injective object, or that $\hom_{\mathcal{O}_X}(*, \mathcal{F})$ is an exact functor (it is already left-exact, but we need to know that it is also right-exact). This follows from:

(1) the functor that maps a sheaf to its stalks at any point is exact, since it is a direct limit and direct-limits are exact (see exercise 20 on page 473).
(2) the functors $\hom_{\mathcal{O}_{X,x}}(\mathcal{G}_x, J_x)$ are exact because J_x are injective modules.

The conclusion follows. □

With this in mind, we define

DEFINITION D.3.2. If $(X, \mathcal{O}_X)$ is a ringed space and $\mathcal{F}$ is a module over $\mathcal{O}_X$, then $\mathcal{F}$ has an injective resolution $\mathscr{I}^*$ and the cohomology groups

$$H^i(X, \mathcal{F}) = H^i(\mathscr{I}^*(X))$$

are well-defined.

REMARK. If we regard evaluating a sheaf on the whole space X as a functor of the sheaf, these cohomology groups are the right derived functors of that. They measure the extent to which an exact sequence of sheaves

$$0 \to \mathcal{F} \to \mathcal{G} \to \mathcal{H} \to 0$$

fails to give an exact sequence

$$0 \to \mathcal{F}(X) \to \mathcal{G}(X) \to \mathcal{H}(X) \to 0$$

— recall the discussion in section D.1.2 on page 526. In other words, they measure the extent to which the phenomena in example B.2.1 on page 498 happens. This is subtly influenced by the geometry of X — in that example, the target space was $\mathbb{C} \setminus \{0\}$ and that "hole" in the space gave rise to sheaf-cohomology and the failure of the sequence of sheaves to be exact.

Unfortunately, the injective sheaves computed in proposition D.3.1 on the preceding page are "monsters" and it is not clear how one would ever compute sheaf-cohomology.

Section D.3.2 on the next page will reduce this computation to resolutions by flasque sheaves, and section D.3.15 on page 553 will further reduce it to sheaves on open affines. At that point, sheaf-cohomology will be a practical tool.

D.3.2. Reduction to flasque sheaves. The interesting thing about flasque sheaves is:

PROPOSITION D.3.3. *If $(X, \mathcal{O}_X)$ is a ringed space and $\mathcal{F}$ is a flasque sheaf over $\mathcal{O}_X$ then*

$$H^i(X, \mathcal{F}) = 0$$

for $i > 0$.

REMARK. It follows from corollary D.1.22 on page 530 that we can use resolutions by *flasque sheaves* to compute sheaf-cohomology. These will turn out to be much more tractable than injective sheaves.

PROOF. Construct a short exact sequence of sheaves

$$0 \to \mathcal{F} \to \mathcal{G} \to \mathcal{H} \to 0$$

where $\mathcal{G}$ is a flasque injective sheaf constructed in proposition D.3.1 on page 545. Exercise 1 on page 501 implies that $\mathcal{H}$ is also flasque.

Then $H^i(X, \mathcal{G}) = 0$ for $i > 0$ and we get a long exact sequence

$$0 \to \mathcal{F}(X) \to \mathcal{G}(X) \to \mathcal{H}(X) \to H^1(X, \mathcal{F}) \to H^1(X, \mathcal{G}) = 0$$

Since

$$0 \to \mathcal{F}(X) \to \mathcal{G}(X) \to \mathcal{H}(X) \to 0$$

is exact, we conclude that $H^1(X, \mathcal{F}) = 0$ for any flasque sheaf $\mathcal{F}$. The rest of the long exact sequence implies that

$$H^i(X, \mathcal{H}) \cong H^{i+1}(X, \mathcal{F})$$

for all $i \geq 1$, so we conclude that all *higher* cohomology groups also vanish. □

This has several interesting consequences:

COROLLARY D.3.4 (Grothendiek's Theorem). *If X is an irreducible space (see definition 2.4.16 on page 67) and $\mathcal{F}$ is a constant sheaf, then*

$$H^i(X, \mathcal{F}) = 0$$

for $i > 0$.

REMARK. Cohomology over constant sheaves (or cohomology with constant coefficients) are vital invariants of a space in algebraic topology. This result implies that the Zariski topology renders cohomology uninteresting and somewhat useless for studying topological properties of schemes.

This motivates the development of étale cohomology — see [**112**].

PROOF. Since X is irreducible, open sets are connected and if $\mathcal{F} = \underline{A}$, where A is an abelian group, it follows that $\mathcal{F}(U) = A$ for *any* open set A (compare with exercise 9 on page 502). This implies that $\mathcal{F}$ is flasque. The conclusion follows from proposition D.3.3. □

Another consequence is:

COROLLARY D.3.5. *Let X be a space with closed subspace $j : Y \hookrightarrow X$. If $\mathcal{F}$ is a sheaf of abelian groups on Y then*

$$H^i(X, j_*\mathcal{F}) = H^i(Y, \mathcal{F})$$

for all i. Here, $i_*\mathcal{F}$ is the direct image sheaf (see exercise 3 on page 498).

PROOF. If

$$0 \to \mathscr{F} \to \mathscr{I}_0 \to \cdots$$

is a flasque resolution of $\mathscr{F}$, then

$$0 \to j_*\mathscr{F} \to j_*\mathscr{I}_0 \to \cdots$$

is also a flasque resolution (see exercise 2 on page 502) and

$$j_*\mathscr{I}_k(X) = \mathscr{I}_k(Y)$$

for all k so the cohomology must be the same. □

D.3.3. Affine schemes. Now we will compute sheaf-cohomology over an affine scheme. We need several algebraic results first.

The Artin-Rees Theorem (A.4.51 on page 438) implies that:

THEOREM D.3.6 (Krull's Theorem). *Let R be a noetherian ring, let $M \subseteq N$ be finitely generated R-modules and let $\mathfrak{a}$ be an ideal of R. For any integer $n > 0$, there exists an integer $n' \geq n$ such that*

$$\mathfrak{a}^n \cdot M \supseteq M \cap \mathfrak{a}^{n'} N$$

We will use this to show that flasque sheaves are very easy to find on an affine scheme.

We need some algebraic results:

LEMMA D.3.7. *Suppose I is an injective module over a noetherian ring, R, $\mathfrak{a} \subset R$ is an ideal and $J \subseteq I$ is the submodule defined by*

$$J = \{x \in I | \exists_n \mathfrak{a}^n \cdot x = 0\}$$

Then J is also an injective module.

PROOF. Using Baer's Criterion (proposition A.5.36 on page 459) it suffices to show that, for any ideal $\mathfrak{b} \subset R$, and any homomorphism

$$f: \mathfrak{b} \to J$$

there exists an extension of f to all of R. The definition of J implies that $f(\mathfrak{b}) \cdot \mathfrak{a}^n = f(\mathfrak{b} \cdot \mathfrak{a}^n) = 0$ for some $n > 0$. Krull's Theorem (D.3.6) implies that there exists an n' such that $\mathfrak{b} \cdot \mathfrak{a}^n \supseteq \mathfrak{b} \cap \mathfrak{a}^{n'}$ so $f(\mathfrak{b} \cap \mathfrak{a}^{n'}) = 0$ and we conclude that f factors through $\mathfrak{b}/\mathfrak{b} \cap \mathfrak{a}^{n'}$ and the diagram

commutes. Since I is injective, the map $\mathfrak{b}/\mathfrak{b} \cap \mathfrak{a}^{n'}$ to I extends to $h: R/\mathfrak{a}^{n'} \to I$. Since this image is annihilated by $\mathfrak{a}^{n'}$, the image of h actually lies in $J \subset I$. The composite $g \circ p: R \to J$ is the required extension of f. □

We also need:

LEMMA D.3.8. *Let I be an injective module over a noetherian ring, R. If $S \subset R$ is a multiplicative set, then the natural map*

$$\theta: I \to S^{-1}I$$

is surjective.

PROOF. We will prove this in the case where S consists of powers of a single element, $s \in R$. Let $\mathfrak{a}_i$ be the ideal that annihilates s^i. Since R is noetherian, the ascending chain

$$\mathfrak{a}_1 \subseteq \cdots \subseteq \mathfrak{a}_n = \mathfrak{a}_{n+1} = \cdots$$

of ideals becomes constant from some point on. If $x \in I_s$, we will show that it is in the image of $y \in I$. The definition of localization implies that there exists an element $z \in I$ such that $x = \theta(z)/s^m$ for some value of m. Define a homomorphism

$$\begin{aligned} \tau: (s^{n+m}) &\to I \\ \alpha \cdot s^{n+m} &\mapsto s^n \cdot z \end{aligned}$$

This is possible since the annihilator, $\mathfrak{a}_{n+m}$ of s^{n+m} is the same as the annihilator, $\mathfrak{a}_n$ of s^n. Since I is injective, τ extends to all of R. Let $\tau(1) = y$ so that $s^{n+m}y = s^n z$. Then $\theta(w) = \theta(y)/s^m = x$. $\qquad\square$

The main result needed to compute the cohomology of sheaves over affine schemes is:

COROLLARY D.3.9. *If $V = \operatorname{Spec} R$, where R is a noetherian ring, and I is any injective module over R, then the sheaf $\mathcal{A}(I)$ is flasque.*

REMARK. See definition 3.5.1 on page 155 for the notation $\mathcal{A}(I)$.
This result means that it is easy to find flasque sheaves on an affine scheme.

PROOF. We will do induction on chains of closed subschemes — R being noetherian implies that there is a finite number of them.
The ground-case: If the support (see definition B.1.5 on page 497) of $\mathcal{A}(I)$ is a *single point,* this sheaf is a skyscraper sheaf, hence flasque.
The induction step: We must show that, if the statement of D.3.9 (our result) is true for all *proper* closed subsets of V it is also true for V. Suppose the support of $\mathcal{A}(I)$ is $Y \subset V$, a fixed closed subset of V. If

$$\operatorname{Supp} \mathcal{A}(I) \subsetneq Y$$

then $\mathcal{A}(I)$ is flasque by induction. We consequently assume that

$$\operatorname{Supp} \mathcal{A}(I) = Y$$

We must show that, for every open set, $U \subset V$ the restriction maps

$$\mathcal{A}(I)(V) \to \mathcal{A}(I)(U)$$

are surjective (lemma D.3.8 implies this for *principal* open sets since $I \to I_f$ is surjective). If $Y \cap U = \varnothing$, there is nothing to prove. Otherwise, there exists

a principal open set $D(f)$ with $D(f) \cap Y \neq \emptyset$. Let $Z = V \setminus D(f)$. We get a commutative diagram

$$
\begin{array}{ccccc}
\mathcal{A}(I)(V) & \longrightarrow & \mathcal{A}(I)(U) & \longrightarrow & \mathcal{A}(I)(D(f)) \\
\uparrow & & \uparrow & & \\
\mathcal{A}(I)(V)_Z & \longrightarrow & \mathcal{A}(I)(U)_Z & &
\end{array}
$$

where $\mathcal{A}(I)(V)_Z$ and $\mathcal{A}(I)(U)_Z$ are the submodules of elements whose support is in Z, or the submodules whose restriction to $D(f)$ vanishes.

If $x \in \mathcal{A}(I)(U)$ then its image $t \in \mathcal{A}(I)(D(f))$ lifts to an element $T \in \mathcal{A}(I)(V)$. We do not know whether $T|U = x$, but we do know that $x - T|U$ maps to 0 in $\mathcal{A}(I)(D(f))$ so that $x - T|U \in \operatorname{im} \mathcal{A}(I)(U)_Z$. The result will be proved if we can show that $\mathcal{A}(I)(V)_Z \to \mathcal{A}(I)(U)_Z$ is surjective.

Note that $i_*(\mathcal{A}(I)|D(f)) = \mathcal{A}(I_f)$, where $i_*\mathcal{A}(I)|D(f)$ is the direct image sheaf (see exercise 3 on page 498 and exercise 5 on page 197) of the restriction. If $\mathscr{F} = \ker \mathcal{A}(I) \to i_*\mathcal{A}(I)|D(f)$, then $\mathscr{F}(U) = \mathcal{A}(I)(U)_Z$ for all open sets, U and $\mathscr{F}$ is also quasi-coherent since it is a kernel of a morphism of quasi-coherent sheaves, hence coherent, by corollary 4.4.18 on page 194.

It follows that $\mathscr{F} = \mathcal{A}(J)$, where $J = \ker I \to I_f$. Since $J = \{x \in I | \exists_n f^n \cdot x = 0\}$ (see definition A.1.89 on page 383), lemma D.3.8 on the preceding page implies that J is an injective module.

Since the support of $\mathcal{A}(J)$ lies on an *proper closed subset* of Y, the inductive hypothesis implies that $\mathcal{A}(J)$ is flasque, whence $\mathcal{A}(J)(V) = \mathcal{A}(I)(V)_Z \to \mathcal{A}(I)(U)_Z$ is surjective. $\qquad\square$

We arrive at our main result:

THEOREM D.3.10. *If $V = \operatorname{Spec} R$ is an affine noetherian scheme and $\mathscr{F}$ is any quasi-coherent sheaf, then*

$$
H^i(V, \mathscr{F}) = \begin{cases} \mathscr{F}(V) & \text{if } i = 0 \\ 0 & \text{otherwise} \end{cases}
$$

REMARK. This means that sheaf-cohomology as developed here is generally not interesting for affine schemes. It also implies that sheaf-cohomology for more general varieties measures how the open affines fit together.

The use of flasque sheaves implies that this result is true in the category of abelian sheaves (see . If we had only been concerned with the category of quasi-coherent sheaves, we could have simply noted that $\mathcal{A}(I)$ is an injective quasi-coherent sheaf and ignored the whole issue of flasque sheaves.

PROOF. First of all, $\mathscr{F} = \mathcal{A}(M)$ for some module, M. Let

$$
0 \to M \to I_0 \to \cdots
$$

be an injective resolution of M. Then

$$
0 \to \mathcal{A}(M) \to \mathcal{A}(I_0) \to \cdots
$$

is a flasque resolution of $\mathscr{F}$. If we evaluate these sheaves on V, we recover the original injective resolution of M, which is acyclic in positive degrees. $\qquad\square$

D.3.4. Čech cohomology. With theorem D.3.10 on the preceding page under our belt, we are in a position to compute cohomology of general varieties and schemes.

We begin by defining a variant on cohomology called Čech cohomology — a "topological" cohomology, as described in section D.1.1 on page 524. This form of cohomology does not require breaking a space into a union of polyhedra — making it suitable for studying certain types of topological spaces (including the ones that occur in algebraic geometry). It was first described by Eduard Čech in [**27**] (where he also gave the first description of the inverse limit).

DEFINITION D.3.11. Let V be a scheme with a finite open cover $\mathcal{U} = \{U_i\}$, $i = 1, \ldots, n$, and sheaf $\mathscr{F}$. Define $U_{i_0, \ldots, i_t} = U_{i_0} \cap \cdots \cap U_{i_t}$ and

$$C(\mathcal{U}, \mathscr{F})^t = \prod_{i_0 < \cdots < i_t} \mathscr{F}(U_{i_0, \ldots, i_t})$$

and let

$$r_j \colon \mathscr{F}(U_{i_0, \ldots, i_{j-1}, i_{j+1}, \ldots i_{t+1}}) \to \mathscr{F}(U_{i_0, \ldots, i_{t+1}})$$

be the restriction-maps. If

$$x = \prod_{i_0 < \cdots < i_t} x_{i_0, \ldots, i_t} \in C(\mathcal{U}, \mathscr{F})^t$$

define

$$\delta^t \colon C(\mathcal{U}, \mathscr{F})^t \to C(\mathcal{U}, \mathscr{F})^{t+1}$$

via

$$\delta^t(x)_{i_0, \ldots, i_{t+1}} = \sum_{j=0}^{t} (-1)^j r_j(x_{i_0, \ldots, i_{j-1}, i_{j+1}, \ldots i_{t+1}})$$

REMARK. It is not hard to verify that this defines a cochain complex — this is left to the reader.

Since we have a cochain complex, we can define cohomology. Although $C(\mathcal{U}, \mathscr{F})$ depends on the open covering, note that if $\mathcal{U}'$ is a refinement of $\mathcal{U}$, we get a homomorphism

$$C(\mathcal{U}, \mathscr{F}) \to C(\mathcal{U}', \mathscr{F})$$

induced by the restriction maps of $\mathscr{F}$. With this in mind, define

DEFINITION D.3.12. If $\mathscr{F}$ is a sheaf on a scheme V with a finite open cover $\{\mathcal{U}\}$, define the *Čech cohomology* of $\mathscr{F}$ via

$$\check{H}^i(V, \mathscr{F}) = \varinjlim H^i(\mathcal{U}, \mathscr{F})$$

where the direct limit is taken over all finite open coverings of V.

REMARK. Eduard Čech defined this in [**27**] because he studied topological spaces that did not lend themselves to being broken up into discrete pieces. Although it is not identical to standard cohomology, the two agree in all cases that will interest us. It will be more suitable to algebraic geometric computations than standard cohomology.

It turns out that the direct limit is easier to compute than one might expect. To see this, define the Čech resolution of $\mathscr{F}$ by

DEFINITION D.3.13. Let V be a scheme with a finite open cover $\mathcal{U} = \{U_i\}$, $i = 1, \ldots, n$, and sheaf $\mathscr{F}$. The *Čech resolution* of $\mathscr{F}$ is defined as follows:
If $U_{i_0,\ldots,i_t} = U_{i_0} \cap \cdots \cap U_{i_t}$ then

$$\mathscr{C}(\mathcal{U}, \mathscr{F})^t = \prod_{i_0 < \cdots < i_t} i_* \mathscr{F}|U_{i_0,\ldots,i_t}$$

where $i_* \mathscr{F}|U_{i_0,\ldots,i_t}$ is the direct image sheaf and let

$$r_j: i_* \mathscr{F}|U_{i_0,\ldots,i_{j-1},i_{j+1},\ldots i_{t+1}} \to i_*|U_{i_0,\ldots,i_{t+1}}$$

be the restriction-maps. Define

$$\delta^t: \mathscr{C}(\mathcal{U}, \mathscr{F})^t \to \mathscr{C}(\mathcal{U}, \mathscr{F})^{t+1}$$

via

$$\delta^t(x)_{i_0,\ldots,i_{t+1}} = \sum_{j=0}^{t} (-1)^j r_j$$

REMARK. This is a cochain-complex of *sheaves* closely related to the Čech complex of *modules* defined above by

$$\mathscr{C}(\mathcal{U}, \mathscr{F})(V) = C(\mathcal{U}, \mathscr{F})$$

We must justify the term "resolution" above:

PROPOSITION D.3.14. *The complex of sheaves constructed in definition D.3.13 is a resolution — i.e. acyclic.*

PROOF. Exactness at the low end

$$\mathscr{F} \to \mathscr{C}(\mathcal{U}, \mathscr{F})^0 \xrightarrow{\delta^0} \mathscr{C}(\mathcal{U}, \mathscr{F})^1$$

follows from the global sections property of a sheaf: the kernel of δ^0 is a sheaf on V constructed from sheaves on the U_i that agree on all of the overlaps $U_i \cap U_j$. To prove acyclicity in higher degrees, it suffices to prove it on every stalk (see exercise 7 on page 502).
If $x \in V$ is an arbitrary point, we will construct a chain-homotopy (as in definition D.1.5 on page 521)

$$\mathscr{C}(\mathcal{U}, \mathscr{F})^n_x \xrightarrow{\Phi} \mathscr{C}(\mathcal{U}, \mathscr{F})^{n-1}_x$$

such that $\delta \circ \Phi + \Phi \circ \delta = 1$, so that the identity map induces the same map in cohomology as the 0-map (see proposition D.1.6 on page 521) — which is only possible if the cohomology vanishes.
If $y \in \mathscr{C}(\mathcal{U}, \mathscr{F})^n_x$, fix an index, J, with $x \in U_J$. If necessary, let $W \subset U_J$ be a smaller open set such that y is the image of $Y \in \mathscr{C}(\mathcal{U}, \mathscr{F})^n(W)$ under the direct limit that defines the stalk. This $Y = Z(W)$, where $Z \in \mathscr{C}(\mathcal{U}, \mathscr{F})^n$ is a product of sheaves over open sets. Now define

$$\Phi'(Z)(U_{i_0,\ldots,i_{n-1}}) = Z(U_{J,i_0,\ldots,i_{n-1}})$$

and define $\Phi(y)$ to be the image of $\Phi'(Z)(W)$ in the stalk, $\mathscr{C}(\mathcal{U}, \mathscr{F})_x^{n-1}$. This is well-defined because U_J is one of the open sets in the open cover and

$$W \cap U_{J,i_0,\dots,i_{n-1}} = W \cap U_{i_0,\dots,i_{n-1}}$$

We must verify that $(\delta \circ \Phi + \Phi \circ \delta)(y) = y$:

$$
\begin{aligned}
\Phi'(\delta(Z)(U_{i_0,\dots,i_m})) &= \delta Z(U_{J,i_0,\dots,i_m}) \\
&= Z(U_{i_0,\dots,i_m}) \\
&\quad - \sum_{j=0}^m (-1)^j Z(U_{J,i_0,\dots,i_{j-1},i_{j+1},\dots,i_m}) \\
&= Z(U_{i_0,\dots,i_m}) - \delta(\Phi'Z)
\end{aligned}
$$

So, evaluating this on W and restricting to the stalk gives the required result. $\square$

Our main result is (see [**150**]):

THEOREM D.3.15. *If V is a noetherian separated scheme and $\mathcal{U} = \{U_i\}$ is any cover by open affines then*

$$H^i(C(\mathcal{U}, \mathscr{F})) = H^i(V, \mathscr{F})$$

for any quasi-coherent sheaf $\mathscr{F}$.

REMARK. There is *more* to this bland statement than meets the eye. It asserts equality between two very different objects. On the *left*, we have an intermediate stage to computing Čech cohomology — a topologically defined cohomology describing functions on a topological space. On the *right*, we have the somewhat mysterious *derived-functor* cohomology expressing properties of the global-sections functor $\Gamma(V, \mathscr{F}) = \mathscr{F}(V)$.

This statement also says that, in computing $\check{H}^i(V, \mathscr{F})$, we can stop right here since the direct limit in definition D.3.12 on page 551 becomes *constant* as soon as the $\mathcal{U}$ consists of open affines — and such coverings are cofinal in the collection of all open covers (see definition A.5.21 and proposition A.5.22). We conclude that

$$\check{H}^i(V, \mathscr{F}) = H^i(C(\mathcal{U}, \mathscr{F})) = H^i(V, \mathscr{F})$$

PROOF. The fact that V is separated implies that any intersection of open affines is *also* affine (see exercise 1 on page 210). It follows that *all* of the open sets used in computing the Čech complex are affine. Now, simply compute the Čech resolution for the open cover given. Theorem D.3.10 on page 550 implies that it will be an F-acyclic resolution, where F is the global-section functor (see from sheaves to modules with

$$F(\mathscr{G}) = \Gamma(V, \mathscr{G}) = \mathscr{G}(V)$$

Corollary D.1.22 on page 530 implies the conclusion. $\square$

COROLLARY D.3.16. *If V is a separated noetherian affine scheme and $\mathscr{F}$ is any quasi-coherent sheaf on V, then*

$$H^0(V, \mathscr{F}) = \mathscr{F}(V)$$

PROOF. Elements of H^0 are sets elements $x_i \in \mathscr{F}(U_i)$ and the conditions $\delta(\{x_i\}) = 0$ is just the requirement that

$$x_i|U_i \cap U_j = x_j|U_i \cap U_j$$

which is the condition that $x_i = x|U_i$ for some $x \in \mathscr{F}(V)$. $\square$

This immediately implies interesting results:

COROLLARY D.3.17. *If V is a separated noetherian scheme that can be covered by $n + 1$ open affines, and $\mathscr{F}$ is any quasi-coherent sheaf on V, then*

$$H^i(V, \mathscr{F}) = 0$$

for $i > n$.

REMARK. For $i > n$ we do not have enough *distinct* open sets to construct higher terms in the Čech complex, so the cohomology vanishes because the Čech *complex* vanishes.

PROOF. This follows immediately from definition D.3.13 on page 552, which shows that $\mathscr{C}(\mathcal{U}, \mathscr{F})^t = 0$ for $t > n$. $\square$

A variation on this result suggests a strong relationship between the dimension of a variety and the cohomology of sheaves on it (see [**150**]):

LEMMA D.3.18. *If $V \subset k\mathbb{P}^n$ is an m-dimensional projective variety and $\mathscr{F}$ is a quasi-coherent sheaf on V, then*

$$H^i(V, \mathscr{F}) = 0$$

for $i > m$.

PROOF. Corollary D.3.5 on page 547 implies that

$$H^i(k\mathbb{P}^n, j_*\mathscr{F}) = H^i(V, \mathscr{F})$$

where $j\colon V \hookrightarrow k\mathbb{P}^n$ is the inclusion. We will find a special type of open covering of $k\mathbb{P}^n$ that implies the result. Lemma 5.6.5 on page 264 implies the existence of $m + 1$ homogeneous polynomials $f_1 \ldots, f_{m+1} \in k[X_0, \ldots, X_n]$ that do not simultaneously vanish on V and exercise 2 on page 245 implies that

$$(D.3.3) \qquad\qquad D_i = k\mathbb{P}^n \setminus \mathcal{P}((f_i))$$

$i = 1, \ldots, m + 1$, open affines of $k\mathbb{P}^n$ whose union contains V. We complete this to a set of open affines for all of $k\mathbb{P}^n$. Let $f_{m+1}, \ldots, f_h$ be a set of homogeneous polynomials that vanish on V but *never* simultaneously vanish on $k\mathbb{P}^n \setminus V$ — for instance a set of homogeneous generators of the ideal that defines V.

Applying equation D.3.3 to these gives an affine open covering

$$\mathcal{U} = \{D_i\}$$

for $i = 1, \ldots, h$ of $k\mathbb{P}^n$ (where $h \geq n + 1$) — and $D_t \cap V = \varnothing$ for $t > m + 1$. Now theorem D.3.15 on the preceding page implies that

$$H^i(\mathcal{U}, j_*\mathscr{F}) = H^i(k\mathbb{P}^n, j_*\mathscr{F}) = H^i(V, \mathscr{F})$$

and, in computing $H^i(\mathcal{U}, j_*\mathscr{F})$, note that $j_*\mathscr{F}(U_{i_0,\ldots,i_t}) = 0$ if any of the subscripts $i_s > m + 1$ (by the way the D_i were constructed). It follows that the cohomology vanishes above degree m. $\square$

Now we can do some computations:

EXAMPLE D.3.19. Consider the example $V = \mathbb{A}^2 \setminus \{(0,0\}$ of a non-affine variety in example 4.6.2 on page 205. We can cover V by *two* open affines $D(X)$ and $D(Y)$, which means that

$$H^i(V, \mathscr{F}) = 0$$

for $i > 1$ (by corollary D.3.17 on the facing page). If we let $\mathscr{F} = \mathcal{O}_V$, we get

$$\begin{aligned} C(\mathcal{U}, \mathcal{O}_V)^0 &= k[D(X)] \oplus k[D(Y)] \\ &= k[X, X^{-1}, Y] \oplus k[X, Y, Y^{-1}] \end{aligned}$$

so the kernel of δ^0 is

$$H^0(V, \mathcal{O}_V) = k[X, Y] = \mathcal{O}_V(V)$$

the coordinate ring.

In degree 1,

$$C(\mathcal{U}, \mathcal{O}_V)^1 = k[X, X^{-1}, Y, Y^{-1}] = k[D(X) \cap D(Y)]$$

and

$$\begin{aligned} \delta^0 : k[X, X^{-1}, Y] \oplus k[X, Y, Y^{-1}] &\rightarrow k[X, X^{-1}, Y, Y^{-1}] \\ p(X, X^{-1}, Y) \oplus q(X, Y, Y^{-1}) &\mapsto p - q \end{aligned}$$

We can think of $k[X, X^{-1}, Y, Y^{-1}]$ as consisting of linear combinations of monomials $X^i Y^j$ where i and j can be arbitrary integers. Monomials with i or $j \geq 0$ lie in the image of δ^0 so

$$H^1(V, \mathcal{O}_V) = \bigoplus_{i,j=1}^{\infty} k \cdot X^{-i} Y^{-j}$$

This distinguishes V from $\mathbb{A}^2$ since the cohomology of $\mathbb{A}^2$ *vanishes* in positive degrees by theorem D.3.10 on page 550.

Another, more relevant example is:

EXAMPLE D.3.20. Let $V = R\mathbb{P}^1$ with open affines $\mathbb{A}_0^1$ and $\mathbb{A}_1^1$ and let $\mathscr{F} = \Omega_V^1$, regarded as a coherent sheaf — i.e. $\mathscr{F}(U) = \Omega_U^1$ (compare corollary 5.9.33 on page 299), so that $\Omega_V^1 = \mathscr{F}(V)$. Then

$$\begin{aligned} \mathscr{F}(\mathbb{A}_0^1) &= R[X] \cdot dX \\ \mathscr{F}(\mathbb{A}_1^1) &= R[Y] \cdot dY \end{aligned}$$

so

$$C(\mathcal{U}, \mathscr{F})^0 = R[X] \cdot dX \oplus R[Y] \cdot dY$$

One the overlap

$$\mathscr{F}(\mathbb{A}_0^1 \cap \mathbb{A}_1^1) = R[X, X^{-1}] \cdot dX = C(\mathcal{U}, \mathscr{F})^1$$

with δ^0 defined by

$$\begin{aligned} \delta^1(f(X) \cdot dX \oplus g(Y) \cdot dY) &= f(X) \cdot dX - \frac{g(1/X)}{X^2} \cdot dX \\ &= \left(f(X) - \frac{g(1/X)}{X^2} \right) \cdot dX \end{aligned}$$

since $Y = 1/X$ when we glue the pieces together. The kernel of δ^0 is clearly 0
so

$$H^0(V, \mathscr{F}) = 0$$

confirming that there are no global differential forms on $R\mathbb{P}^1$. The cokernel
of δ^1 is generated by monomials X^i, where i is either ≥ 0 (derived from the
f-term) or $i \leq -2$ (derived from the g-term). It follows that

$$H^1(V, \mathscr{F}) = R \cdot X^{-1} \cong R$$

Recall the Serre-twist sheaves on a projective space in definition 5.3.2 on
page 234. The isomorphism $\mathcal{O}_{k\mathbb{P}^n}(r) \otimes \mathcal{O}_{k\mathbb{P}^n}(s) = \mathcal{O}_{k\mathbb{P}^n}(r+s)$ induces a coho-
mology homomorphism

$$H^0(k\mathbb{P}^n, \mathcal{O}_{k\mathbb{P}^n}(r)) \otimes H^i(k\mathbb{P}^n, \mathcal{O}_{k\mathbb{P}^n}(s)) \to H^i(k\mathbb{P}^n, \mathcal{O}_{k\mathbb{P}^n}(r+s))$$

— simply take the products of values on open-sets.

This example is relevant for many other results:

THEOREM D.3.21. *If R is a noetherian ring and $V = R\mathbb{P}^n$, then*

(1) $H^0(V, \mathcal{O}_{R\mathbb{P}^n}(d)) = R[X_0, \ldots, X_n]_d$ — *the free module of monomials of de-*
gree d of rank

$$\binom{d+n}{n}$$

so

$$\bigoplus_{d=0}^{\infty} H^0(V, \mathcal{O}_{R\mathbb{P}^n}(d)) = R[X_0, \ldots, X_n]$$

(2) $H^i(V, \mathcal{O}_{R\mathbb{P}^n}(d)) = 0$ *for $0 < i < n$.*

(3) $H^n(V, \mathcal{O}_{R\mathbb{P}^n}(d)) = \begin{cases} 0 & \text{if } d > -n-1 \\ R^{\phi(d)} & \text{otherwise} \end{cases}$ *where $\phi(d) = \binom{-d}{n+1}$.*

(4) *For $m \in \mathbb{Z}$ there is a perfect pairing*

$$H^0(V, \mathcal{O}_{R\mathbb{P}^n}(m)) \otimes H^n(V, \mathcal{O}_{R\mathbb{P}^n}(-m-n-1)) \to H^n(V, \mathcal{O}_{R\mathbb{P}^n}(-n-1)) = R$$

REMARK. "Perfect pairing" means that

$$H^0(V, \mathcal{O}_{R\mathbb{P}^n}(d)) = \hom_R(H^n(V, \mathcal{O}_{R\mathbb{P}^n}(-d-n-1)), R)$$

the dual.

PROOF. We will use the open cover $\mathcal{U} = \{\mathbb{A}_i^n\}$, for $i = 0, \ldots, n$. We will
follow a proof inspired by that in the Stacks Project, [**159**, Lemma 10.1, in co-
herent.pdf].

We proceed by noticing that there is considerably *less* to the
cochain-complex, $C(\mathcal{U}, \mathcal{O}_{k\mathbb{P}^n}(d))$, than meets the eye.

If $R(n_0, \ldots, n_t) = R[X_0, \ldots, X_n][X_{n_0}^{-1}, \ldots, X_{n_t}^{-1}]$ then

$$C(\mathcal{U}, \mathcal{O}_{R\mathbb{P}^n}(d))^q = \bigoplus_{n_0 < \cdots < n_q} R(n_0, \ldots, n_q)$$

and the coboundaries, δ^q, just map monomials, *unchanged,* into modules in
which *more* of the X_i are allowed to have negative exponents. If $\vec{e} = (e_0, \ldots, e_q)$
and $C(\vec{e})$ is a cochain complex containing the single monomial

$$R \cdot X_0^{e_0} \cdots X_n^{e_n}$$

mapped around to the various $R(n_0, \ldots, n_t)$ (via $\pm$the identity map), it follows that

(D.3.4)
$$C(\mathcal{U}, \mathcal{O}_{k\mathbb{P}^n}(d)) = \bigoplus_{|\vec{e}|=d} C(\vec{e})$$

where $|\vec{e}| = e_0 + \cdots + e_t$, a direct sum of *very simple* cochain-complexes. Its cohomology will be the direct sum of the cohomology of these cochain-complexes. Define $\text{NEG}(\vec{e})$ to the set of subscripts i for which $e_i < 0$. This determines which cochain modules can appear in $C(\vec{e})$:

$$C(\vec{e})^q = \bigoplus_{\substack{n_0 < \cdots < n_q \\ \text{NEG}(\vec{e}) \subset \{n_0, \ldots, n_q\}}} R(n_0, \ldots n_q)$$

If $\text{NEG}(\vec{e}) = \varnothing$, then $C(\vec{e})$ is a cochain-complex that simply maps a *single monomial*, $\beta = X_0^{e_0} \cdots X_n^{e_n}$, throughout all of the $R(n_0, \ldots, n_q)$. It is not hard to see that $\beta \in \ker \delta^0$, so $H^0(C(\vec{e})) = R$. We will describe element of $C(\vec{e})^q$ via a function

$$c(i_0, \ldots, i_q) \in R$$

where $c(i_0, \ldots, i_q)$ is the *coefficient* of the monomial β in $R(n_1, \ldots, n_q)$. Then

$$(\delta c)(i_0, \ldots, i_{q+1}) = \sum_{j=0}^{q+1} (-1)^j c(i_0, \ldots, i_{j-1} i_{j+1}, \ldots, i_{q+1})$$

describes the coboundaries.

Define the map

$$\Phi \colon C(\vec{e})^q \to C(\vec{e})^{q-1}$$

by $\Phi(c)(i_0, \ldots, i_q) = c(0, i_0, \ldots, i_q)$, if $q > 0$. Then

$$
\begin{aligned}
\delta\Phi(c)(i_0, \ldots, i_q) &= \delta c(0, i_0, \ldots, i_q) \\
&= c(i_0, \ldots, i_q) \\
&\quad - \sum_{j=0}^{q} c(0, i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, i_q) \\
&= c(i_0, \ldots, i_q) - \Phi\delta c(i_0, \ldots, i_q)
\end{aligned}
$$

so, $\delta \circ \Phi + \Phi \circ \delta = 1$ in positive degrees and the identity map of $C(\vec{e})$ of is homotopic to the 0-map (see definition D.1.5 on page 521). It follows that $C(\vec{e})$ has cohomology equal to R in degree 0 and 0 in all higher degrees. In degree 0, the number of copies of R that can occur is equal to the number of monomials with $|\vec{e}| = d$, which is what was listed.

If $\text{NEG}(\vec{e}) = \{0, \ldots, n\}$ then

$$C(\vec{e})^q = \begin{cases} R \cdot X_0^{e_0} \cdots X_n^{e_n} & \text{if } q = n \\ 0 & \text{otherwise} \end{cases}$$

Since this is the *only* cochain-module, its coboundary maps are 0 and its cohomology is *equal* to it, so equation D.3.4 implies that

$$H^n(C(\mathcal{U}, \mathcal{O}_{R\mathbb{P}^n}(d))) = \begin{cases} \bigoplus_{|\vec{e}|=d} R \cdot X_0^{e_0} \cdots X_n^{e_n} & \text{if } d \leq -n-1 \\ 0 & \text{otherwise} \end{cases}$$

The number of possible summands is clearly $\binom{-d}{n+1}$, as stated.

Statement 2 slightly more complicated: If $\text{NEG}(\vec{e}) \neq \varnothing$ and $\text{NEG}(\vec{e}) \neq \{0, \ldots, n\}$, we claim that $C(\vec{e})$ is acyclic in *all degrees*. Pick a fixed subscript $J \notin \text{NEG}(\vec{e})$ and define

$$\Phi \colon C(\vec{e})^q \to C(\vec{e})^{q-1}$$

via

$$\Phi c(i_0, \ldots, i_q) = (-1)^\alpha c(i_0, \ldots, J, \ldots, i_q)$$

where J occurs in the αth position. As before

$$
\begin{aligned}
\delta \Phi c(i_0, \ldots, i_q) &= (-1)^\alpha c(i_0, \ldots, J, \ldots, i_q) \\
&= c(i_0, \ldots, i_q) \\
&\quad + (-1)^\alpha \sum_{j=0}^{J-1} (-1)^j c(i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, J, \ldots, i_q) \\
&\quad - (-1)^\alpha \sum_{j=J+1}^{q} (-1)^j c(i_0, \ldots, J, \ldots i_{j-1}, i_{j+1}, \ldots, i_q) \\
&= c(i_0, \ldots, i_q) - \Phi \delta c(i_0, \ldots, i_q)
\end{aligned}
$$

so $\delta \circ \Phi + \Phi \circ \delta = 1$ on *all* summands of $C(\vec{e})$. It follows that the identity map is chain-homotopic (see definition D.1.5 on page 521) to the 0-map and $C(\vec{e})$ is acyclic in all degrees.

The final statement follows from the fact that the map

$$\mathcal{O}_{k\mathbb{P}^n}(r) \otimes \mathcal{O}_{k\mathbb{P}^n}(s) = \mathcal{O}_{k\mathbb{P}^n}(r+s)$$

multiplies sections, i.e. monomials so the pairing is defined by

$$r_1 X_0^{e_0} \cdots X_n^{e_n} \otimes r_2 X_0^{-e_0-1} \cdots X_n^{-e_n-1} = r_1 r_2 X_0^{-1} \cdots X_n^{-1}$$

$\square$

This puts us in a position to prove a number of interesting things.

THEOREM D.3.22. *Let V be a projective scheme over a noetherian ring with a very ample invertible sheaf $\mathcal{O}_V(1)$. If $\mathscr{F}$ is a coherent sheaf over V, then*

(1) *for each $i \geq 0$, $H^i(V, \mathscr{F})$ is finitely-generated,*
(2) *there exists an integer n_0 such that, for all $i > 0$ and all $n > n_0$, $H^i(V, \mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(1)^n) = 0$*

REMARK. This shows that $\mathbb{A}^2 \setminus (0,0)$ in example D.3.19 on page 555 cannot be a *projective* scheme because its cohomology is *infinitely* generated.

PROOF. We have a closed immersion

$$i \colon V \to R\mathbb{P}^t$$

If $\mathscr{F}$ is coherent over V, then $i_* \mathscr{F}$ is coherent over $R\mathbb{P}^t$ (see exercise 6 on page 197) and

$$H^j(V, \mathscr{F}) = H^j(R\mathbb{P}^t, i_* \mathscr{F})$$

— see corollary D.3.5 on page 547. It follows that, without loss of generality, we can assume $V = R\mathbb{P}^t$. We can assume that there exists a surjection

$$\bigoplus_{j=1}^{m} \mathcal{O}_{R\mathbb{P}^t}(d) \to \mathscr{F}$$

for some d and m — see lemma 5.3.10 on page 238 so we get a short exact sequence

$$0 \to \mathscr{K} \to \bigoplus_{j=1}^{m} \mathcal{O}_{R\mathbb{P}^t}(d) \to \mathscr{F} \to 0$$

where $\mathscr{K}$ is also coherent. We get a long exact sequence in cohomology

$$\cdots \to \bigoplus_{j=1}^{m} H^j(R\mathbb{P}^t \mathcal{O}_{R\mathbb{P}^t}(d)) \to H^j(R\mathbb{P}^t, \mathscr{K}) \to H^{j+1}(R\mathbb{P}^t, \mathscr{F}) \to \cdots$$

We prove the first statement by downward induction on j. If $j > t$ then cohomology vanishes, by corollary D.3.17 on page 554, so the result is true. Since , the fact that Since $H^{j+1}(R\mathbb{P}^t, \mathscr{F})$ is finitely generated by the inductive hypothesis and $H^j(R\mathbb{P}^t \mathcal{O}_{R\mathbb{P}^t}(d))$ is finitely generated, by theorem D.3.21 on page 556 $H^j(R\mathbb{P}^t, \mathscr{K})$ must a also be finitely generated (here, we're using the fact that R is noetherian).

A similar argument proves the second statement — here, we have the exact sequence

$$0 \to \mathscr{K} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(1)^n \to \bigoplus_{j=1}^{m} \mathcal{O}_{R\mathbb{P}^t}(d) \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(1)^n$$

$$\to \mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(1)^n \to 0$$

or

$$0 \to \mathscr{K} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(n) \to \bigoplus_{j=1}^{m} \mathcal{O}_{R\mathbb{P}^t}(d+n) \to \mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(n) \to 0$$

giving

$$\cdots \to \bigoplus_{j=1}^{m} H^j(R\mathbb{P}^t \mathcal{O}_{R\mathbb{P}^t}(d+n)) \to H^j(R\mathbb{P}^t, \mathscr{K} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(n))$$

$$\to H^{j+1}(R\mathbb{P}^t, \mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(n)) \to \cdots$$

In this case, $H^{j+1}(R\mathbb{P}^t, \mathscr{F} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(n))$ is true by the inductive hypothesis, $\bigoplus_{j=1}^{m} H^j(R\mathbb{P}^t \mathcal{O}_{R\mathbb{P}^t}(d+n))$ vanishes for a suitable choice of n by theorem D.3.21 on page 556 so $H^j(R\mathbb{P}^t, \mathscr{K} \otimes_{\mathcal{O}_{R\mathbb{P}^t}} \mathcal{O}_{R\mathbb{P}^t}(n))$ must vanish. $\square$

This implies that the following makes sense (first defined in [**150**]):

DEFINITION D.3.23. Let V be a projective scheme over a field k with a very ample invertible sheaf $\mathcal{O}_V(1)$. If $\mathscr{F}$ is a coherent sheaf over V, then define

$$\chi(V, \mathscr{F}) = \sum_{i=0}^{\infty} (-1)^i \dim_k H^i(V, \mathscr{F})$$

— the *Euler characteristic* of $\mathscr{F}$ on V.

We have the following result:

PROPOSITION D.3.24. *If V is a projective scheme over a field k with a very ample invertible sheaf $\mathcal{O}_V(1)$ and*

$$0 \to \mathscr{F} \xrightarrow{f} \mathscr{G} \xrightarrow{g} \mathscr{H} \to 0$$

is an exact sequence of coherent sheaves over V, then

$$\chi(V,\mathscr{G}) = \chi(V,\mathscr{F}) + \chi(V,\mathscr{H})$$

PROOF. The short exact sequence of sheaves induces a long exact one in cohomology

$$\cdots \to H^i(V,\mathscr{F}) \xrightarrow{f_i^*} H^i(V,\mathscr{G}) \xrightarrow{g_i^*} H^i(V,\mathscr{H}) \xrightarrow{c_i} H^{i+1}(V,\mathscr{F}) \to \cdots$$

and

$$\begin{aligned}
\dim_k H^i(V,\mathscr{F}) &= \dim_k \operatorname{im} c_{i-1} + \dim_k \operatorname{im} f_i^* \\
\dim_k H^i(V,\mathscr{G}) &= \dim_k \operatorname{im} g_i^* + \dim_k \operatorname{im} f_i^* \\
\dim_k H^i(V,\mathscr{H}) &= \dim_k \operatorname{im} g_i^* + \dim_k \operatorname{im} c_i
\end{aligned}$$

so

$$\dim_k H^i(V,\mathscr{G}) - \dim_k H^i(V,\mathscr{F}) - \dim_k H^i(V,\mathscr{H})$$
$$= -\dim_k \operatorname{im} c_{i-1} - \dim_k \operatorname{im} c_i$$

which implies that

$$\sum_{i=0}^{\infty} (-1)^i \left(\dim_k H^i(V,\mathscr{G}) - \dim_k H^i(V,\mathscr{F}) - \dim_k H^i(V,\mathscr{H}) \right) = 0$$

$\square$

EXERCISES.

1. If V is a projective scheme over a noetherian ring, R, with a very ample invertible sheaf $\mathcal{O}_V(1)$ and

$$0 \to \mathscr{F} \xrightarrow{f} \mathscr{G} \xrightarrow{g} \mathscr{H} \to 0$$

is an exact sequence of coherent sheaves, show that there exists an integer n_0 such that, for any integer $n > n_0$, the induced sequence

$$0 \to (\mathscr{F} \otimes_{\mathcal{O}_V} \mathcal{O}_V(1)^n)(V)$$
$$\to (\mathscr{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(1)^n)(V) \to (\mathscr{H} \otimes_{\mathcal{O}_V} \mathcal{O}_V(1)^n)(V) \to 0$$

of R-modules is exact.

2. If V is an irreducible affine scheme with the Zariski topology and $\mathcal{O}$ is a constant sheaf, give an entirely *heuristic* argument to show that

$$\check{H}^i(V, \mathcal{O}) = 0$$

if $i > 0$ — only using definition D.3.12 on page 551.

D.4. Serre Duality

D.4.1. Preliminaries. We begin with some definitions

DEFINITION D.4.1. If $(X, \mathcal{O}_X)$ is a ringed space with modules $\mathcal{M}$ and $\mathcal{N}$ define

(1) $\operatorname{Ext}^i_{\mathcal{O}_X}(\mathcal{M}, \mathcal{N})$ to be the right derived functor of $\hom_{\mathcal{O}_X}(\mathcal{M}, \mathcal{N})$ (as defined in B.3.4 on page 504) — i.e., take an injective (or flasque) resolution of $\mathcal{N}$ (see definition D.3.2 on page 546), apply $\hom_{\mathcal{O}_X}(\mathcal{M}, *)$ to it, and take cohomology

(2) $\mathscr{E}xt^i(\mathcal{M}, \mathcal{N})$ to be the right derived functor of $\mathscr{H}om(\mathcal{M}, \mathcal{N})$ (as defined in B.3.4 on page 504). These are sheaves.

As mysterious as the functor $\mathscr{H}om(\mathcal{M}, \mathcal{N})$ seems, it is fairly straightforward over an *affine* scheme:

LEMMA D.4.2. *Let $X = \operatorname{Spec} R$ for a ring R and let M and N be modules over R. There exists a natural homomorphism*

$$\mathcal{A}(\hom_R(M, N)) \to \mathscr{H}om(\mathcal{A}(M), \mathcal{A}(N))$$

If M is finitely presented, this is an isomorphism.

PROOF. There exists a natural homomorphism

$$\hom_R(M, N) \to \mathscr{H}om(\mathcal{A}(M), \mathcal{A}(N))(X)$$

that induces the map in question (see proposition 3.5.4 on page 155). If $M = R^t$ (i.e. is a free module) for a finite t, then $\mathcal{A}(M) = \mathcal{O}_X^t$, $\hom_R(M, N) = N^t$, and

$$\mathscr{H}om(\mathcal{A}(M), N) = N^t$$

so the map is an isomorphism in this case.

If M is finitely presented, there exists a short exact sequence

$$R^n \to R^m \to M \to 0$$

inducing

$$\mathcal{O}_X^n \to \mathcal{O}_X^m \to \mathcal{A}(M) \to 0$$

and we get a commutative diagram

$$
\begin{array}{ccccccc}
0 & \longrightarrow & \mathcal{A}(\hom_R(M, N)) & \longrightarrow & \mathcal{A}(\hom_R(R^m, N)) & \longrightarrow & \mathcal{A}(\hom_R(R^n, N)) \\
& & \downarrow{\scriptstyle i} & & \| & & \| \\
0 & \longrightarrow & \mathscr{H}om(\mathcal{A}(M), \mathcal{A}(N)) & \longrightarrow & \mathscr{H}om(\mathcal{O}_X, \mathcal{A}(N)) & \longrightarrow & \mathscr{H}om(\mathcal{O}_X, \mathcal{A}(N))
\end{array}
$$

that implies that i is an isomorphism. $\square$

PROPOSITION D.4.3. *On $V = R\mathbb{P}^t$*

$$\mathcal{H}om(\mathcal{O}_V(n), \mathcal{O}_V(m)) \cong \mathcal{O}_V(m - n) = \mathcal{O}_V(m) \otimes_{\mathcal{O}_V} \mathcal{O}_V(-n)$$

PROOF. The simplest way to see this is to restrict to the open affines $\mathbb{A}_i^t$. The restrictions are free sheaves of rank 1, so within a $\mathbb{A}_i^t$, we simply have

$$\hom(R, R) = R$$

The only thing that matters at this point are the gluing maps: The gluing map for $\mathcal{O}_V(n)$ is

$$
\begin{array}{ccc}
\mathcal{O}_V(n)(\mathbb{A}_i^t)|\mathbb{A}_i^t \cap \mathbb{A}_j^t & \to & \mathcal{O}_V(n)(\mathbb{A}_j^t)|\mathbb{A}_j^t \cap \mathbb{A}_i^t \\
f & \mapsto & f \cdot (X_i/X_j)^n
\end{array}
$$

and since hom is contravariant, the induced maps are applied in the reverse order (i.e., with i and j interchanged in the upper line)

$$
\begin{array}{ccc}
\mathcal{O}_V(n)(\mathbb{A}_i^t)|\mathbb{A}_j^t \cap \mathbb{A}_i^t & \to & \mathcal{O}_V(n)(\mathbb{A}_j^t)|\mathbb{A}_i^t \cap \mathbb{A}_j^t \\
f & \mapsto & f \cdot (X_i/X_j)^n
\end{array}
$$

or (interchanging i and j everywhere)

$$
\begin{array}{ccc}
\mathcal{O}_V(n)(\mathbb{A}_i^t)|\mathbb{A}_i^t \cap \mathbb{A}_j^t & \to & \mathcal{O}_V(n)(\mathbb{A}_j^t)|\mathbb{A}_j^t \cap \mathbb{A}_i^t \\
f & \mapsto & f \cdot (X_i/X_j)^{-n}
\end{array}
$$

$\square$

PROPOSITION D.4.4. *If $(X, \mathcal{O}_X)$ is a ringed space, and $\mathcal{F}$ and $\mathcal{G}$ are $\mathcal{O}_X$-modules, then for any open set $U \subset X$ there exists an isomorphism*

$$\mathcal{E}xt^i(\mathcal{F}, \mathcal{G})|U \cong \mathcal{E}xt^i(\mathcal{F}|U, \mathcal{G}|U)$$

PROOF. If $\mathcal{I}$ is an injective sheaf on X, we claim that $\mathcal{I}|U$ is an injective object of $(U, \mathcal{O}_U)$. Let $j \colon U \to X$ be the inclusion and let $\mathcal{F} \to \mathcal{G}$ be an inclusion of modules over U and suppose $r \colon \mathcal{F} \to \mathcal{I}|U$ is a map. Then we get an injection $j_*\mathcal{F} \to j_*\mathcal{G}$ over X and a map $r_* \colon j_*\mathcal{F} \to j_*(\mathcal{I}|U) \hookrightarrow \mathcal{I}$. Since $\mathcal{I}$ is injective, the composite extends to a map $j_*\mathcal{G} \to \mathcal{I}$. The claim follows by restricting to U.

If $\mathcal{I}^*$ is an injective resolution of $\mathcal{G}$, then $\mathcal{I}^*|U$ is an injective resolution of $\mathcal{G}|U$, so we get a map

$$\mathcal{H}om(\mathcal{F}, \mathcal{I}^*)|U \cong \mathcal{H}om(\mathcal{F}|U, \mathcal{I}^*|U)$$

which implies the result. $\square$

We also have some results on $\mathrm{Ext}^*_{\mathcal{O}_X}$:

PROPOSITION D.4.5. *If $(X, \mathcal{O}_X)$ is a ringed space, and*

$$0 \to \mathcal{F} \to \mathcal{G} \to \mathcal{H} \to 0$$

is a short exact sequence of coherent sheaves and $\mathcal{L}$ is a coherent sheaf, then we get long exact sequences

$$0 \to \hom(\mathcal{H}, \mathcal{L}) \to \hom(\mathcal{G}, \mathcal{L}) \to \hom(\mathcal{F}, \mathcal{L})$$
$$\to \mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{H}, \mathcal{L}) \to \mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{H}, \mathcal{L}) \to \cdots$$

and

(D.4.1) $0 \to \mathscr{H}om(\mathscr{H},\mathscr{L}) \to \mathscr{H}om(\mathscr{G},\mathscr{L}) \to \mathscr{H}om(\mathscr{F},\mathscr{L})$

$$\to \mathscr{E}xt^1(\mathscr{H},\mathscr{L}) \to \mathscr{E}xt^1(\mathscr{H},\mathscr{L}) \to \cdots$$

PROOF. If $\mathscr{I}^*$ is an injective resolution of $\mathscr{L}$, the functors $\hom(*,\mathscr{I}^*)$ and $\mathscr{H}om(*,\mathscr{I}^*)$ are exact (see exercise 6 on page 531), so we get an exact sequence of chain-complexes and the conclusion follows from proposition D.1.9 on page 522. $\square$

We also get:

PROPOSITION D.4.6. *If $\mathscr{F}$ is a sheaf on a ringed space $(X,\mathscr{O}_X)$ then*

$$\mathscr{H}om(\mathscr{O}_X,\mathscr{F}) = \mathscr{F}$$

If $\mathscr{P}$ is a free sheaf of finite rank on X then

$$\mathscr{E}xt^i(\mathscr{P},\mathscr{F}) = 0$$

for $i > 0$.

PROOF. The first statement follows from the fact that $\hom(R,M) = M$ for a module, M, over a ring, R. Apply this to every $\mathscr{F}(U)$ as a module over $\mathscr{O}_X(U)$ to get the conclusion. Since $\mathscr{H}om(\mathscr{O}_X,*)$ is the identity functor, its derived functors vanish. The final statement follows from the fact that the free sheaf in question is a finite direct sum of copies of $\mathscr{O}_X$. $\square$

PROPOSITION D.4.7. *If $(X,\mathscr{O}_X)$ is a ringed space with coherent sheaves $\mathscr{P}$ and $\mathscr{G}$, and $\mathscr{P}$ is locally free, then*

$$\mathscr{E}xt^i(\mathscr{P},\mathscr{G}) = 0$$

for $i > 0$.

It follows that $\mathscr{H}om(\mathscr{P},*)$ is an exact functor.

PROOF. Proposition D.4.4 on the preceding page implies that

$$\mathscr{E}xt^i(\mathscr{P},\mathscr{G})|U \quad = \quad \mathscr{E}xt^i(\mathscr{P}|U,\mathscr{G}|U)$$

for open sets $U \subset X$. Since $\mathscr{P}$ is locally free, $\mathscr{P}|U$ is *free* for sufficiently small open sets, U — so proposition D.4.6 implies that $\mathscr{E}xt^i(\mathscr{P},\mathscr{G})|U = 0$. If $x \in \mathscr{E}xt^i(\mathscr{P},\mathscr{G})|V$ for some open set V, then there is a covering by open sets U_i such that $x|U_i = 0$ for all i. The separation condition for a sheaf (statement 4 in definition B.1.1 on page 495) implies that $x = 0$. $\square$

Compare the following result to corollary D.1.22 on page 530:

COROLLARY D.4.8. *Let $(X,\mathscr{O}_X)$ is a ringed space with coherent sheaves $\mathscr{F}$ and $\mathscr{G}$,*

(1) If $\mathscr{P}_0$ is locally free and

$$0 \to \mathscr{K}_1 \to \mathscr{P}_0 \to \mathscr{F} \to 0$$

is a short exact sequence, then

(D.4.2) $$\mathscr{E}xt^1(\mathscr{F},\mathscr{G}) = \frac{\mathscr{H}om(\mathscr{K}_1,\mathscr{G})}{\text{im } \mathscr{H}om(\mathscr{P}_0,\mathscr{G})}$$

and $\mathscr{E}xt^i(\mathscr{F},\mathscr{G}) = \mathscr{E}xt^{i-1}(\mathscr{K}_1,\mathscr{G})$ for all $i > 1$.

(2) *If*

$$0 \to \mathscr{K}_n \to \mathscr{P}_{n-1} \to \cdots \to \mathscr{P}_0 \to \mathscr{F} \to 0$$

where the $\mathscr{P}_i$ are locally free, then

(D.4.3)
$$\mathscr{E}xt^n(\mathscr{F},\mathscr{G}) = \frac{\mathscr{H}om(\mathscr{K}_n,\mathscr{G})}{\operatorname{im} \mathscr{H}om(\mathscr{P}_{n-1},\mathscr{G})}$$

PROOF. The first statement follows immediately from the exact sequence, D.4.1 on the previous page and proposition D.4.7 on the preceding page.

The second statement follows from induction on short exact sequences

$$0 \to \mathscr{K}_i \to \mathscr{P}_{i-1} \to \mathscr{K}_{i-1} \to 0$$

□

This has an interesting consequence:

THEOREM D.4.9. *Let $(X, \mathcal{O}_X)$ be a ringed space with coherent sheaves $\mathscr{F}$ and $\mathscr{G}$ and a resolution*

$$\cdots \to \mathscr{P}_1 \to \mathscr{P}_0 \to \mathscr{F} \to 0$$

by locally free sheaves $\mathscr{P}_i$. Then

$$\mathscr{E}xt^i(\mathscr{F},\mathscr{G}) = H^i(\mathscr{H}om(\mathscr{P}_*,\mathscr{G}))$$

for all $i \geq 0$.

PROOF. We use the notation of corollary D.4.8 on the previous page. In the short exact sequences there we have an exact sequence

$$\cdots \to \mathscr{P}_{i+1} \to \mathscr{P}_i \to \mathscr{K}_i \to 0$$

inducing (because of the left-exactness of $\mathscr{H}om(*,\mathscr{G})$)

$$0 \to \mathscr{H}om(\mathscr{K}_i,\mathscr{G}) \to \mathscr{H}om(\mathscr{P}_i,\mathscr{G}) \to \mathscr{H}om(\mathscr{P}_{i+1},\mathscr{G})$$

which implies that

$$\mathscr{H}om(\mathscr{K}_i,\mathscr{G}) = \ker \mathscr{H}om(\mathscr{P}_i,\mathscr{G}) \to \mathscr{H}om(\mathscr{P}_{i+1},\mathscr{G})$$

so the conclusion follows from equations D.4.2 on the preceding page and D.4.3, which imply that Let
(D.4.4)
$$\mathscr{E}xt^n(\mathscr{F},\mathscr{G}) = \frac{\ker \mathscr{H}om(\mathscr{P}_i,\mathscr{G}) \to \mathscr{H}om(\mathscr{P}_{i+1},\mathscr{G})}{\operatorname{im} \mathscr{H}om(\mathscr{P}_{n-1},\mathscr{G})} = H^i(\mathscr{H}om(\mathscr{P}_*,\mathscr{G}))$$

□

It follows that the $\mathscr{E}xt^i(*,*)$ have a relatively simple interpretation on an affine variety:

COROLLARY D.4.10. *If $V = \operatorname{Spec} R$, where R is a noetherian ring and M and N are finitely generated, then*

$$\mathcal{A}(\operatorname{Ext}^i_{\mathcal{O}_V}(M,N)) = \mathscr{E}xt^i(\mathcal{A}(M),\mathcal{A}(N))$$

PROOF. Find a resolution

$$\cdots \to F_1 \to F_0 \to M \to 0$$

where the F_i are *free* modules — so that $\mathcal{A}(F_i)$ will also be free. Since R is noetherian, we may assume the F_i are all finitely generated. Theorem D.4.9 on the facing page implies that

$$\begin{aligned} \mathcal{E}xt^i(\mathcal{A}(M), \mathcal{A}(N)) &= H^i(\mathcal{H}om(\mathcal{A}(F_*), \mathcal{A}(N))) \\ &= H^i(\mathcal{A}(\hom_R(F_*, N)) \end{aligned}$$

where the last equality is due to lemma D.4.2 on page 561. The fact that the $\mathcal{A}(*)$-functor is exact (see proposition 3.5.3 on page 155) implies that

$$H^i(\mathcal{A}(\hom_R(F_*, N)) = \mathcal{A}(H^i(\hom_R(F_*, N)))$$

and exercise 9 on page 531 proves the result. $\qquad\square$

We can characterize the somewhat mysterious $\mathcal{E}xt^i$-functor in terms of its behavior on stalks:

PROPOSITION D.4.11. *If V is a noetherian scheme, $\mathcal{F}$ is a coherent sheaf on V, $\mathcal{G}$ is any $\mathcal{O}_V$-module and $p \in V$ is a point, then*

$$\mathcal{E}xt^i(\mathcal{F}, \mathcal{G})_p = \mathrm{Ext}^i_{\mathcal{O}_{V,p}}(\mathcal{F}_p, \mathcal{G}_p)$$

for any $i \geq 0$.

PROOF. Since the statement is local, we may assume that V is affine. In this case, $\mathcal{F}$ has a resolution by locally free sheaves

$$\cdots \to \mathcal{L}_1 \to \mathcal{L}_0 \to \mathcal{F} \to 0$$

that can be used to compute $\mathcal{E}xt^i(\mathcal{F}, \mathcal{G})$, by theorem D.4.9 on the facing page. Taking stalks gives a free resolution

$$\cdots \to \mathcal{L}_{1,x} \to \mathcal{L}_{0,x} \to \mathcal{F}_x \to 0$$

over $\mathcal{O}_{V,x}$. Since

$$\mathcal{H}om(\mathcal{L}_i, \mathcal{G})_x = \hom_{\mathcal{O}_{V,x}}(\mathcal{L}_{i,x}, \mathcal{G}_x)$$

the conclusion follows from the exactness of the stalk-functors (see exercise 7 on page 502). $\qquad\square$

EXERCISES.

1. Show that, if $\mathcal{F}$ is a sheaf on a ringed space $(X, \mathcal{O}_X)$ that

$$\hom(\mathcal{O}_X, \mathcal{F}) = \mathcal{F}(X)$$

2. Show that, on a ringed space $(X, \mathcal{O}_X)$

$$\mathrm{Ext}^i_{\mathcal{O}_X}(\mathcal{O}_X, \mathcal{G}) = H^i(X, \mathcal{G})$$

for all i.

3. If $(X, \mathcal{O}_X)$ is a ringed space and $\mathscr{F}$ is a locally free sheaf (see definition C.2.3 on page 515) of finite rank, define

(D.4.5) $$\mathscr{F}^\vee = \mathscr{H}\!om(\mathscr{F}, \mathcal{O}_X)$$

and show that $\mathscr{F}^{\vee\vee} = \mathscr{F}$.

Show that, for any $\mathcal{O}_X$-module

$$\mathscr{H}\!om(\mathscr{F}, \mathscr{G}) \cong \mathscr{F}^\vee \otimes_{\mathcal{O}_X} \mathscr{G}$$

4. If $(X, \mathcal{O}_X)$ is a ringed space and $\mathscr{A}, \mathscr{B}$, and $\mathscr{C}$ are modules over $\mathcal{O}_X$ show that there are natural isomorphisms

$$\hom(\mathscr{A}, \hom(\mathscr{B}, \mathscr{C})) \to \hom(\mathscr{A} \otimes_{\mathcal{O}_X} \mathscr{B}, \mathscr{C})$$

and

$$\mathscr{H}\!om(\mathscr{A}, \mathscr{H}\!om(\mathscr{B}, \mathscr{C})) \to \mathscr{H}\!om(\mathscr{A} \otimes_{\mathcal{O}_X} \mathscr{B}, \mathscr{C})$$

5. If $\mathscr{F}$ is locally free show that

$$\mathscr{H}\!om(\mathscr{U} \otimes_{\mathcal{O}_X} \mathscr{F}^\vee, \mathscr{W}) = \mathscr{H}\!om(\mathscr{U}, \mathscr{F} \otimes_{\mathcal{O}_X} \mathscr{W})$$

for any coherent sheaves $\mathscr{U}$ and $\mathscr{W}$.

6. If V is a projective scheme over a noetherian ring and

$$\mathscr{F}_1 \to \mathscr{F}_2 \to \cdots \to \mathscr{F}_n$$

is an exact sequence of coherent sheaves, show that there is an integer N such that

$$\left(\mathscr{F}_1 \otimes_{\mathcal{O}_V} \mathcal{O}_V(t)\right)(V) \to \left(\mathscr{F}_2 \otimes_{\mathcal{O}_V} \mathcal{O}_V(t)\right)(V) \to \cdots \to \left(\mathscr{F}_n \otimes_{\mathcal{O}_V} \mathcal{O}_V(t)\right)(V)$$

is exact for all $t \geq N$.

D.4.2. Serre duality. Serre Duality (essentially corollary D.4.20 on page 573) first appeared in Serre's paper [**149**]. Serre later proved an analytic version in [**151**] — using analysis of several complex variable. Grothendieck vastly generalized Serre's original result in [**62**] (to theorem D.4.19 on page 571) and Hartshorne clarified and simplified it in [**69**]. Our treatment will follow the latter two references and [**70**].

We begin with a limited form of the result for projective spaces:

THEOREM D.4.12 (Serre Duality for projective space). *Let* $V = k\mathbb{P}^n$ *over a field k and let* ω_V *be the canonical class (see definition 5.9.36 on page 300) computed in example 5.9.39 on page 301 as* $\mathcal{O}_{k\mathbb{P}^n}(-n-1)$. *Then*

(1) $H^n(V, \omega_V) \cong k$. *Fix an isomorphism.*
(2) For any coherent sheaf $\mathscr{F}$ on V, the natural pairing

$$\hom(\mathscr{F}, \omega_V) \times H^n(V, \mathscr{F}) \to H^n(V, \omega_V) \cong k$$

is a perfect pairing of finite dimensional vector spaces.
(3) for every $i \geq 0$ there is a natural functorial isomorphism

$$\mathrm{Ext}^i_{\mathcal{O}_V}(\mathscr{F}, \omega_V) \xrightarrow{\cong} H^{n-i}(V, \mathscr{F})^*$$

PROOF. The first statement follows from statement 3 in theorem D.3.21 on page 556.

Any homomorphism $f: \mathscr{F} \to \omega_V$ induces a natural homomorphism of cohomology groups

$$H^i(V, \mathscr{F}) \to H^i(V, \omega_V)$$

and this *defines* the map in statement 2. We will prove that it is a perfect pairing (i.e., a map that implies an isomorphism of vector spaces $\hom(\mathscr{F}, \omega_V) \cong H^n(V, \mathscr{F})^*$).

If $\mathscr{F} = \mathcal{O}_V(d)$, proposition D.4.3 on page 562 implies that

$$\mathcal{H}om(\mathscr{F}, \omega_V) = \mathcal{O}_V(-n - d - 1) = \omega_V \otimes_{\mathcal{O}_V} \mathscr{F}$$

so corollary D.3.16 on page 553 implies that

$$\hom(\mathscr{F}, \omega_V) = \mathcal{H}om(\mathscr{F}, \omega_V)(V) = H^0(V, \omega_V \otimes \mathcal{O}_V(-d))$$

So, in this case, statement 2 follows from statement 4 of theorem D.3.21 on page 556.

To prove it for an arbitrary coherent sheaf, note that we have an exact sequence

$$\mathscr{R}_1 \to \mathscr{R}_0 \to \mathscr{F} \to 0$$

where $\mathscr{R}_i$ are direct sums of sheaves of the form $\mathcal{O}_V(q_i)$ (see lemma 5.3.10 on page 238) — which are locally free by construction. chasing the commutative diagram (with exact rows)

$$
\begin{array}{ccccccc}
0 & \longrightarrow & \hom(\mathscr{F}, \omega_V) & \longrightarrow & \hom(\mathscr{R}_0, \omega_V) & \longrightarrow & \hom(\mathscr{R}_1, \omega_V) \\
 & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \cong} & & \downarrow{\scriptstyle \cong} \\
0 & \longrightarrow & H^n(V, \mathscr{F})^* & \longrightarrow & H^n(V, \mathscr{R}_0)^* & \longrightarrow & H^n(V, \mathscr{R}_1)^*
\end{array}
$$

implies that f is an isomorphism.

To prove the second statement, note that $\mathrm{Ext}^i_{\mathcal{O}_V}(\mathscr{F}, \omega_V)$ and $H^{n-i}(V, \mathscr{F})^*$ are both contravariant δ-functors of $\mathscr{F}$ (see definition D.1.23 on page 532). Since every coherent sheaf is a quotient of a direct sum of $\mathcal{O}_V(-d)$'s for arbitrarily large d (see lemma 5.3.10 on page 238), we get an epimorphism

$$\bigoplus \mathcal{O}_V(-d) \to \mathscr{F} \to 0$$

and the induced maps

$$
\begin{array}{rcl}
\mathrm{Ext}^i_{\mathcal{O}_V}(\mathscr{F}, \omega_V) & \to & \mathrm{Ext}^i_{\mathcal{O}_V}(\bigoplus \mathcal{O}_V(-d), \omega_V) \\
H^{n-i}(V, \mathscr{F})^* & \to & H^{n-i}(V, \bigoplus \mathcal{O}_V(-d))^*
\end{array}
$$

vanish for all $i > 0$. Proposition D.1.26 on page 533 implies that $\mathrm{Ext}^i_{\mathcal{O}_V}(\mathscr{F}, \omega_V)$ and $H^{n-i}(V, \mathscr{F})^*$ are universal δ-functors (see definition D.1.24 on page 532). Since they agree for $i = 0$, they must be naturally isomorphic. □

To extend this to projective varieties, we need the concept:

DEFINITION D.4.13. If X is a projective scheme of dimension n over a field k, a *dualizing sheaf* for X is a coherent sheaf ω_X° together with a trace morphism

$$t: H^n(X, \omega_X^\circ) \to k$$

such that, for all coherent sheaves $\mathscr{F}$ on X the natural pairing

$$\hom(\mathscr{F}, \omega_X^\circ) \times H^n(X, \mathscr{F}) \to H^n(V, \omega_X^\circ) \xrightarrow{t} k$$

gives an isomorphism

$$\hom(\mathscr{F}, \omega_X^\circ) \to H^n(X, \mathscr{F})^*$$

REMARK. Note that a dualizing sheaf is one that makes statement duality2 of theorem D.4.12 on page 566 work. To prove the general case, we will have to find a dualizing sheaf *and* a way to make the induction in the last statement work.

Now we will show that such sheaves exist for projective varieties. We need some preliminary results first.

LEMMA D.4.14. *If $(X, \mathcal{O}_X)$ is a ringed space, $\mathscr{P}$ is a locally-free sheaf of finite rank, $\mathscr{I}$ is an injective sheaf, then $\mathscr{P} \otimes_{\mathcal{O}_X} \mathscr{I}$ is also injective.*

PROOF. Exercise 5 on page 566 implies that

$$\mathscr{H}om(* \otimes_{\mathcal{O}_X} \mathscr{P}^\vee, \mathscr{I}) = \mathscr{H}om(*, \mathscr{P} \otimes_{\mathcal{O}_X} \mathscr{I})$$

so the conclusion follows from the fact that $\otimes \mathscr{P}^\vee$ is an exact functor (see proposition D.4.6 on page 563) and $\mathscr{H}om(*, \mathscr{I})$ is exact. $\square$

COROLLARY D.4.15. *If $(X, \mathcal{O}_X)$ is a ringed space, $\mathscr{P}$ is a locally-free sheaf of finite rank, $\mathscr{I}$ is an injective sheaf, and $\mathscr{R}$ and $\mathscr{S}$ are modules over $\mathcal{O}_X$ then*

$$\mathrm{Ext}^i_{\mathcal{O}_X}(\mathscr{R} \otimes_{\mathcal{O}_X} \mathscr{P}, \mathscr{S}) = \mathrm{Ext}^i_{\mathcal{O}_X}(\mathscr{R}, \mathscr{S} \otimes_{\mathcal{O}_X} \mathscr{P}^\vee)$$

where $\mathscr{P}^\vee = \mathscr{H}om(\mathscr{P}, \mathcal{O}_X)$ and

$$\mathscr{E}xt^i(\mathscr{R} \otimes_{\mathcal{O}_X} \mathscr{P}, \mathscr{S}) = \mathscr{E}xt^i(\mathscr{R}, \mathscr{S} \otimes_{\mathcal{O}_X} \mathscr{P}^\vee) = \mathscr{E}xt^i(\mathscr{R}, \mathscr{S}) \otimes_{\mathcal{O}_X} \mathscr{P}^\vee$$

PROOF. Exercise 5 on page 566 implies the result for $i = 0$. Let $\mathscr{I}$ be an injective module over $\mathcal{O}_V$ such that there is a monomorphism

$$\mathscr{S} \to \mathscr{I}$$

(see D.3.1 on page 545). The functors $\mathrm{Ext}^i_{\mathcal{O}_X}(\mathscr{R} \otimes_{\mathcal{O}_X} \mathscr{P}, *)$ and $\mathrm{Ext}^i_{\mathcal{O}_X}(\mathscr{R}, * \otimes_{\mathcal{O}_X} \mathscr{P}^\vee)$ are covariant δ-functors that agree for $i = 0$. Since the maps induced in them by $\mathscr{S} \to \mathscr{I}$ both vanish (by lemma D.4.14), they are both universal by proposition D.1.25 on page 533, hence isomorphic.

The final statement follows from the fact that $* \otimes_{\mathcal{O}_X} \mathscr{P}^\vee$ is exact. $\square$

Now we discuss relations between $\mathscr{E}xt^i(*, *)$ and $\mathrm{Ext}^i_{\mathcal{O}_X}(*, *)$. If $(X, \mathcal{O}_X)$ is any ringed space, $i = 0$, and $\mathscr{F}$ and $\mathscr{G}$ are modules over $\mathcal{O}_X$, we know that

$$\hom(\mathscr{F}, \mathscr{G}) = \mathscr{H}om(\mathscr{F}, \mathscr{G})(X)$$

If $X = \mathrm{Spec}\, R$ is a noetherian affine scheme, corollary D.4.10 on page 564 shows that

$$\mathrm{Ext}^i_{\mathcal{O}_X}(\mathscr{F}, \mathscr{G}) = \mathscr{E}xt^i(\mathscr{F}, \mathscr{G})(X)$$

for all i. On a projective scheme, we must use Serre-twists to get similar results:

PROPOSITION D.4.16. *If V is a projective scheme over a noetherian ring, R, $\mathcal{O}_V(1)$ is a very ample invertible sheaf, $\mathcal{F}$ and $\mathcal{G}$ are coherent sheaves over V, then there exists an integer $N > 0$ depending on $\mathcal{F}, \mathcal{G}$ and i such that for every $n \geq N$*

$$(\text{D.4.6}) \qquad \operatorname{Ext}^i_{\mathcal{O}_V}(\mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(1)^n) \cong \mathscr{E}xt^i(\mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n))(V)$$

for all $i \geq 0$, where $\mathcal{O}_V(n) = \mathcal{O}_V(1)^n$.

REMARK. The sheaf $\mathcal{O}_V(n) = \mathcal{O}_V(1)^n$ is a kind of Serre twist on V.

PROOF. We know that this is true for $i = 0$ (without any twisting). If $\mathcal{F} = \mathcal{O}_V$, then $\operatorname{Ext}^i_{\mathcal{F}}(\mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) = H^i(V, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n))$ (exercise 2) and theorem D.3.22 on page 558 implies that this *vanishes* for sufficiently large n. On the other hand, $\mathscr{E}xt^i(\mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n))$ vanishes for all n and $i > 0$, so equation D.4.6 is true.

If $\mathcal{F}$ is locally free, then we apply corollary D.4.15 on the facing page with $\mathcal{R} = \mathcal{O}_V$ to get

$$\begin{aligned}
\operatorname{Ext}^i_{\mathcal{O}_V}(\mathcal{O}_V \otimes_{\mathcal{O}_V} \mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) &= \operatorname{Ext}^i_{\mathcal{O}_V}(\mathcal{O}_V, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n) \otimes_{\mathcal{O}_X} \mathcal{F}^*) \\
&= H^i(V, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n) \otimes_{\mathcal{O}_X} \mathcal{F}^*)
\end{aligned}$$

which vanishes for n sufficiently large, and

$$\begin{aligned}
\mathscr{E}xt^i(\mathcal{O}_V \otimes_{\mathcal{O}_V} \mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) &= \mathscr{E}xt^i(\mathcal{O}_V, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n) \otimes_{\mathcal{O}_X} \mathcal{F}^*) \\
&= 0
\end{aligned}$$

so, again, equation D.4.6 holds (in that both sides are 0).

In the general case, let $\mathcal{E}$ be a locally free sheaf that fits into an exact sequence

$$0 \to \mathcal{K} \to \mathcal{E} \to \mathcal{F} \to 0$$

— this exists by lemma 5.3.10 on page 238, and $\mathcal{K}$ is also coherent. We develop the long exact sequences:

$$\begin{aligned}
0 \to \hom(\mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) &\to \hom(\mathcal{E}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) \\
&\to \hom(\mathcal{K}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) \to \operatorname{Ext}^1_{\mathcal{O}_V}(\mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) \to 0
\end{aligned}$$

— where $\operatorname{Ext}^1_{\mathcal{O}_V}(\mathcal{E}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) = 0$ by the argument given above, and we get

$$\operatorname{Ext}^i_{\mathcal{O}_V}(\mathcal{K}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n)) = \operatorname{Ext}^{i+1}_{\mathcal{O}_V}(\mathcal{F}, \mathcal{G} \otimes_{\mathcal{O}_V} \mathcal{O}_V(n))$$

We also get a similar long exact sequence for $\mathscr{H}om$ and $\mathscr{E}xt^i$. After taking tensor products with $\mathcal{O}_V(d)$ for some d, there evaluations on V become exact, by exercise 6 on page 566. Comparing the two long exact sequences demonstrates the conclusion. □

PROPOSITION D.4.17. *If $j\colon V \hookrightarrow k\mathbb{P}^n$ is a closed variety of codimension r with a very ample sheaf $\mathcal{O}_V(1)$, then*

$$\mathscr{E}xt^i(\mathcal{O}_V, \omega_{k\mathbb{P}^n}) = 0$$

(evaluated on $k\mathbb{P}^n$) for all $i < r$.

REMARK. Note that $\mathcal{O}_V$ is naturally a module over $\mathcal{O}_{k\mathbb{P}^n}$ via the surjective morphism
$$j^*\colon \mathcal{O}_{k\mathbb{P}^n} \to \mathcal{O}_V$$
and the same is true of any module over $\mathcal{O}_V$. As $\mathcal{O}_{k\mathbb{P}^n}$-modules, $j_*\mathscr{F} = \mathscr{F}$, for any $\mathcal{O}_V$-module, $\mathscr{F}$[3].

PROOF. We claim that $\mathscr{E}^i = \mathscr{E}xt^i(\mathcal{O}_V, \omega_{k\mathbb{P}^n}) = \mathscr{E}xt^i(\mathcal{O}_V, \omega_{k\mathbb{P}^n})$ is a coherent sheaf on $k\mathbb{P}^n$ for all $i \geq 0$. This follows by restricting to open affines and applying lemma D.4.2 on page 561 and corollary D.4.10 on page 564. To show that it vanishes, note that its tensor product with $\mathcal{O}_{k\mathbb{P}^n}(d)$ is generated by global sections (see definition 5.3.6 on page 237 and corollary 5.3.9 on page 237).

It suffices, then, to show that $\left(\mathscr{E}^i \otimes_{\mathcal{O}_{k\mathbb{P}^n}} \mathcal{O}_{k\mathbb{P}^n}(d)\right)(k\mathbb{P}^n) = 0$ for d sufficiently large. Corollary D.4.15 on page 568 implies that
$$\mathscr{E}xt^i(\mathcal{O}_V, \omega_{k\mathbb{P}^n}) \otimes_{\mathcal{O}_{k\mathbb{P}^n}} \mathcal{O}_{k\mathbb{P}^n}(d) \quad = \quad \mathscr{E}xt^i(\mathcal{O}_V, \omega_{k\mathbb{P}^n} \otimes_{\mathcal{O}_X} \mathcal{O}_{k\mathbb{P}^n}(d))$$
and proposition D.4.16 on the previous page implies that
$$\left(\mathscr{E}xt^i(\mathcal{O}_V, \omega_{k\mathbb{P}^n} \otimes_{\mathcal{O}_X} \mathcal{O}_{k\mathbb{P}^n}(d))\right)(k\mathbb{P}^n) \quad = \quad \mathrm{Ext}^i_{\mathscr{R}}(\mathcal{O}_V, \omega_{k\mathbb{P}^n} \otimes_{\mathcal{O}_{k\mathbb{P}^n}} \mathcal{O}_{k\mathbb{P}^n}(d))$$
$$= \quad \mathrm{Ext}^i_{\mathscr{R}}(\mathcal{O}_V \otimes_{\mathcal{O}_{k\mathbb{P}^n}} \mathcal{O}_{k\mathbb{P}^n}(-d), \omega_{k\mathbb{P}^n})$$
where $\mathscr{R} = \mathcal{O}_{k\mathbb{P}^n}$. The projective case of Serre Duality (D.4.12 on page 566) implies that
$$\mathrm{Ext}^i_{\mathscr{R}}(\mathcal{O}_V \otimes_{\mathcal{O}_X} \mathcal{O}_{k\mathbb{P}^n}(-d), \omega_{k\mathbb{P}^n}) \quad = \quad H^{n-i}(k\mathbb{P}^n, \mathcal{O}_V \otimes_{\mathcal{O}_{k\mathbb{P}^n}} \mathcal{O}_{k\mathbb{P}^n}(-d))^*$$
$$= \quad H^{n-i}(V, \mathcal{O}_V(-d))^*$$
which must vanish for $n - i > n - r$ or $i < r$. (see lemma D.3.18 on page 554). $\square$

The sheaf $\mathscr{E}xt^r(\mathcal{O}_V, \omega_{k\mathbb{P}^n})$ will turn out to be our dualizing sheaf:

LEMMA D.4.18. *If $j\colon V \hookrightarrow k\mathbb{P}^n$ is a closed variety of codimension r with a very ample sheaf $\mathcal{O}_V(1)$ and $\omega_V^\circ = \mathscr{E}xt^r(\mathcal{O}_V, \omega_{k\mathbb{P}^n})$, then for any $\mathcal{O}_V$-module $\mathscr{F}$ there exists a functorial isomorphism*
$$\hom_{\mathcal{O}_V}(\mathscr{F}, \omega_V^\circ) \cong \mathrm{Ext}^r_{\mathscr{R}}(\mathscr{F}, \omega_{k\mathbb{P}^n})$$
where $\mathscr{R} = \mathcal{O}_{k\mathbb{P}^n}$.

It follows that $\omega_V^\circ = \mathscr{E}xt^r(\mathcal{O}_V, \omega_{k\mathbb{P}^n})$ is a dualizing sheaf for V, where the $\mathscr{E}xt^i$ is computed over $\mathcal{O}_{k\mathbb{P}^n}$.

REMARK. Note that $\omega_V^\circ = \mathscr{E}xt^r(\mathcal{O}_V, \omega_{k\mathbb{P}^n})$ has a natural $\mathcal{O}_V$-module structure due to the presence of $\mathcal{O}_V$ in the left side of the $\mathscr{E}xt^i(*, *)$.

PROOF. If $0 \to \omega_{k\mathbb{P}^n} \to \mathscr{A}^0 \to \cdots$ is an injective resolution of $\omega_{k\mathbb{P}^n}$, then
$$\mathrm{Ext}^i_{\mathscr{R}}(\mathscr{F}, \omega_{k\mathbb{P}^n}) = H^i(\hom_{k\mathbb{P}^n}(\mathscr{F}, \mathscr{A}^*))$$
where $\mathscr{R} = \mathcal{O}_{k\mathbb{P}^n}$. Since $\mathscr{F}$ is an $\mathcal{O}_V$-module, any $\mathcal{O}_{k\mathbb{P}^n}$-homomorphism $\mathscr{F} \to \mathscr{A}^i$ factors through $\mathscr{B}^i = \mathscr{H}om_{k\mathbb{P}^n}(\mathcal{O}_V, \mathscr{A}^i)$, so
$$\mathrm{Ext}^i_{\mathscr{R}}(\mathscr{F}, \omega_{k\mathbb{P}^n}) = H^i(\hom_V(\mathscr{F}, \mathscr{B}^*))$$

---

[3]One important characteristic of sheaves that are *modules* over $\mathcal{O}_V$ is that push-forwards are unnecessary.

We claim that each $\mathscr{B}^i$ is an injective $\mathcal{O}_V$-module. This follows from the fact that

$$\hom_V(\mathscr{F}, \mathscr{B}^i) = \hom_{k\mathbb{P}^n}(\mathscr{F}, \mathscr{A}^i)$$

so that $\hom_V(*, \mathscr{B}^i)$ is an exact functor.

Now, let us examine the cochain complex $\{\mathscr{B}^*\}$ of $\mathcal{O}_V$-modules. We have

$$
\begin{aligned}
H^i(\mathscr{B}^*) &= H^i(\mathscr{H}om_{k\mathbb{P}^n}(\mathcal{O}_V, \mathscr{A}^*) \\
&= \mathscr{E}xt^i(\mathcal{O}_V, \omega_{k\mathbb{P}^n}) \\
&= 0 \text{ if } i < r
\end{aligned}
$$

The bottom equality follows from proposition D.4.17 on page 569. It follows (since the $\mathscr{B}^j$ are *injective* objects and, consequently, are *direct summands* of any object that contains them) that

$$\mathscr{B}^* = \mathscr{B}_1^* \oplus \mathscr{B}_2^*$$

where $\mathscr{B}_1^*$ runs from degree 0 to r and is acyclic in *all* degrees, and $\mathscr{B}_2^*$ *begins* in degree r. It follows that $\omega_V^\circ = \ker \mathscr{B}_2^r \to \mathscr{B}_2^{r+1}$ and that

$$
\begin{aligned}
H^r(\hom_V(\mathscr{F}, \mathscr{B}_2^*) &= \operatorname{Ext}_{\mathscr{R}}^r(\mathscr{F}, \omega_{k\mathbb{P}^n}) \\
&= \hom(\mathscr{F}, \omega_V^\circ)
\end{aligned}
$$

which proves the first statement. The final statement follows from the projective case of Serre Duality (D.4.12 on page 566) which implies that

$$
\begin{aligned}
\operatorname{Ext}_{\mathcal{O}_V}^r(\mathscr{F}, \omega_{k\mathbb{P}^n}) &\cong H^{n-r}(k\mathbb{P}^n, \mathscr{F})^* \\
&= H^d(V, \mathscr{F})^*
\end{aligned}
$$

where $d = n - r = \dim V$. We get a functorial isomorphism

$$\hom(\mathscr{F}, \omega_V^\circ) \cong H^d(V, \mathscr{F})^*$$

If we set $\mathscr{F} = \omega_V^\circ$, then we have

$$\hom(\omega_V^\circ, \omega_V^\circ) \cong H^d(V, \omega_V^\circ)^*$$

and the image of $1: \omega_V^\circ \to \omega_V^\circ \in \hom(\omega_V^\circ, \omega_V^\circ)$ is a trace map

$$t: H^d(V, \omega_V^\circ) \to k$$

$\square$

Now we are in a position to prove the general duality theorem:

THEOREM D.4.19 (Serre Duality). *Let V be an n-dimensional projective variety with dualizing sheaf ω_V° and very ample sheaf $\mathcal{O}_V(1)$. Then, for all coherent sheaves $\mathscr{F}$ on V there are natural functorial homomorphisms*

$$\theta^i: \operatorname{Ext}_{\mathcal{O}_V}^i(\mathscr{F}, \omega_V^\circ) \to H^{n-i}(V, \mathscr{F})^*$$

where θ^0 is the isomorphism induced by the trace map and $W^* = \hom_k(W, k)$ — the vector-space dual. If V is smooth and all of its components have dimension n, these maps are isomorphisms for all $i \geq 0$.

PROOF. Let $\mathcal{O}_V(-t) = \mathcal{O}_V(1)^{-t}$ — the pullback of $\mathcal{O}_{k\mathbb{P}^N}(-t)$ under some embedding

$$j\colon V \hookrightarrow k\mathbb{P}^N$$

Lemma 5.3.10 on page 238 implies that there is a surjective map

$$\bigoplus_{i=1}^{m} \mathcal{O}_{k\mathbb{P}^N}(-t) \to \mathscr{F}$$

which induces a surjection

(D.4.7)
$$\mathscr{E} = \bigoplus_{i=1}^{m} \mathcal{O}_V(-t) \to \mathscr{F}$$

Then corollary D.4.15 on page 568 implies that

$$
\begin{aligned}
\operatorname{Ext}^i_{\mathcal{O}_V}(\mathscr{E}, \omega_V^\circ) &= \operatorname{Ext}^i_{\mathcal{O}_V}(\mathcal{O}_V \otimes_{\mathcal{O}_V} \mathscr{E}, \omega_V^\circ) \\
&= \operatorname{Ext}^i_{\mathcal{O}_V}(\mathcal{O}_V, \mathscr{E}^* \otimes_{\mathcal{O}_V} \omega_V^\circ) \\
&= \bigoplus_{i=1}^{m} H^i(V, \omega_V^\circ \otimes_{\mathcal{O}_V} \mathcal{O}_V(t))
\end{aligned}
$$

and theorem D.3.22 on page 558 implies that this vanishes for $i > 0$ and t sufficiently large. It follows that $\operatorname{Ext}^i_{\mathcal{O}_V}(*, \omega_V^\circ)$ is a contravariant delta-functor so that the map θ^0 induces all of the θ^i (see proposition D.1.26 on page 533).

To complete the proof, get a surjective map as in equation D.4.7, embed V in $k\mathbb{P}^N$, and consider a point $x \in V$. Then $\mathscr{E}$ is locally free over $\mathcal{O}_V$ so $\mathscr{E}_x$ has depth n over $\mathcal{O}_{V,x}$ and over $\mathcal{O}_{k\mathbb{P}^N,x}$. Since $\mathcal{O}_{k\mathbb{P}^N,x}$ is also a local Cohen-Macaulay ring, the Auslander-Buchsbaum formula D.2.24 on page 545) implies that

$$\operatorname{proj-dim} \mathscr{E}_x = N - n$$

where this is calculated *over* $\mathcal{O}_{k\mathbb{P}^N,x}$. It follows (from proposition D.4.11 on page 565) that

$$\mathscr{E}xt^i(\mathscr{E}, *) = 0$$

for $i > N - n$. For t sufficiently large, proposition D.4.16 on page 569 implies that

$$\mathscr{E}xt^{N-i}(\mathscr{E}, \omega_{k\mathbb{P}^N} \otimes_{\mathcal{O}_{k\mathbb{P}^n}} \mathcal{O}_{k\mathbb{P}^n}(t))(k\mathbb{P}^n) =$$
$$\operatorname{Ext}^{N-i}_{\mathcal{O}_{k\mathbb{P}^n}}(\mathscr{E}, \omega_{k\mathbb{P}^N} \otimes_{\mathcal{O}_{k\mathbb{P}^n}} \mathcal{O}_{k\mathbb{P}^n}(t)) = 0$$

for $i > n$. The Projective case of Serre Duality (D.4.12 on page 566) implies that

$$H^i(k\mathbb{P}^N, \mathscr{E} \otimes_{\mathcal{O}_{k\mathbb{P}^N}} \mathcal{O}_{k\mathbb{P}^N}(-d))$$

is dual to this, so

$$H^i(k\mathbb{P}^N, \mathscr{E} \otimes_{\mathcal{O}_{k\mathbb{P}^N}} \mathcal{O}_{k\mathbb{P}^N}(-d)) = H^i(V, \mathscr{E} \otimes_{\mathcal{O}_V} \mathcal{O}_V(-d)) = 0$$

for $i < n$ (and sufficiently large d). It follows that, replacing $\mathscr{E}$ by $\mathscr{E}' = \mathscr{E} \otimes_{\mathcal{O}_V} \mathcal{O}_V(-d)$ if necessary, we have $\mathscr{F}$ is the surjective image of the locally free $\mathscr{E}'$ with $H^i(V, \mathscr{E}') = 0$ for $i < n$. Proposition D.1.26 on page 533 implies that

$$H^i(V, *)^*$$

is a universal contravariant δ-functor like $\mathrm{Ext}^i_{\mathcal{O}_V}(*, \omega^\circ_V)$ and the natural isomorphism

$$\theta^i \colon \mathrm{Ext}^i_{\mathcal{O}_V}(\mathscr{F}, \omega^\circ_V) \to H^{n-i}(V, \mathscr{F})^*$$

for $i = 0$ induces natural isomorphisms for higher values of i. □

If the sheaf $\mathscr{F}$ is *locally-free,* we can say more:

COROLLARY D.4.20. *If V is a projective variety over k all of whose components have dimension n and $\mathscr{F}$ is a locally free sheaf of finite rank, then*

$$H^i(V, \mathscr{F}) \cong H^{n-i}(V, \mathscr{F}^\vee \otimes_{\mathcal{O}_V} \omega^\circ_V)^*$$

where $\mathscr{F}^\vee = \mathscr{H}om(\mathscr{F}, \mathcal{O}_V)$ and W^* is the dual of W as a vector-space over k.

REMARK. This is often what people mean when they talk about the Serre Duality theorem — since it is a variation of Serre's *original* version in [**149**].

Since locally free sheaves correspond to vector-bundles, this result gives useful information on the behavior of vector-bundles over a variety.

PROOF. Theorem D.4.19 on page 571 implies that

$$\mathrm{Ext}^{n-i}_{\mathcal{O}_V}(\mathscr{F}, \omega^\circ_V) \cong H^i(V, \mathscr{F})^*$$

We have

$$\mathrm{Ext}^{n-i}_{\mathcal{O}_V}(\mathscr{F}, \omega^\circ_V) = \mathrm{Ext}^{n-i}_{\mathcal{O}_V}(\mathcal{O}_V \otimes_{\mathcal{O}_V} \mathscr{F}, \omega^\circ_V)$$

and corollary D.4.15 on page 568 implies that

$$\begin{aligned} \mathrm{Ext}^{n-i}_{\mathcal{O}_V}(\mathcal{O}_V \otimes_{\mathcal{O}_V} \mathscr{F}, \omega^\circ_V) &= \mathrm{Ext}^{n-i}_{\mathcal{O}_V}(\mathcal{O}_V, \mathscr{F}^\vee \otimes_{\mathcal{O}_V} \omega^\circ_V) \\ &= H^{n-i}(V, \mathscr{F}^\vee \otimes_{\mathcal{O}_V} \omega^\circ_V) \end{aligned}$$

and the result follows from theorem D.3.22 on page 558, which shows that all of the modules in question are finite-dimensional vector spaces over k. □

We conclude this chapter by computing the dualizing sheaf:

PROPOSITION D.4.21. *Let $j \colon V \hookrightarrow k\mathbb{P}^N$ be a smooth n-dimensional variety defined by the quasicoherent ideal-sheaf $\mathscr{I} \subset \mathcal{O}_{k\mathbb{P}^N}$. Then the dualizing sheaf of V is given by*

$$\begin{aligned} \omega^\circ_V &= \omega_{k\mathbb{P}^N} \otimes_{\mathcal{O}_{k\mathbb{P}^N}} \mathcal{O}_V \otimes_{\mathcal{O}_V} \left(\Lambda^r \left(\mathscr{I} / \mathscr{I}^2 \right) \right)^\vee \\ &= j^*(\omega_{k\mathbb{P}^N}) \otimes_{\mathcal{O}_V} \left(\Lambda^r \left(\mathscr{I} / \mathscr{I}^2 \right) \right)^\vee \end{aligned}$$

where $r = N - n$ is the codimension of V. The Adjunction Formula (theorem 4.6.25 on page 212) implies that $\omega^\circ_V = \omega_V$, the canonical sheaf (as in definition 4.6.24 on page 212).

PROOF. We have to compute $\mathscr{E}xt^r(\mathcal{O}_V, \omega_{k\mathbb{P}^n})$ over $k\mathbb{P}^N$. Let $U \subset k\mathbb{P}^N$ be an open affine over which $\mathscr{I}$ can be generated by r elements $f_1, \ldots, f_r \in \mathcal{O}_{k\mathbb{P}^N}(U)$ and let $R = \mathcal{O}_{k\mathbb{P}^N}(U)$. If U is sufficiently small, we get

$$\mathscr{E}xt^r(\mathcal{O}_V, \omega_{k\mathbb{P}^N})|U = \mathscr{E}xt^r(\mathcal{O}_V|U, \omega_{k\mathbb{P}^n}|U) = \mathcal{A}(\mathrm{Ext}^i_{\mathcal{O}_{k\mathbb{P}^N}(U)}(\mathcal{O}_V(U \cap V), \omega_{k\mathbb{P}^N}(U)))$$

by proposition D.4.4 on page 562 and corollary D.4.10 on page 564.

If $x \in V \cap U$ is a point with maximal ideal $\mathfrak{m}$. Since V has codimension r and $R_\mathfrak{m}$ is Cohen-Macaulay and $f_1, \ldots, f_r$ form a regular sequence for $R_\mathfrak{m}$. If we "flip" the Koszul complex $\mathbf{K}(f_1, \ldots, , f_r)$ (see D.2.4 on page 545 and D.2.5 on page 545) we get a free resolution, C_*, of $\mathcal{O}_{V,x}$ over $R_\mathfrak{m}$ whose top-dimensional chain modules look like

(D.4.8) $\qquad C_r = \Lambda^0 R_\mathfrak{m}^r = R_\mathfrak{m} \xrightarrow{\partial_r = \begin{bmatrix} \cdot f_1 \\ \vdots \\ \cdot f_r \end{bmatrix}} C_{r-1} = \Lambda^1 R_\mathfrak{m}^r = \bigoplus_{i=1}^{r} R_\mathfrak{m}$

(see definition D.2.18 on page 540). Choosing a smaller neighborhood, if necessary, we get a complex of sheaves giving a resolution of $\mathcal{O}_V(U \cap V)$ over $\mathcal{O}_{k\mathbb{P}^N}(U)$ whose top boundary operators have the form of D.4.8.

We get

$$H^r(\mathscr{H}om(C_*, \omega_{k\mathbb{P}^N}|U) = \operatorname{coker} \partial_r^* : R^r \to R = R/\mathscr{I}(U) = \mathcal{O}_V(U)$$

so $\omega_V^\circ \cong \omega_{k\mathbb{P}^N} \otimes_{\mathcal{O}_{k\mathbb{P}^N}} \mathcal{O}_V$, by an unnatural isomorphism that depends on the choice of $f_1, \ldots, f_r$. A change of generators affects the entire Koszul complex: if

$$g_i = \sum_{j=1}^{t} a_{i,j} f_j$$

then, at the low end of the resolution, we have

$$\cdots \to \Lambda^{r-1} R^r \to \Lambda^r R^r \to \mathcal{O}_V(U) \to 0$$

and the change of basis will have the effect of mapping $\mathcal{O}_V(U)$ via

$$\mathcal{O}_V(U) \xrightarrow{\times \det(a_{i,j})} \mathcal{O}_V(U)$$

Since we want the map of free resolutions to cover the *identity map* of $\mathcal{O}_V(U)$, we must multiply the entire resolution by $\det(a_{i,j})^{-1}$ whenever the basis changes.

We accomplish this by taking the tensor product $C_* \otimes_{\mathcal{O}_V} \left(\Lambda^r \left(\mathscr{I}/\mathscr{I}^2 \right) \right)^\vee$. The quotient $\mathscr{I}/\mathscr{I}^2$ is a vector space of dimension r over $\mathcal{O}_V$ and $\Lambda^r \left(\mathscr{I}/\mathscr{I}^2 \right)$ — a locally free sheaf of rank 1 — gets multiplied by $\det(a_{i,j})$ when the basis changes, so that $\left(\Lambda^r \left(\mathscr{I}/\mathscr{I}^2 \right) \right)^\vee$ gets multiplied by $\det(a_{i,j})^{-1}$ (see exercise 2 on page 514).

It follows that the sheaves $\mathcal{A}(\operatorname{Ext}^i_{\mathcal{O}_{k\mathbb{P}^N}(U)}(\mathcal{O}_V(U \cap V), \omega_{k\mathbb{P}^N}(U)))$ will patch together to form a global sheaf ω_V°. $\square$

Solutions to Selected Exercises

Chapter 1, 1.2 Exercise 1 (p. 9) Just write

$$z^2 \left(\left(\frac{x}{z} \right)^2 + 3 \frac{x}{z} \frac{y}{z} + 25 \right) = 0$$

to get

$$x^2 + 3xy + 25z^2 = 0$$

In the first case, we set $z = 1$ and recover our original equation. The solution-set in $\mathbb{C}^2$ is a kind of quadratic hyperbola:
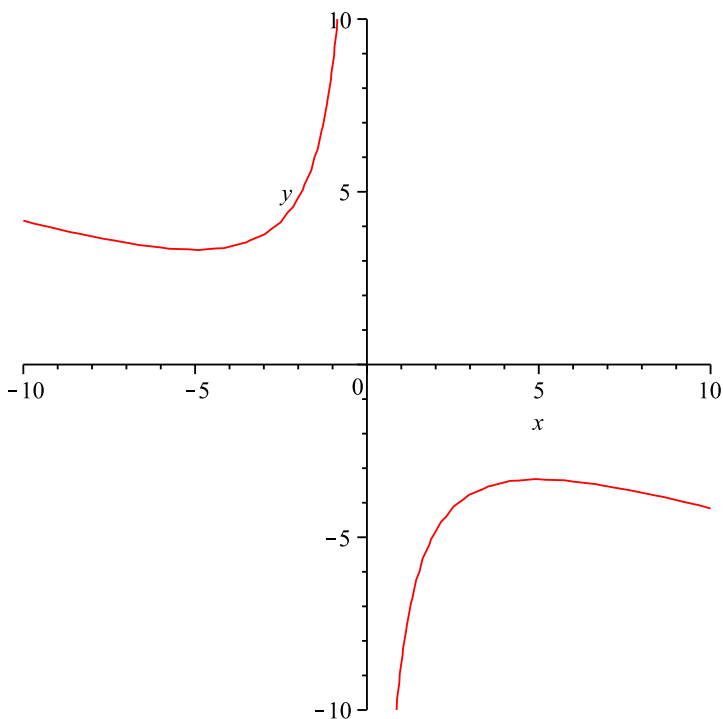
To get the points at infinity, set $z = 0$ to get

$$x^2 + 3xy = 0 = x(x + 3y)$$

which gives several other points:

(1) $x = 0$ and $y \neq 0$, which defines a *single point* in $\mathbb{C}P^1$, $[0\!:\!1\!:\!0]$.
(2) $x \neq 0$ and $x = -3y$, which *also* defines a single point in $\mathbb{C}P^2$, $[-3y\!:\!y\!:\!0]$, or $[-3\!:\!1\!:\!0] = [3\!:\!-1\!:\!0]$.

Chapter 1, 1.2 Exercise 2 (p. 9) We convert them into equations defined over $\mathbb{C}P^2$

$$
\begin{aligned}
y - x^2 - 1 &= 0 \\
y - x^2 - 2 &= 0
\end{aligned}
$$

and

$$
\begin{aligned}
z^2 \left(\frac{y}{z} - \left(\frac{x}{z} \right)^2 - 1 \right) &= 0 \\
z^2 \left(\frac{y}{z} - \left(\frac{x}{z} \right)^2 - 1 \right) &= 0
\end{aligned}
$$

or

$$
\begin{aligned}
yz - x^2 &= z^2 \\
yz - x^2 &= 2z^2
\end{aligned}
$$

The parabolas do not intersect in $\mathbb{C}^2$ but if we go to the $\mathbb{C}P^1$ at infinity (by setting $z = 0$) we get $x = 0$ with no condition on y. This defines the single point $[0 : y : 0] = [0 : 1 : 0] \in \mathbb{C}P^1$.

Chapter 1, 1.2 Exercise 3 (p. 9) This function is not a multiple of y, so we set $y = 1$ to get the polynomial

$$
x^3 + 6x^2 + 11x + 6
$$

which factors to

$$
(x+1)(x+2)(x+3)
$$

giving the factorization

$$
x^3 + 6x^2 + 11x + 6 =
$$

$$
y^3 \left(\left(\frac{x}{y} + 1 \right) \left(\frac{x}{y} + 2 \right) \left(\frac{x}{y} + 3 \right) \right)
$$

$$
= (x+y)(x+2y)(x+3y)
$$

Chapter 1, 1.2 Exercise 4 (p. 9) We start by writing it as

$$
\left(\frac{x}{z} \right)^2 + \left(\frac{y}{z} \right)^2 + 9 = 0
$$

and multiply by z^2 to get

$$
x^2 + y^2 + 9z^2 = 0
$$

This is a circle of radius $3i$ with two points at infinity $(1 : \pm i : 0)$.

Chapter 1, 1.2 Exercise 5 (p. 9) We only have to verify that if

$$
(x_0 : \cdots : x_n) \sim (y_0 : \cdots : y_n)
$$

and

$$
\begin{bmatrix} z_0 \\ \vdots \\ z_n \end{bmatrix} = A \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix}, \quad \begin{bmatrix} w_0 \\ \vdots \\ w_n \end{bmatrix} = A \begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}
$$

then $(z_0 : \cdots : z_n) \sim (w_0 : \cdots : w_n)$. The equivalence above implies that there exists a nonzero $t \in \mathbb{C}$ such that $y_j = t \cdot x_j$ for $j = 0, \ldots, n$. But this implies that $w_j = t \cdot z_j$ for $j = 0, \ldots, n$.

Chapter 1, 1.2 Exercise 6 (p. 9) If the matrix fails to be invertible, then there is a vector

$$
\begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix} \neq 0
$$

such that

$$
A \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix} = 0
$$

so the map $\bar{A}$ ceases to be *well-defined* at the point $(x_0: \cdots : x_n)$.

Chapter 1, 1.2 Exercise 7 (p. 9) We need

$$
A \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \\ 1 \end{bmatrix}
$$

The 1 in the bottom of the right term only depends on the bottom row of A. We have

$$
A_{n+1,1} x_1 + \cdots + A_{n+1,n} x_n + A_{n+1,n+1} = 1
$$

Since the x_i are arbitrary, we conclude that $A_{n+1,1} = \cdots = A_{n+1,n} = 0$ and $A_{n+1,n+1} = 1$.

Chapter 1, 1.2 Exercise 8 (p. 10) It is the mapping

$$
\begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} x_1 + z_1 \\ \vdots \\ x_n + z_n \\ 1 \end{bmatrix}
$$

so it essentially *displaces* everything in $\mathbb{C}^n \subset \mathbb{C}P^n$ by

$$
\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}
$$

Chapter 1, 1.3 Exercise 1 (p. 14) The Sylvester matrix of $t - x(1 + t^2)$ and $t^2 - y(1 - t)$ is

$$
\begin{bmatrix}
-x & 1 & -x & 0 \\
0 & -x & 1 & -x \\
1 & y & -y & 0 \\
0 & 1 & y & -y
\end{bmatrix}
$$

and the determinant is

$$
\text{Res}(t - x(1 + t^2), t^2 - y(1 - t), t) = x^2 + 2yx^2 + 2y^2x^2 + yx - y^2x - y
$$

so the implicit equation is

$$
x^2 + 2yx^2 + 2y^2x^2 + yx - y^2x - y = 0
$$

Chapter 1, 1.3 Exercise 2 (p. 14) These parametric equations are equivalent to $t - x(1 - t^2) = 0$ and $t - y(1 + t^2) = 0$ with a Sylvester matrix of

$$
\begin{bmatrix}
x & 1 & -x & 0 \\
0 & x & 1 & -x \\
-y & 1 & -y & 0 \\
0 & -y & 1 & -y
\end{bmatrix}
$$

and resultant of

$$
r = 4y^2x^2 - x^2 + y^2
$$

so the implicit equation is $r = 0$.

Chapter 1, 1.3 Exercise 3 (p. 15) Our polynomials are $1 - t - x(1+t) = 0$ and $t^2 - y(1 + t^2) = 0$ with a Sylvester matrix of

$$
\begin{bmatrix}
-1 - x & 1 - x & 0 \\
0 & -1 - x & 1 - x \\
1 - y & 0 & -y
\end{bmatrix}
$$

giving the implicit equation

$$-2\,y + 1 - 2\,x - 2\,yx^2 + x^2 = 0$$

Chapter 1, 1.3 Exercise 4 (p. 15) The resultant in question is

$$x^4 + 2\,x^3 + x^2 - 4\,x = x(x - 1)(x^2 + 3x + 4)$$

It follows that x can have one of the 4 values

$$\left\{ 0, 1, \frac{-3 \pm i\sqrt{7}}{2} \right\}$$

Each of these x-values turns out to correspond to a *unique* y-value. Our four solutions are

$$(x, y) = \left\{ (0, 1),\, (1, 0),\, \left(\frac{-3 - i\sqrt{7}}{2}, \frac{3 - i\sqrt{7}}{2} \right),\, \left(\frac{-3 + i\sqrt{7}}{2}, \frac{3 + i\sqrt{7}}{2} \right) \right\}$$

Chapter 1, 1.3 Exercise 5 (p. 15) We get

$$
\begin{aligned}
\mathrm{Res}(s + t - x, s^2 - t^2 - y, s) &= -2xt + x^2 - y \\
\mathrm{Res}(s^2 - t^2 - y, 2s - 3t^2 - z, s) &= 9\,t^4 + 6\,t^2 z - 4\,t^2 - 4\,y + z^2 \\
\mathrm{Res}(s + t - x, 2s - 3t^2 - z, s) &= -3\,t^2 - 2\,t + 2\,x - z
\end{aligned}
$$

and

$$R = \mathrm{Res}(-2xt + x^2 - y, -3\,t^2 - 2\,t + 2\,x - z, t) =$$
$$- 3\,x^4 + 4\,x^3 + 6\,x^2 y - 4\,x^2 z + 4\,yx - 3\,y^2$$

so the implicit equation is

$$3\,x^4 - 4\,x^3 - 6\,x^2 y + 4\,x^2 z - 4\,yx + 3\,y^2 = 0$$

If we compute the resultant of $9\,t^4 + 6\,t^2 z - 4\,t^2 - 4\,y + z^2$ and $-2xt + x^2 - y$ we get

$$9\,x^8 - 36\,x^6 y + 24\,x^6 z - 16\,x^6 + 54\,x^4 y^2$$
$$- 48\,x^4 yz - 32\,x^4 y + 16\,x^4 z^2$$
$$- 36\,x^2 y^3 + 24\,x^2 y^2 z - 16\,x^2 y^2 + 9\,y^4$$

which turns out to be a multiple of R.

Chapter 1, 1.5 Exercise 1 (p. 27) Given 6 points, the first line can start at any of them. Its endpoint has 5 possibilities, and the line from there can go to 4 possible places. The number of paths connecting the 6 points is, therefore, $6! = 720$. A cyclic permutation of the points gives the same hexagon, leaving $6!/6 = 120$ possibilities. Furthermore, traversing the lines in the opposite direction gives the same hexagon, so we are left with $120/2 = 60$ possibilities.

Chapter 1, 1.5 Exercise 2 (p. 27) This is why we prove Pascal's theorem in a projective space, $\mathbb{R}P^2 = \mathbb{R}^2 \cup \mathbb{R}P^1$ (see proposition 1.2.3 on page 4). The one-dimensional projective space $\mathbb{R}P^1$ is the "line at infinity" of $\mathbb{R}P^2$ and is where the intersections lie.

Chapter 1, 1.6 Exercise 1 (p. 33) Rotation in by ϕ in the xy-plane is accomplished by

$$
\begin{bmatrix}
\cos(\pi/4) & -\sin(\pi/4) & 0 \\
\sin(\pi/4) & \cos(\pi/4) & 0 \\
0 & 0 & 1
\end{bmatrix}
=
\begin{bmatrix}
\sqrt{2}/2 & -\sqrt{2}/2 & 0 \\
\sqrt{2}/2 & \sqrt{2}/2 & 0 \\
0 & 0 & 1
\end{bmatrix}
$$

or

$$
R_1 =
\begin{bmatrix}
\sqrt{2}/2 & -\sqrt{2}/2 & 0 & 0 \\
\sqrt{2}/2 & \sqrt{2}/2 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

in $\mathbb{R}P^3$. The displacement is accomplished by the linear transformation

$$
D_1 =
\begin{bmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 2 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

and the composite is

$$
D_1 R_1 =
\begin{bmatrix}
\sqrt{2}/2 & -\sqrt{2}/2 & 0 & 1 \\
\sqrt{2}/2 & \sqrt{2}/2 & 0 & 2 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

The second rotation (in the xz-plane) is done by

$$
R_2 =
\begin{bmatrix}
\cos(\pi/3) & 0 & -\sin(\pi/3) & 0 \\
0 & 1 & 0 & 0 \\
\sin(\pi/3) & 0 & \cos(\pi/3) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
1/2 & 0 & -\sqrt{3}/2 & 0 \\
0 & 1 & 0 & 0 \\
\sqrt{3}/2 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

and the composite of all three is

$$
R_2 D_1 R_1 =
\begin{bmatrix}
\sqrt{2}/4 & -\sqrt{2}/4 & -\sqrt{3}/2 & (1-\sqrt{3})/2 \\
\sqrt{2}/2 & \sqrt{2}/2 & 0 & 2 \\
\sqrt{6}/4 & -\sqrt{6}/4 & 1/2 & (1+\sqrt{3})/2 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

Chapter 1, 1.6 Exercise 2 (p. 33) The two-dimensional image of the scene is just the result of projecting

$$
\mathbb{R}^3 \subset \mathbb{R}P^3
$$

onto $\mathbb{R}P^2$ given by

$$
\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}
\mapsto
\begin{bmatrix} x \\ y \\ z \end{bmatrix}
\in \mathbb{R}P^2
$$

In regular Cartesian coordinates, this is $(x/z, y/z)$.

When the camera is *not* at the origin (pointing in the z-direction), we simply devise a linear transformation, T, that moves it there (and in that direction) and apply it before projecting onto $\mathbb{R}P^2$, so

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix} \in \mathbb{R}P^2$$

Chapter 1, 1.6 Exercise 3 (p. 33) Simply regard this curve as being in $\mathbb{R}P^3$ and project it onto the plane $z = 1$, to get

$$x = \frac{\cos(3t) + t}{t + 3}$$

$$y = \frac{\sin(3t) - t}{t + 3}$$

Plotting this gives the image:
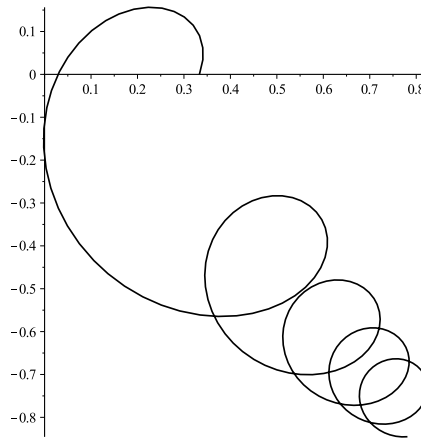
FIGURE E.0.1.   Perspective in projective space

Chapter 2, 2.1 Exercise 1 (p. 38) The closed sets of $\mathbb{A}^1$ are:

   (1) the empty set,
   (2) all of $\mathbb{A}^1$,
   (3) finite sets of point (roots of polynomials).

It follows that the closed sets in the product-topology on $\mathbb{A}^1 \times \mathbb{A}^1$ consist of

   (1) all of $\mathbb{A}^1 \times \mathbb{A}^1$
   (2) {finite set} $\times \mathbb{A}^1$
   (3) $\mathbb{A}^1 \times$ {finite set}
   (4) {finite set} $\times$ {finite set}

and the Zariski topology on $\mathbb{A}^2$ has *many* more closed sets, like the set of points that satisfy

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

or even the diagonal line

$$y = x$$

Chapter 2, 2.1 Exercise 2 (p. 38) Both V and ℓ are closed sets of $\mathbb{A}^n$, so their intersection is also a closed set and a closed subset of $\ell = \mathbb{A}^1$. The only closed sets of ℓ (in the Zariski topology) are:

(1) $\varnothing$
(2) ℓ
(3) finite sets of points.

Since $p \in \ell$ and $p \notin V$, case 2 is ruled out.

Chapter 2, 2.1 Exercise 3 (p. 38) This follows from exercise 40 on page 380, which shows that
$$\mathrm{Ann}(M_1) \cdot \mathrm{Ann}(M_3) \subset \mathrm{Ann}(M_2) \subset \mathrm{Ann}(M_1) \cap \mathrm{Ann}(M_3)$$
and proposition 2.1.2 on page 37.

Chapter 2, 2.1 Exercise 4 (p. 38) We can simplify the ideal $(X_1^2 + X_2^2 - 1, X_1 - 1)$ considerably. Since $X_1^2 - 1 = (X_1 + 1)(X_1 - 1)$, we subtract $X_1 + 1$ times the second generator from the first to get $(X_2^2, X_1 - 1)$. It follows that V consists of the single point $(0, 1)$ and $\mathcal{I}(V) = (X_1 - 1, X_2)$.

Chapter 2, 2.1 Exercise 5 (p. 38) In characteristic 2, $(X_1 + X_2 + X_3)^2 = X_1^2 + X_2^2 + X_3^2$, so V is the plane defined by
$$X_1 + X_2 + X_3 = 0$$
and $\mathcal{I}(V) = (X_1 + X_2 + X_3)$.

Chapter 2, 2.1 Exercise 6 (p. 38) This is XY, since $XY = 0$ implies $X = 0$ or $Y = 0$.

Chapter 2, 2.1 Exercise 7 (p. 38) If we use the results of the previous exercise, $XY = 0$ so $\mathcal{V}((XY))$ is the $P_{XZ} \cup P_{YZ}$ where P_{XZ} denotes the XZ-plane and P_{YZ} denotes the YZ-plane. Similarly, $\mathcal{V}((XZ)) = P_{XY} \cup P_{YZ}$ so that
$$\mathcal{V}((XY, XZ)) = (P_{XZ} \cup P_{YZ}) \cap (P_{XY} \cup P_{YZ}) = P_{YZ}$$
Since $\mathcal{V}((YZ)) = P_{XY} \cup P_{XZ}$, we get
$$\begin{aligned}
\mathcal{V}((XY, XZ, YZ)) &= (P_{XY} \cup P_{XZ}) \cap P_{YZ} \\
&= (P_{XY} \cap P_{YZ}) \cup (P_{XZ} \cap P_{YZ})
\end{aligned}$$
and each of these terms are equal to the union of the axes.

Chapter 2, 2.1 Exercise 8 (p. 39) In $k[V] = k[X, Y]/(Y^2 - X^3)$ the identity $Y^2 = X^3$ holds, so every occurrence of Y^2 can be replaced by X^3.

Chapter 2, 2.1 Exercise 9 (p. 39) It is easily checked to be 1-1 and onto. For it to be an isomorphism, it would have to have a *polynomial inverse,* i.e. function
$$f(X) + g(X)Y$$
such that $f(T^2) + T^3 g(T^2) = T$. This is impossible since degrees of polynomials add, so
$$\deg\left(f(T^2) + T^3 g(T^2)\right) \geq 2$$

Chapter 2, 2.2 Exercise 1 (p. 45) Suppose that $\mathfrak{p} \subset k[X_1, \dots, X_n]$ is prime and suppose that $a^n \in \mathfrak{p}$. If we write $a^n = a \cdot a^{n-1}$, then the defining property of a prime ideal implies that either $a \in \mathfrak{p}$ or $a^{n-1} \in \mathfrak{p}$. In the first case, the claim is proved. In the second case, we do downward induction on n.

Chapter 2, 2.2 Exercise 2 (p. 45) Suppose $\mathfrak{a} \subset k[X_1, \dots, X_n]$ is a proper ideal. The strong form of the Nullstellensatz says that $\mathcal{I}\mathcal{V}(\mathfrak{a}) = \sqrt{\mathfrak{a}}$.

We claim that if $\mathfrak{a} \neq k[X_1, \dots, X_n]$ then the same is true of $\sqrt{\mathfrak{a}}$. The statement that $1 \in \sqrt{\mathfrak{a}}$, is equivalent to saying that $1^n \in \mathfrak{a}$ for some n. But $1^n = 1$ so $1 \in \sqrt{\mathfrak{a}}$ implies that $1 \in \mathfrak{a}$.

Since $\sqrt{\mathfrak{a}} \neq k[X_1, \dots, X_n]$, we conclude that $V(\mathfrak{a}) \neq \varnothing$.

Chapter 2, 2.2 Exercise 3 (p. 45) Set $n = m = 1$. In this case, the Zarski-closed sets are finite sets of points or the empty set or all of $\mathbb{A}^1$. The maps that swaps two points (like 1 and 2) but leaves all other points fixed is Zarski-continuous, but is clearly not regular.

Chapter 2, 2.2 Exercise 4 (p. 45) Since the determinant, z, can never vanish and since it is a polynomial over $X_1, \ldots, X_n$, Hilbert's Nullstellensatz implies that it must be a constant (any nonconstant polynomial has a zero somewhere).

Chapter 2, 2.2 Exercise 5 (p. 45) The second equation implies that

$$XZ - Z = Z(X - 1) = 0$$

so $X = 1$ or $Z = 0$. Plugging each of these cases into the first equation gives:

Case 1:

If $X = 1$ then the first equation becomes $YZ = 1$ which generates a hyperbola.

Case 2:

If $Z = 0$ then the first equation becomes $X = 0$ and Y is unrestricted — i.e., we get the Y-axis. So the two components are the

(1) hyperbola $X = 1$, $YZ = 1$ and
(2) the Y-axis.

Chapter 2, 2.3 Exercise 1 (p. 61) We use $\mathcal{I}(X) = (Y, Z)$, $\mathcal{I}(Y) = (X, Z)$ and $\mathcal{I}(Z) = (X, Y)$. To compute the intersection, we can use the proposition 2.3.14 on page 53:

To compute $\mathcal{I}(X) \cap \mathcal{I}(Y)$, define the ideal

$$(TY, TZ, (1 - T)X, (1 - TZ) \subset k[Y, Z, T]$$

where T is a new variable, and compute the Gröbner basis of this ideal with the Maple command

```
Basis([T*Y, T*Z, (1-T)*X, (1-T)*Z], plex(T,X, Y, Z))
```

We get

$$(Z, YX, TY, -X + XT)$$

Throwing away the terms with T gives

$$\mathcal{I}(X) \cap \mathcal{I}(Y) = (Z, XY)$$

Now we repeat the process, computing

$$\mathcal{I}(X) \cap \mathcal{I}(Y) \cap \mathcal{I}(Z) = (Z, XY) \cap (X, Y)$$

by finding a Gröbner basis of

$$(TZ, TXY, (1 - T)X, (1 - T)Y)$$

that orders T higher than the others. The Maple command is

```
Basis([T*Z, T*X*Y, (1-T)*X, (1-T)*Y], plex(T, X, Y, Z))
```

The result is

$$(ZY, ZX, YX, TZ, -Y + TY, -X + XT)$$

Throwing away all terms containing T gives

$$\mathcal{I}(X) \cap \mathcal{I}(Y) \cap \mathcal{I}(Z) = (XY, YZ, XZ)$$

Chapter 2, 2.3 Exercise 2 (p. 61) We plug $x = 1/2$, $y = 1/2$, and $z = 1 + \sqrt{2}/2$ into the ideal $\mathfrak{P}$ in example 2.3.19 on page 58 to give

$$\mathfrak{P}' = (-1 + 2\,b_5{}^2, -b_5 + a_5, -2 - \sqrt{2} + 2\,b_4 + b_4\sqrt{2},$$
$$-1 + b_4, -2 + 2\,b_4{}^2, a_4, b_3 + b_5 b_4, -\sqrt{2} + 2\,a_3,$$
$$\sqrt{3}b_5 + \sqrt{3}b_5\sqrt{2} - 2\,b_4\sqrt{3}b_5 + 3\,b_2,$$
$$3\,a_2 - 2\sqrt{3} - 1/2\sqrt{2}\sqrt{3} + b_4\sqrt{3})$$

If we take a Gröbner basis of *this,* we get an even simpler representation

$$\mathfrak{P}' = (-1 + 2\,b_5{}^2, -b_5 + a_5, -1 + b_4, a_4, b_3 + b_5, -\sqrt{2} + 2\,a_3,$$

$$3\,b_2 + \left(\sqrt{2}\sqrt{3} - \sqrt{3}\right) b_5, 6\,a_2 - \sqrt{2}\sqrt{3} - 2\,\sqrt{3})$$

from which we conclude

$$
\begin{aligned}
a_5 = b_5 &= \pm 1/\sqrt{2} \\
b_4 &= 1 \\
a_4 &= 0 \\
b_3 &= -b_5 \\
a_3 &= 1/\sqrt{2} \\
b_2 &= -b_5(\sqrt{6} - \sqrt{3})/3 \\
a_2 &= (\sqrt{6} + 2\sqrt{3})/6
\end{aligned}
$$

which gives two solutions:

(1) $\phi_1 = 45°$, $\theta_1 = 90°$, $\theta_2 = 315°$, $\theta_3 = 99.735°$
(2) $\phi_1 = 225°$, $\theta_1 = 90°$, $\theta_2 = 45°$, $\theta_3 = 80.264°$

Note that the two possible values of θ_3 sum up to 180°.

Chapter 2, 2.3 Exercise 3 (p. 61) We start with the same equations as before:

$$
\begin{aligned}
a_5 a_4 a_3 - a_5 b_4 b_3 + a_5 a_4 - x &= 0 \\
b_5 a_4 a_3 - b_5 b_4 b_3 + b_5 a_4 - y &= 0 \\
b_4 a_3 + a_4 b_3 + b_4 - z &= 0 \\
a_3^2 + b_3^2 - 1 &= 0 \\
a_4^2 + b_4^2 - 1 &= 0 \\
a_5^2 + b_5^2 - 1 &= 0
\end{aligned}
$$

(E.0.1)

And we plug in the new directions to get

$$
\begin{aligned}
(a_5 a_4 a_3 - a_5 b_4 b_3)\,a_2 + (-a_5 a_4 b_3 - a_5 b_4 a_3)\,b_2 - 1 &= 0 \\
(b_5 a_4 a_3 - b_5 b_4 b_3)\,a_2 + (-b_5 a_4 b_3 - b_5 b_4 a_3)\,b_2 &= 0 \\
(b_4 a_3 + a_4 b_3)\,a_2 + (a_4 a_3 - b_4 b_3)\,b_2 &= 0 \\
a_2^2 + b_2^2 - 1 &= 0
\end{aligned}
$$

(E.0.2)

The Gröbner basis with lexicographic ordering is

$$(y, b_5, a_5{}^2 - 1,$$

$$4\,x^2 b_4{}^2 - 4\,x^2 + 2\,x^2 z^2 - 4\,x^2 z b_4 + x^4 + z^4 - 4\,z^3 b_4 + 4\,z^2 b_4{}^2,$$

$$z^2 a_5 - 2\,z a_5 b_4 + x^2 a_5 + 2\,x a_4,$$

$$- 4\,x a_5 + 4\,z b_4 a_4 - 2\,z^2 a_4 + 4\,b_4{}^2 x a_5 + z^2 a_5 x - 2\,z a_5 x b_4 + x^3 a_5,$$

$$a_4{}^2 + b_4{}^2 - 1,$$

$$2\,b_4 a_4 - z a_4 + b_3 + x a_5 b_4$$

$$- 2 + 4\,b_4{}^2 - 4\,z b_4 + z^2 + 2\,a_3 + x^2,$$

$$z a_5 - a_5 b_4 + b_2, -a_5 a_4 - x + a_2)$$

from which we conclude that $y = 0$ and $a_5 = \pm 1$. The term next to the last implies that

$$x^2 - 4zb_4 + z^2 + 4b_4^2 = x^2 + (z - 2b_4)^2 = 2 - 2a_3$$

which means x and z lie on a circle of radius $\sqrt{2(1-a_3)}$ and center $(0, 2b_4)$. If we specify that $a_3 = c$, some constant and take a further Gröbner basis (not including c in the list of variables), we get an additional relation between x and z (among other things):

$$(c-1)z^2 + (1+c)x^2 = 0$$

or

$$z = \pm x \sqrt{\frac{1+c}{1-c}}$$

so the reachability set is contained in this pair of lines in the xz-plane (and very small!). The possible values of z are

$$b_4 c + b_4 \pm \sqrt{1 - {b_4}^2 - c^2 + {b_4}^2 c^2}$$

It is interesting that, although the set of points that can be reached is limited, there are many ways to reach each of these points.

Chapter 2, 2.3 Exercise 4 (p. 61) We compute the intersection of the principal ideals generated by these polynomials and take their intersection, using the method of proposition 2.3.14 on page 53: we find a Gröbner basis of the ideal

$$(T(-X^3 - 2YX^2 - XY^2 + 2X),$$
$$(1-T)(4 - 4X^2 - 4Y^2 + X^4 - 2Y^2X^2 + Y^4))$$

using a lexicographic ordering $T \succ X \succ Y$ to get

$$Y^4 X - 2X^3 Y^2 + X^5 - 4XY^2 - 4X^3 + 4X,$$
$$X^4 - 3Y^2 X^2 - 2X^2 + TY^2 X^2 - 2TX^2 - 2XY^3$$
$$+ 4XY + 2XTY^3 - 4TXY + Y^4 T - 4TY^2 + 4T,$$
$$- X^3 - 2YX^2 - XY^2 + 2X + X^3 T + 2YTX^2 + XTY^2 - 2XT$$

Since the only term that does *not* contain T is the top one, it is the answer.

Chapter 2, 2.3 Exercise 5 (p. 61) No. The basis given for $\mathfrak{a}$ is a Gröbner basis with lexicographic ordering and

$$X + Y \to_{\mathfrak{a}} 2Y$$

so $X + Y \notin \mathfrak{a}$.

Chapter 2, 2.3 Exercise 6 (p. 61) Yes. If we compute Gröbner basis of $\mathfrak{a} + (1 - T(X+Y))$ (with respect to *any* ordering) we get (1). Using the `NormalForm` command in Maple gives

$$(X+Y)^2 \quad \to_{\mathfrak{a}} \quad 4Y^2$$
$$(X+Y)^3 \quad \to_{\mathfrak{a}} \quad 0$$

so $(X+Y)^3 \in \mathfrak{a}$.

Chapter 2, 2.3 Exercise 7 (p. 61)

$$X + Y \quad \to_{\mathfrak{a}} \quad X + Y$$
$$(X+Y)^2 \quad \to_{\mathfrak{a}} \quad 4XY$$
$$(X+Y)^3 \quad \to_{\mathfrak{a}} \quad 12XY^2 - 4Y^3$$
$$(X+Y)^4 \quad \to_{\mathfrak{a}} \quad 32XY^3 - 16Y^4$$
$$(X+Y)^5 \quad \to_{\mathfrak{a}} \quad 0$$

so $(X+Y)^5 \in \mathfrak{a}$.

Chapter 2, 2.3 Exercise 8 (p. 61) We find a Gröbner basis for $\mathfrak{a} = (X^2 - 2, Y^3 - 2, A - X - Y)$ with lexicographic order with

$$X \succ Y \succ A$$

to get

$$\mathfrak{a} = (-4 - 24\,A + 12\,A^2 - 6\,A^4 - 4\,A^3 + A^6,$$
$$- 364 + 152\,A - 156\,A^2 + 9\,A^4$$
$$- 160\,A^3 + 24\,A^5 + 310\,Y,$$
$$364 - 462\,A + 156\,A^2 - 9\,A^4$$
$$+ 160\,A^3 - 24\,A^5 + 310\,X)$$

so the minimal polynomial of α is

$$\alpha^6 - 6\alpha^4 - 4\alpha^3 + 12\alpha^2 - 24\alpha - 4 = 0$$

Chapter 2, 2.4 Exercise 1 (p. 70) If we add the two equations, we get

$$2X^2 = -1$$

so

$$X = \pm \frac{i}{\sqrt{2}}$$

If we plug this into the second equation, we get

$$Y^2 + Z^2 = \frac{1}{2}$$

so V consists of two disjoint circles.

Chapter 2, 2.4 Exercise 2 (p. 70) We will prove that the negations of these statements are equivalent. If there exists a nonempty proper subset $S \subset X$ that is both open and closed then $X \setminus S$ is a closed set and

$$X = S \cup (X \setminus S)$$

is a decomposition with $S \cap (X \setminus S) = \varnothing$.

Conversely, if there exists a decomposition

$$X = X_1 \cup X_2$$

with X_1, X_2 nonempty closed sets and $X_1 \cap X_2 = \varnothing$, then $X \setminus X_1 = X_2$, so X_2 is open as well as closed.

Chapter 2, 2.4 Exercise 3 (p. 70) The product $h_1 h_2 \in k$ is nonzero if and only if both factors are nonzero.

Chapter 2, 2.4 Exercise 4 (p. 70) The inclusion $D(f) \subset D(g)$ is equivalent to saying $\mathcal{V}(g) \subset \mathcal{V}(f)$ which is equivalent to saying that $\mathcal{IV}(f) \subset \mathcal{IV}(g)$. The conclusion follows by the strong form of Hilbert's Nullstellensatz.

Chapter 2, 2.4 Exercise 5 (p. 70) Cramer's Rule states that

$$(A^{-1})_{i,j} = \frac{\hat{A}_{i,j}}{\det A}$$

where $\hat{A}_{i,j}$ is the i, j^{th} cofactor of A — a polynomial function of the entries of A. In the coordinate ring of $GL(n, k)$ (see equation 2.4.1 on page 64), $T \cdot \det A = 1$ so our regular map is given by

$$(A^{-1})_{i,j} = T \cdot \hat{A}_{i,j}$$

which are polynomial functions.

Chapter 2, 2.4 Exercise 6 (p. 71) Suppose k is an infinite field with $k \subset R$. The key fact here is that everything under discussion is a *vector-space* over k: all of the ideals and R itself. The hypotheses imply that

$$\mathfrak{J} = \bigcup_{i=1}^{n} (\mathfrak{J} \cap \mathfrak{J}_i)$$

with $\mathfrak{J} \cap \mathfrak{J}_i \subset \mathfrak{J}$ for all i. The conclusion follows from the fact (from linear algebra) that a vector-space over an infinite field cannot be a finite union of *proper* subspaces (see exercise 37 on page 380).

Chapter 2, 2.4 Exercise 7 (p. 71) We do induction on n. The case $n = 1$ is trivial. Now suppose

$$\mathfrak{J} \subset \bigcup_{i=1}^{n} \mathfrak{J}_i$$

but $\mathfrak{J} \not\subset \mathfrak{J}_\alpha$ for all α. We will get a contradiction. Without loss of generality, assume n is minimal. Then, for each i, we can pick $x_i \in \mathfrak{J} \setminus \bigcup_{j \neq i} \mathfrak{J}_j$. If $n = 2$, then $x_1 + x_2$ is not contained in either $\mathfrak{J}_1$ or $\mathfrak{J}_2$ — if $x_1 + x_2 \in \mathfrak{J}_1$, then $x_1 + x_2 - x_1 = x_2 \in \mathfrak{J}_1$ and a similar argument shows that $x_1 + x_2 \notin \mathfrak{J}_2$.

If $n > 2$, assume $\mathfrak{J}_n$ is prime and consider

$$y = x_1 \cdots x_{n-1} + x_n$$

If $y \in \bigcup_{j=1}^{n-1} \mathfrak{J}_j$, then $x_1 \cdots x_{n-1} \in \mathfrak{J}_j$ for $j = 1, \ldots, n-1$ so $y - x_1 \cdots x_{n-1} = x_n \in \bigcup_{j=1}^{n-1} \mathfrak{J}_j$, a contradiction.

If $y \in \mathfrak{J}_n$, then so is $y - x_n = x_1 \cdots x_{n-1} \in \mathfrak{J}_n$. Since $\mathfrak{J}_n$ is prime, one of the $x_i \in \mathfrak{J}_n$ for some $1 \leq i \leq n-1$ — also a contradiction.

Chapter 2, 2.4 Exercise 8 (p. 71) Since

$$\mathcal{V}(\mathfrak{a}) \cap D(b) \subset \mathcal{V}(\mathfrak{a}) \setminus \mathcal{V}(\mathfrak{b})$$

it follows that f vanishes on $\mathcal{V}(\mathfrak{a}) \cap D(b)$ — whose coordinate ring is the localization, $k[\mathcal{V}(\mathfrak{a})]_b = R_b / \mathfrak{a}_b$, (see proposition 2.4.5 on page 63), where

$$\mathfrak{a}_b = \mathfrak{a} \cdot R_b \subset R_b$$

is the image of $\mathfrak{a}$ in the localization. It follows that $f \in \mathfrak{a}_b$. If

$$\mathfrak{a} = (u_1, \ldots, u_t)$$

then we get $f = \sum_{j=1}^{t} a_j u_j / g^{n_i}$ so that clearing denominators implies the conclusion.

Chapter 2, 2.4 Exercise 9 (p. 71) Since R is noetherian, exercise 8 on page 71 shows that

$$\mathcal{I}(\mathcal{V}(\mathfrak{a}) \setminus \mathcal{V}(\mathfrak{b})) \subset (\mathfrak{a} : \mathfrak{b}^\infty)$$

Conversely, suppose $x \in (\mathfrak{a} : \mathfrak{b}^\infty)$, $p \in \mathcal{V}(\mathfrak{a}) \setminus \mathcal{V}(\mathfrak{b})$, and $x(p) \neq 0$. Since $p \notin \mathfrak{b}$, there exists an element $g \in \mathfrak{b}$ such that $g(p) \neq 0$. It follows that $(x \cdot g^n)(p) \neq 0$ for all n. But this contradicts the facts that $x \cdot g^n \in \mathfrak{a}$ for some n and $p \in \mathcal{V}(\mathfrak{a})$.

Chapter 2, 2.5 Exercise 1 (p. 79) $F(\mathfrak{m}_p) \cdot k[V] = k[V]$, representing the *empty set*.

Chapter 2, 2.5 Exercise 2 (p. 79) The induced map of coordinate rings is

$$
\begin{array}{rcl}
k[X, Y] & \rightarrow & k[X, Y] \\
X & \mapsto & X \\
Y & \mapsto & XY
\end{array}
$$

So the maximal ideal $(X - a, Y - b)$ maps to $(X - a, XY - b)$. If $a \neq 0$ this is equivalent to

$$\left(X - a, Y - \frac{b}{a} \right)$$
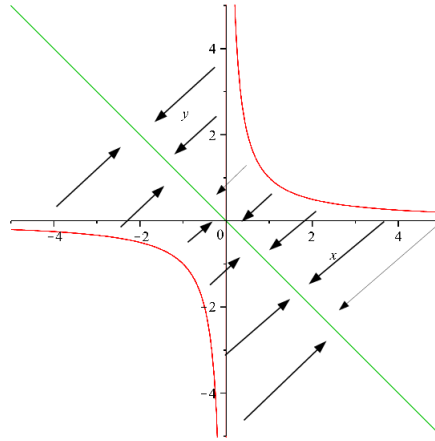
FIGURE E.0.2.   Hyperbola projected onto a line

corresponding to the point $(a, b/a)$. If $a = 0$ and $b = 0$, we get the ideal $(X, XY) = (X)$ which corresponds to the entire Y-axis. If $a = 0$ and $b \neq 0$, we get $(X, XY - b) = (-b) = (1)$ corresponding to the empty set.

Chapter 2, 2.5 Exercise 3 (p. 79) The geometric approach is probably the simplest way to proceed. The hyperbola in figure E.0.2 projects onto the line $Y = -X$. We get a projection onto $\mathbb{A}^1$ by the map

$$(X, Y) \mapsto Y - X$$

and the inverse image of a point $U \in \mathbb{A}^1$ consists of the solutions of

$$X - X^{-1} = U$$

or

$$X^2 - XU - 1 = 0$$

which has ≤ 2 solutions for any U. This induces a map

$$k[U] \to k[X, Y]/(XY - 1) = k[T, T^{-1}]$$

where the image of U is $T - T^{-1}$. As a module over $k[U]$, $k[T, T^{-1}]$ is

$$k[U] \cdot 1 \oplus k[U] \cdot T$$

Chapter 2, 2.5 Exercise 4 (p. 80) This follows from exercise 3 on page 38 and a straight-forward induction. Since the ideals, $\mathfrak{p}_i$, are prime, proposition 2.4.18 on page 68 implies that the $\mathcal{V}(\mathfrak{p}_i)$ are irreducible.

Chapter 2, 2.5 Exercise 6 (p. 90) If $f = p/q$ where $p, q \in k[V]$, simply define $U = D(q)$ (see definition 2.4.4 on page 63 and proposition 2.4.5 on page 63).

Chapter 2, 2.5 Exercise 7 (p. 90) The set U is certainly open, but if it were affine, the maps

$$\begin{aligned}
k[\mathbb{A}^2] &\to k[U] \\
k[U] &\to k[D(X)] = k[X, X^{-1}, Y] \\
k[U] &\to k[D(Y)] = k[X, Y, Y^{-1}]
\end{aligned}$$

would all be injective, by lemma 2.5.20 on page 82, so $k[U] = k[X, X^{-1}, Y] \cap k[X, Y, Y^{-1}] \subset k(X, Y)$ which is precisely

$k[X, Y] = k[\mathbb{A}^2]$. So the inclusion

$$k[\mathbb{A}^2] \to k[U]$$

must be an isomorphism — i.e., a 1-1 and onto mapping of *spaces.* But the inclusion

$$U \hookrightarrow \mathbb{A}^2$$

is not onto.

Chapter 2, 2.5 Exercise 9 (p. 90) Certainly, it is well-defined if $X \neq 0$. We have

$$\frac{Y}{X} = \frac{Y}{X} \cdot \frac{Y}{Y} = \frac{Y^2}{XY} \sim \frac{X^2 + X^3}{XY} = \frac{X + X^2}{Y}$$

which is well-defined if $Y \neq 0$.

Chapter 2, 2.5 Exercise 11 (p. 90) The problem is that the rationalization is defined over $\mathbb{C}$ and its subfields. There is no homomorphism from one of these to $\mathbb{F}_5$. For instance, plugging $t = 1$ into equation 2.5.10 on page 88 gives the point $(4, 3)$ which does not satisfy the equation:

$$4^2 + 4 \cdot 3 + 3^2 \equiv 2 \pmod 5$$

Chapter 2, 2.5 Exercise 13 (p. 91) The identity

$$x \cdot f + y \cdot g = 1 \in F$$

implies that the ideal (x, y) corresponds to the *empty set* in $R[f, g]$. If W is the algebraic set whose coordinate ring is $R[f, g] \subset F$, then the inclusions

$$R \hookrightarrow R[f, g] \hookrightarrow k[D] = R_x \cap R_y$$

induce maps

$$D \to W \xrightarrow{f} V$$

and f misses the origin, i.e. $f(W) \subset D \subset V$. We conclude that $D = W$ and

$$R[f, g] = R_x \cap R_y = k[D]$$

We can describe $R[f, g]$ as an affine ring

$$k[X, Y, U, V, F, G]/(XY + X^2U + Y^2V, FY + U, GX + V)$$

Chapter 2, 2.7 Exercise 1 (p. 98) This follows immediately from diagram 2.7.5 on page 96: Δ is a regular map from W to $\Delta(W)$ and either projection p_1 or p_2 is an inverse regular map.

Chapter 2, 2.7 Exercise 2 (p. 98) If $q \in V$ is any point, let $r_q = 1\colon V \to \{q\} \times V$ be the identification. Now we simply define $g_a = \mu \circ (a \times 1) \circ r_a\colon V \to V$, a regular map. This is just left-multiplication by a. If we compose these types of maps, we get

$$
\begin{aligned}
g_a \circ g_{a'}(v) &= \mu \circ (a \times 1) \circ r_a \circ \mu \circ (a' \times 1) \circ r_{a'}(v) \\
&= \mu \circ (1 \times \mu) \left(\{a\} \times \{a'\} \times \{v\}\right) \\
&= \mu \circ (\mu \times 1) \left(\{a\} \times \{a'\} \times \{v\}\right) \\
&= \mu \left(\{\mu(a, a'\} \times \{v\}\right) \\
&= g_{\mu(a, a')}(v)
\end{aligned}
$$

for any $v \in V$, so $g_a \circ g_{a'} = g_{\mu(a, a')}$ and the inverse to g_a is precisely $g_{\iota(a)}$. Now we can define

$$f_{a,b} = g_b \circ g_{\iota(a)} = g_{\mu(b, \iota(a))} = g_{ba^{-1}}$$

Chapter 2, 2.8 Exercise 1 (p. 107) Any chain of prime ideals in $S^{-1}R$ gives rise to one in R (see corollary A.1.94 on page 385).

Chapter 2, 2.8 Exercise 2 (p. 107) We know that $\dim R_{\mathfrak{m}} \leq \dim R$. Suppose

$$\mathfrak{p}_d \supsetneq \cdots \supsetneq \mathfrak{p}_0$$

is a maximal chain of prime ideals in R, where $\mathfrak{p}_d = \mathfrak{m}$ is maximal. Then corollary A.1.94 on page 385 shows that

$$\mathfrak{p}_d \cdot R_{\mathfrak{m}} \supsetneq \cdots \supsetneq \mathfrak{p}_0 \cdot R_{\mathfrak{m}}$$

is also a chain of prime ideals, so that

$$\dim R_{\mathfrak{m}} = \dim R$$

Chapter 2, 2.8 Exercise 3 (p. 115) Suppose A is a unique factorization domain and $\mathfrak{p}$ is a prime ideal of height 1. If $x \in \mathfrak{p}$ then x has a unique factorization

$$x = \prod a_i^{n_i}$$

where the $a_i \in A$ are irreducible. Since $\mathfrak{p}$ is prime, we must have $a_j \in \mathfrak{p}$ for some j and $(a_j) \subset \mathfrak{p}$. Since a_j is irreducible, lemma A.3.1 on page 411 implies that (a_j) is prime. Since $\mathfrak{p}$ is of height 1, the only prime ideal properly contained within $\mathfrak{p}$ is (0). Since $(a_j) \neq (0)$, it follows that $(a_j) = \mathfrak{p}$. So every height-1 prime ideal is principal.

Conversely, suppose every height-1 prime ideal is principal. If $a \in A$ is an irreducible element, let $\mathfrak{p}$ be the smallest prime ideal such that $(a) \subset \mathfrak{p}$. Theorem 2.8.29 on page 109 implies that $\mathrm{ht}(\mathfrak{p}) = 1$, so $\mathfrak{p} = (b)$ for some $b \in A$ and $(a) \subset (b)$ implies that $b|a$. But a is irreducible so $b = u \cdot a$ for some unit $u \in A$ and $(a) = (b)$. Then (a) is a prime ideal and lemma A.3.1 on page 411 implies that A is a unique factorization domain.

Chapter 2, 2.8 Exercise 4 (p. 115) Since the image of f is dense, the intersection $f(V) \cap W' \neq \varnothing$ for any open set $W' \subset W$. Since $U \subset V$ is dense and f is Zariski-continuous, $f^{-1}(W') \cap U \neq \varnothing$, which implies that $f(U) \cap W' \neq \varnothing$.

Chapter 2, 2.8 Exercise 5 (p. 115) The hypotheses imply that V and W have the same dimension and $[k(V):k(W)] = d$, so that the norm-map $N_{k(V)/k(W)}:k(V) \to k(W)$ is well-defined. If

$$N_{k(V)/k(W)}(x) = y = x \cdot \prod_{j=1}^{d} x_k \in k(W)$$

where the $x_j = p_i/q_i$ are the conjugates of x. Now $y = p/q$ is regular in an open set $D(q) \subset W$ in which case $U'' = D(p) \cap D(q) \subset D(q)$ is open. Since $q = \prod_{j=2}^{d} q_j$ and $p = x \cdot \prod_{j=1}^{d} p_j$, we have $f^{-1}(U'') = D(p) \cap D(q)$ where p, q are regarded as elements of $k[V]$ (see exercise 5 on page 80). It follows that

$$f^{-1}(U'') = U \cap \bigcap_{j=2}^{d} \left(D(p_j) \cap D(q_j) \right) \subset U$$

Chapter 2, 2.8 Exercise 6 (p. 115) Since the X_i generate $k[V]$, their images, x_i, generate $k(V)$ over $k(W)$, i.e.

$$k(V) = k(W)(x_1, \ldots, x_n)$$

The element α exists by the Primitive Element Theorem (A.2.21 on page 394). That it has the required form follows from the *proof* of theorem A.2.21 on page 394, which implies that we can form linear combinations of the x_j to get α as long as the field k is infinite.

Chapter 2, 2.8 Exercise 7 (p. 115) We have $k(V) = k(W)[\alpha]$ where $\alpha = \sum_{j=1}^{n} \beta_j x_j$, so define

$$\begin{aligned} \ell: \mathbb{A}^n &\to \mathbb{A}^1 \\ (X_1, \ldots, X_n) &\mapsto \sum_{j=1}^{n} \beta_j X_j \end{aligned}$$

and set

$$\varphi = (f, \ell) \colon V \to W \times \mathbb{A}^1$$

Now define $\bar{V}$ to be the Zariski closure of im φ. Pulling back rational functions on im φ gives us $k(\bar{V}) = k(W)[\alpha] = k(V)$ so $\varphi \colon V \to \bar{V}$ is a birational equivalence.

Chapter 2, 2.8 Exercise 8 (p. 115) That is so the extension $k(W) \subset k(V)$ is *separable* — required by theorem A.2.21 on page 394.

Chapter 2, 2.8 Exercise 9 (p. 115) Let $p_\alpha(X) \in k(W)[X]$ be the minimum polynomial of α from exercises 7 on page 115 and 6 on page 115 — this polynomial has degree d and coefficients that are elements of $k(W)$. Consider

$$V' = \mathcal{V}(p_\alpha(X)) \subset W \times \mathbb{A}^1$$

We have $\bar{V} \subset V'$ and $k(V') \to k(\bar{V})$ is an isomorphism, so that V' is birationally equivalent to $\bar{V}$ and V — which means that it is an isomorphism on an open set $U \subset V$. Let $W' \subset W$ be the open set where $p_\alpha(X)$ is a regular function — i.e., where the its rational coefficients are regular (see exercise 6 on page 90). Let $\Delta \in k[U]$ be the discriminant of $p_\alpha(X)$ (see corollary A.1.56 on page 366). Then Δ is not identically zero and $D(\Delta) \cap W' \subset W'$ is an open set where $\Delta \neq 0$. By exercise 5 on page 115, $f(U)$ contains an open set $W'' \subset W$.

    It follows that $W_1 = D(\Delta) \cap W' \cap W'' \subset W$ is an open set for which $f^{-1}(w)$ contains d distinct points whenever $w \in W_1$.

Chapter 3, 3.3 Exercise 1 (p. 130) This is easily found by differentiating $Y^2 - X^2 - X^3$. It has one singular point at the origin.

Chapter 3, 3.3 Exercise 2 (p. 130) We begin by computing the Jacobian

$$\mathcal{J} = \begin{bmatrix} 3X^2 & 2X \\ 3Y^2 & 2Y \\ 3Z^2 & 2Z \\ 3W^2 & 2W \end{bmatrix}$$

The second equation in 3.3.7 on page 130 implies that at least one of the variables must be nonzero. We assume that is X. If $X \neq 0$, we get (from lemma 3.3.15 on page 128)

$$\mathrm{rank}(\mathcal{J}) = 1 + \mathrm{rank} \begin{bmatrix} 6X^2Y - 6Y^2X \\ 6X^2Z - 6Z^2X \\ 6X^2W - 6YW^2X \end{bmatrix}$$

The variety $\mathcal{R}(\mathcal{J}, 1)$ (where the rank of $\mathcal{J}$ is less than maximal) is given by

$$\begin{bmatrix} 6X^2Y - 6Y^2X \\ 6X^2Z - 6Z^2X \\ 6X^2W - 6YW^2X \end{bmatrix} = 0$$

or $Y = Z = W = X$. The second equation in 3.3.7 on page 130 then implies that

$$X = \pm\frac{1}{2}, \; Y = \pm\frac{1}{2}, \; Z = \pm\frac{1}{2}, \; W = \pm\frac{1}{2}$$

and the quantity $X^3 + Y^3 + Z^3 + W^3$ in the first equation in 3.3.7 on page 130 takes on the values

$$\pm\frac{1}{2}, \pm\frac{1}{4}, 0$$

If a is not equal to any of these five values, the variety $\mathcal{R}(\mathcal{J}, 1)$ will not intersect V, and V will be smooth and two-dimensional.
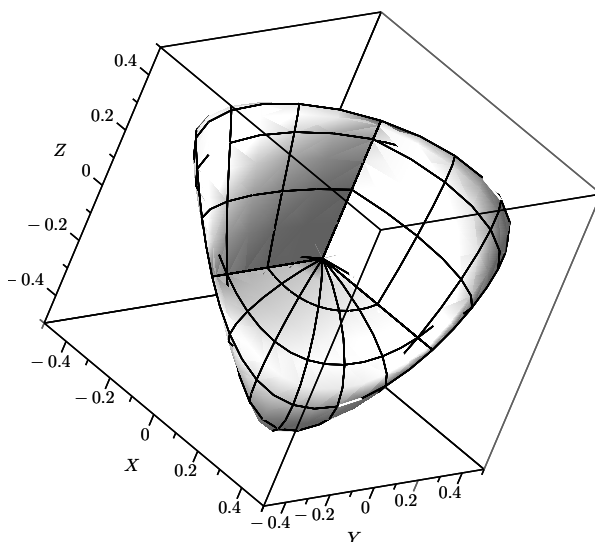
FIGURE E.0.3.   Steiner's Roman surface

Chapter 3, 3.3 Exercise 3 (p. 130) Use the parametric equation for the sphere in example 2.5.31 on page 89 to get a rational parametrization of the Roman surface

$$\begin{aligned}
X &= 2t(s^2 + t^2 - 1)/(1 + s^2 + t^2)^2 \\
Y &= 2s(s^2 + t^2 - 1)/(1 + s^2 + t^2)^2 \\
Z &= 4st/(1 + s^2 + t^2)^2
\end{aligned}$$

As in example 3.3.14 on page 126, we eliminate s and t from these equations to get equations for X, Y, and Z. The simplest way to do this is to find a Gröbner basis for the ideal

$$\begin{aligned}
L = (X(1 + s^2 + t^2)^2 - 2t(s^2 + t^2 - 1), \\
Y(1 + s^2 + t^2)^2 - 2s(s^2 + t^2 - 1), \\
Z(1 + s^2 + t^2)^2 - 4st)
\end{aligned}$$

lexicographically, ordering s and t higher than X, Y, and Z — for instance, issue the Maple command `Basis(L,plex(s,t,X,Y,Z)`. Setting the only term in the Gröbner that does *not* contain s or t to zero gives the implicit equation

$$X^2Y^2 + Y^2Z^2 + X^2Z^2 - XYZ = 0$$

The Jacobian is

$$\mathcal{J} = \begin{bmatrix} 2\,XY^2 + 2\,XZ^2 - YZ \\ 2\,X^2Y + 2\,YZ^2 - XZ \\ 2\,Y^2Z + 2\,X^2Z - XY \end{bmatrix}$$

and singularities occur when it is of rank 0, i.e., when all of its terms vanish.

We form a Gröbner basis of the terms of $\mathcal{J}$ to get

$$(-20\,YZ^4 + YZ^2 + 64\,YZ^6,$$

$$- Y^2Z^2 + 16\,Y^2Z^4,$$

$$44\,YZ^3 - 128\,YZ^5 - 3\,YZ + 12\,Y^3Z,$$

$$32\,YZ^5 - 14\,YZ^3 + 3\,XZ^2,$$

$$-4\,Y^2Z^2 + YXZ, 28\,YZ^3 + 6\,Y^2X - 3\,YZ - 64\,YZ^5,$$

$$2\,Y^2Z + 2\,ZX^2 - YX,$$

$$2\,YX^2 + 2\,YZ^2 - XZ)$$

The term $-Y^2Z^2 + 16\,Y^2Z^4$ implies that, if $Y \neq 0$ then either $Z = 0$ or $Z = \pm 1/2$. If we set $Z = 1/2$ (by adjoining $Z - 1/2$ to the terms in $\mathcal{J}$ and take a Gröbner basis, we get

$$(2Z - 1, Y^2, X - Y)$$

so $Z = 1/2$ implies $X = Y = 0$. In fact, we get the same result if Z is *any* nonzero value. Since the equation for the Roman surface is symmetric in the three variables, we conclude that the Roman surface has precisely three singular lines

(1) $X = Y = 0$, Z arbitrary
(2) $X = Z = 0$, Y arbitrary
(3) $Y = Z = 0$, X arbitrary

all of which meet at a triple point at the origin. This is clear from figure E.0.3 on the previous page.

Chapter 3, 3.3 Exercise 4 (p. 130) It is necessary and sufficient to show that

$$f(X, Y, Z) = f(\bar{X}, \bar{Y}, \bar{Z})$$

if and only if

$$(X, Y, Z) = \pm(\bar{X}, \bar{Y}, \bar{Z})$$

We represent $f(X, Y, Z) = f(\bar{X}, \bar{Y}, \bar{Z})$ on the unit sphere by the ideal

$$(YZ - \bar{Y}\bar{Z}, XZ - \bar{X}\bar{Z}, XY - \bar{X}\bar{Y},$$

$$Y^2 - Z^2 - \bar{Y}^2 + \bar{Z}^2,$$

$$X^2 + Y^2 + Z^2 - 1, \bar{X}^2 + \bar{Y} + \bar{Z} - 1) \in \mathbb{R}[X, Y, Z, \bar{X}, \bar{Y}, \bar{Z}]$$

If we compute a Gröbner basis via lexicographic ordering with $X \succ Y \succ Z \succ \bar{X} \succ \bar{Y} \succ \bar{Z}$, we get

$$(\bar{X}^2 + \bar{Y} + \bar{Z} - 1, Z^2 - \bar{Z}^2, Y\bar{Z} - \bar{Y}Z, YZ - \bar{Y}\bar{Z},$$

$$Y^2 - \bar{Y}^2, X\bar{Z} - \bar{X}Z, X\bar{Y} - \bar{X}Y, XZ - \bar{X}\bar{Z}, XY - \bar{X}\bar{Y}, X^2 + \bar{Y}^2 + \bar{Z}^2 - 1)$$

The basis-elements $Y^2 - \bar{Y}^2$ and $Z^2 - \bar{Z}^2$ imply that $Y = \pm\bar{Y}$ and $Z = \pm\bar{Z}$. If all variables are nonzero, the facts that $XY = \bar{X}\bar{Y}$ and $YZ = \bar{Y}\bar{Z}$ imply that

$$Y = \bar{Y} \implies \begin{cases} X = \bar{X} \\ Z = \bar{Z} \end{cases}$$

and

$$Y = -\bar{Y} \implies \begin{cases} X = -\bar{X} \\ Z = -\bar{Z} \end{cases}$$

Chapter 3, 3.3 Exercise 5 (p. 130) Since W is smooth, it has no singular points — so the partial derivatives do not vanish on W. It follows that

$$W \cap \mathcal{V}\left(\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}\right) = \varnothing$$

The weak form of Hilbert's Nullstellensatz (2.2.3 on page 40) implies the conclusion.

Chapter 3, 3.3 Exercise 6 (p. 132) In this case, we will get $\mathrm{IT}(\mathfrak{I}) = (c)$ where c is a nonzero constant. This defines the empty set.

Chapter 3, 3.3 Exercise 7 (p. 139) Corollary 3.3.32 on page 138 implies that x is contained in a single irreducible component, V_0, of V. It follows that $\mathcal{O}_{V,x} = \mathcal{O}_{V_0,x}$, a localization of $k[V_0]$ which is an integral domain (because V_0 is irreducible).

Chapter 3, 3.3 Exercise 8 (p. 139) This follows from lemma 3.3.25 on page 134 and exercise 7 on page 139 which show that

$$\frac{R}{(u_1, \ldots, u_i)}$$

is an integral domain of dimension $m - i$. If the image of u_{i+1} in it were 0, then the maximal ideal of R could have been generated by fewer than m elements, which is a contradiction.

Chapter 3, 3.3 Exercise 9 (p. 139) Since G is smooth, by lemma 3.3.17 on page 129, it follows that the inversion map

$$\begin{aligned}
\iota: G &\rightarrow G \\
x &\mapsto x^{-1}
\end{aligned}$$

must send G_0 to a single component of G, since it is *continuous*. Since it sends the identity element to itself, that single component must be G_0. A similar argument shows that the right-multiplication map (where $a \in G_0$)

$$\mu_a: G \rightarrow G$$

must send G_0 to itself (since $a^{-1} \in G_0$). It follows that $G_0 \subset G$ is a subgroup.

Chapter 3, 3.4 Exercise 1 (p. 150) Suppose $k[W] = k[X_1, \ldots, X_\ell]/(r_1, \ldots, r_m)$. Since W is smooth, the kernel of the Jacobian,

$$\frac{\partial(r_1, \ldots, r_m)}{\partial(X_1, \ldots, X_\ell)}$$

has a dimension equal to the dimension of W. Since $p(T)$ is a *minimal polynomial*, it is irreducible in $k(W)$ hence it has no zeroes in common with $p'(T)$, so the augmented Jacobian

$$\frac{\partial(r_1, \ldots, r_m, p)}{\partial(X_1, \ldots, X_\ell, T)}$$

has the same nullspace.

Chapter 3, 3.4 Exercise 2 (p. 150)   (1) The map $g: V \rightarrow \bar{V}$ is *finite*. This is because $k[V]$ is a finitely generated module over $k[W]$, so it is a finitely generated module over $k[W][\alpha]$.

(2) It is of *degree 1*. This is because the induced map of function-fields $k(W)[\alpha] \rightarrow k(V)$ is an *isomorphism*.

The conclusion follows from exercise 1 on page 150, which shows that $\bar{V}$ is smooth, hence normal, and lemma 3.4.6 on page 147.

Chapter 3, 3.4 Exercise 3 (p. 150) We need W to be smooth so $\mathfrak{m}_w$ is generated by n local parameters, where $n = \dim W$. At singular points, it can require more elements to generate it.

Chapter 3, 3.4 Exercise 4 (p. 154) We have

$$R = k[W] = \frac{k[X,Y,Z]}{(Z^2 - XY)} = k[X,Y] \oplus Z \cdot k[X,Y]$$

and we can define an injective homomorphism of rings

$$f: R = k[X,Y] \oplus Z \cdot k[X,Y] \to k[U,V]$$
$$X \mapsto U^2$$
$$Y \mapsto V^2$$
$$Z \mapsto UV$$

Since U and V have minimal polynomials $U^2 - X$ and $V^2 - Y$, it follows that they are *integral* over R and corollary A.4.6 on page 421 implies that $F: R \to k[U,V]$ is an integral extension of rings. The homomorphism f also induces an inclusion of fraction-fields

$$\bar{f}: F = k(X,Y)[\sqrt{XY}] \to k(U,V)$$

It is not hard to see that

$$k[U,V] \subset k(X,Y)[\sqrt{XY}][\sqrt{X}] \subset k(U,V)$$

Since $k[U,V]$ is integrally closed in $k(U,V)$ (see corollary A.4.12 on page 423) it is certainly integrally closed in the smaller field F. It follows that $k[U,V]$ is the integral closure of R over $k(W)[\sqrt{X}]$.

Chapter 3, 3.4 Exercise 5 (p. 154) The equations above define an injective homomorphism of rings

$$f: k[W] = \frac{k[X,Y]}{(Y^2 - X^2(X+1))} \to k[T]$$

Since T satisfies

$$T^2 - 1 - X = 0$$

it follows that T is integral over $k[W]$ and that f defines an integral extension of rings. If we pass to fraction-fields, we get

$$\frac{Y}{X} \mapsto T$$

so that the induced map of fraction-fields is an isomorphism. The conclusion follows from the fact that $k[T]$ is integrally closed in $k(T)$ (see corollary A.4.12 on page 423).

Chapter 3, 3.5 Exercise 1 (p. 159) All elements of $\Omega_{k[V]/k}$ are of the form

$$f(X,Y) \cdot dX + g(X,Y) \cdot dY$$

for $f,g \in k[V]$. The relation $dY = 2X \cdot dX$ implies that we can eliminate dY from any expression, so we get

$$\Omega_{k[V]/k} = k[V] \cdot dX$$

This is a free module defining a trivial bundle — which makes sense since V is isomorphic to an affine line $\mathbb{A}^1$ (see example 2.4.9 on page 65).

Chapter 3, 3.5 Exercise 2 (p. 160) Since W is smooth, there exist (see exercise 5 on page 130) functions $r,s \in k[X,Y]$ such that

$$r \cdot \frac{\partial f}{\partial X} + s \cdot \frac{\partial f}{\partial Y} = 1 \quad (\mathrm{mod}\ (f(X,Y)))$$

Now multiply our formula

$$\omega = \frac{dX}{\partial f/\partial Y} = -\frac{dY}{\partial f/\partial X} \in \Omega_{k[W]/k}$$

by this to get

$$\omega = r \cdot dX - s \cdot dY$$

Chapter 3, 3.5 Exercise 3 (p. 160) As before, we have

$$2X \cdot dX + 2Y \cdot dY = 0$$

so we have

$$\omega = \frac{dX}{Y} = -\frac{dY}{X}$$

is regular — and the relation $X^2 + Y^2 = 1$ implies

$$\omega = Y \cdot dX - X \cdot dY$$

so it is regular. We have

$$\begin{aligned} dX &= Y \cdot \omega \\ dY &= -X \cdot \omega \end{aligned}$$

so

$$\Omega_{k[V]/k} = k[V] \cdot \omega$$

Chapter 4, 4.2 Exercise 1 (p. 168) The rings in question are not algebras over fields. For instance consider

$$f \colon \mathbb{Z}[X] \to \mathbb{Q}[X]$$

The ideal $(X) \subset \mathbb{Q}[X]$ is maximal and $\mathbb{Q}[X]/(X) = \mathbb{Q}$. Its inverse image in $\mathbb{Z}[X]$ is also (X) and we get an inclusion

$$\mathbb{Z} = \frac{\mathbb{Z}[X]}{(X)} \hookrightarrow \frac{\mathbb{Q}[X]}{(X)} = \mathbb{Q}$$

but $(X) \subset \mathbb{Z}[X]$ is not a maximal idea. Since the inclusion implies that $\mathbb{Z}[X]/(X)$ is an integral domain, we conclude that (X) is a prime ideal.

Chapter 4, 4.2 Exercise 2 (p. 168) A function satisfying the properties of the exercise has a global square root that is continuous. It follows that the image of f in

$$\frac{\mathfrak{m}_p}{\mathfrak{m}_p^2}$$

is zero.

Chapter 4, 4.2 Exercise 3 (p. 168) This immediately follows from the solution to exercise 1 on page 494, which shows that dx and $d(e^x)$ are linearly independent in $\Omega_{M/\mathbb{C}}$.

Chapter 4, 4.2 Exercise 4 (p. 168) If $\alpha \in \mathbb{C}$ is an algebraic number, let $p(X)$ be its minimal polynomial over $\mathbb{Q}$. After clearing denominators, we get a polynomial $\bar{p}(X)$ (even a primitive polynomial — see definition A.3.4 on page 413). The principle ideal $(\bar{p}(X)) \subset \mathbb{Z}[X]$ is prime and defines a point with evaluation field the field of fractions of $\mathbb{Z}[X]/(\bar{p}(X))$ and the function X restricts to α at this point.

Chapter 4, 4.2 Exercise 5 (p. 168) If x is a unit, it is a unit in every quotient of R, hence its restriction is also a unit and nonzero. If x fails to be a unit, $(x) \neq R$ so that it is contained in a maximal ideal $\mathfrak{m} \subset R$ (by proposition A.1.28 on page 353). It follows that the restriction of x to the (closed) point of Spec R defined by $\mathfrak{m}$ is 0.

Chapter 4, 4.2 Exercise 6 (p. 168) This follows immediately from definition 4.2.1 on page 163 and theorem A.1.47 on page 360, which implies that the intersection of all prime ideals in a ring is the set of nilpotent elements.

Chapter 4, 4.2 Exercise 7 (p. 168) This follows from the basic properties of an ideal. For all prime ideals $\mathfrak{p} \subset R$, $r \in \mathfrak{p}$ implies that $(r) \subset \mathfrak{p}$, which in turn implies that $rs \in \mathfrak{p}$. Taking the contrapositive, we conclude that $rs \notin \mathfrak{p}$ implies that $r \notin \mathfrak{p}$, which is the conclusion.

Chapter 4, 4.2 Exercise 8 (p. 168) This follows from the special properties of a *prime* ideal. The solution to exercise 7 on page 168 implies that $D(rs) \subseteq D(r) \cap D(s)$. To see the reverse inclusion note that for all prime ideals $\mathfrak{p} \subset R$, $rs \in \mathfrak{p} \subset R$, implies that $r \in \mathfrak{p}$ or $s \in \mathfrak{p}$. It follows that $\mathfrak{p} \notin D(rs)$ implies that $\mathfrak{p} \notin D(r)$ or $\mathfrak{p} \notin D(s)$ or $\mathfrak{p} \notin D(r) \cap D(s)$. This implies that $D(r) \cap D(s) \subseteq D(rs)$.

Chapter 4, 4.2 Exercise 9 (p. 168) If r is not a zero-divisor, then $rs \neq 0$ for all $s \in R$ so $D(r) \cap D(s) = D(rs)$. Since rs is also not a zero-divisor, it is not nilpotent so $D(rs) \neq \varnothing$ and the conclusion follows. If r is a zero-divisor, then $r \cdot s = 0$ for some $s \in R$ so $D(r) \cap D(s) = D(rs) = \varnothing$, so $D(r)$ fails to be dense.

Chapter 4, 4.2 Exercise 10 (p. 168) The ring, $k[X,Y]_{(X,Y)}$ has a unique maximal ideal (i.e. it is a local ring) so it only has *one* closed point — given by the ideal (X,Y). All of the other prime ideals of $k[X,Y]$ have been classified in example 2.8.35 on page 112 beside $(X-a, Y-b)$, they are principle ideals $(f(X,Y))$ where $f(X,Y)$ is an irreducible polynomial. After localizing at (X,Y) the primes that survive are the ones that are a subset of (X,Y): these are (X,Y) itself, and any principle ideal $(f(X,Y))$ where f is irreducible and of the form $X \cdot g_1(X,Y) + Y \cdot g_2(X,Y)$. These will all be non-closed points. The ideal (0) is the generic point.

Chapter 4, 4.2 Exercise 11 (p. 168) Since $\mathbb{R}[X]$ is a principle ideal domain, every prime ideal is of the form $(X-a)$, where $a \in \mathbb{R}$ or $(aX^2 + bX + c)$, where $b^2 - 4ac < 0$. Every one of these ideals is maximal, so all points are closed. The ideals of the form $(X-a)$ give rise to all of the points of $\mathbb{R}$. Those of the form $(aX^2 + bX + c)$ give rise to complex-conjugation pairs — so the points defined are in a 1-1 correspondence with the *upper half-plane* in $\mathbb{C}$.

Chapter 4, 4.2 Exercise 12 (p. 168) This is equivalent to requiring each nonempty open set to contain a closed point. It suffices for every open set of the form $D(f) \subset \operatorname{Spec} R$, where f is not *nilpotent* (see exercise 6), to have a closed point in it. This is equivalent to saying that there exists a maximal ideal $\mathfrak{m} \subset R$ such that $f \notin \mathfrak{m}$. The contrapositive of this is that the intersection of all maximal ideals consists of *only* the nilpotent elements of R, i.e,

$$\mathfrak{N}(R) = \mathfrak{J}(R)$$

— see definition A.1.46 on page 360 and definition A.4.21 on page 427 (theorem A.1.47 on page 360 implies that $\mathfrak{N}(R)$ is the intersection of all prime ideals, hence contained in $\mathfrak{J}(R)$).

Chapter 4, 4.2 Exercise 13 (p. 168) Prime ideals of R correspond to irreducible subschemes of $\operatorname{Spec} R$ with the smallest corresponding to the largest subschemes. It follows that the irreducible components of $\operatorname{Spec} R$ correspond to minimal prime ideals and $\operatorname{Spec} R$ is irreducible if and only if it has a single minimal prime ideal. If R is an integral domain, its unique minimal prime ideal is (0) so $\operatorname{Spec} R$ is irreducible (as in the theory of affine varieties). In scheme-theory it is possible for R to have nilpotent elements and *still* be irreducible.

Chapter 4, 4.2 Exercise 14 (p. 168) They are the points corresponding to the images in R of the prime ideals (Y) and (X,Y) — which annihilate Y. These prime ideals correspond to the *whole space* and the *origin,* respectively. The ideal (X) is not prime since $0 \in (X)$ and $Y^2 = 0$.

Chapter 4, 4.2 Exercise 15 (p. 168) *Vanishing* at an associated *point* is equivalent to being *in* an associated *prime* of R. The conclusion follows from corollary A.1.76 on page 375.

Chapter 4, 4.3 Exercise 1 (p. 177) In this case, the ring of fractions $R(X_1, \ldots, X_n)$ does not exist (it vanishes) but we can replace it with

$$R[X_1, \ldots, X_n]_{\mathfrak{N}}$$

where $\mathfrak{N} = \mathfrak{N}(R[X_1,\dots,X_n])$, the nilradical (see definition A.1.46 on page 360). Everything else in example 4.3.15 on page 174 goes through with this substitution. Theorem A.1.47 on page 360 implies that the nilradical is contained in every prime ideal of $R[X_1,\dots,X_n]$, so that

$$R[X_1,\dots,X_n]_{\mathfrak{p}} \subset R[X_1,\dots,X_n]_{\mathfrak{N}}$$

for all prime ideals $\mathfrak{p} \subset R[X_1,\dots,X_n]$.

Chapter 4, 4.3 Exercise 2 (p. 177) In this case,

$$R[X_1,\dots,X_n] = R_1[X_1,\dots,X_n] \oplus R_2[X_1,\dots,X_n]$$

and if $f = f_1 \oplus f_2$ then

$$R[X_1,\dots,X_n]_f = R_1[X_1,\dots,X_n]_{f_1} \oplus R_2[X_1,\dots,X_n]_{f_2}$$

It is not hard to see that the analysis of example 4.3.15 on page 174 goes through after we substitute

$$R_1(X_1,\dots,X_n) \oplus R_2(X_1,\dots,X_n)$$

for $R(X_1,\dots,X_n)$.

Chapter 4, 4.3 Exercise 3 (p. 177) This follows from basic properties of Artinian rings, described in section A.1.7 on page 381. They have a finite number of prime ideals, all of which are maximal — see lemma A.1.84 on page 381. The last statement follows from the proof of theorem A.1.88 on page 383, which shows that R has a *unique* decomposition

$$R = \prod_{i=1}^n R_i$$

where R_i is an Artinian local ring. Uniqueness implies that $R_i = \mathcal{O}_{V,p_i}$.

Chapter 4, 4.3 Exercise 4 (p. 177) This is essentially a tautology:

$$\mathcal{A}(R)(U) = R \otimes_R \mathcal{O}_V(U) = \mathcal{O}_V(U)$$

Chapter 4, 4.3 Exercise 5 (p. 177) The coordinate ring at p is $R_{\mathfrak{p}} = S^{-1}R$ where $S = R \setminus \mathfrak{p}$. This is also equal to

$$\varinjlim R_x$$

as x runs over products of elements of S — see corollary A.5.25 on page 452. But this second formulation is the definition of $\mathcal{O}_{V,p}$ — see definition B.1.1 on page 495.

Chapter 4, 4.3 Exercise 6 (p. 177) If R is a ring and $f \in R$ is a function with support $S \subset \operatorname{Spec} R$, we claim that $C = \operatorname{Spec} R \setminus S$ is open. If $\mathfrak{p} \subset R$ defines a point of C then the image of f in $R_{\mathfrak{p}}$ is 0. But $R_{\mathfrak{p}} = \varinjlim R_x$ as x runs over all products of elements of $R \setminus \mathfrak{p}$ (see corollary A.5.25 on page 452). The definition of direct limit implies that there exists a *specific* such x for which the image of f in R_x is zero. This implies that $f|D(x) = 0$ so that every point of C has an *open neighborhood* that is also in C. It follows that C is open and S is closed.

Chapter 4, 4.3 Exercise 7 (p. 180) We have

$$k[W \times_V U] = k[W] \otimes_{k[V]} k[U]$$

with homomorphisms

$$\iota_1 \colon k[W] \to k[W] \otimes_{k[V]} k[U]$$
$$x \mapsto x \otimes 1$$
$$\iota_2 \colon k[U] \to k[W] \otimes_{k[V]} k[U]$$
$$y \mapsto 1 \otimes y$$

If $\mathfrak{m} \subset k[W] \otimes_{k[V]} k[U]$ is the maximal ideal corresponding to p, then $\mathfrak{m}_1 = \iota_1^{-1}(\mathfrak{m})$ and $\mathfrak{m}_2 = \iota_2^{-1}(\mathfrak{m})$ are the maximal ideals corresponding to p_1 and p_2, respectively (see proposition 2.5.3 on page 72). If $S = k[W] \otimes_{k[V]} k[U] \setminus \mathfrak{m}$, then $\mathcal{O}_{W \times_V U, p} = \left(k[W] \otimes_{k[V]} k[U] \right)_S$ and $\iota_1^{-1}(S) = k[W] \setminus \mathfrak{m}_1$ and $\iota_2^{-1}(S) = k[U] \setminus \mathfrak{m}_2$. It follows that we get a morphism

$$\mathcal{O}_{W,p_1} \otimes_{\mathcal{O}_{V,p_3}} \mathcal{O}_{U,p_2} \to \mathcal{O}_{W \times_V U, p}$$

The target is the result of inverting elements of $k[W \times_V U] \setminus \mathfrak{m}$ that are not derived from the $\mathfrak{m}_i$.

Chapter 4, 4.3 Exercise 8 (p. 181) No. Although our definition of product has the structural properties of a product (as in definition A.5.1 on page 439), it is *not* a Cartesian product (except on the points of affine varieties) and its topology is *not* the product-topology (see exercise 1 on page 38)

Chapter 4, 4.3 Exercise 9 (p. 181) There exists an inverse $\bar{j}: j(X') \to X'$ such that $\bar{j} \circ j = 1: X' \to X'$. The varieties $Y \times_X X'$ and $f^{-1}(j(X'))$ both have the same universal properties: Given any morphisms $u: A \to Y$, $v: A \to X'$ that makes the solid arrows in the diagram

commute, there exists a unique dotted arrow that makes the whole thing commute. The diagram implies that $u(A) \subset f^{-1}(j(X')) \subset Y$, so A also has a unique map to $f^{-1}(j(X'))$. In fact, the image of the projection $p_1: Y \times_X X' \to Y$ must lie within $f^{-1}(j(X'))$.

On the other hand, there exists a morphism

$$f^{-1}(j(X')) \to Y \times_X X'$$

that makes the diagram

commute and the composites of these morphism must be the identity maps.

Chapter 4, 4.4 Exercise 1 (p. 192) The scheme

$$\operatorname{Proj} R[X_0, \ldots, X_n]$$

has a covering by open affines $D(X_i)$, $i = 0, \ldots, n$. Each of these is a copy of $\mathbb{A}^n = \operatorname{Spec} R[Y_1, \ldots, Y_n]$ and proposition 4.4.14 on page 190 shows that, on the overlaps, $D(X_i) \cap D(X_j)$, $Y_t = X_t X_i^{-1}$ on $D(X_i)$ and $Y_t = X_t X_j^{-1}$. It follows that the gluing maps connecting the $D(X_i)$ are compatible with those in definition 4.4.8 on page 186.

Chapter 4, 4.4 Exercise 3 (p. 197) The hypotheses immediately imply that Y is locally closed in X. Since Y is closed in each of the U_α, it is closed in their union — i.e., X. We can patch together the quasi-coherent ideals $\mathfrak{I}_\alpha$ defining $(Y \cap U_\alpha, \mathcal{O}_Y|Y \cap U_\alpha)$ as a

closed subscheme of $(U, \mathcal{O}_X|U)$ to get a global quasi-coherent sheaf, $\mathfrak{I}$ defining $(Y, \mathcal{O}_Y)$ as a subscheme of $(X, \mathcal{O}_X)$.

Chapter 4, 4.4 Exercise 4 (p. 197) On each of the open affines, $\mathbb{A}^n_i$, the sheaf is given by the computations in example 4.3.15 on page 174. An argument like that used in the proof of proposition 4.4.9 shows that

$$\mathcal{O}_{R\mathbb{P}^n}(\mathbb{A}^n_{i_1} \cup \cdots \cup \mathbb{A}^n_{i_t}) = R[X_{j_1}, \ldots, X_{j_{n-t+1}}]$$

where $\{j_1, \ldots, j_{n-t+1}\} = \{0, \ldots, n\} \setminus \{i_1, \ldots, i_t\}$
and

$$\mathcal{O}_{R\mathbb{P}^n}(\mathbb{A}^n_{i_1} \cap \cdots \cap \mathbb{A}^n_{i_t}) = R[X_0, \ldots, X_n][X_{i_1}^{-1}, \ldots, X_{i_t}^{-1}]$$

Chapter 4, 4.4 Exercise 5 (p. 197) Without loss of generality, we can assume $V = \operatorname{Spec} R$, in which case $\mathscr{F} = \mathcal{A}(M)$ for some R-module (see definition 3.5.1 on page 155), M. We have

(E.0.3) $$i_*(\mathscr{F}|U)(U') = \mathscr{F}(U \cap U')$$

where $U' \subset V$ is some open set.

Since $U \subset V$ is an affine open set, then $U = \operatorname{Spec} S^{-1}R$ for some multiplicative set $S \subset R$ and $\mathscr{F}|U = \mathcal{A}(S^{-1}M)$. It is not hard to see that $i_*(\mathscr{F}|U) = \mathcal{A}(S^{-1}M)$, where $S^{-1}M$ is regarded as a module over R, satisfies equation E.0.3 for any affine U'.

Chapter 4, 4.4 Exercise 6 (p. 197) Since f is a closed immersion, it induces a surjection

$$\mathcal{O}_Y \to \mathcal{O}_X$$

which implies surjective maps

$$\mathcal{O}_Y(U) \to \mathcal{O}_X(f^{-1}(U))$$

for all open affines of Y. if we choose an open cover $\{U_i\}$ for Y such that $\mathscr{F}|f^{-1}(U_i) = \mathcal{A}(M_i)$, where M_i is a module over $\mathcal{O}_X(f^{-1}(U_i))$, then $f_*(\mathscr{F})|U_i = \mathcal{A}(M_i)$, where M_i is now regarded as a module over $\mathcal{O}_Y(U)$ by composition with the surjection above. This means that $f_*\mathscr{F}$ is coherent over Y.

Chapter 4, 4.4 Exercise 7 (p. 197) We can use the correspondence between coherent sheaves and modules over the coordinate ring $R = k[X]$ in section C on page 507. An exact sequence of coherent sheaves corresponds to an exact sequence of modules. The Quillen-Suslin theorem (3.5.11 on page 158) shows that all modules over R are free and uniquely defined by their rank. The exact sequence

$$0 \to R^n \to R^{n+1} \to R \to 0$$

shows that $K(X)$ is generated by $[R]$.

Chapter 4, 4.4 Exercise 8 (p. 197) Since R is noetherian and $\mathscr{M}$ is coherent, it follows that $\mathscr{M}$ is finitely generated. Since its generators all vanish on $D(s)$ (because the support is limited to W), lemma 4.4.17 on page 193 implies that there exists an integer n such that $s^n \cdot \mathscr{M} = 0$ — this is the same n as in the statement of the exercise. Now define

$$\mathscr{M}_i = s^{n-i} \cdot \mathscr{M}$$

Chapter 4, 4.5 Exercise 1 (p. 203) (1) If (p) is a point of $\operatorname{Spec} \mathbb{Z}$ with $p \neq 0$, we get $(p) \cdot \mathbb{C} = \mathbb{C}$, so the fiber of (p) is the empty set. The fiber of the generic point (0) is the one point of $\operatorname{Spec} \mathbb{C}$.

(2) The point (p) in $\operatorname{Spec} \mathbb{Z}$ has a fiber in $\mathbb{Z}[X]$ consisting of the point $(p) \cdot \mathbb{Z}[X]$. The fiber of the generic point of $\operatorname{Spec} \mathbb{Z}$ is that of $\operatorname{Spec} \mathbb{Z}[X]$.

Chapter 4, 4.6 Exercise 1 (p. 210) If S is separated, then the diagonal

$$\Delta: S \to S \times S$$

sends closed sets to closed sets and

$$A_1 \cap A_2 \cong \Delta(S) \cap A_1 \times A_2 \subset A_1 \times A_2$$

Since $A_1 \times A_2$ is an affine scheme and $\Delta(S) \cap A_1 \times A_2 \subset A_1 \times A_2$ is a closed subset, lemma 4.4.22 on page 196 implies that it is *also* affine.

Chapter 4, 4.6 Exercise 2 (p. 210) Let N be the union of two copies of $\mathbb{A}^1$ with a "doubled origin" — see example 4.4.7 on page 186. If A_1 and A_2 are these two copies of $\mathbb{A}^1$, then $A_1 \cap A_2 = \mathbb{A}^1 \setminus \{0\}$. This is affine but not open.

If N is a similar union of two copies of $\mathbb{A}^2$ with a doubled origin, then the corresponding $A_1 \cap A_2 = \mathbb{A}^2 \setminus \{(0,0)\}$, which is not even affine (see example D.3.19 on page 555).

Chapter 4, 4.6 Exercise 4 (p. 218) We use the same type of argument as in the remark following definition 4.6.35 on page 216. For every transcendental number $\alpha \in \mathbb{C}$, we can define a morphism

$$
\begin{aligned}
f_\alpha \colon \mathbb{Q}(X) \otimes_{\mathbb{Q}} \mathbb{C} &\rightarrow \mathbb{C} \\
g(X) \otimes z &\mapsto g(\alpha) \cdot z
\end{aligned}
$$

and the kernel will be a maximal ideal, $\mathfrak{m}_\alpha$, defining a closed point. If α is algebraic, there exists a $g(X) \neq 0$ such that $g(\alpha) = 0$ so the map f_α is identically 0 (since $\mathbb{Q}(X)$ is a *field*).

Chapter 5, 5.1 Exercise 1 (p. 225) This follows immediately from lemma 5.1.6 on page 221, which implies that

$$\tau = k\mathbb{P}^n \times \mathbb{A}^{n+1}/\eta$$

Chapter 5, 5.1 Exercise 3 (p. 225) Nonzero *homogeneous* ideals of R are of the form $R \cdot X^n$ for various values of n so they are *all* irrelevant. The only relevant homogeneous ideal is (0), which means that $\operatorname{Proj} R$ has a single point.

Chapter 5, 5.1 Exercise 5 (p. 225) The points in $k\mathbb{P}^{19}$ corresponding to singular hypersurfaces are those satisfying polynomial equations in coordinates (expressing the fact that the tangent space is singular). They constitute a closed subvariety, Δ, of $k\mathbb{P}^{19}$ and $k\mathbb{P}^{19} \setminus \Delta$ parametrizes the smooth hypersurfaces.

Chapter 5, 5.1 Exercise 6 (p. 226) Suppose V is defined by the vanishing of complex-analytic functions $\{f_0, \ldots, f_t\}$. If $B \subset \mathbb{C}^{n+1}$ is a sufficiently small ball centered at 0, then the f_j are equal to their power-series on $V \cap B$, so

$$f_j = \sum_{i=0}^{\infty} f_{i,j}$$

where $f_{i,j}$ is the homogeneous component of degree i — a polynomial — and

$$f_j(tx_0, \ldots, tx_n) = \sum_{j=0}^{\infty} t^i f_{i,j}$$

It follows that the set of zeros of the $\{f_j\}$ is also the set of zeros of the $\{f_{i,j}\}$. But the Hilbert Basis Theorem (A.1.50 on page 362) implies that the ideal generated by them is generated by a *finite* set of homogeneous polynomials.

Chapter 5, 5.1 Exercise 7 (p. 226) Since at least one homogeneous coordinate must be nonzero, it follows that $X_1^2 - X_2^2 = 0$ in $\mathcal{P}(\mathfrak{a})$.

Chapter 5, 5.1 Exercise 9 (p. 226) The statement that f vanishes on $\mathcal{P}(\mathfrak{a})$ is equivalent to saying that

$$f\left(\frac{X_0}{X_j},\ldots,\frac{X_{j-1}}{X_j},\frac{X_{j+1}}{X_j},\ldots,\frac{X_n}{X_j}\right)$$

vanishes on $\mathcal{P}(\mathfrak{a})\cap\mathbb{A}_j^n$ for all j. If $\mathfrak{a}=(g_1,\ldots,g_t)$, the Nullstellensatz implies that

$$f\left(\frac{X_0}{X_j},\ldots,\frac{X_{j-1}}{X_j},\frac{X_{j+1}}{X_j},\ldots,\frac{X_n}{X_j}\right)=\sum_{i=1}^t c_i\cdot g_i\left(\frac{X_0}{X_j},\ldots,\frac{X_{j-1}}{X_j},\frac{X_{j+1}}{X_j},\ldots,\frac{X_n}{X_j}\right)$$

which implies the conclusion after multiplying by a suitable power of X_j.

Chapter 5, 5.1 Exercise 10 (p. 226) Exercise 8 on page 226 implies that $\mathcal{P}(\mathfrak{a})=\mathcal{P}((\mathfrak{a}:\mathfrak{i}^\infty))$, so $(\mathfrak{a}:\mathfrak{i}^\infty)=(\mathfrak{b}:\mathfrak{i}^\infty)$ implies that $\mathcal{P}(\mathfrak{a})=\mathcal{P}(\mathfrak{b})$. Conversely, if $\mathcal{P}((\mathfrak{a}:\mathfrak{i}^\infty))=\mathcal{P}((\mathfrak{b}:\mathfrak{i}^\infty))$, exercise 9 on page 226 implies that an element of $k[X_0,\ldots,X_n]$ that vanishes on $\mathcal{P}((\mathfrak{a}:\mathfrak{i}^\infty))$ must be in $(\mathfrak{a}:\mathfrak{i}^\infty)$ so that $(\mathfrak{b}:\mathfrak{i}^\infty)\subseteq(\mathfrak{a}:\mathfrak{i}^\infty)$. Symmetry implies that they are equal.

Chapter 5, 5.2 Exercise 1 (p. 234) Since $\Lambda^n V=F\cdot e_1\wedge\cdots\wedge e_n\cong F$, we have

$$\Lambda^k V^*\otimes_F\Lambda^n V\cong\Lambda^k V^*$$

Let $\{e_1,\ldots,e_n\}$ be a basis for V and let $\{e^1,\ldots,e^n\}$ be a dual basis for V^*.
    Then an inductive application of lemma 5.2.5 on page 228 implies that

$$(e^{j_1}\wedge\cdots\wedge e^{j_k})\llcorner e_1\wedge\cdots\wedge e_n=\pm e_{i_1}\wedge\cdots\wedge e_{i_{n-k}}$$

where $j_1<\cdots<j_k$, and the set $\{i_1,\ldots,i_{n-k}\}$ with $i_1<\cdots<i_{n-k}$ is the set-difference $\{1,\ldots,n\}\setminus\{j_1,\ldots,j_k\}$. Since the map $\llcorner$ sends basis elements of $\Lambda^k V^*$ to $\pm$basis-elements of $\Lambda^{n-k}V$, it defines an isomorphism.

Chapter 5, 5.2 Exercise 2 (p. 234) If (x,y) is a point on S^1 then

$$x=\frac{X_0^2-X_1^2}{X_0^2+X_1^2},\ y=\frac{2X_0X_1}{X_0^2+X_1^2}$$

so

$$1-x=\frac{2X_1^2}{X_0^2+X_1^2}$$

and

$$\frac{1-x}{y}=\frac{2X_1^2}{2X_0X_1}=\frac{X_1}{X_0}$$

This is well-defined where $y\neq0$ and, on S^1, is equal to

$$\frac{1-x}{y}\cdot\frac{1+x}{1+x}=\frac{1-x^2}{y(1+x)}=\frac{y}{1+x}$$

otherwise — compare to example 2.5.19 on page 82. This function does blow up when $x\to-1$, but we can simply invert it to conclude that the inverse is given by:

$$(X_0:X_1)=g(x,y)=\begin{cases}\left(1:\frac{1-x}{y}\right)&\text{if }y\neq0\\(1:0)&\text{if }y=0,x\neq-1\\(0:1)&\text{if }x=-1\end{cases}$$

is the proper inverse map.

Chapter 5, 5.2 Exercise 3 (p. 234) This set of hyperplanes is $\mathbb{G}_{n-1,n} \cong \mathbb{G}_{1,n} \cong k\mathbb{P}^n$, by exercise 1 on page 234. Alternatively, we could have reasoned that a hyperplane in $k\mathbb{P}^n$ is defined by an equation

$$\sum_{i=0}^{n} a_i X_i = 0$$

and two such equations define the same hyperplane if and only if their coefficients are proportional — i.e., if they define the same point of $k\mathbb{P}^n$.

Chapter 5, 5.2 Exercise 4 (p. 234) A line in $k\mathbb{P}^3$ is a plane in $\mathbb{A}^4$, so it is parametrized by $\mathbb{G}_{2,4}$.

Chapter 5, 5.3 Exercise 1 (p. 241) In the notation of definition 5.3.6 on page 237, there exists a surjection of presheaves

$$\bigoplus_{i \in I} \mathcal{O}_X \to \mathscr{F}$$

which induces a surjection of stalks. That implies that the map to the *completion* of $\mathscr{F}$

$$\bigoplus_{i \in I} \mathcal{O}_X \to \widehat{\mathscr{F}}$$

is surjective as well (see exercise 6 on page 502). If $x = \sum s_i \cdot f_i \in \widehat{\mathscr{F}}$, then this sum is also well-defined in $\mathscr{F}$ so the inclusion of presheaves

$$\mathscr{F} \to \widehat{\mathscr{F}}$$

must be *surjective* as well as injective.

Chapter 5, 5.3 Exercise 2 (p. 242) The existence of $\mathscr{L}$ implies the existence of a closed immersion

$$f \colon V \to k\mathbb{P}^t$$

for some value of t. Since $\mathscr{L} = f^* \mathcal{O}_{k\mathbb{P}^t}(1)$, the standard elements $X_i \in \mathcal{O}_{k\mathbb{P}^t}(k\mathbb{P}^t)$ (see proposition 5.3.3 on page 235) pull back to functions f_i that equal the values of the X_i on the image of f. The conclusion follows.

Chapter 5, 5.4 Exercise 1 (p. 244) Suppose L be defined by the equation $\sum_{j=0}^{n} a_j X_j = 0$. Let M be an $(n+1) \times (n+1)$ invertible matrix with the property that

$$M \cdot \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Then M induces an *isomorphism*

$$\mathbb{P}(M) \colon \mathbb{P}(\mathbb{A}^{n+1}) \to \mathbb{P}(\mathbb{A}^{n+1})$$

whose restriction to $\mathbb{P}(\mathbb{A}^{n+1}) \setminus L$ is an isomorphism

$$\mathbb{P}(M)|\mathbb{P}(\mathbb{A}^{n+1}) \setminus L \colon \mathbb{P}(\mathbb{A}^{n+1}) \setminus L \to \mathbb{P}(\mathbb{A}^{n+1}) \setminus \mathbb{A}_0^{n+1} = \mathbb{A}^n$$

Chapter 5, 5.4 Exercise 3 (p. 245) Compare to exercise 3 on page 234. If we map

$$v_{n,d} \colon k\mathbb{P}^n \to k\mathbb{P}^M$$

via the degree-d Veronese map, the hyper*planes* in $k\mathbb{P}^M$ correspond to hyper*surfaces* in $k\mathbb{P}^n$. The dimension of this space is $M = \binom{n+d}{n} - 1$.

Chapter 5, 5.5 Exercise 1 (p. 248) The identity $X_0X_3 - X_1X_2 = 0$ is precisely the Segre relations in equation 5.5.2 on page 247 for the map

$$k\mathbb{P}^1 \times k\mathbb{P}^1 \to k\mathbb{P}^3$$

with

$$
\begin{aligned}
w_{0,0} &= X_0 \\
w_{1,1} &= X_3 \\
w_{0,1} &= X_1 \\
w_{1,0} &= X_2
\end{aligned}
$$

To see that it is birationally equivalent to $k\mathbb{P}^2$, consider the open $V \cap \mathbb{A}_0^3$ defined by $X_0 = 1$. The intersection is $\mathcal{V}\left((X_3 - X_1X_2)\right) \subset \mathbb{A}^3$ which is easily seen to be isomorphic to $\mathbb{A}^2$. It follows that V has an open affine *isomorphic* to an open affine of $k\mathbb{P}^2$ so they are birationally equivalent.

Chapter 5, 5.5 Exercise 2 (p. 254) The Zariski closure of a set S is the zero-set of *all* algebraic equations that vanish on S. If the f_i satisfy an algebraic relation like

$$\sum_{i=1}^{t} a_i(X) \cdot f_i = 0$$

we have to augment equations 5.5.4 on page 250 with

$$\sum_{i=1}^{t} a_i(X) \cdot Y_i = 0$$

Chapter 5, 5.5 Exercise 3 (p. 254) Since X_1^2 and X_2 are algebraically independent, we can use equations 5.5.4 on page 250 to see that the blowup in question is given by

$$X_1^2 Y_2 = X_2 Y_1$$

where $[Y_1{:}Y_2]$ are the homogeneous coordinates of $k\mathbb{P}^1$. If we consider the chart with $Y_2 = 1$, we get $X_1^2 = X_2 Y_1$ which has a Jacobian that vanishes at $X_1 = X_2 = Y_1 = 0$.

Chapter 5, 5.5 Exercise 4 (p. 254) This follows immediately from equation 5.5.5 on page 250 — carefully compute the grades or levels of the graded ring

$$G = \frac{k[X_1, \ldots, X_n, Y_1, \ldots, Y_t]}{(\{f_i(X) \cdot Y_j - f_j(X) \cdot Y_i\})}$$

where the X_i have degree 0 and the Y_i have degree 1. Clearly, $G_0 = R$ and

$$G_n = \frac{\sum_{1 \le i_1, \ldots, < i_n \le t} R \cdot Y_{i_1} \cdots Y_{i_n}}{(\{f_i(X) \cdot Y_j - f_j(X) \cdot Y_i\})}$$

and we define surjective homomorphisms

$$G_n \to \sum_{1 \le i_1, \ldots, < i_n \le t} R \cdot f_{i_1} \cdots f_{i_n} = \mathfrak{a}^n \qquad \text{regarded as degree } n$$

$$Y_i \mapsto f_i$$

which are also *injective* since the f_i are algebraically independent.

If the f_i are not independent, then the ideal $(\{f_i(X) \cdot Y_j - f_j(X) \cdot Y_i\})$ must be augmented with all of the relations they satisfy (as in the solution of exercise 2 on page 254), so the corresponding map is *still* an isomorphism.

Chapter 5, 5.5 Exercise 5 (p. 254) *Existence* follows immediately from definitions 5.5.8 on page 252 and 5.5.10 on page 253. *Uniqueness* follows from the fact the maps σ_1 and σ_2 are *isomorphisms* away from the exceptional fiber.

Chapter 5, 5.5 Exercise 6 (p. 254) This follows from exercise 5 on page 254 and the fact that $\mathbb{B}_{f^*(\mathfrak{b})}(V) = V$, since $f^*(\mathfrak{b})$ is a principal ideal. This is a form of a "universal property" that blowups have.

Chapter 5, 5.6 Exercise 1 (p. 266) Simply embed V in $k\mathbb{P}^3$ via theorem 5.6.9 on page 265. Since it is of codimension 2, it is defined by two forms.

Chapter 5, 5.6 Exercise 2 (p. 266) Couple proposition 5.6.1 on page 261 with the Veronese embedding (see section 5.4.2 on page 244).

Chapter 5, 5.7 Exercise 1 (p. 273) The answer is

$$
\begin{aligned}
\mathcal{H}_{R/\mathfrak{J}}(s) &= \binom{s+3}{3} - \binom{s+3-d_1}{3} - \binom{s+3-d_2}{3} + \binom{s+3-d_1-d_2}{3} \\
&= s \cdot d_1 d_2 - \frac{d_1^2 d_2 + d_1 d_2^2}{2} + 2 d_1 d_2
\end{aligned}
$$

Chapter 5, 5.7 Exercise 2 (p. 273) If $\deg f = r$, then $\mathfrak{a} = \mathfrak{a} + (f)$ in degrees $\geq r + t$ so the Hilbert functions agree in higher degrees and the Hilbert polynomials are the same.

Chapter 5, 5.7 Exercise 3 (p. 273) This follows from exercise 10 on page 226, which implies that $\bar{\mathfrak{a}} = \bar{\mathfrak{b}}$ and exercise 2 on page 273, which implies that $\mathcal{H}_{R/\mathfrak{a}}(s) = \mathcal{H}_{R/\bar{\mathfrak{a}}}(s)$.

Chapter 5, 5.9 Exercise 1 (p. 294) This follows from the definition of sheaves and their quotients: for each open $U \subset V$, the quotient

$$
\frac{\mathcal{K}_V^*(U)}{\mathcal{O}_V^*(U)}
$$

is the sheaf of local Cartier divisors. Ones that fit together *globally* define Cartier divisors — and elements of the quotient-sheaf

$$
\left(\frac{\mathcal{K}_V^*}{\mathcal{O}_V^*} \right)(V)
$$

The ring $\mathcal{K}_V^*(V)$ is globally defined rational functions, hence principal divisors.

Chapter 5, 5.9 Exercise 2 (p. 295)   (1) If $f(x) = x$, then $\operatorname{ord}_0(f) = 1$ and $\operatorname{ord}_f(\infty) = -1$ (i.e., it has a pole at ∞), so $(f) = \langle 0 \rangle - \langle \infty \rangle$.
    (2) In this case, $\operatorname{ord}_1(f) = 1$ and $\operatorname{ord}_0(f) = -1$ so $(f) = \langle 1 \rangle - \langle 0 \rangle$.
    (3) $\operatorname{ord}_0(f) = -2$ and $\operatorname{ord}_\infty(f) = 2$, so $(f) = 2 \cdot \langle \infty \rangle - 2 \cdot \langle 0 \rangle$
    (4) $(f) = \langle 1 \rangle + \langle 4 \rangle - 2 \cdot \langle \infty \rangle$

Chapter 5, 5.9 Exercise 3 (p. 295)

Chapter 5, 5.9 Exercise 4 (p. 295) This follows from the fact that $(x) = (x)_0 - (x)_\infty$.

Chapter 5, 5.9 Exercise 5 (p. 295) The canonical projection

$$
f: W \to V
$$

is an isomorphism outside the center of the blow-up. Since this center has codimension ≥ 2 the center contributes nothing to $\mathrm{Cl}(V)$. The exceptional fiber is a projective space, which contributes the $\mathbb{Z}$-summand.

Chapter 5, 5.9 Exercise 6 (p. 299) Since $f \in \mathcal{L}(D)$, we have

$$
\operatorname{ord}_Q(f) \geq -\operatorname{ord}_Q(D)
$$

for all prime divisors, Q. If P is a prime divisor that does *not* appear in D (which is a finite linear combination of prime divisors) and with the property that $\operatorname{ord}_P(f) = 0$, then

$$
\operatorname{ord}_P(f) = 0 \not\geq \operatorname{ord}_P(D - P) = 1
$$

Chapter 6, 6.1 Exercise 1 (p. 313) According to corollary 6.1.17 on page 311, there exists a curve L such that

$$V \bullet L = V \bullet W - V \bullet X$$

Bézout's Theorem implies that L must be a *linear* curve.

Chapter 6, 6.1 Exercise 2 (p. 313) Let ℓ_1, ℓ_2, ℓ_3 denote three non-adjacent sides of the hexagon and m_1, m_2, m_3 denote the other three. Then

$$
\begin{aligned}
F &= \ell_1 \cdot \ell_2 \cdot \ell_3 \\
G &= m_1 \cdot m_2 \cdot m_3
\end{aligned}
$$

are two cubic functions such that $\mathcal{P}(F) \bullet \mathcal{P}(G) = \sum_{i=1}^{9} p_i$. The conclusion follows immediately from exercise 1 on page 313 above.

Chapter 6, 6.1 Exercise 3 (p. 313) Let ℓ_1, ℓ_2, ℓ_3 denote three non-adjacent sides of the hexagon and m_1, m_2, m_3 denote the other three. Then

$$
\begin{aligned}
F &= \ell_1 \cdot \ell_2 \cdot \ell_3 \\
G &= m_1 \cdot m_2 \cdot m_3
\end{aligned}
$$

are two cubic functions such that $\mathcal{P}(F) \bullet \mathcal{P}(G) = \sum_{i=1}^{9} p_i$. Suppose p_7, p_8, p_9 are the intersections of the extended sides of the hexagon. Since they lie on a line ℓ, it follows that $\mathcal{P}(F) \bullet \mathcal{P}(\ell) = p_7 + p_8 + p_9$. Corollary 6.1.17 on page 311 implies that there exists a curve $C = \mathcal{P}(H)$ such that

$$\mathcal{P}(F) \bullet C = \mathcal{P}(F) \bullet \mathcal{P}(G) - \mathcal{P}(F) \bullet \mathcal{P}(\ell)$$

Bézout's Theorem implies that C must be of degree 2.

Chapter 6, 6.1 Exercise 4 (p. 313) By proposition 6.1.2 on page 305, it must be 1.

Chapter 6, 6.1 Exercise 5 (p. 313) By proposition 6.1.2 on page 305, it must be defined by a cubic equation.

Chapter 6, 6.2 Exercise 2 (p. 321) We compute the order of

$$\frac{1}{x-1} = \frac{x_3}{x_1 - x_3}$$

which vanishes at $\odot$. Multiply by y/x to get

$$\frac{y}{x^2 - x} = \frac{x_2 x_3^2}{x_1^2 - x_1 x_3}$$

which still vanishes at $\odot$. Now do this a second time to get

$$\frac{y^2}{x^3 - x^2} = \frac{x^3 + ax + b}{x^3 - x^2} = \frac{x_1^3 + ax_1 x_3^2 + bx_3^3}{x_1^3 - x_1^2 x_3}$$

which has a value of 1 at $\odot$. It follows that

$$\frac{1}{x-1} = \left(\frac{x}{y}\right)^2 \cdot \frac{x^3 + ax + b}{x^3 - x^2}$$

A similar argument shows that

$$\frac{1}{y} = \left(\frac{x}{y}\right)^3 \cdot \frac{x^3 + ax + b}{x^3}$$

Chapter 6, 6.2 Exercise 3 (p. 321) This is an example of where the strict correspondence between linear functions and the lines they define breaks down to some extent.

The function $x - a$ defines a line in the finite portion of V and the *function* "blows up" as we approach $\odot$, so it has a valuation of -2. On the other hand, the *line* in $k\mathbb{P}^2$ it defines is given by

$$x_1 - ax_3$$

in homogeneous coordinates and clearly *vanishes* at $\circledcirc = (0\!:\!1\!:\!0)$ so it *passes through* this point, restoring our faith in Bézout's theorem.

Chapter 6, 6.2 Exercise 4 (p. 322) We consider points of degree 2 first: $\eta_1 + \eta_2 = \eta_3$.

If $a_1 = a_2$ and the points are distinct, then $b_1 = -b_2$ and the sum is 0. If the points are the same, then the sum is $2 \cdot (a_1, b_1)$ and that is given by the proof of 6.2.8 on page 318:

$$a_3 = \frac{a_1^4 - 2a_1^2 A - 8a_1 B + A^2}{4(a_1^3 + Aa_1 + B)}$$

$$b_3 = \frac{3a_1^2 + A}{2b_1}(a_3 - a_1) - b_1$$

If $a_1 \neq a_2$, the line passing through (a_1, b_1) and (a_2, b_2) is given by

$$y = t \cdot (x - a_1) + b_1$$

where

$$t = \frac{b_2 - b_1}{a_2 - a_1}$$

To find the third intersection of this line with V, plug the equation for the line into that for V to get

$$(t \cdot (x - a_1) + b_1)^2 = x^3 + Ax + B$$
$$t^2(x - a_1)^2 + 2t(x - a_1)b_1 + b_1^2 = x^3 + Ax + B$$

or

$$x^3 + Ax + b - t^2(x - a_1)^2 - 2t(x - a_1)b_1 - b_1^2 = 0$$

We divide this by $x - a_1$ and ignore the remainder (since we know it evaluates to 0) to get

$$x^2 + a_1 x + a_1^2 + A - t^2(x - a_1) - 2t$$

Now we divide by $x - a_2$ (and throw out the remainder) to get

$$x + a_1 + a_2 - t^2 = 0$$

so

$$a_3 = t^2 - a_1 - a_2$$
$$y = t(a_3 - a_1) + b_1$$
$$= t^3 - 2a_1 t - a_2 t + b_1$$

This is the negative of the answer so we negate y to get

$$b_3 = t(2a_1 + a_2) - b_1 - t^3$$

Chapter 6, 6.2 Exercise 5 (p. 322) If A is an abelian variety, consider the conjugation-map

$$c: A \times A \rightarrow A$$
$$f \times g \mapsto f \cdot g \cdot f^{-1}$$

This is a surjective regular map, therefore a family of maps in the sense of definition 5.5.18 on page 257 with $c_1(A) = 1$ (i.e., set $g = 1$). Lemma 5.5.19 on page 258 implies that $c_g(A) = 1$ for all $g \in A$ so A must be commutative.

Chapter 6, 6.2 Exercise 6 (p. 323) Simply *double x*, k times — using proposition 6.2.8 on page 318.

Chapter 6, 6.2 Exercise 7 (p. 323) Write n as a sum of powers of 2 — i.e., consider its binary representation:

$$n = \sum_{j=1}^{k} 2^{n_j}$$

Now it is easy to compute $2^{n_j} \cdot x$ using the algorithm in the previous problem and then *add up* the results using the solution to exercise 4 on page 322.

Chapter 6, 6.4 Exercise 1 (p. 333) This follows immediately from Serre Duality, which says that

$$H^1(V, \mathcal{O}_V) = \left(H^0(V, K_V \otimes_{\mathcal{O}_V} \mathcal{O}_V^\vee) \right)^* = H^0(V, K_V)^*$$

Chapter 6, 6.5 Exercise 1 (p. 337) The *cokernel* of df is a highly *incoherent* sheaf — see equation 3.5.1 on page 156 — so it does not correspond to any vector-bundle over V.

Chapter 6, 6.5 Exercise 2 (p. 337) This is because Ω_V^{-1} differs from Ω_V and $\mathcal{O}_V$ only in its transition-functions. Since R is a finite set of disconnected points, the transition-functions never come into play.

Chapter 6, 6.6 Exercise 1 (p. 339) We first expand $X(X-1)(X-\lambda)$ to get

$$X^3 - (\lambda + 1)X^2 + \lambda X$$

Now the substitution $X = U + (\lambda + 1)/3$ gives

$$U^3 + U(\lambda - \lambda^2 - 1)/3 + \lambda^2/9 - 2\lambda^3/27 + \lambda/9 - 2/27$$

from which we conclude $A = (\lambda - \lambda^2 - 1)/3$ and $B = \lambda^2/9 - 2\lambda^3/27 + \lambda/9 - 2/27$, and

$$4A^3 + 27B^2 = -\lambda^2(\lambda - 1)^2$$

which gives

$$j = 256 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}$$

Chapter A, A.1 Exercise 1 (p. 348) Just do this

$$
\begin{aligned}
s_1 + (r + s_1) &= & s_1 + (r + s_2) \\
= (s_1 + r) + s_1 &= & (s_1 + r) + s_2 \\
= 0 + s_1 &= & 0 + s_2 \\
= s_1 &= & s_2
\end{aligned}
$$

Chapter A, A.1 Exercise 2 (p. 348) They are ± 1.

Chapter A, A.1 Exercise 3 (p. 348) They are numbers k such that $k \cdot \ell \equiv 1 \pmod{m}$ for some ℓ such that $0 < \ell < m$, or

$$k \cdot \ell = 1 + n \cdot m$$

or

$$k \cdot \ell - n \cdot m = 1$$

The proof of lemma A.1.12 on page 345 implies that the smallest positive value attained by linear combinations of ℓ and m is their greatest common divisor — in this case, 1. It follows that an integer $0 < k < m$ is a unit in $\mathbb{Z}_m$ if and only if it is relatively prime to m.

Chapter A, A.1 Exercise 4 (p. 348) We compute $123 = 27q_1 + r_1$ with

$$
\begin{aligned}
q_1 &= & 4 \\
r_1 &= & 15
\end{aligned}
$$

Now $27 = 15q_2 + r_2$ with

$$
\begin{aligned}
q_2 &= & 1 \\
r_2 &= & 12
\end{aligned}
$$

In stage 3, $15 = 12q_3 + r_3$ with

$$
\begin{aligned}
q_3 &= 1 \\
r_3 &= 3
\end{aligned}
$$

The process terminates at this point so $\gcd(27, 123) = 3$. Now we apply A.1.16 on page 347 to calculate a and b: $x_0 = 0$, $y_0 = 1$, $x_1 = 1$, $y_1 = -q_1 = -4$ and

$$
\begin{aligned}
x_2 &= x_0 - q_2 x_1 = -1 \\
y_2 &= y_0 - q_2 y_1 = 5 \\
x_3 &= x_1 - q_3 x_2 = 2 \\
y_3 &= y_1 - q_3 y_2 = -9
\end{aligned}
$$

So

$$
3 = 2 \cdot 123 - 9 \cdot 27
$$

Chapter A, A.1 Exercise 5 (p. 348) This follows immediately from the fact that integers uniquely factor into primes. Suppose

$$
x = \frac{a}{b} = \frac{\prod p_i^{\alpha_i}}{\prod q_j^{\beta_j}}
$$

is an expression where the p_i, q_j are primes and no p_i is equal to any q_j. Suppose x^x is rational, i.e.

$$
x^x = \frac{\prod r_k^{\gamma_k}}{\prod s_\ell^{\delta_\ell}}
$$

where the r_k and s_ℓ are also disjoint sets of primes. Then we get

$$
x^x = \left(\frac{a}{b}\right)^{\frac{a}{b}} = \frac{\prod r_k^{\gamma_k}}{\prod s_\ell^{\delta_\ell}}
$$

or

$$
a^a \cdot \prod s_\ell^{b\delta_\ell} = b^a \cdot \prod r_k^{b\gamma_k}
$$

or

$$
\prod p_i^{a \cdot \alpha_i} \cdot \prod s_\ell^{b\delta_\ell} = \prod q_j^{a \cdot \beta_j} \cdot \prod r_k^{b\gamma_k}
$$

This is a contradiction because the set of primes $\{p_i, s_\ell\}$ is disjoint from $\{q_j, r_k\}$.

Chapter A, A.1 Exercise 6 (p. 353) The ideal $\mathfrak{g} \subset A$ will be a sub-vector-space. If it includes $k \cdot 1$, then $\mathfrak{g} = A$. Otherwise, the quotient will still be a vector space over k with a unit, $k \cdot 1$.

Chapter A, A.1 Exercise 7 (p. 353) If $r \in R$ and $x \in f^{-1}(\mathfrak{a})$ then $r \cdot x$ maps to $f(r) \cdot f(x) \in \mathfrak{a}$ since $f(x) \in \mathfrak{a}$. It follows that $r \cdot x \in f^{-1}(\mathfrak{a})$. If we compose f with the projection

$$
S \to S/\mathfrak{a}
$$

to get

$$
R \xrightarrow{f} S \to S/\mathfrak{a}
$$

the kernel of this map is precisely $f^{-1}(\mathfrak{a})$, so we get a well-defined map

$$
\frac{R}{f^{-1}(\mathfrak{a})} \to \frac{S}{\mathfrak{a}}
$$

It follows that $x \in R$ maps to zero in $S/\mathfrak{a}$ if and only if it maps to zero in $R/f^{-1}(\mathfrak{a})$.

Chapter A, A.1 Exercise 8 (p. 353) Since $(x, y) = R$ it follows that there exist $a, b \in R$ such that

$$ax + by = 1$$

or $ax = 1$ in $R/(b)$. This means that $a^n x^n = 1$ in $R/(b)$ so that $(x^n, y) = R$. A similar argument with the image of b in $R/(a)$ implies the conclusion.

Chapter A, A.1 Exercise 9 (p. 353) Suppose α has a multiplicative inverse

$$\beta = \sum_{j=0}^{\infty} b_j X^j$$

Then the product is

$$\alpha \cdot \beta = \sum_{n=0}^{\infty} c_n X^n$$

where $c_0 = a_0 \cdot b_0 = 1$.

Chapter A, A.1 Exercise 11 (p. 354) If $\mathfrak{a} \not\subset \mathfrak{p}$ and $\mathfrak{b} \not\subset \mathfrak{p}$ then there exists $x \in \mathfrak{a}$ with $x \notin \mathfrak{p}$ and $y \in \mathfrak{b}$ with $y \notin \mathfrak{p}$. The product, xy, will be in $\mathfrak{a} \cdot \mathfrak{b}$ so $xy \in \mathfrak{p}$. This contradicts the definition of a prime ideal (see definition A.1.19 on page 349).

Chapter A, A.1 Exercise 12 (p. 354) Suppose $x \cdot y \in \mathfrak{p}$ but $x \notin \mathfrak{p}$. Then there exists an integer n such that, for $i > n$, $x \notin \mathfrak{p}_i$. The fact that $\mathfrak{p}_i$ is prime implies that $y \in \mathfrak{p}_i$ for $i > n$, so $y \in \mathfrak{p}$.

Chapter A, A.1 Exercise 13 (p. 354) The ideal (X) consists of polynomials such that the X-degree of every monomial is ≥ 1. If $a(X) \cdot b(X) \in (X)$, each monomial of $a(X) \cdot b(X)$ must have X-degree ≥ 1. If $a(X)$ and $b(X)$ both contain a monomial of X-degree 0, the product of those monomials will also have X-degree zero and $a(X) \cdot b(X) \notin (X)$.

This ideal is not maximal because it is contained in the proper ideal (X, Y).

Chapter A, A.1 Exercise 14 (p. 354) This map clearly preserves addition. It remains to show that $f(x \cdot y) = f(x) \cdot f(y)$ for all $x, y \in \mathbb{Q}[\sqrt{2}]$. If

$$\begin{aligned} x &= a + b\sqrt{2} \\ y &= c + d\sqrt{2} \end{aligned}$$

are two elements, then

$$xy = ac + 2bd + (ad + bc)\sqrt{2}$$

and

$$\begin{aligned} f(x) &= a - b\sqrt{2} \\ f(y) &= c - d\sqrt{2} \end{aligned}$$

so

$$f(x) \cdot f(y) = ac + 2bd - (ad + bc)\sqrt{2} = f(x \cdot y)$$

Chapter A, A.1 Exercise 15 (p. 354) If

$$\begin{aligned} x &= a + b\sqrt{2} \\ y &= c + d\sqrt{2} \end{aligned}$$

are two elements of $\mathbb{Q}[\sqrt{2}]$, then

$$xy = ac + 2bd + (ad + bc)\sqrt{2}$$

If we set $c = a$ and $d = -b$, so $y = a - b\sqrt{2}$, then we get

$$xy = a^2 - 2b^2 \in \mathbb{Q}$$

and the $\sqrt{2}$ term is zero. It follows that

$$(a + b\sqrt{2})^{-1} = \frac{a - b\sqrt{2}}{a^2 - 2b^2}$$

If $a + b\sqrt{2} \neq 0$, denominator is nonzero since $\sqrt{2}$ is irrational.

Chapter A, A.1 Exercise 16 (p. 354) If r is *not* a unit, $(r) \subset R$ is a proper ideal and there exists a *maximal* ideal $\mathfrak{m}$ such that $(r) \subset \mathfrak{m}$. But $r - 1 \in \mathfrak{J} \subset \mathfrak{m}$ (since $\mathfrak{J}$ is the intersection of all maximal ideals), so r and $r - 1$ are *both* contained in $\mathfrak{m}$. This implies that $r - (r - 1) = 1 \in \mathfrak{m}$, which contradicts the fact that $\mathfrak{m}$ is a proper ideal of R.

Chapter A, A.1 Exercise 17 (p. 354) Without loss of generality assume $i = 1$ and write

$$R = (\mathfrak{a}_1 + \mathfrak{a}_2)(\mathfrak{a}_1 + \mathfrak{a}_3) \cdots (\mathfrak{a}_1 + \mathfrak{a}_n)$$

When carrying out the multiplication, all but one term in the product has a factor of $\mathfrak{a}_1$, hence is *contained* in $\mathfrak{a}_1$ (by the defining property of an ideal — see definition A.1.19 on page 349). The *one* exception is the term $\mathfrak{a}_1 + \prod_{j=2}^{n} \mathfrak{a}_j$ — and this contains $\mathfrak{a}_1$ and *all* of the other terms. It follows that

$$R = (\mathfrak{a}_1 + \mathfrak{a}_2) \cdots (\mathfrak{a}_1 + \mathfrak{a}_n) \subseteq \mathfrak{a}_1 + \prod_{j=2}^{n} \mathfrak{a}_j$$

so

$$\mathfrak{a}_1 + \prod_{j=2}^{n} \mathfrak{a}_j = R$$

Chapter A, A.1 Exercise 18 (p. 354) We get a natural map

$$(E.0.4) \qquad\qquad R \to \prod_{i=1}^{n} \frac{R}{\mathfrak{a}_i}$$

that sends $x \in R$ to $(p_1(x), \ldots, p_n(x))$, where p_i is the natural projection

$$p_i \colon R \to \frac{R}{\mathfrak{a}_i}$$

The kernel of the map in E.0.4 is clearly $\mathfrak{a}$. It only remains to show that this map is *surjective.* Use the solution to exercise 17 on page 354 to conclude that

$$\mathfrak{a}_i + \prod_{j \neq i} \mathfrak{a}_j = R$$

for all i. This means that, for each i, there is an element $u_i \in \mathfrak{a}_i$ and $v_i \in \prod_{j \neq i} \mathfrak{a}_j$ such that $u_i + v_i = 1$. It is not hard to see that

$$
\begin{aligned}
p_i(v_i) &= 1 \\
p_j(v_i) &= 0
\end{aligned}
$$

for any $j \neq i$. If

$$(x_1, \ldots, x_n) \in \prod_{i=1}^{n} \frac{R}{\mathfrak{a}_i}$$

is an arbitrary element, set

$$x = \sum_{i=1}^{n} x_i v_i$$

Then $p_i(x) = p_i(v_i x_i) = x_i$ so the map is surjective.

Chapter A, A.1 Exercise 19 (p. 358) We work in the ring $\mathbb{F}[x]$. Definition A.1.31 on page 356 and example A.1.32 on page 356 implies that $\alpha^n = 1$ in $\mathbb{F}$ if and only if $x - \alpha | (x^n - 1)$. Each such $x - \alpha$ is an irreducible factor of $x^n - 1$ and we get

$$x^n - 1 = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_k) p(x)$$

where $p(x)$ is a product of the other irreducible factors. Corollary A.1.35 on page 357 implies that this factorization is *unique,* so $k \leq n$.

Chapter A, A.1 Exercise 20 (p. 358) Two functions $g_1, g_2 \in C[0,1]$ map to the same element of the quotient $C[0,1]/\mathfrak{f}_a$ if and only if $g_1(a) = g_2(a)$. It follows that $C[0,1]/\mathfrak{f}_a \cong \mathbb{R}$. Since this is a field, lemma A.1.30 on page 355 implies that $\mathfrak{f}_a$ must be maximal.

Chapter A, A.1 Exercise 21 (p. 358) We start by dividing the larger polynomial by the smaller one to get

$$
\begin{aligned}
b(X) &= q_1(X) \cdot a(X) + r_1(X) \\
q_1(X) &= X - 3 \\
r_1(X) &= 10X^3 - 7X^2 + 3X + 8
\end{aligned}
$$

Now we compute

$$
\begin{aligned}
a(X) &= q_2(X) \cdot r_1(X) + r_2(X) \\
q_2(X) &= \frac{1}{10}X^2 - \frac{7}{100}X - \frac{81}{1000} \\
r_2(X) &= -\frac{1577}{1000}X^2 + \frac{683}{1000}X - \frac{706}{125}
\end{aligned}
$$

Now we divide $r_1(X)$ by $r_2(X)$ to get

$$
\begin{aligned}
r_1(X) &= q_3(X) \cdot r_2(X) + r_3(X) \\
q_3(X) &= \frac{10000}{1577}X + \frac{4209000}{2486929} \\
r_3(X) &= \frac{93655000}{2486929}X - \frac{3877000}{2486929}
\end{aligned}
$$

We finally divide $r_2(X)$ by $r_3(X)$ to get

$$
\begin{aligned}
r_2(X) &= q_4(X)r_3(X) + r_4(X) \\
q_4(X) &= -\frac{3921887033}{93655000000}X + \frac{8992165757259}{548203689062500} \\
r_4(X) &= \frac{6220545559984}{1096407378125}
\end{aligned}
$$

Since this is a *unit* of $\mathbb{Q}[X]$, it shows that $a(X)$ and $b(X)$ are relatively prime.

Chapter A, A.1 Exercise 22 (p. 358) The monomials X^5 and X^6 have no gcd. Their divisors are 1, X^2, and X^3 — none of which is divisible *in R* by the other two.

Chapter A, A.1 Exercise 23 (p. 363) This is just the pigeonhole Principle: Suppose R is an integral domain with n elements and $x \in R$ is nonzero. Multiply all of the nonzero elements of R by x:

$$\{x \cdot y_1, \ldots, x \cdot y_{n-1}\}$$

We claim that these products must all be distinct. If $x \cdot y_i = x \cdot y_j$ then $x \cdot (y_i - y_j) = 0$ and the only way this can happen in an integral domain is for $y_i = y_j$. It follows that 1 must be in this set of products, so $1 = x \cdot y_k$ for some k and $y_k = x^{-1}$.

Chapter A, A.1 Exercise 24 (p. 363) Suppose every increasing sequence of ideals is eventually constant and

$$\mathfrak{a} = (x_1, \ldots) \subset R$$

is some ideal. Then we have the following increasing sequence of ideals

$$(x_1) \subset (x_1, x_2) \subset \cdots \subset \mathfrak{a}$$

and the ascending chain condition implies that, for some finite n,

$$(x_1, \ldots, x_i) = (x_1, \ldots, x_{i+1})$$

for all $i \geq n$. So $\mathfrak{a} = (x_1, \ldots, x_n)$ is finitely generated.

Conversely, suppose all ideals of R are finitely generated and

$$\mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$$

is an ascending sequence of ideals in R. Then

$$\mathfrak{a} = \bigcup_{i=1}^{\infty} \mathfrak{a}_i$$

is also an ideal in R and $\mathfrak{a} = (x_1, \ldots, x_n)$. Each of the x_i must be contained in one of the $\mathfrak{a}_{j_i}$ and all ideals in the sequence following it. If $m = \max(j_i)$ then $\mathfrak{a}_m = \mathfrak{a}$ and the sequence becomes constant.

Chapter A, A.1 Exercise 25 (p. 363) This follows immediately from the Ascending Chain Condition and lemma A.1.25 on page 352.

Chapter A, A.1 Exercise 26 (p. 363) The ring $\mathbb{Q}[X, Y]$ is certainly an integral domain. So see that it is not Euclidean note that the two variables X and Y have no common divisors other than 1.

If $\mathbb{Q}[X, Y]$ was a Euclidean ring, it would be possible to find polynomials $a(X, Y)$ and $b(X, Y)$ such that

$$1 = a(X, Y) \cdot X + b(X, Y) \cdot Y$$

This is impossible since we could make the right side of the equation equal to 0 by setting $X = 0$ and $Y = 0$, so we would get

$$1 = 0$$

It is interesting that $\mathbb{Q}[X, Y]$ has unique factorization — see Lemma A.3.2 on page 412.

Chapter A, A.1 Exercise 27 (p. 364) If $n = \max(n_1, \ldots, n_k)$ then $(p_1 \cdots p_k)^n \in (x)$ so that $(p_1 \cdots p_k) \in \sqrt{(x)}$.

Chapter A, A.1 Exercise 28 (p. 364) Suppose an ideal $\mathfrak{I}$ contains polynomials $p(X), q(X)$. If these polynomials are relatively prime in $\mathbb{Q}[X]$ then there is a linear combination

$$a(X)p(X) + b(X)q(X) = 1$$

in $\mathbb{Q}[X]$, and after clearing out the denominators, we get

$$n \cdot a(X)p(X) + n \cdot b(X)q(X) = n$$

so this ideal also contains an integer, n. If an ideal *does not* contain any integer then it is not maximal.

The requirement that $\mathbb{Z}[X]/\mathfrak{I}$ is a field (see lemma A.1.30 on page 355) implies that n is a prime, p. We can compute the quotient of $\mathbb{Z}[X]/\mathfrak{I}$ in two stages:

Form the quotient with respect to (p), forming

$$\mathbb{Z}_p[X]$$

and then taking the quotient by the image of the polynomials in $\mathfrak{I}$. Since $\mathbb{Z}_p[X]$ is a principal ideal domain, we can assume that the image of $\mathfrak{I}$ in $\mathbb{Z}_p[X]$ is a principal ideal $(q(X))$. The quotient

$$\mathbb{Z}_p[X]/(q(X))$$

is a field if and only if $q(X)$ is irreducible. It follows that our maximal ideals of $\mathbb{Z}[X]$ are all of the form

$$(p, q_p(X))$$

where $p \in \mathbb{Z}$ is a prime and $q_p(X)$ has an irreducible image in $\mathbb{Z}_p[X]$. Two such ideals

$$(p, a_p(X)), (p, b_p(X))$$

will be equal if and only if $(a_p(X)) = (b_p(X)) \subset \mathbb{Z}_p[X]$.

Chapter A, A.1 Exercise 29 (p. 364) This follows by straightforward induction on n and proposition A.1.7 on page 343.

Chapter A, A.1 Exercise 30 (p. 364) Since R is noetherian, $\mathfrak{N}(R) = (x_1, \ldots, x_n)$ for some finite set of elements of $\mathfrak{N}(R)$. Each of these elements must be nilpotent, i.e.

$$x_i^{\alpha_i} = 0$$

for suitable values of α_i. If $\alpha = \max(\alpha_1, \ldots, \alpha_n)$ then the Pigeonhole Principle implies that

$$\mathfrak{N}(R)^{n \cdot \alpha} = 0$$

Chapter A, A.1 Exercise 31 (p. 364) The localization, $R_{\mathfrak{p}}$, only has one prime ideal, $\mathfrak{p} \cdot R_{\mathfrak{p}}$, and theorem A.1.47 on page 360 implies that all of the elements of $\mathfrak{p} \cdot R_{\mathfrak{p}}$ are nilpotent. If $x \in \mathfrak{p}$, then $x/1 \in \mathfrak{p} \cdot R_{\mathfrak{p}}$ is nilpotent so that there exists an element, $y \in R \setminus \mathfrak{p}$ such that $y \cdot x^n = 0$ for some n.

Chapter A, A.1 Exercise 32 (p. 367) The presence of X^3 implies that we should start with

$$\sigma_1^3 = X^3 + 3X^2Y + 3X^2Z + 3Y^2Z + Y^3 + Z^3 + 6XYZ$$

so

$$X^3 + Y^3 + Z^3 - \sigma_1^3 = -3X^2Y - 3X^2Z - 3Y^2Z - 6XYZ$$

The highest ordered monomial is $-3X^2Y$, which is the highest ordered monomial of $-3\sigma_1\sigma_2$ (see equation A.1.10 on page 365). We get

$$
\begin{aligned}
X^3 + Y^3 + Z^3 - \sigma_1^3 + 3\sigma_1\sigma_2 &= 3XYZ \\
&= 3\sigma_3
\end{aligned}
$$

so

$$X^3 + Y^3 + Z^3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$$

Chapter A, A.1 Exercise 34 (p. 367) We use induction on n. If $n = 2$, the conclusion is clear. Now we assume the conclusion for n, and we will prove it for $n + 1$. We have

$$
V_{n+1} = \begin{bmatrix}
1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^n \\
1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^n \\
1 & \alpha_3 & \alpha_3^2 & \cdots & \alpha_3^n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \alpha_{n+1} & \alpha_{n+1}^2 & \cdots & \alpha_{n+1}^n
\end{bmatrix}
$$

and replace α_1 by X. The determinant is a polynomial $p(X)$ that vanishes if $X = \alpha_2, \ldots, \alpha_{n+1}$. It follows that

$$f(X) = C \cdot (X - \alpha_2) \cdots (X - \alpha_{n+1})$$

where the coefficient of X^n is precisely C. Expanding the determinant of V_{n+1} by minors in the first row shows that the coefficient of X^n, or C, is equal to the $(-1)^n \times$ the determinant of

$$
V = \begin{bmatrix}
1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\
1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\
1 & \alpha_3 & \alpha_3^2 & \cdots & \alpha_3^{n-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1}
\end{bmatrix}
$$

so

$$\det V_{n+1} = (-1)^n \det V_n \cdot \prod_{j=2}^{n+1} (\alpha_1 - \alpha_j)$$

$$= \det V_n \cdot \prod_{j=2}^{n+1} (\alpha_j - \alpha_1)$$

$$= \prod_{1 \le i < j \le n+1} (\alpha_j - \alpha_i)$$

Chapter A, A.1 Exercise 35 (p. 367) The map

$$\frac{k[X_1, \ldots, X_n]}{\mathfrak{I}} \quad \to \quad \frac{k[X_1, \ldots, X_{n+1}]}{\mathfrak{I} + (X_{n+1} - f(X_1, \ldots, X_n))}$$

$$X_i \quad \mapsto \quad X_i$$

and the inverse is defined similarly, except that

$$X_{n+1} \mapsto f(X_1, \ldots, X_n)$$

Chapter A, A.1 Exercise 36 (p. 380) This is basic linear algebra: U is the nullspace of the linear map $V \to W$ and the image is all of W.

Chapter A, A.1 Exercise 37 (p. 380) Suppose V is a vector-space over an infinite field, k, and

$$V = \bigcup_{i=1}^{n} V_i$$

where the V_i are proper subspaces. Without loss of generality, assume this decomposition is *minimal* (none of the V_i's are contained in a union of the others).

    If $x \in V_1$ and $y \in V \setminus V_1$, then $x + r \cdot y \in V$ as $r \in k$ runs over an (infinite) number of nonzero values. The Pigeonhole Principle implies that there is a j such that

$$x + r \cdot y \in V_j$$

for an infinite number of values of r. This means that there exist $r_1 \ne r_2 \in k$ with $x + r_1 \cdot y, x + r_2 \cdot y \in V_j$ which implies that $(x + r_1 \cdot y) - (x + r_2 \cdot y) = (r_1 - r_2) \cdot y \in V_j$ so $y \in V_j$. We conclude that $x \in V_j$ as well. Since x was an *arbitrary* element of V_1, this means that

$$V_1 \subset \bigcup_{i=2}^{n} V_i$$

which contradicts the assumption that the original decomposition was *minimal.*

Chapter A, A.1 Exercise 38 (p. 380) A finite-dimensional vector-space over a finite field has a finite number of elements, hence is the (finite) union of the one-dimensional subspaces generated by these elements,

Chapter A, A.1 Exercise 39 (p. 380) Just apply proposition A.1.72 on page 373 to the diagram

Chapter A, A.1 Exercise 40 (p. 380) If $x \in R$ annihilates M_2, it annihilates any submodule and quotient so

$$\mathrm{Ann}(M_2) \subset \mathrm{Ann}(M_1) \cap \mathrm{Ann}(M_3)$$

If $x \in \mathrm{Ann}(M_1)$, $y \in \mathrm{Ann}(M_3)$, and $m \in M_2$, then $y \cdot m \in M_1$, since its image in M_3 is zero. Then $x \cdot (y \cdot m) = 0$, so

$$\mathrm{Ann}(M_1) \cdot \mathrm{Ann}(M_3) \subset \mathrm{Ann}(M_2)$$

Chapter A, A.1 Exercise 41 (p. 380) We claim that the map

$$(q, h) : U \oplus W \to V$$

is an isomorphism. Suppose (u, v) maps to 0 in V, so $q(u) = -h(v)$. If we map this via p, we get $p \circ q(u) = -p \circ h(v)$. Since $p \circ q = 0$, we get $p \circ h(v) = 0$ which implies that $v = 0$ (since $p \circ h = 1$). Since q is injective, this also implies that $u = 0$. So the map, (q, h), is *injective.*

Suppose $v \in V$ is any element and let $z = v - h \circ p(v)$. Then $p(z) = p(v) - p \circ h \circ p(v) = p(v) - p(v) = 0$. This implies that $z = q(u)$ for some $u \in U$ and

$$v = (q, h)(z, h(v))$$

so the map is also *surjective.*

Chapter A, A.1 Exercise 42 (p. 380) This follows immediately from exercises 39 on page 380 and 41 on page 380.

Chapter A, A.1 Exercise 44 (p. 385) This follows immediately from the Ascending Chain Condition (in exercise 24 on page 363) and corollary A.1.94 on page 385.

Chapter A, A.1 Exercise 46 (p. 386) Let $\mathrm{ann}(m) \subset R$ be the annihilator of m — an ideal. Then m goes to 0 in $M_{\mathfrak{a}}$ if and only if $\mathrm{ann}(m) \not\subset \mathfrak{a}$ (see definition A.1.95 on page 385). But proposition A.1.28 on page 353 shows that *every* ideal is contained in *some* maximal ideal.

Chapter A, A.1 Exercise 47 (p. 386) This follows immediately from exercise 46 on page 386.

Chapter A, A.1 Exercise 48 (p. 386) The general element of F is of the form

$$f = \frac{a_0 + a_1 X + \cdots}{b_0 + b_1 X + \cdots}$$

Suppose a_i is the lowest indexed coefficient in the numerator that is nonzero and b_j is the corresponding one in the denominator. Then

$$f = \frac{X^i (a_i + a_{i+1} X + \cdots)}{X^j (b_j + b_{j+1} X + \cdots)} = X^{i-j} \frac{a_i + a_{i+1} X + \cdots}{b_j + b_{j+1} X + \cdots}$$

where $b_j \neq 0$ so that the denominator is a unit in R (see proposition A.1.7 on page 343). Set $\alpha = i - j$ and $r = (a_i + a_{i+1} X + \cdots) \left(b_j + b_{j+1} X + \cdots \right)^{-1}$.

Chapter A, A.2 Exercise 1 (p. 390) Let $p(X) \in F[X]$ be the minimum polynomial of α. Its being of degree n implies that $F[X]/(p(X)) = F[\alpha] = F(\alpha)$ is a degree-n extension of F. The conclusion follows from proposition A.2.7 on page 388.

Chapter A, A.2 Exercise 2 (p. 395) We claim that $F(X) = F(X^2) \cdot 1 \oplus F(X^2) \cdot X$ as a vector space. If

$$u = \frac{p(X)}{q(X)}$$

we can write

$$q(X) = a(X^2) + X \cdot b(X^2)$$

— just separate the terms with odd powers of X from the others. Now, we get

$$u = \frac{p(X)}{q(X)} \cdot \frac{a(X^2) - X \cdot b(X^2)}{a(X^2) - X \cdot b(X^2)} = \frac{p(X)(a(X^2) - X \cdot b(X^2))}{a(X^2)^2 - X^2 \cdot b(X^2)^2}$$

Now, write the numerator as

$$p(X) = c(X^2) + X \cdot d(X^2)$$

so we get

$$u = \frac{R(Y)}{S(Y)} + X \cdot \frac{T(Y)}{S(Y)}$$

where

$$
\begin{aligned}
S(Y) &= a(Y)^2 - Y \cdot b(Y)^2 \\
R(Y) &= c(Y)a(Y) - Y \cdot b(Y)d(Y) \\
T(Y) &= a(Y)d(Y) - c(Y)b(Y)
\end{aligned}
$$

Chapter A, A.2 Exercise 3 (p. 395) The number $2^{1/3}$ satisfies the equation

$$X^3 - 2 = 0$$

and Eisenstein's Criterion (theorem A.3.8 on page 414) shows that this is irreducible. It follows that $X^3 - 2$ is the minimal polynomial of $2^{1/3}$.

Set $\mathbb{Q}(2^{1/3}) = \mathbb{Q}[X]/(X^3 - 2)$. We would like to find the multiplicative inverse of the polynomial

$$X^2 - X + 1$$

modulo $X^3 - 2$. We can use the extended Euclidean algorithm (algorithm A.1.16 on page 347) for this. Dividing $X^3 - 2$ by $X^2 - X + 1$ gives a quotient of $q_1(X) = X + 1$ and a remainder of -3. We're done since

$$(X + 1) \cdot (X^2 - X + 1) - 1 \cdot (X^3 - 2) = 3$$

or

$$\frac{1}{3}(X + 1) \cdot (X^2 - X + 1) - \frac{1}{3} \cdot (X^3 - 2) = 1$$

so, modulo $X^3 - 2$, we get $\frac{1}{3}(X + 1) \cdot (X^2 - X + 1) = 1$ which implies that

$$\frac{1}{3}(2^{1/3} + 1) = \frac{1}{2^{2/3} - 2^{1/3} + 1} \in \mathbb{Q}(2^{1/3}) = \mathbb{Q}[2^{1/3}]$$

Chapter A, A.2 Exercise 4 (p. 395) Just follow the proof of theorem A.2.21 on page 394. The minimal polynomials of $\sqrt{2}$ and $\sqrt{3}$ are, respectively, $X^2 - 2$ and $X^2 - 3$. Their roots (i.e., the α_i and β_j in the proof) are

$$\pm\sqrt{2}, \pm\sqrt{3}$$

and the set of elements of $\mathbb{Q}$ we must avoid are

$$\frac{\sqrt{2}}{\sqrt{3}} \in \mathbb{Q}[\sqrt{2}, \sqrt{3}] \setminus \mathbb{Q}$$

Since this is not a rational number, it follows that we can pick *any* nonzero rational number for our c. We pick $c = 1$ and $\gamma = \sqrt{2} + \sqrt{3}$.

So

$$\mathbb{Q}[\sqrt{2}, \sqrt{3}] = \mathbb{Q}[\sqrt{2} + \sqrt{3}]$$

To find the minimal polynomial, we refer to example 2.3.20 on page 60.

Chapter A, A.2 Exercise 5 (p. 395) One obvious root of $X^3 - 2 \in \mathbb{Q}[X]$, is $X = 2^{1/3}$, so we try the field extension

$$\mathbb{Q}[2^{1/3}]$$

Since $2^{1/3}$ is a root of $X^3 - 2$, we get $(X - 2^{1/3}) | (X^3 - 2)$ with a quotient of

$$X^2 + 2^{1/3} \cdot X + 2^{2/3}$$

and if we set $X = Y \cdot 2^{1/3}$, this becomes

$$2^{2/3} \cdot (Y^2 + Y + 1)$$

The roots of

$$Y^2 + Y + 1 = 0$$

are

$$\omega, \omega^2 = \frac{-1 \pm \sqrt{-3}}{2}$$

which are the cube-roots of 1 (other than 1 itself). So our splitting field is

$$\mathbb{Q}[2^{1/3}, \omega]$$

of degree 6 over $\mathbb{Q}$.

Chapter A, A.2 Exercise 6 (p. 399) The minimal polynomial is $X^3 - 2$ and we get a basis of $\{1, 2^{1/3}, 2^{2/3}\}$. If $\gamma = a + b 2^{1/3} + c 2^{2/3}$, then the effect of γ on the basis is given by

$$
\begin{aligned}
\gamma \cdot 1 &= a + b 2^{1/3} + c 2^{2/3} \\
\gamma \cdot 2^{1/3} &= 2c + a 2^{1/3} + b 2^{2/3} \\
\gamma \cdot 2^{2/3} &= 2b + 2c 2^{1/3} + a 2^{2/3}
\end{aligned}
$$

which gives a matrix

$$
m_\gamma = \begin{bmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{bmatrix}
$$

with a determinant

$$N_{H/F}(\gamma) = a^3 - 6\,acb + 2\,b^3 + 4\,c^3$$

and characteristic polynomial

$$\chi_\gamma(X) = X^3 - 3\,aX^2 - \left(6\,cb - 3\,a^2\right) X - a^3 + 6\,acb - 2\,b^3 - 4\,c^3$$

Chapter A, A.2 Exercise 7 (p. 399) In this case, our basis for H over $\mathbb{Q}$ is $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$. If $\gamma = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ is a general element, its effect on a basis is

$$
\begin{aligned}
\gamma \cdot 1 &= a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \\
\gamma \cdot \sqrt{2} &= 2b + a\sqrt{2} + 2d\sqrt{3} + c\sqrt{6} \\
\gamma \cdot \sqrt{3} &= 3c + 3d\sqrt{2} + a\sqrt{3} + b\sqrt{6} \\
\gamma \cdot \sqrt{6} &= 6d + 3c\sqrt{2} + 2b\sqrt{3} + a\sqrt{6}
\end{aligned}
$$

which gives a matrix

$$
m_\gamma = \begin{bmatrix} a & 2b & 3c & 6d \\ b & a & 3d & 3c \\ c & 2d & a & 2b \\ d & c & b & a \end{bmatrix}
$$

with a determinant

$$
\begin{aligned}
N_{H/F}(\gamma) = a^4 &- 4\,a^2 b^2 + 48\,adbc - 12\,d^2 a^2 - 6\,a^2 c^2 \\
&+ 4\,b^4 - 24\,d^2 b^2 - 12\,b^2 c^2 + 9\,c^4 - 36\,c^2 d^2 + 36\,d^4
\end{aligned}
$$

and characteristic polynomial

$$\chi_{H/F}(\gamma) = X^4 - 4\,aX^3 + \left(-12\,d^2 - 6\,c^2 - 4\,b^2 + 6\,a^2\right)X^2$$
$$+ \left(-48\,dbc + 24\,d^2a + 12\,ac^2 + 8\,ab^2 - 4\,a^3\right)X$$
$$a^4 - 4\,a^2b^2 + 48\,adbc - 12\,d^2a^2 - 6\,a^2c^2$$
$$+ 4\,b^4 - 24\,d^2b^2 - 12\,b^2c^2 + 9\,c^4 - 36\,c^2d^2 + 36\,d^4$$

Chapter A, A.2 Exercise 8 (p. 403) Suppose

$$f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$$

(after dividing by a_n if necessary) and embed F in its algebraic closure, $\bar{F}$. We get

$$f(X) = \prod_{j=1}^{n}(X - \alpha_j)$$

and the splitting field of $f(X)$ is just

$$F[\alpha_1, \ldots, \alpha_n] \subset \bar{F}$$

which is *unique* in $\bar{F}$. The conclusion follows from theorem A.2.28 on page 399.

Chapter A, A.2 Exercise 9 (p. 403) The characteristic polynomial of γ was computed in the solution to 6 on page 399 (setting $c = 0$)

$$\chi_\gamma(X) = X^3 - 3\,aX^2 + 3\,a^2X - a^3 - 2\,b^3$$

and this is also the minimal polynomial of γ. One factor of this must be $X - \gamma$, so we take the quotient

$$\frac{X^3 - 3\,aX^2 + 3\,a^2X - a^3 - 2\,b^3}{X - \gamma}$$
$$= X^2 + X(\gamma - 3a) + \gamma(\gamma - 3a) + 3a^2$$
$$= X^2 + (-2a + b\,2^{1/3})X + a^2 - ab\,2^{1/3} + b^2\,2^{2/3}$$

The roots of this quadratic equation are the two conjugates of γ:

$$X = a + b\left(\frac{-1 \pm \sqrt{-3}}{2}\right)2^{1/3} = a + b2^{1/3}\omega^j$$

where ω is a primitive cube root of 1 and $j = 1, 2$.

Chapter A, A.2 Exercise 10 (p. 410) Suppose π is algebraic. Then so is πi and $e^{\pi i}$ should be transcendental, by the Lindemann–Weierstrass theorem. But $e^{\pi i} = -1$, which is algebraic. This is a contradiction.

Chapter A, A.2 Exercise 11 (p. 410) The algebraic closure of $\mathbb{Q}$ is the algebraic numbers, $\bar{\mathbb{Q}}$, which is countable. If F is a countable field $F(X)$ is also a countable field and a simple induction shows that

$$F(X_1, \ldots, X_n)$$

is also countable for any n. If $\mathbf{S} = \{X_1, \ldots\}$ is a countable set of indeterminates, then

$$F(\mathbf{S}) = \bigcup_{i=1}^{\infty} F(X_1, \ldots, X_i)$$

is also countable. It follows that an uncountable field like $\mathbb{C}$ must have an uncountable degree of transcendence over $\mathbb{Q}$.

Chapter A, A.4 Exercise 1 (p. 427) Suppose $s^{-1}x \in S^{-1}T$ is integral over $S^{-1}R$. Then it satisfies an equation

$$(s^{-1}x)^n + a_{n-1}(s^{-1}x)^{n-1} + \cdots + a_0 = 0$$

where the $a_i \in S^{-1}R$. Let $\bar{s} \in S$ be able to clear the denominators of all of the a_i. Multiplying this equation by $(s\bar{s})^n$ gives

$$(\bar{s}x)^n + a_{n-1}s\bar{s}(\bar{s}x)^{n-1} + \cdots + s^n\bar{s}a_0 = 0$$

so $\bar{s}x \in R$ is integral over R, and $(s\bar{s})^{-1}(\bar{s}x) = s^{-1}x$ is integral over $S^{-1}R$.

Chapter A, A.4 Exercise 2 (p. 427) Let

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

be the minimal polynomial (see definition A.2.9 on page 388) of x with $a_i \in F$. If $s \in R$ can clear the denominators of the a_i, multiply this equation by s^n to get

$$\begin{aligned}s^n x^n + s^n a_{n-1}x^{n-1} + \cdots + s^n a_0 &= (sx)^n + a_{n-1}s(sx)^{n-1} + \cdots + s^n a_0 \\ &= 0\end{aligned}$$

so sx is integral over R.

Chapter A, A.4 Exercise 3 (p. 427) Clearly, any element $x \in F$ that is integral over R is also integral over T. On the other hand, if x is integral over T, it is also integral over R because of statement 2 of proposition A.4.5 on page 421 (the degree of the monic polynomial over R will usually be higher than that over T). It follows that the integral closures will be the same.

Chapter A, A.5 Exercise 1 (p. 442) This is literally a direct restatement of the definition of a product: every pair of morphisms $f \in \hom_{\mathscr{C}}(W, A)$ and $g \in \hom_{\mathscr{C}}(W, B)$ induces a unique morphism $f \times g \in \hom_{\mathscr{C}}(W, A \times B)$ that makes the diagram A.5.2 on page 439 commute.

Chapter A, A.5 Exercise 3 (p. 442) Suppose $g_1, g_2 \in \hom_{\mathscr{C}}(C, A)$ map to the same element of $\hom_{\mathscr{C}}(C, B)$, then $f \circ g_1 = f \circ g_2 \colon C \to B$, which implies (by the definition of monomorphism) that $g_1 = g_2$.

Chapter A, A.5 Exercise 4 (p. 442) If

$$f\colon A \to B$$

has a kernel, the inclusion of distinct elements of A that differ by an element of the kernel are distinct morphisms whose composite with f are the same. The other conclusion follows by a similar argument.

Chapter A, A.5 Exercise 5 (p. 445) It is A^{tr}, the transpose of A.

Chapter A, A.5 Exercise 6 (p. 445) It is an equivalence because of the natural isomorphism in example A.5.11 on page 444. It is not an isomorphism of categories because the finite-dimensional vector-space V^{**} is not *identical* to V. If V is infinite-dimensional, it is not even isomorphic.

Chapter A, A.5 Exercise 7 (p. 447) Every pair of morphisms

$$\begin{aligned} y &\to x \\ z &\to x \end{aligned}$$

— i.e., every morphism $(y, z) \to \Delta x$ — corresponds to a *unique* morphism

$$y \coprod z \to x$$

so we get an equivalence

$$\hom_{\mathscr{C} \times \mathscr{C}}((y, z), \Delta x) = \hom_{\mathscr{C}}(y \coprod z, x)$$

Chapter A, A.5 Exercise 8 (p. 457) Definition A.5.26 on page 453 implies that there is a natural equivalence

$$\hom_{\mathscr{C}_\infty}(\Delta_\infty x, y) = \hom_{\mathscr{C}}(x, \varprojlim y)$$

for all $x \in \mathscr{C}$ and $y \in \mathscr{C}_\infty$.

Chapter A, A.5 Exercise 9 (p. 457) This follows immediately from the definition of the equivalence relation $\sim$ in definition A.5.17 on page 448.

Chapter A, A.5 Exercise 10 (p. 461) If $a \in A$ is torsion-free then $a \mapsto x$ for any $x \neq 0$ defines a nonzero homomorphism. If a is of order n then $a \mapsto 1/n$ defines a nonzero map.

Chapter A, A.5 Exercise 11 (p. 461) The sub-object $\operatorname{im} A \subset B$ maps to A in a straightforward way. The injective property of A implies that this extends to all of B.

Chapter A, A.5 Exercise 12 (p. 461) We already know that

$$0 \to \hom_{\mathscr{A}}(D, A) \xrightarrow{\hom_{\mathscr{A}}(1,r)} \hom_{\mathscr{A}}(D, B)$$

is exact, by exercise 3 on page 442, so we must still show that

$$\hom_{\mathscr{A}}(D, A) \xrightarrow{\hom_{\mathscr{A}}(1,r)} \hom_{\mathscr{A}}(D, B) \xrightarrow{\hom_{\mathscr{A}}(1,s)} \hom_{\mathscr{A}}(D, C)$$

is exact. If $f \in \hom_{\mathscr{A}}(D, B)$ maps to 0 in $\hom_{\mathscr{A}}(D, C)$, then $s \circ f = 0$. Since $r = \ker s$, we have a *unique* morphism $D \to A$ that makes

$$
\begin{array}{ccc}
D & \xrightarrow{\ f\ } & B \\
 & \searrow{\scriptstyle v} & \uparrow{\scriptstyle r} \\
 & & A
\end{array}
$$

commute. This is precisely the element of $\hom_{\mathscr{A}}(D, A)$ that maps to f.

Chapter A, A.5 Exercise 13 (p. 472) If $Q = M/N$, we get a short exact sequence

$$0 \to N \to M \to Q \to 0$$

and the conclusion follows from the fact that the sequence

$$0 \to A \otimes_R N \to A \otimes_R M \to A \otimes_R Q \to 0$$

is also exact (because A is flat).

Chapter A, A.5 Exercise 14 (p. 472) In the top formula, multilinear maps

$$M \times N \times T \to A$$

where A is an arbitrary R-module factor through $M \otimes_R (N \otimes_R T)$ and $(M \otimes_R N) \otimes_R T$ so the universal property of $\otimes$ implies that they are isomorphic.

　　To see the second equality, regard $M \otimes_R N$ and $N \otimes_R M$ as quotients of $\mathbb{Z}[M \times N]$ by the ideal generated by the identities in definition A.5.44 on page 462 and noting that this ideal is symmetric with respect to factors.

Chapter A, A.5 Exercise 15 (p. 472) Corollary A.5.49 on page 465 implies that

$$k^n \otimes_k k^m = k^{n \cdot m}$$

so the dimensions are as claimed.

　　If $\{e_i\}$ is a basis for V and $\{f_j\}$ is a basis for W then it is not hard to see that $\{e_i \otimes f_j\}$, $i = 1, \dots, n$, $j = 1, \dots, m$ *spans* $V \otimes_k W$ — just use the identities in definition A.5.44 on page 462 to express any $v \otimes w$ in terms of them. The fact that $V \otimes_k W$ is $n \cdot m$-dimensional shows that these elements must be linearly independent too.

To prove the final statement, we must show that the set $\{e_i \otimes f_j\}$ is linearly independent even if there are an infinite number of basis elements. Suppose we have some linear combination

$$(E.0.5) \qquad \sum_{t=1}^{n} a_t(e_{i_t} \otimes f_{k_t}) = 0$$

for $a_t \in k$. Since only a finite number of terms are involved, this equation really involves finite-dimensional subspaces of V and W, namely the span of the $\{e_{i_t}\}$ in V and the span of the $\{f_{j_t}\}$ in W. We have already seen that the $\{e_{i_t} \otimes f_{k_t}\}$ are linearly independent in this case, so all of the $a_t = 0$ in equation E.0.5.

Chapter A, A.5 Exercise 16 (p. 472) This is an $ns \times mt$ matrix called the *Kronecker product* of A and B. If

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

Then

$$A \otimes B = \begin{bmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \cdots & a_{m,n}B \end{bmatrix}$$

Chapter A, A.5 Exercise 17 (p. 473) This follows immediately from proposition A.5.52 on page 466:

$$\frac{R}{\mathfrak{a}} \otimes_R \frac{R}{\mathfrak{b}} = \left(\frac{R}{\mathfrak{a}}\right) \bigg/ \mathfrak{b} \cdot \left(\frac{R}{\mathfrak{a}}\right) = \frac{R}{\mathfrak{a} + \mathfrak{b}}$$

Chapter A, A.5 Exercise 18 (p. 473) We always have a surjective natural map

$$\mathfrak{a} \otimes_R M \quad \to \quad \mathfrak{a} \cdot M$$
$$a \otimes m \quad \mapsto \quad a \cdot m$$

If M is flat, this map is also *injective* since it is

$$\mathfrak{a} \hookrightarrow R$$

$\otimes M$.

Chapter A, A.5 Exercise 19 (p. 473) For each i, take the natural maps

$$z_i : M_i \quad \to \quad \varinjlim M_j$$
$$Z_i : M_i \otimes_R N \quad \to \quad \varinjlim \left(M_j \otimes_R N\right)$$

and form the tensor product of M_i with N to get

$$z_i \otimes 1 : M_i \otimes_R N \to \left(\varinjlim M_j\right) \otimes_R N$$

The universal property of direct limits implies the existence of a *unique* map

$$v : \varinjlim \left(M_j \otimes_R N\right) \to \left(\varinjlim M_j\right) \otimes_R N$$

that makes the diagram

$$\begin{array}{ccc} \varinjlim \left(M_j \otimes_R N\right) & \xrightarrow{\;\;v\;\;} & \left(\varinjlim M_j\right) \otimes_R N \\ {\scriptstyle Z_i}\big\uparrow & & \big\uparrow{\scriptstyle z_i \otimes 1} \\ M_i \otimes_R N & =\!\!=\!\!=\!\!= & M_i \otimes_R N \end{array}$$

commute. If $m \otimes n \neq 0 \in \left(\varinjlim M_j\right) \otimes_R N$, then m is the image of some $m_i \in M_i$ and $m \otimes n = v \circ Z_i(m_i \otimes n)$. It follows that v is surjective.

If $w \in \varinjlim \left(M_j \otimes_R N \right)$ is in the kernel of v, then $w = Z_i(m_i \otimes n)$ for some i. the commutativity of the diagram implies that $(z_i \otimes 1)(m_i \otimes n) = z_i(m_i) \otimes n = 0$, which implies that $Z_i(m_i \otimes n) = 0$ — so v is injective.

Chapter A, A.5 Exercise 20 (p. 473) The hypotheses imply that

$$0 \longrightarrow \varinjlim A_i \xrightarrow{\varinjlim f_i} \varinjlim B_i \xrightarrow{\varinjlim g_i} \varinjlim C_i \longrightarrow 0$$

$$\bar{a}_i \uparrow \qquad \bar{b}_i \uparrow \qquad \bar{c}_i \uparrow$$

$$0 \longrightarrow A_i \xrightarrow{f_i} B_i \xrightarrow{g_i} C_i \longrightarrow 0$$

commutes for all i, where we don't know whether the top row is exact. If $x \in \varinjlim C_i$, then x is the image of some $x_i \in C_i$. The commutativity of this diagram implies that x is in the image of $\bar{b}_i(g_i^{-1}(x_i))$. It follows that $\varinjlim g_i$ is surjective. If $x \in \varinjlim B_i$ such that $x \in \ker \varinjlim g_i$, then $x = \bar{b}_i(x_i)$ for some i and $\bar{c}_i \circ g_i(x_i) = 0$. The definition of direct limit implies that there exists $N > i$ such that $c_N \circ \cdots \circ c_i(x) = 0$, so $b_N \circ \cdots \circ b_i(x_i) \in \ker g_N$. The exactness of the original sequences implies that $b_N \circ \cdots \circ b_i(x_i) = f_N(y_N)$ and the commutativity of the diagram above implies that $\bar{a}_N(y_N) = x$. A similar argument implies that the left end of the upper row is also exact.

Chapter A, A.5 Exercise 21 (p. 473) Consider the map $g \colon (\alpha) \hookrightarrow R$. This is an inclusion and, since S is flat over R

$$(\alpha) \otimes_R S \xrightarrow{g \otimes 1} R \otimes_R S = S$$

is also an inclusion. Since α is a non-zero-divisor in R, $(\alpha) \cong R$ and the isomorphism $R \to R$ induced by the inclusion is multiplication by α. This implies that

$$S = R \otimes_R S \xrightarrow{(\times \alpha) \otimes 1 = \times f(\alpha)} R \otimes_R S = S$$

is also injective, which implies that $f(\alpha) \in S$ is a non-zero-divisor.

Chapter A, A.5 Exercise 22 (p. 473) Let

(E.0.6) $$0 \to U_1 \to U_2 \to U_3 \to 0$$

be a short exact sequence of modules over S. Since we can compose the action of S on these modules with the homomorphism, f, it is also a short exact sequence of modules over R. If we take the tensor product with $M \otimes_R S$, we get

$$0 \longrightarrow U_1 \otimes_S (M \otimes_R S) \longrightarrow U_2 \otimes_S (M \otimes_R S) \longrightarrow U_3 \otimes_S (M \otimes_R S) \longrightarrow 0$$

$$\| \qquad\qquad \| \qquad\qquad \|$$

$$U_1 \otimes_R M \qquad\qquad U_2 \otimes_R M \qquad\qquad U_3 \otimes_R M$$

which is exact since equation E.0.6 is an exact sequence of R-modules.

Chapter A, A.5 Exercise 23 (p. 473)

Chapter A, A.5 Exercise 24 (p. 473) For any commutative ring U and any module A over U, $U \otimes_U A = A$. This (and the associativity of tensor products) implies that

$$\begin{aligned}
(S^{-1}R \otimes_R M) \otimes_{S^{-1}R} (S^{-1}R \otimes_R N) &= (S^{-1}R \otimes_R M) \otimes_{S^{-1}R} S^{-1}R \otimes_R N \\
&= (S^{-1}R \otimes_R M) \otimes_R N \\
&= S^{-1}R \otimes_R M \otimes_R N \\
&= S^{-1}R \otimes_R (M \otimes_R N)
\end{aligned}$$

Chapter A, A.5 Exercise 25 (p. 473) Since M is projective, it is a direct summand of a free module, F, so $F = M \oplus N$ for some other projective module, N. Then

$$F^* = M^* \oplus N^*$$

Chapter A, A.6 Exercise 1 (p. 481) First of all, note that the map $V \to W$ is split, i.e. there exists a left-inverse $t\colon W \to V$ so that $g \circ t = 1$. Any two such splitting maps differ by a map from V to U. Now define $1 - t \circ g\colon W \to \ker g = \operatorname{im} U$ or $f^{-1} \circ (1 - t \circ g)\colon W \to U$. We get an isomorphism

$$(f^{-1} \circ (1 - t \circ g), g)\colon W \cong U \oplus V$$

Given a commutative diagram like the one in the statement of the problem, we can lift a map $t_2\colon W_2 \to V_2$ to get a map $t_1\colon W_1 \to V_2$ so we get a natural isomorphism from W_i to $U_i \oplus V_i$. The conclusion follows from proposition A.6.8 on page 476.

Chapter A, A.6 Exercise 2 (p. 482)

Chapter A, A.6 Exercise 3 (p. 482) Just compute

$$
\begin{aligned}
(e_2 + 2e_4) \wedge e_3 \wedge e_1 \wedge (e_2 - e_4) &= -e_2 \wedge e_3 \wedge e_1 \wedge e_4 \\
&\quad -2e_4 \wedge e_3 \wedge e_1 \wedge e_2 \\
&= -e_1 \wedge e_2 \wedge e_3 \wedge e_4 \\
&\quad +2e_1 \wedge e_2 \wedge e_3 \wedge e_4 \\
&= e_1 \wedge e_2 \wedge e_3 \wedge e_4
\end{aligned}
$$

so the determinant is 1.

Chapter A, A.7 Exercise 1 (p. 493) If $G \in K[Y_1, \ldots, Y_m]$ has the property that $G(f_1, \ldots, f_m) = 0$, then we get a linear dependence

$$\left(\frac{\partial G}{\partial Y_1}\right)_{Y_1 = f_1} df_1 + \cdots + \left(\frac{\partial G}{\partial Y_m}\right)_{Y_m = f_m} df_m = 0$$

We have not used the fact that the characteristic of k is 0 yet (and this condition is unnecessary for this part of the proof).

Conversely, if $\{f_1, \ldots, f_m\}$ is algebraically *independent,* then there exists a transcendence basis for K over k that *includes* $\{f_1, \ldots, f_m\}$ (apply lemma A.2.49 on page 408 m times) so K is a finite extension of

$$L = k(f_1, \ldots, f_m, X_1, \ldots)$$

where $X_1, \ldots$ is a (possibly infinite) set of indeterminates. It follows that (see the remark following proposition A.7.9 on page 485):

$$\Omega_{L/k} = L \cdot df_1 \oplus \cdots \oplus L \cdot df_m \oplus \bigoplus_i L \cdot dX_i$$

Corollary A.7.16 on page 490 (this is where we use the characteristic of the fields being 0) implies that

$$\Omega_{K/k} = K \cdot df_1 \oplus \cdots \oplus k \cdot df_m \oplus \bigoplus_i K \cdot dX_i$$

so there exists a basis for $\Omega_{K/k}$ in which the $\{df_i\}$ are linearly independent.

Chapter A, A.7 Exercise 2 (p. 493) The field F is uncountable. Now, note that $\mathbb{Q}$ is countable, as is $\mathbb{Q}(X)$. In fact

$$\mathbb{Q}(X_1, \ldots, X_n)$$

is countable for all n. It follows that

$$\mathbb{Q}(X_1, \ldots) = \bigcup_{n=1}^{\infty} \mathbb{Q}(X_1, \ldots, X_n)$$

is countable. This means that F has an uncountable *transcendence-degree* over $\mathbb{Q}$.

Chapter B, B.1 Exercise 1 (p. 498) Suppose $u, v \in \mathscr{F}(U)$ map to the same element $\prod_{x \in U} s_x \in \prod_{x \in U} \mathscr{F}_x$. For each x there exists an open set U_x such that $u|U_x = v|U_x$, since they become equal in the direct limit – see exercise 9 on page 457. Since the $\{U_x\}$ cover U, statement 4 in definition B.1.1 on page 495 implies the conclusion.

Chapter B, B.1 Exercise 2 (p. 498) Suppose $f: \mathscr{F} \to \mathscr{G}$ is a morphism of sheaves on a space X, the diagram

$$
\begin{array}{ccc}
\mathscr{F}(U) & \xrightarrow{\ f\ } & \mathscr{G}(U) \\
\downarrow & & \downarrow \\
\prod_{x \in U} \mathscr{F}_x & \longrightarrow & \prod_{x \in U} \mathscr{G}_x
\end{array}
$$

implies the conclusion.

Chapter B, B.1 Exercise 3 (p. 498) The fact that f is continuous implies that $f^{-1}(U)$ is an open set. All of the conditions in definition B.1.1 on page 495 are satisfied because they are already satisfied for $\mathcal{O}_X$.

Chapter B, B.2 Exercise 1 (p. 501) This follows from the fact that a quotient of surjective homomorphisms is surjective. The final statement is always true for an exact sequence of presheaves (and is used to define an exact sequence). The only way it fails for an exact sequence of sheaves is that $\mathscr{H}$ is usually a *completion* of $\mathscr{G}/\mathscr{F}$. For flasque sheaves, this is unnecessary.

Chapter B, B.2 Exercise 2 (p. 502) This follows immediately from the definition of direct image sheaf in exercise 3 on page 498. If

$$U_2 \subset U_1 \subset X$$

then $i^{-1}(U_2) \subset i^{-1}(U_1)$ so the restriction map of sections will be surjective.

Chapter B, B.2 Exercise 3 (p. 502) It is straightforward to show that morphisms of presheaves have kernels and cokernels. A morphism of sheaves has a well-defined kernel but the cokernel might only be a presheaf, as example B.2.1 on page 498 shows.

Chapter B, B.2 Exercise 4 (p. 502) If f is injective, it induces an injection on all stalks, by the solution to exercise 20 on page 473. The converse follows immediately from the diagram

$$
\begin{array}{ccc}
\mathscr{F}(U) & \xrightarrow{\ f\ } & \mathscr{G}(U) \\
\downarrow & & \downarrow \\
\prod_{x \in U} \mathscr{F}_x & \lhook\joinrel\longrightarrow & \prod_{x \in U} \mathscr{G}_x
\end{array}
$$

in the solution to exercise 1 on page 498.

Chapter B, B.2 Exercise 5 (p. 502) We already know that f is injective by exercise 4 on page 502. If $s \in \mathscr{G}(U)$, for some open set U, then there exist $t_x \in \mathscr{F}_x$ that map to s_x. Each x is contained in an open set U_x such that t_x is the image of $T_x \in \mathscr{F}(U_x)$ and s_x is in the image of $f(T_x)$. By shrinking U_x if necessary, we can assume $f(T_x)|U_x = s|U_x$. The U_x form an open cover of U. Since we already know that f is *injective*, it follows that

$$T_x|U_x \cap U_y = T_y|U_x \cap U_y$$

Now the sheaf-condition (statement 5 in definition B.1.1 on page 495 implies that the T_x patch together to form an element $T \in \mathscr{F}(U)$.

Chapter B, B.2 Exercise 6 (p. 502) If f is surjective, it induces a surjection on all stalks, by the solution to exercise 20 on page 473. Conversely, suppose $f_x: \mathscr{F}_x \to \mathscr{G}_x$ is surjective for all $x \in X$. Let $\mathscr{H} \subset \mathscr{G}$ be the image of f, regarded as a map of presheaves. Then $\operatorname{im} f$ is the completion of $\mathscr{H}$ (via lemma B.2.3 on page 500) — see definition B.2.5 on page 501).

Since the process of completion preserves stalks, we have isomorphisms

$$(\operatorname{im} f)_x \to \mathcal{G}_x$$

and the solution to exercise 5 on page 502 implies that $\mathcal{G} = \operatorname{im} f$.

Chapter B, B.2 Exercise 7 (p. 502) The fact that stalks are direct limits and direct limits preserve exactness (see exercise 20 on page 473) implies that the corresponding sequence of stalks is exact.

Exercises 4 on page 502 and 6 on page 502 prove exactness at the ends. We must verify

$$\mathcal{F}(U) \xrightarrow{f} \mathcal{G}(U) \xrightarrow{g} \mathcal{H}(U)$$

is exact for open sets $U \subset X$. That $g \circ f = 0$ follows from the diagram in the solution to exercise 2 on page 498. If $\mathcal{X} = \ker g$, this is a sheaf and we have

$$\operatorname{im} f \subset \mathcal{X}$$

where the image of a sheaf-map is as in definition B.2.5 on page 501 Since the sequences of stalk-maps is exact, we have

$$(\operatorname{im} f)_x = \mathcal{X}_x$$

for all $x \in U$. The conclusion follows from the solution to exercise 5 on page 502.

Chapter B, B.2 Exercise 8 (p. 502) The universal property of direct limits (see definition A.5.17 on page 448) implies that $f^{-1}(\mathcal{F})$ behaves well with respect to restrictions and morphisms of sheaves.

Chapter B, B.2 Exercise 9 (p. 502) If U is an connected open set with subset $U' \subset U$ where U' is a union of disjoint open sets U'_1 and U'_2, then

$$
\begin{aligned}
\underline{A}(U) &= A \\
\underline{A}(U') &= A \oplus A
\end{aligned}
$$

and the restriction map

$$\underline{A}(U) \to \underline{A}(U')$$

is not surjective.

Chapter B, B.3 Exercise 1 (p. 505) By restricting to the stalks at a point p, every element $x \in \mathcal{O}_V(U)$ gives rise to an element of $\mathcal{O}_V(p)$. Since this is a local ring, the quotient by its maximal ideal is an algebraic extension of k, hence k itself. We will call this the *value of x at p* or $x(p)$.

We can consequently regard elements of $\mathcal{O}_V(U)$ as *functions $U \to k$* and the restriction maps as simple restrictions of functions. Conditions 1, 2, and 3 in definition B.1.1 on page 495 are automatically satisfied. We must require that $\mathcal{O}_V(U)$ be closed under addition and multiplication, though (which was automatic in definition B.1.1 on page 495).

Chapter B, B.3 Exercise 3 (p. 505) Simply define a homomorphism of presheaves

$$\mathcal{F}(U) \otimes_{\mathcal{O}_X(U)} \mathcal{H}om(\mathcal{F}, \mathcal{O}_V)(U) = \hom_{\mathcal{O}_X(U)}(\mathcal{F}(U), \mathcal{O}_X(U)) \to \mathcal{O}_X(U)$$

over open sets $U \subset X$ by plugging $\mathcal{F}(U)$ into elements of $\hom_{\mathcal{O}_X(U)}(\mathcal{F}(U), \mathcal{O}_X(U))$. This defines a homomorphism of presheaves which becomes a homomorphism of sheaves after completion.

Chapter B, B.3 Exercise 4 (p. 505) Simply consider the effect of this map on stalks. We get

$$\mathcal{O}_{X,x} \otimes_{\mathcal{O}_{X,x}} \hom(\mathcal{O}_{X,x}, \mathcal{O}_{X.x}) = \mathcal{O}_{X.x}$$

The conclusion follows from exercise 5 on page 502.

Chapter C, C.1 Exercise 1 (p. 513) If ξ is trivial, there exists an isomorphism

$$F\colon \xi \to X \times \mathbb{A}^n$$

The vector bundle $X \times \mathbb{A}^n$ has n linearly independent sections $\{t_1, \ldots, t_n\}$ defined by

$$t_i(x) = x \times e_i$$

where

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where the 1 occurs in the ith row. We can clearly define $s_i = F^{-1} \circ t_i$ and get n linearly independent sections of ξ.

Conversely, if $p\colon \xi \to X$ has n linearly independent sections $\{s_1, \ldots, s_n\}$, and $v \in p^{-1}(x)$, then

$$v = \begin{bmatrix} \alpha_1(x) \\ \vdots \\ \alpha_n(x) \end{bmatrix}$$

in the basis formed by $\{s_1(x), \ldots, s_n(x)\}$ and the map, $F\colon \xi \to X \times \mathbb{A}^n$ that sends $v \in p^{-1}(x)$ to $x \times (\alpha_1(x), \ldots, \alpha_n(x)) \in X \times \mathbb{A}^n$ is clearly an isomorphism.

Chapter C, C.1 Exercise 2 (p. 513) In this case, the transition functions, $\varphi_{\alpha,\beta}$ are just nonzero scalar functions so $\varphi_{\alpha,\beta}^{\text{tr}} = \varphi_{\alpha,\beta}$ and the conclusion follows from the fact that $\varphi_{\beta,\alpha} = \varphi_{\alpha,\beta}^{-1}$.

Chapter C, C.1 Exercise 3 (p. 514) The transition-function for $\xi \otimes \xi$ are $\varphi_{1,2}^2 = 1$, where $\varphi_{1,2}$ is defined by equation C.1.4 on page 509. The only possible invariant for a vector bundle over S^1 is the number of times it "wraps around" S^1. But all bundles that "wrap" an odd number of times are isomorphic and all bundles that "wrap" an even number of times are isomorphic to the trivial bundle.

Chapter C, C.1 Exercise 4 (p. 514) This follows from the fact that $\oplus$ is a coproduct in the category of vector-spaces (after forming a trivializing cover for the vector-bundles involved).

Chapter C, C.1 Exercise 5 (p. 514) This follows immediately from exercise 1 on page 481, which implies the conclusion for

$$0 \to E|U \to F|U \to G|U \to 0$$

with U is a trivializing cover. *Naturality* of the isomorphism in exercise 1 on page 481 implies that this commutes with transition functions and defines a global isomorphism of vector bundles.

Chapter C, C.2 Exercise 1 (p. 518) We could work this out in detail by using trivializing covers or we could "cheat" and use category theory: since $\xi_1 \oplus \xi_2$ is a coproduct in the category of vector-bundles, $\mathscr{S}_{\xi_1 \oplus \xi_2}$ has the same universal property as the coproduct in the category of modules over $\mathcal{O}_X$, so it must equal $\mathscr{S}_{\xi_1} \oplus \mathscr{S}_{\xi_2}$.

Chapter C, C.2 Exercise 2 (p. 518) This follows immediately from theorem C.2.7 on page 518 and exercise 5 on page 514.

Chapter D, D.1 Exercise 1 (p. 525) The cohomology groups of the third cochain complex will appear in the long exact sequence in proposition D.1.9 on page 522 sandwiched between the zero-groups of the other two.

Chapter D, D.1 Exercise 2 (p. 525) In the long exact sequence

$$\cdots \to H^i(C) \xrightarrow{f^*} H^i(D) \xrightarrow{g^*} H^i(E) \xrightarrow{c} H^{i+1}(C) \to H^{i+1}(D) \to \cdots$$

we have $H^i(D) = H^{i+1}(D)$ so the exact sequence reduces to

$$\cdots \to H^i(C) \xrightarrow{f^*} 0 \xrightarrow{g^*} H^i(E) \xrightarrow{c} H^{i+1}(C) \to 0 \to \cdots$$

Chapter D, D.1 Exercise 3 (p. 525) If the chain-homotopy between f and g is Φ, simply use $F(\Phi)$ as the chain-homotopy between $F(f)$ and $F(g)$.

Chapter D, D.1 Exercise 4 (p. 525) If $b \in B$ maps to 0 under v, then its image under s must map to 0 under w. But w is an isomorphism so $s(b) = 0$. The exactness of the top row implies that $b = r(a)$ for some $a \in A$. If $a \neq 0$, then it maps to something nonzero under u (since u is an isomorphism) and therefore to something nonzero under r', which gives a contradiction. It follows that b must have been 0 to start with. So v is injective. Proof of surjectivity is left to the reader.

Chapter D, D.1 Exercise 5 (p. 531) This follows from the corresponding property of hom.

Chapter D, D.1 Exercise 6 (p. 531) In this case N is its *own* injective resolution. And all others are chain-homotopy equivalent to it, so they have the same cohomology.

Chapter D, D.1 Exercise 7 (p. 531) We can identity $\hom_{\mathcal{M}_R}(R, M) = \hom_R(R, M) = M$. It follows that, applied to any resolution of M, we just recover the resolution.

Chapter D, D.1 Exercise 8 (p. 531) A projective module is a direct summand of a free module, so we get

$$\mathrm{Ext}_R^i(P \oplus Q, M) = \mathrm{Ext}_R^i(P, M) = \mathrm{Ext}_R^i(Q, M) = \mathrm{Ext}_R^i(F, M)$$

and $\mathrm{Ext}_R^i(F, M) = 0$, by exercise 7 on page 531.

Chapter D, D.1 Exercise 9 (p. 531) The proof is very much like that of corollary D.1.22 on page 530 except that you "reverse the arrows."

Chapter D, D.1 Exercise 10 (p. 531) This is just

$$\mathbb{Q} \to \mathbb{Q}/\mathbb{Z}$$

Chapter D, D.1 Exercise 11 (p. 531) This follows immediately from proposition A.5.38 on page 460.

Chapter D, D.1 Exercise 12 (p. 531) First note that $\hom_{\mathscr{A}b}(\mathbb{Z}/n \cdot \mathbb{Z}, \mathbb{Q}/\mathbb{Z}) = \mathbb{Z}/n \cdot \mathbb{Z}$. The conclusion follows from the finite direct sum

$$A = \bigoplus \frac{\mathbb{Z}}{n_i \cdot \mathbb{Z}}$$

so

$$\hom_{\mathscr{A}b}(A, \mathbb{Q}/\mathbb{Z}) = \prod \frac{\mathbb{Z}}{n_i \cdot \mathbb{Z}} = A$$

since the product is *finite*.

The second statement follows from looking at the injective resolution of $\mathbb{Z}$ in exercise 10 on page 531.

Chapter D, D.1 Exercise 13 (p. 532) Corollary D.1.20 on page 529 shows that an extension like that above induces a long exact sequence

$$0 \to \hom_{\mathscr{A}}(A, B) \to \hom_{\mathscr{A}}(A, E) \to \hom_{\mathscr{A}}(A, A) \xrightarrow{\delta} \mathrm{Ext}_R^1(A, B)$$

We will associate the extension to $\delta(1) \in \text{Ext}^1_(A, B)$, where $1 \in \hom_\mathscr{A}(A, A)$ is the identity map. The fact that the long exact sequence is *natural* means that an equivalence of extensions gives rise to a commutative diagram

$$
\begin{array}{ccccc}
\hom_\mathscr{A}(A, E_1) & \longrightarrow & \hom_\mathscr{A}(A, A) & \xrightarrow{\ \delta\ } & \text{Ext}^1_(A, B) \\
\downarrow & & \| & & \| \\
\hom_\mathscr{A}(A, E_2) & \longrightarrow & \hom_\mathscr{A}(A, A) & \xrightarrow{\ \delta\ } & \text{Ext}^1_(A, B)
\end{array}
$$

so equivalent extensions give rise to the same element of $\text{Ext}^1_R(A, B)$.

Given $x \in \text{Ext}^1_R(A, B)$ and an injective resolution for B, I^*, represent x by a homomorphism $x\colon A \to I^1$ whose image lies in the kernel of $\delta^1\colon I^1 \to I^2$. This means it is in the image of I^0 which is isomorphic to I^0/B. We get a commutative diagram

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & B & \xrightarrow{\ r\ } & E_1 & \xrightarrow{\ s\ } & A & \longrightarrow & 0 \\
& & \| & & v\downarrow & & \downarrow x & & \\
0 & \longrightarrow & B & \xrightarrow{\ r'\ } & I^0 & \xrightarrow{\ s'\ } & I^0/B & \longrightarrow & 0
\end{array}
$$

inducing a diagram

$$
\begin{array}{ccccccc}
\hom_\mathscr{A}(A, E_1) & \longrightarrow & \hom_\mathscr{A}(A, A) & \xrightarrow{\ \delta\ } & \text{Ext}^1_R(A, B) & \longrightarrow & \text{Ext}^1_R(A, E_1) \\
\downarrow & & \downarrow \hom_\mathscr{A}(1,x) & & \| & & \downarrow \\
\hom_\mathscr{A}(A, I^0) & \longrightarrow & \hom_\mathscr{A}(A, I^0/B) & \xrightarrow{\ \delta\ } & \text{Ext}^1_R(A, B) & \longrightarrow & \text{Ext}^1_R(A, I^0)
\end{array}
$$

Since $\text{Ext}^1_R(A, I^0) = 0$, it is clear that the identity map of A maps to x. The proof that split exact sequences give 0 is left to the reader.

Chapter D, D.3 Exercise 1 (p. 560) This follows immediately from the second statement of theorem D.3.22 on page 558: simply take n_0 to be the maximum of the n_0's that come from theorem D.3.22 on page 558 and recall the meaning of cohomology *vanishing* (see the discussion in section D.1.2 on page 526 and the remarks following definition D.3.2 on page 546).

Chapter D, D.3 Exercise 2 (p. 560) The "modus operandi" of *topological* cohomology (as opposed to sheaf cohomology) can be summarized by:

> Break a topological space into tiny, "elementary" fragments (here "elementary"=contractible) and analyze how they fit together to form the space. If the fragments are not contractible, treat them as though they are (i.e., give them vanishing cohomology in positive dimensions).
>
> In simplicial cohomology, the fragments are polyhedra called simplices, and in the Čech form, they are tiny contractible open sets.

When we use this apparatus on *schemes* with the Zariski topology, a problem becomes evident: no tiny fragments exist! For instance the smallest open set in an irreducible affine scheme containing a point is the *entire scheme itself!*

The *rigorous* proof of this statement uses corollary D.3.4 on page 547 and theorem D.3.15 on page 553.

Chapter D, D.4 Exercise 1 (p. 565) This follows immediately from exercise D.4.6 on page 563 and the fact that

$$
\mathscr{H}om(\mathcal{O}_X, \mathscr{F})(X) = \hom(\mathcal{O}_X, \mathscr{F})
$$

Chapter D, D.4 Exercise 2 (p. 565) This follows immediately from exercise 1 on page 565 — since $\hom(\mathcal{O}_X, \mathscr{F}) = \mathscr{F}(X)$, the functors $\hom(\mathcal{O}_X, *)$ and $e(*)$ are the same (where e is the functor defined by $e(\mathscr{F}) = \mathscr{F}(X)$), they have the same derived functors.

Chapter D, D.4 Exercise 3 (p. 566) We can define a map

$$\mathscr{F} \to \mathscr{F}^{\vee\vee}$$

by sending

$$\mathscr{F}|U \mapsto (g\colon \mathscr{F}|U \mapsto \mathcal{O}_V|U) \mapsto g(\mathscr{F}|U)$$

Since $\mathscr{F}$ is locally free its stalks are all free of the same rank, and the stalks of $\mathscr{F}^{**}$ will have the property that

$$\mathscr{F}^{**}_x = (\mathscr{F}_x)^{**}$$

which are isomorphisms if they are free modules of finite rank (or finite dimensional vector spaces). The conclusion follows from exercise 5 on page 502. To prove the second statement, note that there is a natural map

$$\begin{aligned}
\mathscr{F}^* \otimes_{\mathcal{O}_X} \mathscr{G} &\to \mathscr{H}om(\mathscr{F}, \mathscr{G}) \\
f \otimes g &\mapsto f \cdot g
\end{aligned}$$

where $f \in \mathscr{F}^{\vee} = \mathscr{H}om(\mathscr{F}, \mathcal{O}_X)$ has its value in $\mathcal{O}_X$, which acts on g. If we restrict to stalks, $\mathscr{F}^*_x$ is a free module and the restriction is an isomorphism. Again, the conclusion follows from exercise 5 on page 502.

Chapter D, D.4 Exercise 4 (p. 566) Proposition A.5.51 on page 466 implies that, for R-modules A, B, C

$$\hom_R(A, \hom_R(B, C)) = \hom_R(A \otimes_R B, C)$$

Since this is a natural equality, it commutes with all restriction maps and defines an isomorphism

$$\hom(\mathscr{A}, \hom(\mathscr{B}, \mathscr{C})) \cong \hom(\mathscr{A} \otimes_{\mathcal{O}_X} \mathscr{B}, \mathscr{C})$$

of modules over $\mathcal{O}_X$. Applying this to all *restrictions* of the sheaves implies the second isomorphism.

Chapter D, D.4 Exercise 5 (p. 566) This follows immediately from exercises 4 on page 566 and 3 on page 566.

Chapter D, D.4 Exercise 6 (p. 566) Exactness of the original sequence simply means that

$$\widehat{\operatorname{im}\mathscr{F}_i} = \ker \mathscr{F}_{i+1} \to \mathscr{F}_{i+2}$$

for all i, where $\widehat{\ast}$ denote the completion-operation described in lemma B.2.3 on page 500 when evaluated on any open set. The only thing that prevents the sequence of global evaluations from being exact is this completion-operation. Since n is finite, corollary 5.3.9 on page 237 implies that there exists an in integer d such that $\mathscr{F}_i \otimes_{\mathcal{O}_V} \mathcal{O}_V(t)$ is generated by global sections, for all i and all $t \geq d$, Now exercise 1 on page 242 implies that im $\mathscr{F}_i \otimes_{\mathcal{O}_V} \mathcal{O}_V(t)$ is already a sheaf (no completion needed). The conclusion follows.

Glossary

⊚   The zero-element in an elliptic curve, regarded as an algebraic group. See 6.2.12 on page 320.

$\mathscr{A}b$ The category of abelian groups.

diffeomorphic   Two topological spaces, X and Y, are *diffeomorphic* if there exist smooth maps $f\colon X \to Y$ and $g\colon Y \to X$ whose composites are the *identity maps* of the two spaces. Note the similarity to homeomorphism.

$\mathbb{A}^n$ This is n-dimensional affine space. See chapter 2 on page 35.

$\mathcal{V}(\mathfrak{a})$ The affine variety defined by an ideal. See definition 2.5.4 on page 73.

homeomorphic   Two topological spaces, X and Y, are *homeomorphic* if there exist continuous maps $f\colon X \to Y$ and $g\colon Y \to X$ whose composites are the *identity maps* of the two spaces.

$\operatorname{Ann}(M)$ The annihilator of a module. See definition A.1.73 on page 374.

$\operatorname{ann}(m)$ The annihilator of an element of a module. See definition A.1.73 on page 374.

$\mathcal{S}M$ The symmetric algebra generated by a module. See definition A.6.4 on page 475.

$\operatorname{Assoc}(R)$ The set of associated primes of a ring R. See definition A.1.73 on page 374.

$\mathbb{B}_{\mathfrak{a}}(V)$ The blow-up of a variety via an ideal. See definition 5.5.8 on page 252.

$\mathbb{C}$ The field of complex numbers.

$\operatorname{Cart}(*)$ The group of Cartier divisors on a variety. See 5.9.13 on page 290.

$\mathscr{C}h$ The category of chain-complexes. See definition D.1.1 on page 519.

$\mathscr{C}o$ The category of cochain-complexes. See definition D.1.2 on page 520.

$\check{H}(V)$ The Čech cohomology of a variety. See definition D.3.12 on page 551.

$\operatorname{coker} f$ The cokernel of a homomorphism. See definition A.1.63 on page 369.

$\underline{A}$ If A is an object, $\underline{A}$ is the constant sheaf equal to A on every open set.

dense subset   A subset $S \subset X$ of a topological space is *dense* if, for any open set $U \subset X$, $S \cap U \neq \varnothing$.

$\operatorname{depth} M$ The depth of a module. See definition D.2.14 on page 539.

$\operatorname{Der}_k(R, M)$ The module of derivations. See definition A.7.2 on page 482.

Ω^i_V The sheaf of regular differential forms of degree i. See proposition 4.6.19 on page 210.

$\Omega^i(V)$ The vector space of rational differential forms of degree i. See definition 5.9.34 on page 299.

$\operatorname{Cl}(V)$ The group divisor-classes of a variety. See definition 5.9.4 on page 286.

$\operatorname{Div}(V)$ The group of Weil divisors of a variety. See definition 5.9.1 on page 284.

$\langle W \rangle$ A codimension-1 subvariety, regarded as a divisor. See section 5.9 on page 284.

$\mathscr{F}^{\vee}$ The dual of a sheaf of modules. See equation D.4.5 on page 566.

$\mathrm{Ext}_R^i(A, B)$ The Ext-functor — see section D.2.1 on page 534.

$\mathscr{E}xt^i(\mathscr{F}, \mathscr{G})$ The sheaf-version of the Ext-functor — see definition D.4.1 on page 561.

$\Lambda^i M$ The ith exterior power of a module, M. See definition A.6.7 on page 476.

$\mathbb{F}_{p^n}$ The unique finite field of order p^n, where p is a prime number. See section A.2.5 on page 403.

$x^{\underline{n}}$ The falling factorial or Pochhammer symbol, defined by $x^{\underline{n}} = x(x - 1) \cdots (x - n + 1)$ for $n \geq 1$.

$\bar{\mathbb{F}}_p$ The algebraic closure of the field $\mathbb{F}_p$. See theorem A.2.42 on page 405.

$\mathcal{F}_p$ The Frobenius homomorphism of a finite field of characteristic p. See definition A.2.40 on page 404.

gl-dim R The global dimension of a ring. See definition D.2.9 on page 537.

Γ_f The graph of a regular morphism of schemes or varieties. See 4.6.5 on page 206.

$\mathbb{G}_n(V)$ The Grassmann variety whose points are n-dimensional subspaces of a vector-space V. See definition 5.2.10 on page 230.

$\mathbb{G}_{n,m}$ This is just $\mathbb{G}_n(k^m)$ — i.e., we are not interested in the functorial properties of the vector-space.

ht($\mathfrak{p}$)   The height of a prime ideal in a ring. See definition 2.8.7 on page 100.

$h_R(n)$ The Hilbert function of a graded ring. See definition 5.7.1 on page 267.

$\mathcal{H}_R(n)$ The Hilbert polynomial of a graded ring. For n sufficiently large $\mathcal{H}_R(n) = h_R(n)$.

$\hom_{k-\mathrm{alg}}(*, *)$ Set of homomorphisms of affine k-algebras. See definition 2.4.13 on page 67.

$\hom_{\mathrm{Scheme}}(*, *)$ The set of morphisms between two schemes. See definition 4.6.36 on page 217.

$\mathscr{H}om(\mathscr{A}, \mathscr{B})$ The sheaf of homomorphisms between sheaves. See definition B.3.4 on page 504.

$\mathcal{I}(X)$ The ideal of polynomials vanishing on a set of points. See definition 2.1.4 on page 38.

inj-dim M The injective dimension of a module. See definition D.1.12 on page 526.

$\mathfrak{J}(R)$ The Jacobson radical of a ring. See definition A.4.21 on page 427.

$k_H[V]$ The homogeneous coordinate ring of a projective variety — see definition 5.1.7 on page 222.

$k[[X]]$ Power series ring, defined in definition A.1.5 on page 343.

$k\{\{X\}\}$ Field of Puiseux series, defined in example A.4.35 on page 432.

$\Omega_{R/k}$ The module of Kaehler differentials. See definition A.7.5 on page 483.

$\mathbf{K}(x_1, \ldots, x_n)$ The Koszul cochain complex associated to a sequence of elements of a ring. See definition D.2.18 on page 540.

$\ell(D)$ This is dimension of the space of a divisor, D. See definition 5.9.27 on page 296.

$\mathcal{L}(D)$ This is the space of a divisor, D. See definition 5.9.27 on page 296.

$\ell^S(D)$ This is dimension of the relative space of a divisor, D. See definition 6.1.10 on page 309.

$\mathcal{L}^S(D)$ This is the relative space of a divisor, D. See definition 6.1.10 on page 309.

$\varinjlim A_n$ The direct limit of a sequence of objects and morphisms. See definition A.5.17 on page 448.

$\varprojlim R_n$ The inverse limit of a sequence of objects and morphisms. See definition A.5.26 on page 453.

$\mathrm{LT}(*)$ Leading monomial of a polynomial with respect to some ordering. Used in computing Gröbner bases. See section 2.3 on page 45.

$\mathbf{A}(f)$ The algebraic mapping cone of a chain-map, f. See definition D.1.11 on page 524.

$\mathcal{K}_X^*$ The sheaf of nonzero meromorphic functions. See definition 4.3.21 on page 178

$\mathcal{K}_X$ The sheaf of meromorphic functions on a ringed space. See definition 4.3.21 on page 178.

$\backslash$ A difference between *sets*, so $A \backslash B$ is the elements of A that are not contained in B.

$\mathscr{M}_R$ The category of modules over a ring, R. See statement 7 on page 441.

$\hom_C(A, B)$ The set of morphisms between objects of a category C. See definition A.5.2 on page 439.

$\hom_{\mathrm{Var}}(A, B)$ The set of regular maps of algebraic varieties. See definition 2.2.10 on page 44.

$\mathfrak{N}(R)$ The nilradical of a ring. See definition A.1.46 on page 360.

$\mathrm{ord}_W(f)$ The order of a function at a divisor. See definition 5.9.3 on page 286.

$p_g(X)$ The geometric genus of a variety. See definition 5.9.40 on page 303.

$\mathrm{Pic}(V)$ The Picard group of line-bundles on a variety. See definition C.1.13 on page 513.

$\mathrm{Proj}\, R$ The Proj-construction. See definition 4.4.12 on page 190.

$\mathbb{C}P^n$ Complex projective space. See definition 1.2.2 on page 4.

$k\mathbb{P}^n$ Projective space over a field k.

$\mathrm{proj\text{-}dim}(M)$ The projective dimension of a module. See definition D.1.13 on page 526.

$\mathbb{P}(V)$ Projective space derived from a vector space. This is essentially the same as $k\mathbb{P}^{\dim V}$, but we are concerned with the functorial properties of V,

$\mathcal{P}(\mathfrak{a})$ Closed subscheme of projective space defined by a homogeneous ideal, $\mathfrak{a}$. See definitions 4.4.15 on page 191 and 5.1.7 on page 222.

$R\mathbb{P}^n$ A scheme that is like a projective space over a ring R. See definition 4.4.8 on page 186.

$\mathbb{R}P^n$ real projective space. See definition 1.2.2 on page 4.

$\mathbb{Z}P^n$ This is just $R\mathbb{P}^n$ with the ring equal to $\mathbb{Z}$.

$\mathbb{Q}$ The field of rational numbers.

quasicompact A topological space with the property that every open cover has a finite subcover. This is very similar to the definition of compact, but the Bourbaki people require a compact space to be Hausdorff.

$\mathbb{R}$ The field of real numbers.

$R^i F$ Right derived functors of F. See definition D.1.17 on page 528.

$\sqrt{*}$ Radical of an ideal. See definition A.1.43 on page 359.

$\mathrm{rank}(A)$ Rank of a matrix, A.

$\Gamma(\mathfrak{a})$ The Rees algebra of a ring with respect to an ideal. See definition A.4.43 on page 436.

$\mathrm{Res}(f,g)$ Resultant of two polynomials. See definition 1.3.2 on page 13.

$\mathbf{s}(D)$ The invertible sheaf associated with a Cartier divisor. See proposition 5.9.20 on page 292.

$\mathcal{O}_V(m)$ If V is a projective space and m is an integer, this denotes the Serre twist of degree m — see definition 5.3.1 on page 234.

$(\mathfrak{a}\colon\mathfrak{b}^\infty)$ The saturation of one ideal with respect to another. See exercise 9 on page 71.

$\mathscr{S}_\xi$ The sheaf of sections of a vector-bundle ξ. See definition C.2.1 on page 514.

$\Gamma(V,\xi)$ The sections of a vector bundle over a space V. See definition C.1.10 on page 511. Also occasionally used when ξ is a sheaf, in which case $\Gamma(V,\xi) = \xi(V)$

$\mathcal{A}(M)$ The Serre functor associating a sheaf to a module over the ring of regular functions. See definition 3.5.1 on page 155.

$\mathcal{O}_V$ The sheaf of regular functions on a variety or scheme. See section 4.3.1 on page 169.

$\mathrm{Spec}\,R$ The affine scheme defined by a ring R. See definition 4.2.1 on page 163.

$\mathrm{specm}\,A$ The maximum spectrum of a ring. See section 2.5 on page 71.

$\mathrm{Supp}\,\mathscr{F}$ The support of a sheaf of modules. See definition B.1.5 on page 497.

$T_{V,p}$ The tangent space of a variety at a point. See definition 3.3.2 on page 119.

$\mathrm{Tor}_R^i(A,B)$ The Tor-functor — see section D.2.1 on page 534.

$\mathrm{Trace}\,A$ The trace of a square matrix (the sum of its diagonal elements).

$R^\times$ The group of units of a ring or field.

$\mathbb{Z}$ The ring of integers.

$\mathbb{Z}_{(p)}$ p-adic numbers. See example A.5.30 on page 456.

Index

Bibliography

[1] Niels Abel. *Œuvres complètes*. Christiania, 1881.

[2] Selman Akbulut and Henry C. King. Real algebraic variety structures on p.l. manifolds. *Bull. Amer. Math. Soc.*, 83:281–281, 1977.

[3] Emil Artin. Zur Arithmetik hyperkomplexer Zahlen. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 5:261–289, 1928.

[4] Emil Artin. Zur Theorie der hyperkomplexer Zahlen. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 5:251–260, 1928.

[5] W. Arveson. *An Invitation to C*-Algebras*. Springer-Verlag, 1981.

[6] Michael Francis Atiyah. *K-theory*. Advanced Book Classics. Addison-Wesley, 2nd edition, 1989.

[7] Maurice Auslander and David A. Buchsbaum. Homological dimension in local rings. *Transactions of the American Mathematical Society*, 85:390–405, 1957.

[8] Maurizio Avellone, Aldo Brigaglia, and Carmela Zappulla. The foundations of projective geometry in Italy from De Paolis to Pieri. *Archive for History of Exact Sciences*, 56:363–425, 2002.

[9] Gorô Azumaya. On maximally central algebras. *Nagoya Math. J.*, 2:119–150, 1951. Available online from http://projecteuclid.org.

[10] R. Baer. Abelian groups that are direct summands of every containing abelian group. *Bull. Amer. Math. Soc.*, 46:800–806, 1940.

[11] John Baez. Higher-Dimensional Algebra II: 2-Hilbert Spaces. *Adv. Math.*, 127:125–189, 1997. available from arXiv:q-alg/9609018v2.

[12] Alan Baker. *Transcendental Number Theory*. Cambridge University Press, 1975.

[13] W. W. Rouse Ball. *A short account of the history of mathematics*. Project Gutenberg, http://www.gutenberg.org, e-book number 31246 edition, 2010.

[14] Elaine Barker, Don Johnson, and Miles Smid. Recommendation for Pair-Wise Key Establishment Schemes Using Discrete Logarithm Cryptography. Technical Report NIST Special Publication 800-56A, National Insitute of Standards and Technology (NIST), 2007.

[15] Étienne Bézout. Sur le degré des équations résultantes de l'évanouissement des inconnues et sur les moyens qu'il convient d'employer pour trouver ces équations. *Histoire de l'académie royale des sciences*, pages 288–338, 1764.

[16] Étienne Bézout. *Theorie général des équations algébrique*. Paris, 1770.

[17] Étienne Bézout. *General Theory of Algebraic Equations*. Princeton University Press, 2006. Translated by Eric Feron.

[18] Jacek Bochnak, Michel Coste, and Marie-Francoise Roy. *Real Algebraic Geometry*, volume 3 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. A Series of Modern Surveys in Mathematics*. Springer, 1998.

[19] Alexander Brill and Max Noether. Über die algebraischen Functionen und ihre Anwendungen in der Geometrie. *Math. Ann.*, 7:293–310, 1874.

[20] R.L. Bryant, S.S. Chern, R.B. Gardner, H.L. Goldschmidt, and P.A. Griffiths. *Exterior differential systems*. Springer, 1991.

[21] Bruno Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal*. PhD thesis, Johannes Kepler University of Linz (JKU), RISC Institute., 1965.

[22] Bruno Buchberger. Some properties of Gröbner bases for polynomial ideals. *ACM SIGSAM Bulletin*, 10:19–24, 1976.

[23] Bruno Buchberger. A criterion for detecting unnecessary reductions in the construction of Gröbner bases. In *Proceedings of the International Symposium on Symbolic and Algebraic Manipulation (EUROSAM '79)*, 1979.

[24] Paola Cantù. *Giuseppe Veronese e i fondamenti della geometria*, volume 10 of *Biblioteca di cultura filosofica*. Unicopli, 1999.

[25] Dustin Cartwright and Bernd Sturmfels. The number of eigenvalues of a tensor. arXiv:1004.4953. Linear Algebra and its Applications, special issue on Tensors and Multilinear Algebra.

[26] Dustin Cartwright and Bernd Sturmfels. The number of eigenvalues of a tensor. *Linear Algebra and its Applications*, 2011. In press.

[27] Eduard Čech. Théorie générale de l'homologie dans un espace quelconque. *Fund. Math.*, 19:149–183, 1932.

[28] Certicom Research. *SEC 2: Recommended Elliptic Curve Domain Parameters*, 2000.

[29] Leonard S. Charlap and Raymond Coley. An Elementary Introduction to Elliptic Curves II. Technical report, IDA Center for Communications Research - Princeton, July 1990. URL: http://www.idaccr.org/reports/reports.html.

[30] Leonard S. Charlap and David P. Robbins. An Elementary Introduction to Elliptic Curves. Technical report, IDA Center for Communications Research - Princeton, December 1988. URL: http://www.idaccr.org/reports/reports.html.

[31] Ciro Ciliberto, Antony V. Geramita, Brian Harbourne, Rosa Maria Miró-Roig, and Kristian Ranestad, editors. *Projective varieties with unexpected properties: a volume in memory of Giuseppe Veronese*, Proceedings in Mathematics. De Gruyter, 2004.

[32] Ciro Ciliberto, Friedrich Hirzebruch, Rick Miranda, and Mina Teicher. *Applications of Algebraic Geometry to Coding Theory, Physics and Computation*. NATO Science Series II: Mathematics, Physics and Chemistry. Springer, 2001.

[33] Thierry Coquand and Henri Lombardi. A short proof for the Krull dimension of a polynomial ring. *American Math. Monthly.*, 112:826–829, 2005.

[34] Jesús A. De Loera, Bernd Sturmfels, and Cynthia Vinzant. The central curve in linear programming. arXiv:1012.3978.

[35] P.G.L. Dirichlet and R. Dedekind. *Lectures on Number Theory*. American Mathematical Society, 1999. translated by John Stillwell. Original scanned version available online at http://gdz.sub.uni-goettingen.de.

[36] B. Eckmann and A. Schopf. Über injektive Modulen. *Archiv der Math.*, 3:75–78, 1953.

[37] Samuel Eilenberg and Saunders MacLane. General theory of natural equivalences. *Trans. of the Amer. Math. Soc*, 58(2):231–294, 1945. available from Google Scholar.

[38] David Eisenbud. *Commutative Algebra: with a View Toward Algebraic Geometry*. Springer, 1995.

[39] David Eisenbud, Daniel Grayson, Michael Stillman, and Berndt Sturmfels, editors. *Computations in algebraic geometry with Macaulay 2*, volume 8 of *Algorithms and Computations in Mathematics*. Springer, 2001. Can be downloaded from http://www.math.uiuc.edu/Macaulay2/Book/.

[40] R. Engelking. *Theory of Dimensions. Finite and Infinite*. Heldermann Verlag, 1995.

[41] Arno van den Essen. *Polynomial Automorphisms: and the Jacobian Conjecture*. Progress in Mathematics. Birkhäuser Basel, 2000.

[42] Henri Cartan et al., editor. *Séminaire de Topologie algébrique*, 1948-49.

[43] Euclid. *The Thirteen Books of the Elements, Vol. 2*. Dover Publications, 1956.

[44] Euclid. *The Thirteen Books of the Elements, Vol. 3*. Dover Publications, 1956.

[45] Leonhard Euler. *Introductio in analysin infinitorum*. Berlin Academy, 1748.

[46] Howard W. Eves. *Mathematical Circles Adieu and Return to Mathematical Circles*. The Mathematical Association of America, 2002.

[47] Gerd Faltings. Diophantine approximation on abelian varieties. *Annals of Mathematics*, 1991:549–576, 133.

[48] Edward FitzGerald. *The Rubaiyat of Omar Khayyam*. Number 246 in ebook. Project Gutenberg, http://www.gutenberg.org/ebooks/246, 1995.

[49] William Fulton. Algebraic curves.
URL: http://www.math.lsa.umich.edu/ wfulton/CurveBook.pdf.

[50] William Fulton and Joe Harris. *Representation theory. A first course*, volume 129 of *Graduate Texts in Mathematics*. Springer, 1991.

[51] Évariste Galois. Sur la théorie des nombres. *Bulletin des Sciences mathématiques, physiques et chimiques*, XIII:428–435, 1830.

[52] Johann Carl Friedrich Gauss. *Disquisitiones Arithmeticae*. Yale University Press, 1965. Translated by Arthur A. Clarke. Original available online at http://resolver.sub.uni-goettingen.de/purl?PPN235993352.

[53] Walter Gautschi. Leonhard Euler: His Life, the Man, and His Works. *SIAM Review*, 2008.

[54] I. M. Gelfand and M. A. Naimark. Normed rings with involutions and their representations. *Izv. Akad. Nauk SSSR Ser. Mat.*, 12(5):445–480, 1948.

[55] A. O. Gelfond. *Transcendental and Algebraic Numbers*. Dover Phoenix Editions. Dover Publications, 2003.

[56] Sophie Germain. *Recherches sur la théorie des surfaces élastiques*. Courcier, 1821. Available online from http://books.google.com.

[57] Larry J. Gerstein. A new algorithm for computing the rank of a matrix. *The American Mathematical Monthly*, 95:950–952, 1988.

[58] Paul Gordan. Neuer Beweis des Hilbertschen Satzes über homogene Funktionen. *Nachrichten König. Ges. der Wiss. zu Gött.*, 3:240–242, 1899.

[59] Ronald Graham, Donald Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison–Wesley, 1994.

[60] Hermann Grassmann. *Die Lineale Ausdehnungslehre — Ein neuer Zweig der Mathematik*. Otto Wigand Verlag, 1844. Available from Google Books.

[61] R. Grone. Decomposable tensors as a quadratic variety. *Proc. Amer. Math. Soc.*, 64:227–230, 1977.

[62] Alexander Grothendiek. Théorèms de dualité pour faisceaux algébriques. Sémainaire Bourbaki, 1957.

[63] Alexander Grothendiek. *Éléments de géométrie algébrique (rédigés avec la collaboration de Jean Dieudonné) : I. Le langage des schémas*, volume 4 of *Publications Mathématiques*. Institut des Haute Études Scientifiques, 1960. Available from http://www.numdam.org/.

[64] Alexander Grothendiek. *Éléments de géométrie algébrique (rédigés avec la collaboration de Jean Dieudonné) : II. Étude globale élémentaire de quelques classes de morphismes*, volume 8 of *Publications Mathématiques*. Institut des Haute Études Scientifiques, 1960. available from http://www.numdam.org/.

[65] Alexander Grothendiek. *Éléments de géométrie algébrique (rédigés avec la collaboration de Jean Dieudonné) : III. Étude cohomologique des faisceaux cohérents, Première partie*, volume 11 of *Publications Mathématiques*. Institut des Haute Études Scientifiques, 1961. Available from http://www.numdam.org/.

[66] Alexander Grothendiek. *Éléments de géométrie algébrique (rédigés avec la collaboration de Jean Dieudonné) : III. Étude cohomologique des faisceaux cohérents, Seconde partie*, volume 16 of *Publications Mathématiques*. Institut des Haute Études Scientifiques, 1963. Available from http://www.numdam.org/.

[67] Alexander Grothendiek. *Éléments de géométrie algébrique (rédigés avec la collaboration de Jean Dieudonné) : IV. Étude locale des schémas et des morphismes de schémas. Quatrième partie*, volume 32 of *Publications Mathématiques*. Institut des Haute Études Scientifiques, 1967. Available from http://www.numdam.org/.

[68] Wolfgang Gröbner. Über die Eliminationstheorie. *Monatshefte für Mathematik*, 54:71–78, 1950.

[69] Robin Hartshorne. *Residues and duality, Lecture Notes of a Seminar on the Work of A. Grothendieck*, volume 20 of *Lecture Notes in Mathematics*. Springer, 1966/1967.

[70] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer, 1977.

[71] Robin Hartshorne. *Foundations of Projective Geometry*. Ishi Press, 2009.

[72] Allen Hatcher. Vector bundles & K-theory. Available online from http://www.math.cornell.edu/ hatcher/VBKT/VBpage.html.

[73] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[74] Herwig Hauser. The Hironaka theorem on resolution of singularities (Or: A proof we always wanted to understand). *Bull. Amer. Math. Soc.*, 40:323–403, 2003.

[75] Kurt Hensel. Über eine neue Begründung der Theorie der algebraischen Zahlen. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 6:83–88, 1897. Available online from http://www.digizeitschriften.de.

[76] Ivan Herman. *The Use of Projective Geometry in Computer Graphics*, volume 564 of *Lecture Notes in Computer Science*. Springer-Verlag, 1992.

[77] David Hilbert. Über die Theorie von algebraischen Formen. *Math. Ann.*, 36:473–534, 1893.

[78] Heisuke Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero. I. *Ann. of Math.*, 79:109–203, 1964.

[79] Heisuke Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero. II. *Ann. of Math.*, 79:205–326, 1964.

[80] Friedrich Hirzebruch. *Topological Methods in Algebraic Geometry*. Classics in Mathematics. Springer, 1995.

[81] Witold Hurewicz and Henry Wallman. *Dimension Theory*, volume 4 of *Princeton Mathematical Series*. Princeton University Press, 1996 (Revised edition).

[82] Lars Hörmander. *An Introduction to Complex Analysis in Several Variables*. North-Holland Publishing Company, 1973.

[83] Maurice Janet. Les systèmes d'équations aux dérivées partielles. *Journal de Mathématiques Pures et. Appliquées,*, 3:65, 1920.

[84] Daniel Kan. Adjoint functors. *Trans. Amer. Math. Soc.*, 87:294–329, 1958.

[85] I. Kaplansky. *Commutative Rings*. Allyn and Bacon, Inc., Boston, MA., 1970.

[86] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM Symposium on the Theory of Computing - STOC '84*, pages 302–311. ACM, 1984.

[87] Steven Kleiman. Misconceptions about K_X. *L'Enseignement Mathématique*, 25:203–206, 1979. Can be downloaded from http://www.unige.ch/math/EnsMath/EM_fr/welcome.html.

[88] Felix Klein. *Elementary Mathematics from an Advanced Standpoint: Geometry*. Dover Publications, 2004.

[89] M. Kline. *Mathematical Thought from Ancient to Modern Times*, volume 1. Oxford University Press, 1972.

[90] Neal Koblitz. Elliptic curve cryptosystems. *Mathematics of Computation*, 48:203–209, 1987.

[91] Helmut Koch. *Number theory: algebraic numbers and functions*. American Mathematical Society, 2000.

[92] L. Kronecker. *Grundzüge einer arithmetischen Theorie der algebraischen Grössen*. G. Reimer, 1882. Available online from Google Books.

[93] Wolfgang Krull. Über eine Hauptsatz der allgemeinen Idealtheorie. *Sitzungsberichte Heidelberg Ak. Wiss.*, pages 11–16, 1929.

[94] Wolfgang Krull. Allgemeine Bewertungstheorie. *J. Reine Angew. Math.*, 167:160–196, 1932.

[95] Wolfgang Krull. *Idealtheorie*. Springer, 1935. Second edition 1968.

[96] Wolfgang Krull. Beiträge zur Arithmetik kommutativer Integritätsbereiche III. Zum Dimensionsbegriff der Idealtheorie. *Math. Z*, 42:745–766, 1937.

[97] Wolfgang Krull. Beiträge zur Arithmetik kommutativer Integritätsbereiche. *Math. Z.*, 43:768–782, 1938.

[98] Wolfgang Krull. Dimensionstheorie in Stellenringen. *J. Reine Angew. Math.*, 179:204–226, 1938.

[99] Wolfgang Krull. Jacobsonsche Ringe, Hilbertscher Nullstellensatz, Dimensionstheorie. *Math. Z.*, 54:354–387, 1951.

[100] Casimir Kuratowski. Une méthode d'élimination des nombres transfinis des raisonnements mathématiques. *Fundamenta Mathematicae*, 3:76–108, 1922.

[101] Y. S. Kwoh, J. Hou, E. A. Jonckheere, and S. Hayall. A robot with improved absolute positioning accuracy for CT guided stereotactic brain surgery. *IEEE Trans. Biomed. Engng.*, 35:153–161, 1988.

[102] Serge Lang. *Algebra*. Graduate Texts in Mathematics. Springer-Verlag, 2002.

[103] Steven M. LaValle. *Planning Algorithms*. Cambridge University Press, 2006.

[104] John M. Lee. *Introduction to Topological Manifolds*, volume 202 of *Graduate Texts in Mathematics*. Springer, 2000.

[105] Lek-Heng Lim. Singular values and eigenvalues of tensors: a variational approach. In *Proc. IEEE Internat. Workshop on Comput. Advances in Multi-Sensor Adaptive Processing*, pages 129–132, 2005.

[106] Saunders MacLane. *Categories for the Working Mathematician*. Graduate Texts in Mathematics. Springer, second edition, 2010.

[107] Collin MacLauren. *Geometria organica sive descripto linearum curverum universalis*. London, 1720.

[108] Ernst W. Mayr. Some complexity results for polynomial ideals. *Journal of Complexity*, 13:301–384, 1997.

[109] John Theodore Merz. *A history of European thought in the nineteenth century*. W. Blackwood and sons, 1912. Available online at
http://www.archive.org/details/ahistoryeuropea06merzgoog.

[110] Victor Miller. Use of elliptic curves in cryptography. In *Advances in Cryptology — CRYPTO '85 Proceedings*, volume 218 of *Lecture Notes in Computer Science*, pages 417–426. Springer, 1985.

[111] James S. Milne. Abelian varieties. URL: http://www.jmilne.org/math.

[112] James S. Milne. *Étale cohomology*, volume 33 of *Princeton Mathematical Series*. Princeton University Press, 1980.

[113] James S. Milne. *Elliptic Curves*. BookSurge Publishers, 2006.

[114] John Milnor. *Introduction to Algebraic K-Theory*. Princeton University Press, 1972.

[115] John W. Milnor and James D. Stasheff. *Characteristic classes*, volume 76 of *Annals of Mathematics Studies*. Princeton University Press, 1974.

[116] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. Physics. W. H. Freeman, 1973.

[117] Teo Mora. An algorithm to compute the equations of tangent cones. In *Proc. EUROCAM 82*, number 144 in Lecture Notes in Computer Science, pages 158–165. Springer, 1982.

[118] Teo Mora. An algorithmic approach to local rings. In *Proc. EUROCAL 85*, volume 204 of *Lecture Notes in Computer Science*, pages 518–525. Springer, 1985.

[119] Louis Mordell. On the rational solutions of the indeterminate equations of the third and fourth degrees. *Proc. Cam. Phil. Soc.*, 21:179, 1922.

[120] David Mumford. *Abelian Varieties*. Tata Institute of Fundamental Research, 1970.

[121] David Mumford. *Algebraic Geometry I. Complex projective varieties*. Springer, 1976.

[122] David Mumford. *The Red Book of Varieties and Schemes*, volume 1358 of *Lecture Notes in Mathematics*. Springer-Verlag, second expanded edition, 1999.

[123] Trygve Nagell. Sur les propriétés arithmétiques des cubiques planes du premier genre. *Acta Mathematica*, 52:93–126, 1929.

[124] Tadashi Nakayama. A remark on finitely generated modules, II. *Nagoya Math. J.*, 9:21–23, 1955. Available online from http://projecteuclid.org.

[125] John Nash. Real algebraic manifolds. *Annals of Mathematics, Second Series,*, 56:405–421, 1952. Available online at http://www.jstor.org/pss/1969649.

[126] Dwight E. Neuenschwander. *Emmy Noether's Wonderful Theorem*. The Johns Hopkins University Press, 2010.

[127] Isaac Newton. *The correspondence of Isaac Newton. II.* Cambridge University Press, 1960. Letter to Oldenburg dated Oct 24, 1676.

[128] Emmy Noether. Idealtheorie in Ringbereichen. *Mathematische Annalen*, 83:24–66, 1921.

[129] Emmy Noether. Der Endlichkeitsatz der Invarianten endlicher linearer Gruppen der Charakteristik p. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, 1926:28–35, 1926. Available online from http://gdz.sub.uni-goettingen.de.

[130] Emmy Noether. Abstrakter Aufbau der Idealtheorie in algebraischen Zahl und Funktionenkörpern. *Math. Ann.*, 92:26–61, 1927.

[131] Krzysztof Jan Nowak. A simple proof of Puiseux's theorem. *Univ. Iagel. Acta. Math.*, 32:199–200, 1995.

[132] Christof Paar, Jan Pelzl, and Bart Preneel. *Understanding Cryptography: A Textbook for Students and Practitioners*. Springer, 2010.

[133] Lior Pachter and Bernd Sturmfels. The mathematics of phylogenomics. *SIAM Review*, 49:3–31, 2007.

[134] Richard S. Palais. *Seminar on the Atiyah–Singer Index Theorem*, volume 57 of *Annals of Mathematics Studies*. Princeton Univ Press, 1965.

[135] Carol Parikh. *The Unreal Life of Oscar Zariski*. Springer-Verlag, 2009.

[136] Michael E. Peskin and Dan V. Schroeder. *An Introduction To Quantum Field Theory*. Frontiers in Physics. Westview Press, 1995.

[137] Benjamin C. Pierce. *Basic Category Theory for Computer Scientists*. The MIT Press, 1991.

[138] Nicholas Proudfoot and David Speyer. A broken circuit ring. *Beiträge Algebra Geom.*, 47:161–166, 2006.

[139] Victor Alexandre Puiseux. Recherches sur les fonctions algébriques. *J. Math. Pures Appl.*, 15:365–480, 1850.

[140] Daniel Quillen. Projective Modules over Polynomial Rings. *Invent. Math.*, 36:167–171, 1976.

[141] S. Rabinowich. Zum Hilbertschen Nullstellensatz. *Math. Ann.*, 102:520, 1929.

[142] Bernhard Riemann. Theorie der Abel'schen Functionen. *J. Reine Angew. Math.*, 54:101–155, 1857.

[143] Gustav Roch. Über die Anzahl der willkürlichen Constanten in algebraischen Functionen. *J. Reine Angew. Math.*, 65:372–376, 1865.

[144] Edward Rosen. *Kepler's Conversation with Galileo's Sidereal messenger. First Complete Translation, with an Introduction and notes*. Johnson Reprint Corp., 1965.

[145] Jonathan Rosenberg. *Algebraic K-Theory and Its Applications*, volume 147 of *Graduate Texts in Mathematics*. Springer.

[146] Joseph Rotman. *Galois Theory*. Springer, 2nd edition, 1998.

[147] Pierre Samuel. On unique factorization domains. *Illinois J. of Math.*, 5:1–17, 1961.

[148] Friedrich Karl Schmidt. Analytische Zahlentheorie in Körpern der Charakteristik p. *Math. Z.*, 33:1–32, 1931.

[149] Jean-Pierre Serre. Cohomologie et géométrie algébrique. In *ICM Proceedings*, volume III, pages 515–520, 1954.

[150] Jean-Pierre Serre. Faisceaux algébriques cohérents. *Annals of Mathematics*, 61(2):197–278, 1955. JSTOR:1969915.

[151] Jean-Pierre Serre. Un théorème de dualité. *Commentarii Mathematici Helvetici*, 29:9–26, 1955.

[152] Jean-Pierre Serre. Géométrie algébrique et géométrie analytique. *Université de Grenoble. Annales de l'Institut Fourier*, 6:1–42, 1956.

[153] Igor Shafarevich. *Basic Algebraic Geometry*, volume 213 of *Die Grundlehren der mathematischen Wissenshaften*. Springer, 1974.

[154] David Eugene Smith. *History of Modern Mathematics*. J. Wiley & Sons, 1906.

[155] Justin R. Smith. *Abstract Algebra*. Five Dimensions Press, 2016.

[156] Oswald Spengler. *The Decline of the West: Form and Actuality*. Arktos Media Ltd.;, arktos 2021 edition edition, 2021.

[157] Michael Spivak. *A Comprehensive Introduction to Differential Geometry*, volume 1. Publish or Perish, 3rd edition, 1999.

[158] Michael Spivak. *A Comprehensive Introduction to Differential Geometry, Volume 2*. Publish or Perish, 1999.

[159] The Stacks Project Authors. *Stacks Project*. http://stacks.math.columbia.edu.

[160] Ernst Steinitz. Algebraische Theorie der Körper. *Journal für die reine und angewandte Mathematik*, 137:167–309, 1910. Available online at http://gdz.sub.uni-goettingen.de/.

[161] Neal Stephenson. *Cryptonomicon*. Avon Books, 1999.

[162] Andrei Suslin. Проективные модули над кольцами многочленов свободны. *Doklady Akademii Nauk SSSR*, 229:1063–1066, 1976.

[163] Yves R. Talpaert. *Tensor Analysis and Continuum Mechanics*. Kluwer Academic Publishers, 2002.

[164] Günter Tamme. *Introduction to étale cohomology*. Universitext. Springer, 1994.

[165] Jean-Pierre Tignol. *Galois' Theory of Algebraic Equations*. World Scientific, Singapore, 2001.

[166] A. Tognoli. Su una congettura di Nash. *Ann. Scuola Norm. Sup. Pisa.*, 27:167–185, 1973.

[167] Robert J. Vanderbei. *Linear Programming: Foundations and Extensions. Third edition*. Number 114 in International Series in Operations Research & Management Science,. Springer, 2008.

[168] B.L Van der Waerden. Eine Verallgemeinerung des Bezoutschen Theorems. *Math. Ann.*, 99:297–541, 1928.

[169] B.L Van der Waerden. Eine Verallgemeinerung des Bezoutschen Theorems. *Math. Ann.*, 100:752, 1928.

[170] Karl Weierstrass. *Mathematische Werke. II. Abhandlungen 2*. Mayer & Müller, 1895.

[171] André Weil. L'arithmétique sur les courbes algébriques. *Acta Math.*, 52:281–315, 1929.

[172] André Weil. *Foundations of Algebraic Geometry*, volume 29 of *Colloquium Publications*. American Mathematical Society, 1946.

[173] Charles Wieibel. *An introduction to homological algebra*. Cambridge University Press, 1994.

[174] Andrew Wiles. Modular elliptic curves and Fermat's Last Theorem. *Annals of Mathematics*, 142:443–551, 1995.

[175] Stephen Willard. *General Topology*. Dover Publications, 2004.

[176] Oscar Zariski. The concept of a simple point of an abstract algebraic variety. *Trans. Amer. Math. Soc.*, 62:1–52, 1947.

[177] Oscar Zariski. *Algebraic surfaces*. Classics in mathematics. Springer-Verlag, second edition, 2004. originally published in 1935.

[178] Max Zorn. A remark on method in transfinite algebra. *Bulletin of the American Mathematical Society*, 41:667–670, 1935.