STUDENT MATHEMATICAL LIBRARY
IAS/PARK CITY MATHEMATICAL SUBSERIES
Volume 66

Algebraic Geometry

A Problem Solving Approach

Thomas Garrity
Richard Belshoff
Lynette Boos
Ryan Brown
Carl Lienert
David Murphy
Junalyn Navarra-Madsen
Pedro Poitevin
Shawn Robinson
Brian Snyder
Caryn Werner

American Mathematical Society
Institute for Advanced Study

Algebraic Geometry

A Problem Solving Approach

STUDENT MATHEMATICAL LIBRARY
IAS/PARK CITY MATHEMATICAL SUBSERIES
Volume 66

Algebraic Geometry

A Problem Solving Approach

Thomas Garrity
Richard Belshoff
Lynette Boos
Ryan Brown
Carl Lienert
David Murphy
Junalyn Navarra-Madsen
Pedro Poitevin
Shawn Robinson
Brian Snyder
Caryn Werner


American Mathematical Society, Providence, Rhode Island
Institute for Advanced Study, Princeton, New Jersey

2010 *Mathematics Subject Classification.* Primary 14–01.

Cover image: Visualization of "Seepferdchen" produced and designed by Herwig Hauser, University of Vienna and University of Innsbruck, Austria.

For additional information and updates on this book, visit
www.ams.org/bookpages/stml-66

Library of Congress Cataloging-in-Publication Data

Garrity, Thomas A., 1959–
Algebraic geometry : a problem solving approach / Thomas Garrity, Richard Belshoff, Lynette Boos, Ryan Brown, Carl Lienert, David Murphy, Junalyn Navarra-Madsen, Pedro Poitevin, Shawn Robinson, Brian Snyder, Caryn Werner.
pages cm. – (Student mathematical library ; volume 66. IAS/Park City mathematical subseries)
Includes bibliographical references and index.
ISBN 978-0-8218-9396-8 (alk. paper)
1. Geometry, Algebraic. I. Title.

QA564.G37 2013
516.3′5–dc23 2012037402

---

10 9 8 7 6 5 4 3 2 1 18 17 16 15 14 13

Dedicated to the Memory of Sidney James Drouilhet II

Contents

IAS/Park City Mathematics Institute

The IAS/Park City Mathematics Institute (PCMI) was founded in 1991 as part of the "Regional Geometry Institute" initiative of the National Science Foundation. In mid-1993 the program found an institutional home at the Institute for Advanced Study (IAS) in Princeton, New Jersey. The PCMI continues to hold summer programs in Park City, Utah.

The IAS/Park City Mathematics Institute encourages both research and education in mathematics and fosters interaction between the two. The three-week summer institute offers programs for researchers and postdoctoral scholars, graduate students, undergraduate students, mathematics teachers, mathematics education researchers, and undergraduate faculty. One of PCMI's main goals is to make all of the participants aware of the total spectrum of activities that occur in mathematics education and research: we wish to involve professional mathematicians in education and to bring modern concepts in mathematics to the attention of educators. To that end the summer institute features general sessions designed to encourage interaction among the various groups. In-year activities at sites around the country form an integral part of the Secondary School Teacher Program.

Each summer a different topic is chosen as the focus of the Research Program and Graduate Summer School. Activities in the Undergraduate Program deal with this topic as well. Lecture notes from the Graduate Summer School are published each year in the IAS/Park City Mathematics Series. Course materials from the Undergraduate Program, such as the current volume, are now being published as part of the IAS/Park City Mathematical Subseries in the Student Mathematical Library. We are happy to make available more of the excellent resources which have been developed as part of the PCMI.

John Polking, Series Editor
October 2012

Preface

Algebraic Geometry

As the name suggests, algebraic geometry is the linking of algebra to geometry. For example, the unit circle, a geometric object, can be

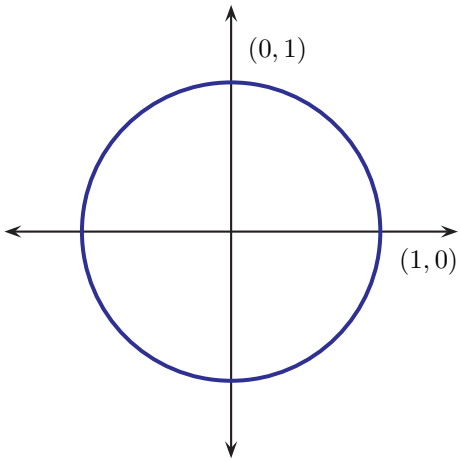


Figure 1. The unit circle centered at the origin.

described as the points (x, y) in the plane satisfying the polynomial

equation

$$x^2 + y^2 - 1 = 0,$$

an algebraic object. Algebraic geometry is thus often described as the study of those geometric objects that can be defined by polynomials. Ideally, we want a complete correspondence between the geometry and the algebra, allowing intuitions from one to shape and influence the other.

The building up of this correspondence has been at the heart of much of mathematics for the last few hundred years. It touches on area after area of mathematics. By now, despite the humble beginnings of the circle

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 - 1 = 0\},$$

algebraic geometry is not an easy area to break into.

Hence this book.

Overview

Algebraic geometry is amazingly useful, yet much of its development has been guided by aesthetic considerations. Some of the key historical developments in the subject were the result of an impulse to achieve a strong internal sense of beauty.

One way of doing mathematics is to ask bold questions about concepts you are interested in studying. Usually this leads to fairly complicated answers having many special cases. An important advantage of this approach is that the questions are natural and easy to understand. A disadvantage is that the proofs are hard to follow and often involve clever tricks, the origins of which are very hard to see.

A second approach is to spend time carefully defining the basic terms, with the aim that the eventual theorems and their proofs are straightforward. Here the difficulty is in understanding how the definitions, which often initially seem somewhat arbitrary, ever came to be. The payoff is that the deep theorems are more natural, their insights more accessible, and the theory is more aesthetically pleasing. It is this second approach that has prevailed in much of the development of algebraic geometry.

This second approach is linked to solving *equivalence problems*. By an equivalence problem, we mean the problem of determining, within a certain mathematical context, when two mathematical objects are *the same*. What is meant by *the same* differs from one mathematical context to another. In fact, one way to classify different branches of mathematics is to identify their equivalence problems.

Solving an equivalence problem, or at least setting up the language for a solution, frequently involves understanding the functions defined on an object. Since we will be concerned with the algebra behind geometric objects, we will spend time on correctly defining natural classes of functions on these objects. This in turn will allow us to correctly describe what we will mean by equivalence.

Now for a bit of an overview of this text. In Chapter 1 our motivation will be to find the natural context for being able to state that all nonsingular conics are the same. The key will be the development of the complex projective plane $\mathbb{P}^2$. We will say that two curves in this new space $\mathbb{P}^2$ are *the same* (we will use the term "isomorphic") if one curve can be transformed into the other by a projective change of coordinates (which we will define). We will also see that our conic "curves" can actually be thought of as spheres.

Chapter 2 will look at when two cubic curves are the same in $\mathbb{P}^2$, meaning again that one curve can be transformed into the other by a projective change of coordinates. Here we will see that there are many different cubics. We will further see that the points on a cubic have incredible structure; technically they form an abelian group. Finally, we will see that cubic curves are actually one-holed surfaces (tori).

Chapter 3 turns to higher degree curves. From our earlier work, we still think of these curves as "living" in the space $\mathbb{P}^2$. The first goal of this chapter is to see that these "curves" are actually surfaces. Next we will prove Bézout's Theorem. If we stick to curves in the real plane $\mathbb{R}^2$, which would be the naive first place to work, we can prove that a curve that is the zero locus of a polynomial of degree d will intersect another curve of degree e in at most de points. In our claimed more natural space of $\mathbb{P}^2$, we will see that these two curves will intersect in exactly de points, with the additional subtlety of needing to give the correct definition for intersection multiplicity.

The other major goal of Chapter 3 is the Riemann-Roch Theorem, which connects the geometry and topology of a curve to its function theory. We will also define on a curve its natural class of functions, which will be called the curve's *ring of regular functions*.

In Chapter 4 we look at the geometry of more general objects than curves. We will be treating the zero loci of collections of polynomials in many variables, and hence looking at geometric objects in $\mathbb{C}^n$ and in fact in k^n, where k is any algebraically closed field. Here the function theory plays an increasingly important role and the exercises work out how to bring much more of the full force of ring theory to bear on geometry. With this language we will see that there are actually two different but natural equivalence problems: isomorphism and birationality.

Chapter 5 develops the true natural ambient space, projective n-space $\mathbb{P}^n$, and the corresponding ring theory.

Chapter 6 increases the level of mathematics, providing an introduction to the more abstract, and more powerful, developments in algebraic geometry from the 1950s and 1960s.

Problem Book

This is a book of problems. We envision three possible audiences.

The first audience consists of students who have taken courses in multivariable calculus and linear algebra. The first three chapters are appropriate for a semester-long course for these students. If you are in this audience, here is some advice. You are at the stage of your mathematical career where you are shifting from merely solving homework exercises to proving theorems. While working the problems ask yourself what the big picture is. After working a few problems, close the book and try to think of what is going on. Ideally you would try to write down in your own words the material that you just covered. Most likely the first few times you try this, you will be at a loss for words. This is normal. Use this as an indication that you are not yet mastering this section. Repeat this process until you can describe the mathematics with confidence and feel ready to lecture to your friends.

The second audience consists of students who have had a course in abstract algebra. Then the whole book is fair game. You are at the stage where you know that much of mathematics is the attempt to prove theorems. The next stage of your mathematical development involves coming up with your own theorems, with the ultimate goal to become creative mathematicians. This is a long process. We suggest that you follow the advice given in the previous paragraph, and also occasionally ask yourself some of your own questions.

The third audience is what the authors refer to as "mathematicians on an airplane." Many professional mathematicians would like to know some algebraic geometry, but jumping into an algebraic geometry text can be difficult. We can imagine these professionals taking this book along on a long flight, and finding most of the problems just hard enough to be interesting but not so hard so that distractions on the flight will interfere with thinking. It must be emphasized that we do not think of these problems as being easy for student readers.

History of the Book

This book, with its many authors, had its start in the summer of 2008 at the Park City Mathematics Institute's Undergraduate Faculty Program on Algebraic and Analytic Geometry. Tom Garrity led a group of mathematicians on the basics of algebraic geometry, with the goal being for the participants to be able to teach algebraic geometry to undergraduates at their own college or university.

Everyone knows that you cannot learn math by just listening to someone lecture. The only way to learn is by thinking through the math on your own. Thus we decided to write a new beginning text on algebraic geometry, based on the reader solving many exercises. This book is the result.

Other Texts

There are a number of excellent introductions to algebraic geometry, at both the undergraduate and graduate levels. The following is a brief list, taken from the first few pages of Chapter 8 of [**Fowler04**].

Undergraduate texts. Bix's *Conics and Cubics: A Concrete Introduction to Algebraic Geometry* [**Bix98**] concentrates on the zero loci of second degree (conics) and third degree (cubics) two-variable polynomials. This is a true undergraduate text. Bix shows the classical fact, as we will see, that smooth conics (i.e., ellipses, hyperbolas, and parabolas) are all equivalent under a projective change of coordinates. He then turns to cubics, which are much more difficult, and shows in particular how the points on a cubic form an abelian group. For even more leisurely introductions to second degree curves, see Akopyan and Zaslavsky's *Geometry of Conics* [**AZ07**] and Kendig's *Conics* [**Ken**].

Reid's *Undergraduate Algebraic Geometry* [**Rei88**] is another good text, though the undergraduate in the title refers to British undergraduates, who start to concentrate in mathematics at an earlier age than their U.S. counterparts. Reid starts with plane curves, shows why the natural ambient space for these curves is projective space, and then develops some of the basic tools needed for higher dimensional varieties. His brief history of algebraic geometry is also fun to read.

Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra by Cox, Little, and O'Shea [**CLO07**] is almost universally admired. This book is excellent at explaining Groebner bases, which is the main tool for producing algorithms in algebraic geometry and has been a major theme in recent research. It might not be the best place for the rank beginner, who might wonder why these algorithms are necessary and interesting.

An Invitation to Algebraic Geometry by K. Smith, L. Kahanpaa, P. Kekaelaeinen, and W. N. Traves [**SKKT00**] is a wonderfully intuitive book, stressing the general ideas. It would be a good place to start for any student who has completed a first course in algebra that included ring theory.

Gibson's *Elementary Geometry of Algebraic Curves: An Undergraduate Introduction* [**Gib98**] is also a good place to begin.

There is also Hulek's *Elementary Algebraic Geometry* [**Hul03**], though this text might be more appropriate for German undergraduates (for whom it was written) than U.S. undergraduates.

The most recent of these books is Hassett's *Introduction to Algebraic Geometry*, [**Has07**], which is a good introductory text for students who have taken an abstract algebra course.

Graduate texts. There are a number, though the first two on the list have dominated the market for the last 35 years.

Hartshorne's *Algebraic Geometry* [**Har77**] relies on a heavy amount of commutative algebra. Its first chapter is an overview of algebraic geometry, while chapters four and five deal with curves and surfaces, respectively. It is in chapters two and three that the heavy abstract machinery that makes much of algebraic geometry so intimidating is presented. These chapters are not easy going but vital to get a handle on the Grothendieck revolution in mathematics. This should not be the first source for learning algebraic geometry; it should be the second or third source. Certainly young budding algebraic geometers should spend time doing all of the homework exercises in Hartshorne; this is the profession's version of paying your dues.

Principles of Algebraic Geometry by Griffiths and Harris [**GH94**] takes a quite different tack from Hartshorne. The authors concentrate on the several complex variables approach. Chapter zero in fact is an excellent overview of the basic theory of several complex variables. In this book analytic tools are freely used, but an impressive amount of geometric insight is presented throughout.

Shafarevich's *Basic Algebraic Geometry* is another standard, longtime favorite, now split into two volumes, [**Sha94a**] and [**Sha94b**]. The first volume concentrates on the relatively concrete case of subvarieties in complex projective space, which is the natural ambient space for much of algebraic geometry. Volume II turns to schemes, the key idea introduced by Grothendieck that helped change the very language of algebraic geometry.

Mumford's *Algebraic Geometry I: Complex Projective Varieties* [**Mum95**] is a good place for a graduate student to get started. One of the strengths of this book is how Mumford will give a number of

definitions, one right after another, of the same object, forcing the reader to see the different reasonable ways the same object can be viewed.

Mumford's *The Red Book of Varieties and Schemes* [**Mum99**] was for many years only available in mimeograph form from Harvard's Mathematics Department, bound in red (hence its title "The Red Book"), though it is now actually yellow. It was prized for its clear explanation of schemes. It is an ideal second or third source for learning about schemes. This new edition includes Mumford's delightful book *Curves and their Jacobians*, which is a wonderful place for inspiration.

Fulton's *Algebraic Curves* [**Ful69**] is a good brief introduction. When it was written in the late 1960s, it was the only reasonable introduction to modern algebraic geometry.

Miranda's *Algebraic Curves and Riemann Surfaces* [**Mir95**] is a popular book, emphasizing the analytic side of algebraic geometry.

Harris's *Algebraic Geometry: A First Course* [**Har95**] is chockfull of examples. In a forest versus trees comparison, it is a book of trees. This makes it difficult as a first source, but ideal as a reference for examples.

Ueno's two volumes, *Algebraic Geometry 1: From Algebraic Varieties to Schemes* [**Uen99**] and *Algebraic Geometry 2: Sheaves and Cohomology* [**Uen01**], will lead the reader to the needed machinery for much of modern algebraic geometry.

Bump's *Algebraic Geometry* [**Bum98**], Fischer's *Plane Algebraic Curves* [**Fis01**] and Perrin's *Algebraic Geometry: An Introduction* [**Per08**] are all good introductions for graduate students.

Another good place for a graduate student to get started, a source that we used more than once for this book, is Kirwan's *Complex Algebraic Curves* [**Kir92**].

Kunz's *Introduction to Plane Algebraic Curves* [**Kun05**] is another good beginning text; as an added benefit, it was translated into English from the original German by one of the authors of this book (Richard Belshoff).

Holme's *A Royal Road to Algebraic Geometry* [**Hol12**] is a quite good recent beginning graduate text, with the second part a serious introduction to schemes.

An Aside on Notation

Good notation in mathematics is important but can be tricky. It is often the case that the same mathematical object is best described using different notations depending on context. For example, in this book we will sometimes denote a curve by the symbol C, while at other times denote the curve by the symbol $\mathrm{V}(P)$ when the curve is the zero locus of the polynomial $P(x, y)$. Both notations are natural and both will be used.

Acknowledgments

The authors are grateful to many people and organizations.

From Hillsdale College, Jennifer Falck, Aaron Mortier, and John Walsh, and from Williams College, Jake Levinson, Robert Silversmith, Liyang Zhang, and Josephat Koima provided valuable feedback. We would also like to thank the students in the Spring 2009 and Fall 2011 special topics courses at Georgia College and State University. In particular, Reece Boston, Madison Hyer, Joey Shackelford, and Chris Washington provided useful contributions.

We would like to thank our editor Ed Dunne of the AMS for his support and guidance from almost the beginning.

We would like to thank Alexander Izzo for suggesting the title.

We would like to thank L. Pedersen for a careful reading of the galley proofs.

We would like to thank the Institute for Advanced Study and the Park City Mathematics Institute for their support.

Finally, we dedicate this book to the memory of Sidney James Drouilhet II from Minnesota State University Moorhead. Jim joined us at Park City and was planning to collaborate with us on the text. Unfortunately, Jim passed away while we were in the early stages of

writing this book, though the section on duality in Chapter 1 was in part inspired by his insights. We are grateful for having had the chance to work with him, if ever so briefly.

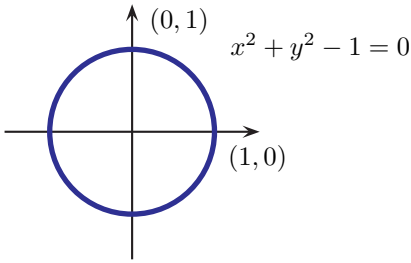
Chapter 1

Conics

Linear algebra studies the simplest type of geometric objects, such as straight lines and planes. Straight lines in the plane are the zero sets of linear, or first degree, polynomials, such as $\{(x, y) \in \mathbb{R}^2 : 3x + 4y - 1 = 0\}$. However, there are far more plane curves than just straight lines. Higher degree polynomials define other plane curves, and these are where we begin our exploration of algebraic geometry.

We start by looking at conics, which are the zero sets of second degree polynomials. The quintessential conic is the circle:

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 - 1 = 0\}.$$



Despite their apparent simplicity, an understanding of second degree equations and their solution sets is the beginning of much of algebraic geometry. By the end of the chapter, we will have developed some beautiful mathematics.

1.1. Conics over the Reals

The goal of this section is to understand the basic properties of conics in the real plane $\mathbb{R}^2$. In particular, we will see how to graph these conics.

For second degree polynomials, you can usually get a fairly good graph of the corresponding curve by just drawing it "by hand." The first series of exercises will lead you through this process. Our goal is to develop basic techniques for thinking about curves without worrying about too many technical details.

We start with the polynomial $P(x, y) = y - x^2$ and want to look at its zero set

$$C = \{(x, y) \in \mathbb{R}^2 : P(x, y) = 0\}.$$

We also denote this set by $\mathrm{V}(P)$.

Exercise 1.1.1. Show that for any $(x, y) \in C$, we also have

$$(-x, y) \in C.$$

Thus the curve C is symmetric about the y-axis.

Exercise 1.1.2. Show that if $(x, y) \in C$, then we have $y \geq 0$.

Exercise 1.1.3. Show that for every $y \geq 0$, there is a point $(x, y) \in C$ with this y-coordinate. Now, for points $(x, y) \in C$, show that if y goes to infinity, then one of the corresponding x-coordinates also approaches infinity while the other corresponding x-coordinate must approach negative infinity.

The last two exercises show that the curve C is unbounded in the positive and negative x-directions, unbounded in the positive y-direction, but bounded in the negative y-direction. This means that we can always find $(x, y) \in C$ so that x is arbitrarily large, in either the positive or negative directions, y is arbitrarily large in the positive direction, but that there is a number M (in this case 0) such that $y \geq M$ (in this case $y \geq 0$).

Exercise 1.1.4. Sketch the curve $C = \{(x, y) \in \mathbb{R}^2 : P(x, y) = 0\}$.

Conics that have these symmetry and boundedness properties and look like this curve C are called *parabolas*. Of course, we could have analyzed the curve $\{(x, y) : x - y^2 = 0\}$ and made similar observations, but with the roles of x and y reversed. In fact, we could have shifted, stretched, and rotated our parabola many ways and still retained these basic features.

We now perform a similar analysis for the plane curve

$$C = \left\{ (x, y) \in \mathbb{R}^2 : \frac{x^2}{4} + \frac{y^2}{9} - 1 = 0 \right\}.$$

Exercise 1.1.5. Show that if $(x, y) \in C$, then the three points $(-x, y)$, $(x, -y)$, and $(-x, -y)$ are also on C. Thus the curve C is symmetric about both the x- and y-axes.

Exercise 1.1.6. Show that for every $(x, y) \in C$, we have $|x| \leq 2$ and $|y| \leq 3$.

This shows that the curve C is bounded in both the positive and negative x- and y-directions.

Exercise 1.1.7. Sketch $C = \left\{ (x, y) \in \mathbb{R}^2 : \frac{x^2}{4} + \frac{y^2}{9} - 1 = 0 \right\}$.

Conics that have these symmetry and boundedness properties and look like this curve C are called *ellipses*.

There is a third type of conic. Consider the curve

$$C = \{(x, y) \in \mathbb{R}^2 : x^2 - y^2 - 4 = 0\}.$$

Exercise 1.1.8. Show that if $(x, y) \in C$, then the three points $(-x, y)$, $(x, -y)$, and $(-x, -y)$ are also on C. Thus the curve C is also symmetric about both the x- and y-axes.

Exercise 1.1.9. Show that if $(x, y) \in C$, then we have $|x| \geq 2$.

These exercises show that the curve C has two connected components. Intuitively, this means that C is composed of two distinct pieces that do not touch.

Exercise 1.1.10. Show that the curve C is unbounded in the positive and negative x-directions and also unbounded in the positive and negative y-directions.

Exercise 1.1.11. Sketch $C = \left\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 - 4 = 0\right\}$.

Conics that have these symmetry, connectedness, and boundedness properties are called *hyperbolas*.

In the following exercise, the goal is to sketch many concrete conics.

Exercise 1.1.12. Sketch the graph of each of the following conics in $\mathbb{R}^2$. Identify which are parabolas, ellipses, or hyperbolas.

(1) $\mathrm{V}(x^2 - 8y)$
(2) $\mathrm{V}(x^2 + 2x - y^2 - 3y - 1)$
(3) $\mathrm{V}(4x^2 + y^2)$
(4) $\mathrm{V}(3x^2 + 3y^2 - 75)$
(5) $\mathrm{V}(x^2 - 9y^2)$
(6) $\mathrm{V}(4x^2 + y^2 - 8)$
(7) $\mathrm{V}(x^2 + 9y^2 - 36)$
(8) $\mathrm{V}(x^2 - 4y^2 - 16)$
(9) $\mathrm{V}(y^2 - x^2 - 9)$

A natural question arises in the study of conics. If we have a second degree polynomial, how can we determine whether its zero set is an ellipse, hyperbola, parabola, or something else in $\mathbb{R}^2$? Suppose we have the following polynomial:

$$P(x, y) = ax^2 + bxy + cy^2 + dx + ey + h.$$

Are there conditions on a, b, c, d, e, h that determine what type of conic $\mathrm{V}(P)$ is?

Whenever we have a polynomial in more than one variable, a useful technique is to treat P as a polynomial in a single variable whose coefficients are themselves polynomials in the remaining variables.

Exercise 1.1.13. Express the polynomial $P(x, y) = ax^2 + bxy + cy^2 + dx + ey + h$ in the form

$$P(x, y) = Ax^2 + Bx + C$$

where A, B, and C are polynomials in y. What are A, B, and C?

Since we are interested in the zero set $\mathrm{V}(P)$, we want to find the roots of $Ax^2 + Bx + C = 0$ in terms of y. As we know from high school algebra the roots of the quadratic equation $Ax^2 + Bx + C = 0$ are

$$\frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

To determine the number of real roots, we need to look at the sign of the discriminant

$$\Delta_x = B^2 - 4AC.$$

Exercise 1.1.14. Treating $P(x, y) = ax^2 + bxy + cy^2 + dx + ey + h$ as a polynomial in the variable x, show that the discriminant is

$$\Delta_x(y) = (b^2 - 4ac)y^2 + (2bd - 4ae)y + (d^2 - 4ah).$$

Exercise 1.1.15.

(1) Suppose $\Delta_x(y_0) < 0$. Explain why there is no point on $\mathrm{V}(P)$ whose y-coordinate is y_0.

(2) Suppose $\Delta_x(y_0) = 0$. Explain why there is exactly one point on $\mathrm{V}(P)$ whose y-coordinate is y_0.

(3) Suppose $\Delta_x(y_0) > 0$. Explain why there are exactly two points on $\mathrm{V}(P)$ whose y-coordinate is y_0.

This exercise demonstrates that in order to understand the set $\mathrm{V}(P)$ we need to understand the set $\{y : \Delta_x(y) \geq 0\}$. We will first see how for parabolas we expect the scalar $b^2 - 4ac$ to be zero.

Exercise 1.1.16. Suppose $b^2 - 4ac = 0$. Suppose further that $2bd - 4ae > 0$.

(1) Show that $\Delta_x(y) \geq 0$ if and only if $y \geq \dfrac{4ah - d^2}{2bd - 4ae}$.

(2) Conclude that if $b^2 - 4ac = 0$ and $2bd - 4ae > 0$, then $\mathrm{V}(P)$ is a parabola.

Notice that if $b^2 - 4ac \neq 0$, then $\Delta_x(y)$ is itself a quadratic function in y, and the features of the set over which $\Delta_x(y)$ is nonnegative is determined by its quadratic coefficient.

Exercise 1.1.17. Suppose $b^2 - 4ac < 0$.

(1) Show that one of the following occurs:
   (a) $\{y \mid \Delta_x(y) \geq 0\} = \emptyset$,
   (b) $\{y \mid \Delta_x(y) \geq 0\} = \{y_0\}$,
   (c) there exist real numbers α and β, $\alpha < \beta$, such that

$$\{y \mid \Delta_x(y) \geq 0\} = \{y \mid \alpha \leq y \leq \beta\}.$$

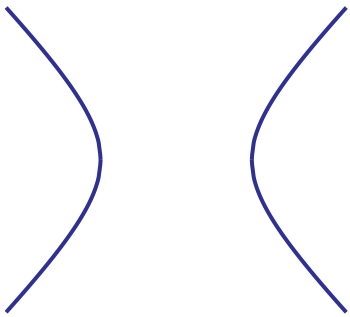
(2) Conclude that $\mathrm{V}(P)$ is either empty, a point, or an ellipse.

Exercise 1.1.18. Suppose $b^2 - 4ac > 0$.

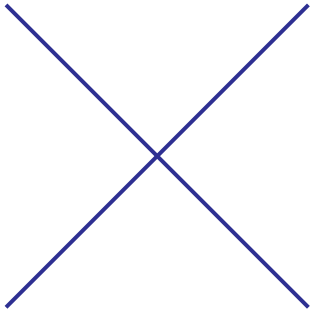
(1) Show that one of the following occurs:
   (a) $\{y \mid \Delta_x(y) \geq 0\} = \mathbb{R}$ and $\Delta_x(y) \neq 0$,
   (b) $\{y \mid \Delta_x(y) = 0\} = \{y_0\}$ and $\{y \mid \Delta_x(y) > 0\} = \{y \mid y \neq y_0\}$,
   (c) there exist real numbers α and β, $\alpha < \beta$, such that

$$\{y \mid \Delta_x(y) \geq 0\} = \{y \mid y \leq \alpha\} \cup \{y \mid y \geq \beta\}.$$

(2) If $\{y \mid \Delta_x(y) \geq 0\} = \mathbb{R}$, show that $\mathrm{V}(P)$ is a hyperbola opening left and right:



(3) If $\{y \mid \Delta_x(y) = 0\} = \{y_0\}$ and there is a point on $\mathrm{V}(P)$ with y-coordinate equal to y_0, show that $\mathrm{V}(P)$ is two lines intersecting in a point:



(4) If there are two real numbers α and β, $\alpha < \beta$, such that

$$\{y \mid \Delta_x(y) \geq 0\} = \{y \mid y \leq \alpha\} \cup \{y \mid y \geq \beta\},$$

show that $\mathrm{V}(P)$ is a hyperbola opening up and down:

Above we decided to treat P as a function of x, but we could have treated P as a function of y, $P(x, y) = A'y^2 + B'y + C'$, each of whose coefficients is now a polynomial in x.

Exercise 1.1.19. Show that the discriminant of $A'y^2 + B'y + C' = 0$ is

$$\Delta_y(x) = (b^2 - 4ac)x^2 + (2be - 4cd)x + (e^2 - 4ch).$$

Note that the quadratic coefficient is again $b^2 - 4ac$, so our observations from above are the same in this case as well. In the preceding exercises we were intentionally vague about some cases. For example, we do not say anything about what happens when $b^2 - 4ac = 0$ and $2bd - 4ae = 0$. This is an example of a "degenerate" conic. We treat degenerate conics later in this chapter, but for now it suffices to note

that if $b^2 - 4ac = 0$, then $\mathrm{V}(P)$ is neither an ellipse nor a hyperbola. If $b^2 - 4ac < 0$, then $\mathrm{V}(P)$ is neither a parabola nor a hyperbola. And if $b^2 - 4ac > 0$, then $\mathrm{V}(P)$ is neither a parabola nor an ellipse. This leads us to the following theorem.

Theorem 1.1.20. Suppose $P(x, y) = ax^2 + bxy + cy^2 + dx + ey + h$. If $\mathrm{V}(P)$ is a parabola in $\mathbb{R}^2$, then $b^2 - 4ac = 0$; if $\mathrm{V}(P)$ is an ellipse in $\mathbb{R}^2$, then $b^2 - 4ac < 0$; and if $\mathrm{V}(P)$ is a hyperbola in $\mathbb{R}^2$, then $b^2 - 4ac > 0$.

In general, it is not immediately clear whether a given conic $C = \mathrm{V}(ax^2 + bxy + cy^2 + dx + e + h)$ is an ellipse, hyperbola, or parabola. When the coefficient $b = 0$, then it is much easier to determine what type of curve C is.

Corollary 1.1.1. Suppose $P(x, y) = ax^2 + cy^2 + dx + ey + h$. If $\mathrm{V}(P)$ is a parabola in $\mathbb{R}^2$, then $ac = 0$; if $\mathrm{V}(P)$ is a hyperbola in $\mathbb{R}^2$, then $ac < 0$, i.e. a and c have opposite signs; and if $\mathrm{V}(P)$ is a ellipse in $\mathbb{R}^2$, then $ac > 0$, i.e. a and c have the same sign.

1.2. Changes of Coordinates

The goal of this section is to show that, in $\mathbb{R}^2$, any ellipse can be transformed into any other ellipse, any hyperbola into any other hyperbola, and any parabola into any other parabola.

Here we start to investigate what it could mean for two conics to be *the same*; thus we start to solve an equivalence problem for conics. Intuitively, two curves are the same if we can shift, stretch, or rotate one to obtain the other. Cutting or gluing, however, is not allowed.

Our conics live in the real plane $\mathbb{R}^2$. In order to describe conics as the zero sets of second degree polynomials, we first must choose a coordinate system for the plane. Different choices for these coordinates will give different polynomials, even for the same curve. (To make this concrete, imagine 10 people separately go to a blank blackboard, put a dot on it to correspond to an origin, and then draw two axes. There will be 10 quite different coordinate systems chosen.)

Consider the two coordinate systems:

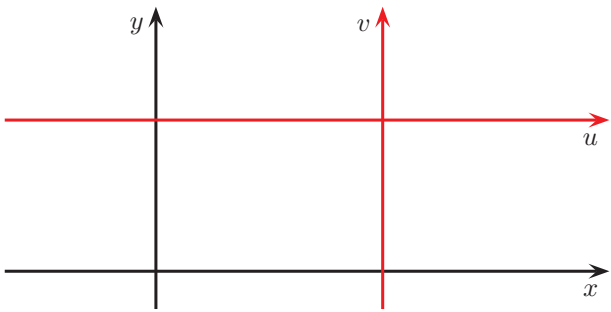


Figure 1. xy- and uv-coordinate systems.

Suppose there is a dictionary between these coordinate systems, given by

$$u = x - 3,$$
$$v = y - 2.$$

Then the circle of radius 4 has either the equation

$$u^2 + v^2 - 16 = 0$$

or the equation

$$(x-3)^2 + (y-2)^2 - 16 = 0,$$

which is the same as $x^2 - 6x + y^2 - 4y - 3 = 0$.

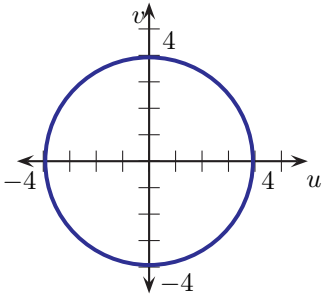


Figure 2. Circle of radius 4 centered at the origin in the uv-coordinate system.

These two coordinate systems differ only by where you place the origin.

Coordinate systems can also differ in their orientation. Consider two coordinate systems where the dictionary between the coordinate systems is:

$$u = x - y$$
$$v = x + y.$$

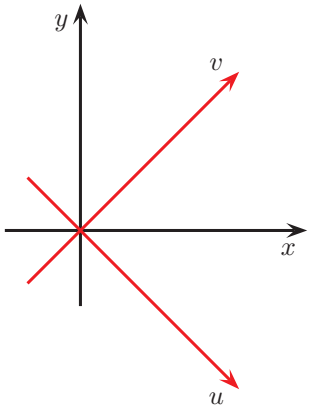


Figure 3. xy- and uv-coordinate systems with different orientations.

Coordinate systems can also vary by the chosen units of length. Consider two coordinate systems where the dictionary between the coordinate systems is:

$$u = 2x$$
$$v = 3y.$$

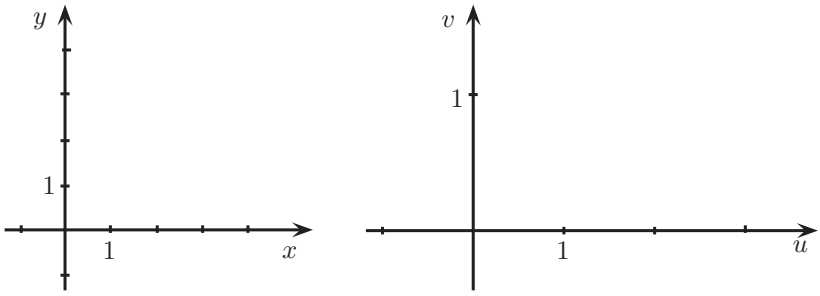


Figure 4. xy- and uv-coordinate systems with different units.

All of these possibilities are captured in the following.

Definition 1.2.1. A *real affine change of coordinates* in the real plane, $\mathbb{R}^2$, is given by

$$u = ax + by + e$$
$$v = cx + dy + f,$$

where $a, b, c, d, e, f \in \mathbb{R}$ and

$$ad - bc \neq 0.$$

In matrix language, we have

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix},$$

where $a, b, c, d, e, f \in \mathbb{R}$, and

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq 0.$$

Exercise 1.2.1. Show that the origin in the xy-coordinate system agrees with the origin in the uv-coordinate system if and only if $e = f = 0$. Thus the constants e and f describe translations of the origin.

Exercise 1.2.2. Show that if $u = ax + by + e$ and $v = cx + dy + f$ is a change of coordinates, then the inverse change of coordinates is

$$x = \left(\frac{1}{ad - bc}\right)(du - bv) - \left(\frac{1}{ad - bc}\right)(de - bf)$$
$$y = \left(\frac{1}{ad - bc}\right)(-cu + av) - \left(\frac{1}{ad - bc}\right)(-ce + af).$$

(This is why we require that $ad - bc \neq 0$.) There are two ways of working this problem. One method is to just start fiddling with the equations. The second is to translate the change of coordinates into the matrix language and then use a little linear algebra.

It is also common for us to change coordinates multiple times, but we need to ensure that a composition of real affine changes of coordinates is a real affine change of coordinates.

Exercise 1.2.3. Show that if

$$u = ax + by + e$$
$$v = cx + dy + f$$

and

$$s = Au + Bv + E$$
$$t = Cu + Dv + F$$

are two real affine changes of coordinates from the xy-plane to the uv-plane and from the uv-plane to the st-plane, respectively, then the composition from the xy-plane to the st-plane is a real affine change of coordinates.

We frequently go back and forth between using a change of coordinates and its inverse. For example, suppose we have the ellipse

$\mathrm{V}(x^2+y^2-1)$ in the xy-plane. Under the real affine change of coordinates

$$u = x + y$$
$$v = 2x - y,$$

this ellipse becomes $\mathrm{V}(5u^2 - 2uv + 2v^2 - 9)$ in the uv-plane (verify this). To change coordinates from the xy-plane to the uv-plane, we use the inverse change of coordinates

$$x = \frac{1}{3}u + \frac{1}{3}v$$
$$y = \frac{2}{3}u - \frac{1}{3}v.$$

Since any affine transformation has an inverse transformation, we will not worry too much about whether we are using a transformation or its inverse in our calculations. When the context requires care, we will make the distinction.

Exercise 1.2.4. For each pair of ellipses, find a real affine change of coordinates that maps the ellipse in the xy-plane to the ellipse in the uv-plane.

(1) $\mathrm{V}(x^2+y^2-1)$, $\mathrm{V}(16u^2+9v^2-1)$
(2) $\mathrm{V}((x-1)^2+y^2-1)$, $\mathrm{V}(16u^2+9(v+2)^2-1)$
(3) $\mathrm{V}(4x^2+y^2-6y+8)$, $\mathrm{V}(u^2-4u+v^2-2v+4)$
(4) $\mathrm{V}(13x^2-10xy+13y^2-1)$, $\mathrm{V}(4u^2+9v^2-1)$

We can apply a similar argument for hyperbolas.

Exercise 1.2.5. For each pair of hyperbolas, find a real affine change of coordinates that maps the hyperbola in the xy-plane to the hyperbola in the uv-plane.

(1) $\mathrm{V}(xy-1)$, $\mathrm{V}(u^2-v^2-1)$
(2) $\mathrm{V}(x^2-y^2-1)$, $\mathrm{V}(16u^2-9v^2-1)$
(3) $\mathrm{V}((x-1)^2-y^2-1)$, $\mathrm{V}(16u^2-9(v+2)^2-1)$
(4) $\mathrm{V}(x^2-y^2-1)$, $\mathrm{V}(v^2-u^2-1)$
(5) $\mathrm{V}(8xy-1)$, $\mathrm{V}(2u^2-2v^2-1)$

Now we move on to parabolas.

Exercise 1.2.6. For each pair of parabolas, find a real affine change of coordinates that maps the parabola in the xy-plane to the parabola in the uv-plane.

(1) $\mathrm{V}(x^2 - y)$, $\mathrm{V}(9v^2 - 4u)$

(2) $\mathrm{V}((x-1)^2 - y)$, $\mathrm{V}(u^2 - 9(v+2))$

(3) $\mathrm{V}(x^2 - y)$, $\mathrm{V}(u^2 + 2uv + v^2 - u + v - 2)$

(4) $\mathrm{V}(x^2 - 4x + y + 4)$, $\mathrm{V}(4u^2 - (v+1))$

(5) $\mathrm{V}(4x^2 + 4xy + y^2 - y + 1)$, $\mathrm{V}(4u^2 + v)$

The preceding three problems suggest that we can transform ellipses to ellipses, hyperbolas to hyperbolas, and parabolas to parabolas by way of real affine changes of coordinates. This turns out to be the case. Suppose $C = \mathrm{V}(ax^2 + bxy + cy^2 + dx + ey + h)$ is a conic in $\mathbb{R}^2$. Our goal in the next several exercises is to show that if C is an ellipse, we can transform it to $\mathrm{V}(x^2 + y^2 - 1)$; if C is a hyperbola, we can transform it to $\mathrm{V}(x^2 - y^2 - 1)$; and if C is a parabola, we can transform it to $\mathrm{V}(x^2 - y)$. We will pass through a series of real affine transformations and appeal to Exercise 1.2.3. This result ensures that the final composition of our individual transformations is the real affine transformation we seek. This composition is, however, a mess, so we won't write it down explicitly. We will see in Section 1.11 that we can organize this information much more efficiently by using tools from linear algebra.

We begin with ellipses. Suppose $\mathrm{V}(ax^2 + bxy + cy^2 + dx + ey + h)$ is an ellipse in $\mathbb{R}^2$. Our first transformation will be to remove the xy term, i.e. to find a real affine transformation that will align our given curve with the coordinate axes. By Theorem 1.1.20 we know that $b^2 - 4ac < 0$.

Exercise 1.2.7. Explain why if $b^2 - 4ac < 0$, then $ac > 0$.

Exercise 1.2.8. Show that under the real affine transformation

$$x = \sqrt{\frac{c}{a}}u + v$$
$$y = u - \sqrt{\frac{a}{c}}v,$$

the ellipse $\mathrm{V}(ax^2 + bxy + cy^2 + dx + ey + h)$ in the xy-plane becomes an ellipse in the uv-plane whose defining equation is $Au^2 + Cv^2 + Du + Ev + H = 0$. Find A and C in terms of a, b, c. Show that if $b^2 - 4ac < 0$, then $A \neq 0$ and $C \neq 0$.

Now we have a new ellipse $\mathrm{V}(Au^2 + Cv^2 + Du + Ev + H)$ in the uv-plane. If our original ellipse already had $b = 0$, then we would have skipped the previous step and gone directly to this one.

Exercise 1.2.9. Show that there exist constants R, S, and T such that the equation

$$Au^2 + Cv^2 + Du + Ev + H = 0$$

can be rewritten in the form

$$A(u - R)^2 + C(v - S)^2 - T = 0.$$

Express R, S, and T in terms of A, C, D, E, and H.

To simplify notation, we revert to using x and y as our variables instead of u and v, but we keep in mind that we are not really still working in our original xy-plane. This is a convenience to avoid subscripts. Without loss of generality we can assume $A, C > 0$, since if $A, C < 0$ we could simply multiply the above equation by -1 without affecting the conic. Note that we assume that our original conic is an ellipse, i.e., it is nondegenerate. A consequence of this is that $T > 0$.

Exercise 1.2.10. Suppose $A, C > 0$. Find a real affine change of coordinates that maps the ellipse

$$\mathrm{V}(A(x - R)^2 + C(y - S)^2 - T),$$

to the circle

$$\mathrm{V}(u^2 + v^2 - 1).$$

Hence, we have found a composition of real affine changes of coordinates that transforms any ellipse $\mathrm{V}(ax^2+bxy+cy^2+dx+ey+h)$ to the circle $\mathrm{V}(u^2+v^2-1)$.

We want a similar process for parabolas. Suppose $\mathrm{V}(ax^2+bxy+cy^2+dx+ey+h)$ is a parabola in $\mathbb{R}^2$. We want to show, by direct algebra, that there is a change of coordinates that takes this parabola to

$$\mathrm{V}(u^2-v).$$

By Theorem 1.1.20 we know that $b^2-4ac=0$. As before our first task is to eliminate the xy term. Suppose first that $b\neq 0$. Since $b^2>0$ and $4ac=b^2$ we know $ac>0$, so we can repeat Exercise 1.2.8.

Exercise 1.2.11. Consider the values A and C found in Exercise 1.2.8. Show that if $b^2-4ac=0$, then either $A=0$ or $C=0$, depending on the signs of a,b,c. [Hint: Recall, $\sqrt{\alpha^2}=-\alpha$ if $\alpha<0$.]

Since either $A=0$ or $C=0$ we can assume $C=0$ without loss of generality. Then $A\neq 0$, for our curve is a parabola and not a straight line, so our transformed parabola is $\mathrm{V}(Au^2+Du+Ev+H)$ in the uv-plane. If our original parabola already had $b=0$, then we also know, since $b^2-4ac=0$, that either $a=0$ or $c=0$, so we could have skipped ahead to this step.

Exercise 1.2.12. Show that there exist constants R and T such that the equation

$$Au^2+Du+Ev+H=0$$

can be rewritten as

$$A(u-R)^2+E(v-T)=0.$$

Express R and T in terms of A,D,E, and H.

As above we revert our notation to x and y with the same caveat as before. Multiplying our equation by -1 if necessary, we may assume $A>0$.

Exercise 1.2.13. Suppose $A>0$ and $E\neq 0$. Find a real affine change of coordinates that maps the parabola

$$\mathrm{V}(A(x-R)^2-E(y-T))$$

to the parabola

$$\mathrm{V}(u^2 - v).$$

Hence, we have found a real affine change of coordinates that transforms any parabola $\mathrm{V}(ax^2 + bxy + cy^2 + dx + ey + h)$ to the parabola $\mathrm{V}(u^2 - v)$.

Finally, suppose $\mathrm{V}(ax^2 + bxy + cy^2 + dx + ey + h)$ is a hyperbola in $\mathbb{R}^2$. We want to show that there is a change of coordinates that takes this hyperbola to

$$\mathrm{V}(u^2 - v^2 - 1).$$

By Theorem 1.1.20 we know that $b^2 - 4ac > 0$. Suppose first that $b \neq 0$. Unlike before, we can now have $ac > 0$, $ac < 0$, or $ac = 0$.

Exercise 1.2.14. Suppose $ac > 0$. Use the real affine transformation in Exercise 1.2.8 to transform C to a conic in the uv-plane. Find the coefficients of u^2 and v^2 in the resulting equation and show that they have opposite signs.

Now for the $ac < 0$ case.

Exercise 1.2.15. Suppose $ac < 0$ and $b \neq 0$. Use the real affine transformation

$$x = \sqrt{-\frac{c}{a}}u + v$$
$$y = u - \sqrt{-\frac{a}{c}}v$$

to transform C to a conic in the uv-plane of the form

$$Au^2 + Cv^2 + Du + Ev + H = 0.$$

Find the coefficients of u^2 and v^2 in the resulting equation and show that they have opposite signs.

Note that in the case when $ac < 0$ and $b = 0$, then a and c have opposite signs and the hyperbola is already of the form

$$ax^2 + cy^2 + dx + ey + f = 0.$$

Exercise 1.2.16. Suppose $ac = 0$ (so $b \neq 0$). Since either $a = 0$ or $c = 0$, we can assume $c = 0$. Use the real affine transformation

$$x = u + v$$
$$y = \left(\frac{1-a}{b}\right) u - \left(\frac{1+a}{b}\right) v$$

to transform $\mathrm{V}(ax^2 + bxy + dx + ey + h)$ to a conic in the uv-plane of the form

$$\mathrm{V}(u^2 - v^2 + Du + Ev + H).$$

In all three cases we find that the hyperbola can be transformed to $\mathrm{V}(Au^2 - Cv^2 + Du + Ev + H)$ in the uv-plane, with both A and C positive. We can now complete the transformation of the hyperbola as we did above with parabolas and ellipses.

Exercise 1.2.17. Show that there exist constants R, S and T so that

$$Au^2 - Cv^2 + Du + Ev + H = A(u - R)^2 - C(v - S)^2 - T.$$

Express R, S, and T in terms of A, C, D, E, and H.

We are assuming that we have a hyperbola. Hence $T \neq 0$, since otherwise we would have just two lines through the origin. If $T < 0$, then we can multiply the equation $A(u - R)^2 - C(v - S)^2 - T = 0$ through by -1 and then interchange u with v. Thus we can assume that our original hyperbola has become

$$\mathrm{V}(A(u - R)^2 - C(v - S)^2 - T)$$

with A, C and T all positive.

Exercise 1.2.18. Suppose $A, C, T > 0$. Find a real affine change of coordinates that maps the hyperbola

$$\mathrm{V}(A(x - R)^2 - C(y - S)^2 - T),$$

to the hyperbola

$$\mathrm{V}(u^2 - v^2 - 1).$$

We have now shown that in $\mathbb{R}^2$ we can find a real affine change of coordinates that will transform any ellipse to $\mathrm{V}(x^2 + y^2 - 1)$, any hyperbola to $\mathrm{V}(x^2 - y^2 - 1)$, and any parabola to $\mathrm{V}(x^2 - y)$. Thus we

have three classes of smooth conics in $\mathbb{R}^2$. Our next task is to show that these are distinct, that is, that we cannot transform an ellipse to a parabola and so on.

Exercise 1.2.19. Give an intuitive argument, based on the number of connected components, for the fact that no ellipse can be transformed into a hyperbola by a real affine change of coordinates.

Exercise 1.2.20. Show that there is no real affine change of coordinates

$$u = ax + by + e$$
$$v = cx + dy + f$$

that transforms the ellipse $\mathrm{V}(x^2 + y^2 - 1)$ to the hyperbola $\mathrm{V}(u^2 - v^2 - 1)$.

Exercise 1.2.21. Give an intuitive argument, based on boundedness, for the fact that no parabola can be transformed into an ellipse by a real affine change of coordinates.

Exercise 1.2.22. Show that there is no real affine change of coordinates that transforms the parabola $\mathrm{V}(x^2 - y)$ to the circle $\mathrm{V}(u^2 + v^2 - 1)$.

Exercise 1.2.23. Give an intuitive argument, based on the number of connected components, for the fact that no parabola can be transformed into a hyperbola by a real affine change of coordinates.

Exercise 1.2.24. Show that there is no real affine change of coordinates that transforms the parabola $\mathrm{V}(x^2 - y)$ to the hyperbola $\mathrm{V}(u^2 - v^2 - 1)$.

Definition 1.2.2. Two conics are *equivalent under a real affine change of coordinates* if the defining polynomial for one of the conics can be transformed via a real affine change of coordinates into the defining polynomial of the other conic.

Combining all of the work in this section, we have just proven the following theorem.

Theorem 1.2.25. Under a real affine change of coordinates, all ellipses in $\mathbb{R}^2$ are equivalent, all hyperbolas in $\mathbb{R}^2$ are equivalent, and all

parabolas in $\mathbb{R}^2$ are equivalent. Further, these three classes of conics are distinct; no conic of one class can be transformed via a real affine change of coordinates to a conic of a different class.

In Section 1.11 we will revisit this theorem using tools from linear algebra. This approach will yield a cleaner and more straightforward proof than the one we have in the current setting. The linear algebraic setting will also make all of our transformations simpler, and it will become apparent how we arrived at the particular transformations.

1.3. Conics over the Complex Numbers

The goal of this section is to see how, under a complex affine changes of coordinates, all ellipses and hyperbolas are equivalent, while parabolas are still distinct.

While it is certainly natural to begin with the zero set of a polynomial $P(x, y)$ as a curve in the real plane $\mathbb{R}^2$, polynomials also have roots over the complex numbers. In fact, throughout mathematics it is almost always easier to work over the complex numbers than over the real numbers. This can be seen in the solutions given by the quadratic equation $x^2 + 1 = 0$, which has no solutions if we require $x \in \mathbb{R}$ but does have the two solutions, $x = \pm i$, in the complex numbers $\mathbb{C}$.

Exercise 1.3.1. Show that the set

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 + 1 = 0\}$$

is empty but that the set

$$C = \{(x, y) \in \mathbb{C}^2 : x^2 + y^2 + 1 = 0\}$$

is not empty. In fact, show that given any complex number x there must exist a $y \in \mathbb{C}$ such that

$$(x, y) \in C.$$

Then show that if $x \neq \pm i$, then there are two distinct values $y \in \mathbb{C}$ such that $(x, y) \in C$, while if $x = \pm i$ there is only one such y.

Thus if we use only real numbers, some zero sets of second degree polynomials will be empty. This does not happen over the complex numbers.

Exercise 1.3.2. Let

$$P(x, y) = ax^2 + bxy + cy^2 + dx + ey + f,$$

with $a \neq 0$. Show that for any value $y \in \mathbb{C}$, there must be at least one $x \in \mathbb{C}$, but no more than two such x's, such that

$$P(x, y) = 0.$$

[Hint: Write $P(x, y) = Ax^2 + Bx + C$ as a function of x whose coefficients A, B, and C are themselves functions of y, and use the quadratic formula. As mentioned before, this technique will be used frequently.]

Thus for any second order polynomial, its zero set is non-empty provided we work over the complex numbers.

But even more happens. We start with:

Exercise 1.3.3. Let $C = \mathrm{V}\left(\frac{x^2}{4} + \frac{y^2}{9} - 1\right) \subset \mathbb{C}^2$. Show that C is unbounded in both x and y. (Over the complex numbers $\mathbb{C}$, being unbounded in x means, given any number M, there will be a point $(x, y) \in C$ such that $|x| > M$. Compare this result to Exercise 1.1.6.)

Hyperbolas in $\mathbb{R}^2$ come in two pieces. In $\mathbb{C}^2$, it can be shown that hyperbolas are connected, meaning there is a continuous path from any point to any other point. The following shows this for a specific hyperbola.

Exercise 1.3.4. Let $C = \mathrm{V}(x^2 - y^2 - 1) \subset \mathbb{C}^2$. Show that there is a continuous path on the curve C from the point $(-1, 0)$ to the point $(1, 0)$, despite the fact that no such continuous path exists in $\mathbb{R}^2$. (Compare this exercise with Exercise 1.1.9.)

These two exercises demonstrate that in $\mathbb{C}^2$ ellipses are unbounded (just like hyperbolas and parabolas) and hyperbolas are connected (just like ellipses and parabolas). Thus the intuitive arguments in Exercises 1.2.19, 1.2.21, and 1.2.23 no longer work in $\mathbb{C}^2$. We have even more.

Exercise 1.3.5. Show that if $x = u$ and $y = iv$, then the circle $\{(x, y) \in \mathbb{C}^2 : x^2 + y^2 = 1\}$ transforms into the hyperbola $\{(u, v) \in \mathbb{C}^2 : u^2 - v^2 = 1\}$.

Definition 1.3.1. A *complex affine change of coordinates* in the complex plane $\mathbb{C}^2$ is given by

$$\begin{aligned} u &= ax + by + e \\ v &= cx + dy + f, \end{aligned}$$

where $a, b, c, d, e, f \in \mathbb{C}$ and

$$ad - bc \neq 0.$$

Exercise 1.3.6. Show that if $u = ax + by + e$ and $v = cx + dy + f$ is a change of coordinates, then the inverse change of coordinates is

$$\begin{aligned} x &= \left(\frac{1}{ad - bc}\right)(du - bv) - \left(\frac{1}{ad - bc}\right)(de - bf) \\ y &= \left(\frac{1}{ad - bc}\right)(-cu + av) - \left(\frac{1}{ad - bc}\right)(-ce + af). \end{aligned}$$

This proof should look almost identical to the solution of Exercise 1.2.2.

Definition 1.3.2. Two conics are *equivalent under a complex affine change of coordinates* if the defining polynomial for one of the conics can be transformed via a complex affine change of coordinates into the defining polynomial for the other conic.

Exercise 1.3.7. Use Theorem 1.2.25 together with the new result of Exercise 1.3.5 to conclude that all ellipses and hyperbolas are equivalent under complex affine changes of coordinates.

Parabolas, though, are still different.

Exercise 1.3.8. Show that the circle $\{(x, y) \in \mathbb{C}^2 : x^2 + y^2 - 1 = 0\}$ is not equivalent under a complex affine change of coordinates to the parabola $\{(u, v) \in \mathbb{C}^2 : u^2 - v = 0\}$.

We now want to look more directly at $\mathbb{C}^2$ in order to understand more clearly why the class of ellipses and the class of hyperbolas are different as real objects but the same as complex objects. We start by looking at $\mathbb{C}$. Algebraic geometers regularly use the variable x for

a complex number. Complex analysts more often use the variable z, which allows a complex number to be expressed in terms of its real and imaginary parts

$$z = x + iy,$$

where x is the real part of z and y is the imaginary part.

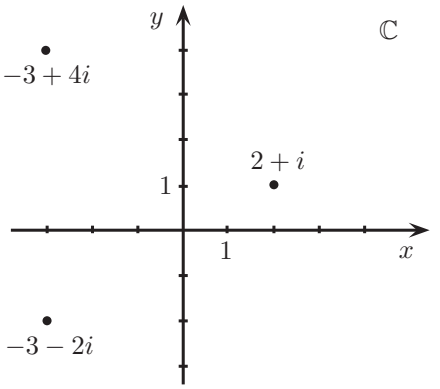


Figure 5. Points in the complex plane.

Similarly, an algebraic geometer will usually use (x, y) to denote points in the complex plane $\mathbb{C}^2$ while a complex analyst will instead use (z, w) to denote points in the complex plane $\mathbb{C}^2$. Here the complex analyst will write

$$w = u + iv.$$

There is a natural bijection from $\mathbb{C}^2$ to $\mathbb{R}^4$ given by

$$(z, w) = (x + iy, u + iv) \to (x, y, u, v).$$

In the same way, there is a natural bijection from $\mathbb{C}^2 \cap \{(x, y, u, v) \in \mathbb{R}^4 : y = 0, v = 0\}$ to the real plane $\mathbb{R}^2$, given by

$$(x + 0i, u + 0i) \to (x, 0, u, 0) \to (x, u).$$

Likewise, there is a similar natural bijection from $\mathbb{C}^2 = \{(z, w) \in \mathbb{C}^2\} \cap \{(x, y, u, v) \in \mathbb{R}^4 : y = 0, u = 0\}$ to $\mathbb{R}^2$, given this time by

$$(x + 0i, 0 + vi) \to (x, 0, 0, v) \to (x, v).$$

One way to think about conics in $\mathbb{C}^2$ is to consider two-dimensional slices of $\mathbb{C}^2$. Let

$$C = \{(z, w) \in \mathbb{C}^2 : z^2 + w^2 = 1\}.$$

Exercise 1.3.9. Give a bijection from

$$C \cap \{(x + iy, u + iv) : x, u \in \mathbb{R}, y = 0, v = 0\}$$

to the real circle of unit radius in $\mathbb{R}^2$. (Thus a real circle in the plane $\mathbb{R}^2$ can be thought of as a real slice of the complex curve C.)

Taking a different real slice of C will yield not a circle but a hyperbola.

Exercise 1.3.10. Give a bijection from

$$C \cap \{(x + iy, u + iv) \in \mathbb{R}^4 : x, v \in \mathbb{R}, y = 0, u = 0\}$$

to the hyperbola $\mathrm{V}(x^2 - v^2 - 1)$ in $\mathbb{R}^2$.

Thus the single complex curve C contains both real circles and real hyperbolas.

1.4. The Complex Projective Plane $\mathbb{P}^2$

The goal of this section is to introduce the complex projective plane $\mathbb{P}^2$, which is the natural ambient space (with its higher dimensional analog $\mathbb{P}^n$) for much of algebraic geometry. In $\mathbb{P}^2$, we will see that all ellipses, hyperbolas, and parabolas are equivalent.

In $\mathbb{R}^2$ all ellipses are equivalent, all hyperbolas are equivalent, and all parabolas are equivalent under real affine changes of coordinates. Further, these classes of conics are distinct in $\mathbb{R}^2$. When we move to $\mathbb{C}^2$, ellipses and hyperbolas are equivalent under complex affine changes of coordinates, but parabolas remain distinct. The next step is to describe a larger plane in which all three classes are equivalent.

First, we will define the complex projective plane $\mathbb{P}^2$ and discuss some of its basic properties. While it may not be immediately clear from this definition, we will see how $\mathbb{C}^2$ naturally lives in $\mathbb{P}^2$. Further, the extra points in $\mathbb{P}^2$ that are not in $\mathbb{C}^2$ can be viewed as "points

at infinity." Then we will look at the projective analogue of change of coordinates and see how we can view all ellipses, hyperbolas, and parabolas as equivalent.

Definition 1.4.1. Define a relation $\sim$ on points in $\mathbb{C}^3 - \{(0,0,0)\}$ as follows: $(x,y,z) \sim (u,v,w)$ if and only if there exists $\lambda \in \mathbb{C} - \{0\}$ such that $(x,y,z) = (\lambda u, \lambda v, \lambda w)$.

Exercise 1.4.1.

(1) Show that $(2, 1+i, 3i) \sim (2-2i, 2, 3+3i)$.

(2) Show that $(1,2,3) \sim (2,4,6) \sim (-2,-4,-6) \sim (-i,-2i,-3i)$.

(3) Show that $(2, 1+i, 3i) \not\sim (4, 4i, 6i)$.

(4) Show that $(1,2,3) \not\sim (3,6,8)$.

Exercise 1.4.2. Show that $\sim$ is an equivalence relation. (Recall that an *equivalence relation* $\sim$ on a set X satisfies the conditions (i) $a \sim a$ for all $a \in X$, (ii) $a \sim b$ implies $b \sim a$, and (iii) $a \sim b$ and $b \sim c$ implies $a \sim c$.)

Exercise 1.4.3. Suppose that $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ and that $x_1 = x_2 \neq 0$. Show that $y_1 = y_2$ and $z_1 = z_2$.

Exercise 1.4.4. Suppose that $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ with $z_1 \neq 0$ and $z_2 \neq 0$. Show that

$$(x_1, y_1, z_1) \sim \left(\frac{x_1}{z_1}, \frac{y_1}{z_1}, 1\right) = \left(\frac{x_2}{z_2}, \frac{y_2}{z_2}, 1\right) \sim (x_2, y_2, z_2).$$

Let $(x : y : z)$ denote the equivalence class of (x,y,z), i.e. $(x : y : z)$ is the following set

$$(x : y : z) = \left\{(u,v,w) \in \mathbb{C}^3 - \{(0,0,0)\} : (x,y,z) \sim (u,v,w)\right\}.$$

Exercise 1.4.5.

(1) Find the equivalence class of $(0,0,1)$.

(2) Find the equivalence class of $(1,2,3)$.

Exercise 1.4.6. Show that the equivalence classes $(1 : 2 : 3)$ and $(2 : 4 : 6)$ are equal as sets.

Definition 1.4.2. The *complex projective plane* $\mathbb{P}^2(\mathbb{C})$ is the set of equivalence classes of the points in $\mathbb{C}^3 - \{(0,0,0)\}$. This is often written as

$$\mathbb{P}^2(\mathbb{C}) = \left(\mathbb{C}^3 - \{(0,0,0)\}\right) \Big/ \sim .$$

The set of points $\{(x : y : z) \in \mathbb{P}^2(\mathbb{C}) : z = 0\}$ is called the *line at infinity*. We will write $\mathbb{P}^2$ to mean $\mathbb{P}^2(\mathbb{C})$ when the context is clear.

Let $(a, b, c) \in \mathbb{C}^3 - \{(0,0,0)\}$. Then the complex line through this point and the origin $(0,0,0)$ can be defined as all points, (x, y, z), satisfying

$$x = \lambda a,\ y = \lambda b,\ \text{and } z = \lambda c$$

for any complex number λ. Here λ can be thought of as an independent parameter.

Exercise 1.4.7. Explain why the elements of $\mathbb{P}^2$ can intuitively be thought of as complex lines through the origin in $\mathbb{C}^3$.

Exercise 1.4.8. If $c \neq 0$, show, in $\mathbb{C}^3$, that the line $x = \lambda a,\ y = \lambda b,\ z = \lambda c$ intersects the plane $\{(x, y, z) : z = 1\}$ in exactly one point. Show that this point of intersection is $\left(\frac{a}{c}, \frac{b}{c}, 1\right)$.

In the next several exercises we will use

$$\mathbb{P}^2 = \{(x : y : z) \in \mathbb{P}^2 : z \neq 0\} \cup \{(x : y : z) \in \mathbb{P}^2 : z = 0\}$$

to show that $\mathbb{P}^2$ can be viewed as the union of $\mathbb{C}^2$ with the line at infinity.

Exercise 1.4.9. Show that the map $\phi : \mathbb{C}^2 \to \{(x : y : z) \in \mathbb{P}^2 : z \neq 0\}$ defined by $\phi(x, y) = (x : y : 1)$ is a bijection.

Exercise 1.4.10. Find a map from $\{(x : y : z) \in \mathbb{P}^2 : z \neq 0\}$ to $\mathbb{C}^2$ that is the inverse of the map ϕ in Exercise 1.4.9.

The maps ϕ and ϕ^{-1} in Exercises 1.4.9 and 1.4.10 show us how to view $\mathbb{C}^2$ inside $\mathbb{P}^2$. Now we show how the set $\{(x : y : z) \in \mathbb{P}^2 : z = 0\}$ corresponds to directions towards infinity in $\mathbb{C}^2$.

Exercise 1.4.11. Consider the line $\ell = \{(x, y) \in \mathbb{C}^2 : ax + by + c = 0\}$ in $\mathbb{C}^2$. Assume $a, b \neq 0$. Explain why, as $|x| \to \infty$, $|y| \to \infty$. (Here, $|x|$ is the modulus of x.)

Exercise 1.4.12. Consider again the line ℓ. We know that a and b cannot both be 0, so we will assume without loss of generality that $b \neq 0$.

(1) Show that the image of ℓ in $\mathbb{P}^2$ under ϕ is the set

$$\{(bx : -ax - c : b) : x \in \mathbb{C}\}.$$

(2) Show that this set equals the following union.

$$\{(bx : -ax - c : b) : x \in \mathbb{C}\} = \{(0 : -c : b)\} \cup \left\{ \left(1 : -\frac{a}{b} - \frac{c}{bx} : \frac{1}{x}\right) \right\}.$$

(3) Show that as $|x| \to \infty$, the second set in the above union becomes

$$\{(1 : -\frac{a}{b} : 0)\}.$$

Thus, the points $(1 : -\dfrac{a}{b} : 0)$ are directions toward infinity and the set $\{(x : y : z) \in \mathbb{P}^2 : z = 0\}$ is the *line at infinity*.

If a point $(a : b : c)$ in $\mathbb{P}^2$ is the image of a point $(x, y) \in \mathbb{C}^2$ under the map from $\phi : \mathbb{C}^2 \to \mathbb{P}^2$, we say that (a, b, c) are *homogeneous coordinates* for (x, y). Notice that homogeneous coordinates for a point $(x, y) \in \mathbb{C}^2$ are not unique. For example, the points $(2 : -3 : 1)$, $(10 : -15 : 5)$, and $(2 - 2i : -3 + 3i : 1 - i)$ all provide homogeneous coordinates for $(2, -3)$.

In order to consider zero sets of polynomials in $\mathbb{P}^2$, a little care is needed. We start with:

Definition 1.4.3. A polynomial is *homogeneous* if every monomial term has the same total degree, that is, if the sum of the exponents in every monomial is the same. The *degree* of the homogeneous polynomial is the total degree of any of its monomials. An equation is homogeneous if every non-zero monomial has the same total degree.

Exercise 1.4.13. Explain why the following polynomials are homogeneous, and find each degree.

(1) $x^2 + y^2 - z^2$

(2) $xz - y^2$

(3) $x^3 + 3xy^2 + 4y^3$

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

Exercise 1.4.14. Explain why the following polynomials are not homogeneous.

(1) $x^2 + y^2 - z$

(2) $xz - y$

(3) $x^2 + 3xy^2 + 4y^3 + 3$

(4) $x^3 + x^2y^2 + x^2$

Exercise 1.4.15. Show that if the homogeneous equation $Ax+By+Cz = 0$ holds for the point (x, y, z) in $\mathbb{C}^3 - \{(0, 0, 0)\}$, then it holds for every point of $\mathbb{C}^3$ that belongs to the equivalence class $(x : y : z)$ in $\mathbb{P}^2$.

Exercise 1.4.16. Show that if the homogeneous equation $Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz = 0$ holds for the point (x, y, z) in $\mathbb{C}^3 - \{(0, 0, 0)\}$, then it holds for every point of $\mathbb{C}^3$ that belongs to the equivalence class $(x : y : z)$ in $\mathbb{P}^2$.

Exercise 1.4.17. State and prove the generalization of the previous two exercises for any degree n homogeneous equation $P(x, y, z) = 0$.

Exercise 1.4.18. Consider the non-homogeneous equation $P(x, y, z) = x^2+2y+2z = 0$. Show that $(2, -1, -1)$ satisfies this equation. Find a point of the equivalence class $(2 : -1 : -1)$ that does not satisfy the equation.

Thus the zero set of a non-homogeneous polynomial is not well-defined in $\mathbb{P}^2$. These exercises demonstrate that the only zero sets of polynomials that are well-defined on $\mathbb{P}^2$ are homogeneous polynomials.

To study the behavior at infinity of a curve in $\mathbb{C}^2$, we would like to extend the curve to $\mathbb{P}^2$. Thus, we want to be able to pass from zero sets of polynomials in $\mathbb{C}^2$ to zero sets of homogeneous polynomials in $\mathbb{P}^2$. This motivates our next step, a method to *homogenize* polynomials.

We start with an example. For any point $(x : y : z) \in \mathbb{P}^2$ with $z \neq 0$ we have $(x : y : z) = \left(\frac{x}{z} : \frac{y}{z} : 1\right)$, which we identify, via ϕ^{-1} from Exercise 1.4.10, with the point $\left(\frac{x}{z}, \frac{y}{z}\right) \in \mathbb{C}^2$.

Under this identification, the polynomial $P(x, y) = y - x - 2$ maps to $P(x, y, z) = \frac{y}{z} - \frac{x}{z} - 2$. Since $P(x, y, z) = 0$ and $zP(x, y, z) = 0$ have the same zero set if $z \neq 0$ we clear the denominator and, with an abuse of notation, consider the homogeneous polynomial $P(x, y, z) = y - x - 2z$. The zero set of $P(x, y, z) = y - x - 2z$ in $\mathbb{P}^2$ corresponds to the zero set of $P(x, y) = y - x - 2 = 0$ in $\mathbb{C}^2$ precisely when $z = 1$.

Similarly, the polynomial $x^2 + y^2 - 1$ maps to $\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 - 1$. Again, clear the denominators to obtain the homogeneous polynomial $x^2 + y^2 - z^2$, whose zero set $\mathrm{V}(x^2 + y^2 - z^2) \subset \mathbb{P}^2$ corresponds to the zero set $\mathrm{V}(x^2 + y^2 - 1) \subset \mathbb{C}^2$ when $z = 1$.

Definition 1.4.4. Let $P(x, y)$ be a degree n polynomial defined over $\mathbb{C}^2$. The corresponding homogeneous polynomial defined over $\mathbb{P}^2$ is

$$P(x, y, z) = z^n P\left(\frac{x}{z}, \frac{y}{z}\right).$$

This method is called the *homogenization* of $P(x, y)$.

In a similar manner, we can homogenize an equation.

Exercise 1.4.19. Homogenize the following equations. Then find the point(s) where the curves intersect the line at infinity.

(1) $ax + by + c = 0$
(2) $x^2 + y^2 = 1$
(3) $y = x^2$
(4) $x^2 + 9y^2 = 1$
(5) $y^2 - x^2 = 1$

Exercise 1.4.20. Show that in $\mathbb{P}^2$, any two distinct lines will intersect in a point. Notice this implies that parallel lines in $\mathbb{C}^2$, when embedded in $\mathbb{P}^2$, intersect at the line at infinity.

Exercise 1.4.21. Once we have homogenized an equation, the original variables x and y are no more important than the variable z. Suppose we regard x and z as the original variables in our homogenized equation. Then the image of the xz-plane in $\mathbb{P}^2$ would be $\{(x : y : z) \in \mathbb{P}^2 : y = 1\}$.

(1) Homogenize the equations for the parallel lines $y = x$ and $y = x + 2$.

(2) Now regard x and z as the original variables, and set $y = 1$ to sketch the image of the lines in the xz-plane.

(3) Explain why the lines in part (2) meet at the x-axis.

1.5. Projective Changes of Coordinates

The goal of this section is to define a projective change of coordinates and then show that all ellipses, hyperbolas, and parabolas are equivalent under projective changes of coordinates.

Earlier we described a complex affine change of coordinates from points $(x, y) \in \mathbb{C}^2$ to points $(u, v) \in \mathbb{C}^2$ by setting $u = ax + bx + e$ and $v = cx + dy + f$. We will define the analogue for changing homogeneous coordinates $(x : y : z) \in \mathbb{P}^2$ to homogeneous coordinates $(u : v : w) \in \mathbb{P}^2$. We need the change of coordinates equations to be both homogeneous and linear.

Definition 1.5.1. A *projective change of coordinates* is given by

$$\begin{aligned} u &= a_{11}x + a_{12}y + a_{13}z \\ v &= a_{21}x + a_{22}y + a_{23}z \\ w &= a_{31}x + a_{32}y + a_{33}z, \end{aligned}$$

where the $a_{ij} \in \mathbb{C}$ and

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \neq 0.$$

In matrix language

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

where $A = (a_{ij})$, $a_{ij} \in \mathbb{C}$, and $\det A \neq 0$.

For any affine change of coordinates, there is a corresponding projective change of coordinates as seen in the following:

Exercise 1.5.1. For the complex affine change of coordinates

$$\begin{aligned} u &= ax + by + e \\ v &= cx + dy + f, \end{aligned}$$

where $a, b, c, d, e, f \in \mathbb{C}$ and $ad - bc \neq 0$, show that

$$\begin{aligned} u &= ax + by + ez \\ v &= cx + dy + fz \\ w &= z \end{aligned}$$

is the corresponding projective change of coordinates.

Definition 1.5.2. Two conics in $\mathbb{P}^2$ are *equivalent under a projective change of coordinates*, or *projectively equivalent*, if the defining homogeneous polynomial for one of the conics can be transformed into the defining polynomial for the other conic via a projective change of coordinates.

By Exercise 1.5.1, if two conics in $\mathbb{C}^2$ are equivalent under a complex affine change of coordinates, then the corresponding conics in $\mathbb{P}^2$ will still be equivalent, but now under a projective change of coordinates.

Exercise 1.5.2. Let $C_1 = \mathrm{V}(x^2 + y^2 - 1)$ be an ellipse in $\mathbb{C}^2$ and let $C_2 = \mathrm{V}(u^2 - v)$ be a parabola in $\mathbb{C}^2$. Homogenize the defining polynomials for C_1 and C_2 and show that the projective change of coordinates

$$\begin{aligned} u &= ix \\ v &= y + z \\ w &= y - z \end{aligned}$$

transforms the ellipse in $\mathbb{P}^2$ into the parabola in $\mathbb{P}^2$.

Exercise 1.5.3. Use the results of Section 1.3, together with the above problem, to show that, under a projective change of coordinates, all ellipses, hyperbolas, and parabolas are equivalent in $\mathbb{P}^2$.

1.6. The Complex Projective Line $\mathbb{P}^1$

The goal of this section is to define the complex projective line $\mathbb{P}^1$ and show that it can be viewed topologically as a sphere. In the next section we will use this to show that ellipses, hyperbolas, and parabolas are also topologically spheres.

We start with the definition of $\mathbb{P}^1$.

Definition 1.6.1. Define an equivalence relation $\sim$ on points in $\mathbb{C}^2 - \{(0,0)\}$ as follows: $(x,y) \sim (u,v)$ if and only if there exists $\lambda \in \mathbb{C} - \{0\}$ such that $(x,y) = (\lambda u, \lambda v)$. Let $(x:y)$ denote the equivalence class of (x,y). The *complex projective line* $\mathbb{P}^1$ is the set of equivalence classes of points in $\mathbb{C}^2 - \{(0,0)\}$. That is,

$$\mathbb{P}^1 = \big(\mathbb{C}^2 - \{(0,0)\}\big) \Big/ \sim .$$

The point $(1:0)$ is called the *point at infinity*.

The next series of problems are direct analogues of problems for $\mathbb{P}^2$.

Exercise 1.6.1. Suppose that $(x_1, y_1) \sim (x_2, y_2)$ and that $x_1 = x_2 \neq 0$. Show that $y_1 = y_2$.

Exercise 1.6.2. Suppose that $(x_1, y_1) \sim (x_2, y_2)$ with $y_1 \neq 0$ and $y_2 \neq 0$. Show that

$$(x_1, y_1) \sim \left(\frac{x_1}{y_1}, 1\right) = \left(\frac{x_2}{y_2}, 1\right) \sim (x_2, y_2).$$

Exercise 1.6.3. Explain why the elements of $\mathbb{P}^1$ can intuitively be thought of as complex lines through the origin in $\mathbb{C}^2$.

Exercise 1.6.4. If $b \neq 0$, show that the line $x = \lambda a$, $y = \lambda b$ will intersect the line $\{(x,y) : y = 1\}$ in exactly one point. Show that this point of intersection is $\left(\frac{a}{b}, 1\right)$.

We have that

$$\mathbb{P}^1 = \big\{(x:y) \in \mathbb{P}^1 : y \neq 0\big\} \cup \{(1:0)\}.$$

Exercise 1.6.5. Show that the map $\phi : \mathbb{C} \to \{(x : y) \in \mathbb{P}^1 : y \neq 0\}$ defined by $\phi(x) = (x : 1)$ is a bijection.

Exercise 1.6.6. Find a map from $\{(x : y) \in \mathbb{P}^1 : y \neq 0\}$ to $\mathbb{C}$ that is the inverse of the map ϕ in Exercise 1.6.5.

The maps ϕ and ϕ^{-1} in Exercises 1.6.5 and 1.6.6 show us how to view $\mathbb{C}$ inside $\mathbb{P}^1$. Now we want to see how the extra point $(1 : 0)$ will correspond to the point at infinity of $\mathbb{C}$.

Exercise 1.6.7. Consider the map $\phi : \mathbb{C} \to \mathbb{P}^1$ given by $\phi(x) = (x : 1)$. Show that as $|x| \to \infty$, we have $\phi(x) \to (1 : 0)$.

Hence we can think of $\mathbb{P}^1$ as the union of $\mathbb{C}$ and a single point at infinity. Now we want to see how we can regard $\mathbb{P}^1$ as a sphere, which means we want to find a homeomorphism between $\mathbb{P}^1$ and a sphere. A *homeomorphism* is a continuous map with a continuous inverse. Two spaces are topologically equivalent, or homeomorphic, if we can find a homeomorphism from one to the other. We know that the points of $\mathbb{C}$ are in one-to-one correspondence with the points of the real plane $\mathbb{R}^2$, so we will first work in $\mathbb{R}^2 \subset \mathbb{R}^3$. Specifically, identify $\mathbb{R}^2$ with the xy-plane in $\mathbb{R}^3$ via $(x, y) \mapsto (x, y, 0)$. Let S^2 denote the unit sphere in $\mathbb{R}^3$ centered at the origin. This sphere is given by the equation

$$x^2 + y^2 + z^2 = 1.$$

Exercise 1.6.8. Let p denote the point $(0, 0, 1) \in S^2$, and let ℓ denote the line through p and the point $(x, y, 0)$ in the xy-plane, whose parametrization is given by

$$\gamma(t) = (1 - t)(0, 0, 1) + t(x, y, 0),$$

i.e.,

$$\ell = \{(tx, ty, 1 - t) \mid t \in \mathbb{R}\}.$$

(1) ℓ clearly intersects S^2 at the point p. Show that there is exactly one other point of intersection q.

(2) Find the coordinates of q.

(3) Define the map $\psi : \mathbb{R}^2 \to S^2 - \{p\}$ to be the map that takes the point (x, y) to the point q. Show that ψ is a continuous bijection.

(4) Show that as $\sqrt{x^2+y^2} \to \infty$, we have $\psi(x,y) \to p$. Thus as we move away from the origin in $\mathbb{R}^2$, $\psi(x,y)$ moves toward the North Pole.

The above argument does establish a homeomorphism, but it relies on coordinates and an embedding of the sphere in $\mathbb{R}^3$. We now give an alternative method for showing that $\mathbb{P}^1$ is a sphere that does not rely as heavily on coordinates.

If we take a point $(x : y) \in \mathbb{P}^1$, then we can choose a representative for this point of the form $\left(\frac{x}{y} : 1\right)$, provided $y \neq 0$, and a representative of the form $\left(1 : \frac{y}{x}\right)$, provided $x \neq 0$.

Exercise 1.6.9. Determine which point(s) in $\mathbb{P}^1$ do **not** have two representatives of the form $(x : 1) = \left(1 : \frac{1}{x}\right)$.

Our construction needs two copies of $\mathbb{C}$. Let U denote the first copy of $\mathbb{C}$, whose elements are denoted by x. Let V be the second copy of $\mathbb{C}$, whose elements we'll denote y. Further let $U^* = U - \{0\}$ and $V^* = V - \{0\}$.

Exercise 1.6.10. Map $U \to \mathbb{P}^1$ via $x \mapsto (x : 1)$ and map $V \to \mathbb{P}^1$ via $y \mapsto (1 : y)$. Show that $(x : 1) \mapsto \left(1 : \frac{1}{x}\right)$ is a natural one-to-one map from U^* onto V^*.

The next two exercises have quite a different flavor than most of the problems in the book. The emphasis is not on calculations but on the underlying intuitions.

Exercise 1.6.11. A sphere can be split into a neighborhood of its northern hemisphere and a neighborhood of its southern hemisphere. Show that a sphere can be obtained by correctly gluing together two copies of $\mathbb{C}$.

Exercise 1.6.12. Put together the last two exercises to show that $\mathbb{P}^1$ is topologically equivalent to a sphere.

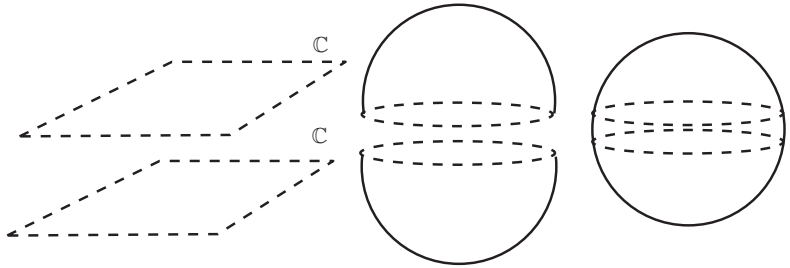


Figure 6. Gluing copies of $\mathbb{C}$ together.

1.7. Ellipses, Hyperbolas, and Parabolas as Spheres

The goal of this section is to show that there is always a bijective polynomial map from $\mathbb{P}^1$ to any ellipse, hyperbola, or parabola. Since we showed in the last section that $\mathbb{P}^1$ is topologically equivalent to a sphere, this means that all ellipses, hyperbolas, and parabolas are spheres.

We start with rational parameterizations of conics. While we will consider conics in the complex plane $\mathbb{C}^2$, we often draw these conics in $\mathbb{R}^2$. Part of learning algebraic geometry is developing a sense for when the real pictures capture what is going on in the complex plane.

Consider a conic $C = \{(x, y) \in \mathbb{C}^2 : P(x, y) = 0\} \subset \mathbb{C}^2$, where $P(x, y)$ is a second degree polynomial. Our goal is to parametrize C with polynomial or rational maps. This means we want to find a map $\phi : \mathbb{C} \to C \subset \mathbb{C}^2$, given by $\phi(\lambda) = (x(\lambda), y(\lambda))$ such that $x(\lambda)$ and $y(\lambda)$ are polynomials or rational functions. In the case of a parabola, for example when $P(x, y) = x^2 - y$, it is easy to find a bijection from $\mathbb{C}$ to the conic C.

Exercise 1.7.1. Find a bijective polynomial map from $\mathbb{C}$ to the conic $C = \{(x, y) \in \mathbb{C}^2 : x^2 - y = 0\}$.

Sometimes it may be easy to find a parametrization but not one that is rational.

Exercise 1.7.2. Let $C = \mathrm{V}(x^2 + y^2 - 1)$ be an ellipse in $\mathbb{C}^2$. Find a trigonometric parametrization of C. [Hint: Think high school trigonometry.]

This exercise gives a parameterization for the circle, but in algebraic geometry we restrict our parameterizations to polynomial or rational maps. We develop a standard method, similar to the method developed in Exercise 1.6.8, to find such a parameterization below.

Exercise 1.7.3. Consider the ellipse $C = \mathrm{V}(x^2 + y^2 - 1) \subset \mathbb{C}^2$ and let p denote the point $(0, 1) \in C$.

(1) Parametrize the line segment from p to the point $(\lambda, 0)$ on the complex line $y = 0$ as in Exercise 1.6.8.

(2) This line segment clearly intersects C at the point p. Show that if $\lambda \neq \pm i$, then there is exactly one other point of intersection. Call this point q.

(3) Find the coordinates of $q \in C$.

(4) Show that if $\lambda = \pm i$, then the line segment intersects C only at p.

Define the map $\widetilde{\psi} : \mathbb{C} \to C \subset \mathbb{C}^2$ by

$$\widetilde{\psi}(\lambda) = \left(\frac{2\lambda}{\lambda^2 + 1}, \frac{\lambda^2 - 1}{\lambda^2 + 1} \right).$$

But we want to work in projective space. This means that we have to homogenize our map.

Exercise 1.7.4. Show that the above map can be extended to the map

$$\psi : \mathbb{P}^1 \to \{(x : y : z) \in \mathbb{P}^2 : x^2 + y^2 - z^2 = 0\}$$

given by

$$\psi(\lambda : \mu) = (2\lambda\mu : \lambda^2 - \mu^2 : \lambda^2 + \mu^2).$$

Exercise 1.7.5.

(1) Show that the map ψ is one-to-one.

(2) Show that ψ is onto. [Hint: Consider two cases: $z \neq 0$ and $z = 0$. For $z \neq 0$ follow the construction given above. For

$z = 0$, find values of λ and μ to show that these points are given by ψ. How does this relate to Part 4 of Exercise 1.7.3?]

Since we already know that every ellipse, hyperbola, and parabola is projectively equivalent to the conic defined by $x^2 + y^2 - z^2 = 0$, we have, by composition, a one-to-one and onto map from $\mathbb{P}^1$ to any ellipse, hyperbola, or parabola.

However, we can construct such maps directly. Here is what we can do for any conic C. Fix a point p on C, and parametrize the line segment through p and the point $(\lambda, 0)$. We use this to determine another point on the curve C, and the coordinates of this point give us our map.

Exercise 1.7.6. For the following conics and the given point p, follow what we did for the conic $x^2 + y^2 - 1 = 0$ to find a rational map from $\mathbb{C}$ to the curve in $\mathbb{C}^2$ and then a one-to-one map from $\mathbb{P}^1$ onto the conic in $\mathbb{P}^2$.

(1) $x^2 + 2x - y^2 - 4y - 4 = 0$ with $p = (0, -2)$
(2) $3x^2 + 3y^2 - 75 = 0$ with $p = (5, 0)$
(3) $4x^2 + y^2 - 8 = 0$ with $p = (1, 2)$

1.8. Links to Number Theory

The goal of this section is to see how geometry can be used to find all primitive Pythagorean triples, a classic problem from number theory.

Overwhelmingly, in this book we are interested in working over the complex numbers. If instead we work over the integers or the rational numbers, some of the deepest questions in mathematics appear.

We want to see this approach in the case of conics. In particular we want to link the last section to the search for primitive Pythagorean triples. A *Pythagorean triple* is a triple, (x, y, z), of integers that satisfies the equation

$$x^2 + y^2 = z^2.$$

Exercise 1.8.1. Suppose (x_0, y_0, z_0) is a solution to $x^2 + y^2 = z^2$. Show that (mx_0, my_0, mz_0) is also a solution for any scalar m.

A *primitive Pythagorean triple* is a Pythagorean triple that cannot be obtained by multiplying another Pythagorean triple by an integer. The simplest example, after the trivial solution $(0, 0, 0)$, is $(3, 4, 5)$. These triples get their name from the attempt to find right triangles with integer length sides, x, y, and z. We will see that the previous section gives us a method to compute all possible primitive Pythagorean triples.

We first see how to translate the problem of finding integer solutions of $x^2+y^2 = z^2$ to finding rational number solutions to $x^2+y^2 = 1$.

Exercise 1.8.2. Let $(a, b, c) \in \mathbb{Z}^3$ be a solution to $x^2+y^2 = z^2$. Show that $c = 0$ if and only if $a = b = 0$.

This means that we can assume $c \neq 0$, since there is only one solution when $c = 0$.

Exercise 1.8.3. Show that if (a, b, c) is a Pythagorean triple with $c \neq 0$, then the pair of rational numbers $\left(\frac{a}{c}, \frac{b}{c}\right)$ is a solution to $x^2 + y^2 = 1$.

Exercise 1.8.4. Let $\left(\frac{a}{c_1}, \frac{b}{c_2}\right) \in \mathbb{Q}^2$ be a rational solution to $x^2 + y^2 = 1$. Find a corresponding Pythagorean triple.

Thus to find Pythagorean triples, we want to find the rational points on the curve $x^2 + y^2 = 1$. We denote these points as

$$C(\mathbb{Q}) = \{(x, y) \in \mathbb{Q}^2 : x^2 + y^2 = 1\}.$$

Recall from the last section, the parameterization

$$\widetilde{\psi} : \mathbb{Q} \to \{(x, y) \in \mathbb{Q}^2 : x^2 + y^2 = 1\}$$

given by

$$\lambda \mapsto \left(\frac{2\lambda}{\lambda^2 + 1}, \frac{\lambda^2 - 1}{\lambda^2 + 1}\right).$$

Exercise 1.8.5. Show that the above map $\widetilde{\psi}$ sends $\mathbb{Q} \to C(\mathbb{Q})$.

Extend this to a map $\psi : \mathbb{P}^1(\mathbb{Q}) \to C(\mathbb{Q}) \subset \mathbb{P}^2(\mathbb{Q})$ by

$$(\lambda : \mu) \mapsto (2\lambda\mu : \lambda^2 - \mu^2 : \lambda^2 + \mu^2),$$

where $\lambda, \mu \in \mathbb{Z}$. Since we know already that the map ψ is one-to-one by Exercise 1.7.5, this gives us a way to produce an infinite number of integer solutions to $x^2 + y^2 = z^2$.

We now want to show that the map ψ is onto, so that we actually obtain all Pythagorean triples.

Exercise 1.8.6.

(1) Show that $\psi : \mathbb{P}^1(\mathbb{Q}) \to C(\mathbb{Q}) \subset \mathbb{P}^2(\mathbb{Q})$ is onto.

(2) Show that every primitive Pythagorean triple is of the form $(2\lambda\mu, \lambda^2 - \mu^2, \lambda^2 + \mu^2)$.

Exercise 1.8.7. Find a rational point on the conic $x^2 + y^2 - 2 = 0$. Develop a parameterization and conclude that there are infinitely many rational points on this curve.

Exercise 1.8.8. By mimicking the above, find four rational points on each of the following conics.

(1) $x^2 + 2x - y^2 - 4y - 4 = 0$ with $p = (0, -2)$

(2) $3x^2 + 3y^2 - 75 = 0$ with $p = (5, 0)$

(3) $4x^2 + y^2 - 8 = 0$ with $p = (1, 2)$

Exercise 1.8.9. Show that the conic $x^2 + y^2 = 3$ has no rational points.

Diophantine problems are those where you try to find integer or rational solutions to a polynomial equation. The work in this section shows how we can approach such problems using algebraic geometry. For higher degree equations the situation is quite different and leads to the heart of a great deal of the current research in number theory.

1.9. Degenerate Conics

The goal of this section is to extend our study of conics from ellipses, hyperbolas, and parabolas to the "degenerate" conics: crossing lines and double lines.

Let $f(x, y, z)$ be any homogeneous second degree polynomial with complex coefficients. The overall goal of this chapter is to understand curves

$$C = \{(x : y : z) \in \mathbb{P}^2 : f(x, y, z) = 0\}.$$

Most of these curves will be various ellipses, hyperbolas, and parabolas. Now consider the second degree polynomial

$$f(x, y, z) = (-x+y+z)(2x+y+3z) = -2x^2+y^2+3z^2+xy-xz+4yz.$$

Exercise 1.9.1. Dehomogenize $f(x, y, z)$ by setting $z = 1$. Graph the curve

$$C(\mathbb{R}) = \{(x : y : z) \in \mathbb{P}^2 : f(x, y, 1) = 0\}$$

in the real plane $\mathbb{R}^2$.

The zero set of a second degree polynomial could be the union of crossing lines.

Exercise 1.9.2. Consider the two lines given by

$$(a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z) = 0,$$

and suppose

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \neq 0.$$

Show that the two lines intersect at a point where $z \neq 0$.

Exercise 1.9.3. Dehomogenize the equation in the previous exercise by setting $z = 1$. Give an argument that, as lines in the complex plane $\mathbb{C}^2$, they have distinct slopes.

Exercise 1.9.4. Again consider the two lines

$$(a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z) = 0.$$

Suppose that

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = 0$$

but that

$$\det \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \neq 0 \text{ or } \det \begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix} \neq 0.$$

Show that the two lines still have one common point of intersection, but that this point must have $z = 0$.

There is one other possibility. Consider the zero set

$$C = \{(x : y : z) \in \mathbb{P}^2 : (ax + by + cz)^2 = 0\}.$$

As a zero set, the curve C is geometrically the line

$$ax + by + cz = 0$$

but due to the exponent 2, we call C a *double line.*

Exercise 1.9.5. Let

$$f(x, y, z) = (a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z),$$

where at least one of $a_1, b_1,$ or c_1 is non-zero and at least one of the $a_2, b_2,$ or c_2 is non-zero. Show that the curve defined by $f(x, y, z) = 0$ is a double line if and only if

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = 0, \quad \det \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} = 0, \quad \text{and } \det \begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix} = 0.$$

We now want to show that any two crossing lines are equivalent under a projective change of coordinates to any other two crossing lines and any double line is equivalent under a projective change of coordinates to any other double line. This will yield that there are precisely three types of conics: the ellipses, hyperbolas, and parabolas; crossing lines; and double lines.

For the exercises that follow, assume that at least one of $a_1, b_1,$ or c_1 is non-zero and at least one of $a_2, b_2,$ or c_2 is non-zero.

Exercise 1.9.6. Consider the crossing lines

$$(a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z) = 0,$$

with

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \neq 0.$$

Find a projective change of coordinates from xyz-space to uvw-space so that the crossing lines become

$$uv = 0.$$

Exercise 1.9.7. Consider the crossing lines $(a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z) = 0$, with

$$\det \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \neq 0.$$

Find a projective change of coordinates from xyz-space to uvw-space so that the crossing lines become

$$uv = 0.$$

Exercise 1.9.8. Show that there is a projective change of coordinates from xyz-space to uvw-space so that the double line $(ax+by+cz)^2 = 0$ becomes the double line

$$u^2 = 0.$$

Exercise 1.9.9. Argue that there are three distinct classes of conics in $\mathbb{P}^2$.

1.10. Tangents and Singular Points

The goal of this section is to develop the idea of singularity. We'll show that all ellipses, hyperbolas, and parabolas are smooth, while crossing lines and double lines are singular.

So far, we have not explicitly needed to use calculus; that changes in this section. We will use the familiar differentiation rules from real calculus.

Let $f(x, y)$ be a polynomial. Recall that if $f(a, b) = 0$, then a normal vector for the curve $f(x, y) = 0$ at the point (a, b) is given by the gradient vector

$$\nabla f(a, b) = \left(\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b) \right).$$

A tangent vector to the curve at the point (a, b) is perpendicular to $\nabla f(a, b)$ and hence must have a dot product of zero with $\nabla f(a, b)$. This observation shows that the tangent line is given by

$$\left\{ (x, y) \in \mathbb{C}^2 : \left(\frac{\partial f}{\partial x}(a, b) \right)(x - a) + \left(\frac{\partial f}{\partial y}(a, b) \right)(y - b) = 0 \right\}.$$

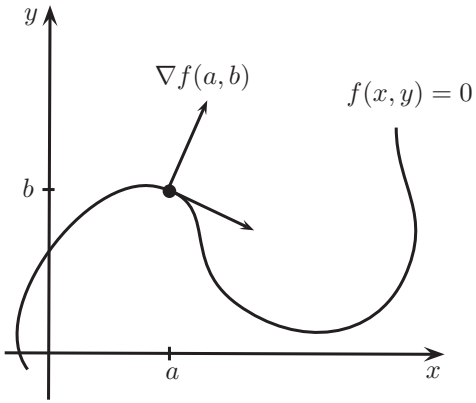


Figure 7. Gradient and tangent vectors.

Exercise 1.10.1. Explain why if both $\dfrac{\partial f}{\partial x}(a,b) = 0$ and $\dfrac{\partial f}{\partial y}(a,b) = 0$, then the tangent line is not well-defined at (a,b).

This exercise motivates the following definition.

Definition 1.10.1. A point $p = (a,b)$ on a curve $C = \{(x,y) \in \mathbb{C}^2 : f(x,y) = 0\}$ is said to be *singular* if

$$\frac{\partial f}{\partial x}(a,b) = 0 \text{ and } \frac{\partial f}{\partial y}(a,b) = 0.$$

A point that is not singular is called *smooth*. If there is at least one singular point on C, then the curve C is called *singular*. If there are no singular points on C, the curve C is called *smooth*.

Exercise 1.10.2. Show that the curve

$$C = \{(x,y) \in \mathbb{C}^2 : x^2 + y^2 - 1 = 0\}$$

is smooth.

Exercise 1.10.3. Show that the pair of crossing lines

$$C = \{(x,y) \in \mathbb{C}^2 : (x + y - 1)(x - y - 1) = 0\}$$

has exactly one singular point. Give a geometric interpretation of this singular point.

Exercise 1.10.4. Show that every point on the double line

$$C = \{(x, y) \in \mathbb{C}^2 : (2x + 3y - 4)^2 = 0\}$$

is singular.

These definitions can also be applied to curves in $\mathbb{P}^2$.

Definition 1.10.2. A point $p = (a : b : c)$ on a curve $C = \{(x : y : z) \in \mathbb{P}^2 : f(x, y, z) = 0\}$, where $f(x, y, z)$ is a homogeneous polynomial, is said to be *singular* if

$$\frac{\partial f}{\partial x}(a, b, c) = 0, \quad \frac{\partial f}{\partial y}(a, b, c) = 0, \text{ and } \quad \frac{\partial f}{\partial z}(a, b, c) = 0.$$

We have similar definitions, as before, for smooth point, smooth curve, and singular curve.

Exercise 1.10.5. Show that the curve

$$C = \{(x : y : z) \in \mathbb{P}^2 : x^2 + y^2 - z^2 = 0\}$$

is smooth.

Exercise 1.10.6. Show that the pair of crossing lines

$$C = \{(x : y : z) \in \mathbb{P}^2 : (x + y - z)(x - y - z) = 0\}$$

has exactly one singular point.

Exercise 1.10.7. Show that every point on the double line

$$C = \{(x : y : z) \in \mathbb{P}^2 : (2x + 3y - 4z)^2 = 0\}$$

is singular.

For homogeneous polynomials, there is a simple relationship among f, $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, and $\frac{\partial f}{\partial z}$, which is the goal of the next few exercises.

Exercise 1.10.8. For

$$f(x, y, z) = x^2 + 3xy + 5xz + y^2 - 7yz + 8z^2,$$

show that

$$2f = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z}.$$

Exercise 1.10.9. For

$$f(x, y, z) = ax^2 + bxy + cxz + dy^2 + eyz + hz^2,$$

show that

$$2f = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z}.$$

Exercise 1.10.10. Let $f(x, y, z)$ be a homogeneous polynomial of degree n. Show that

$$nf = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z}.$$

(This problem is quite similar to the previous two, but working out the details takes some work.)

Exercise 1.10.11. Use Exercise 1.10.10 to show that if $p = (a : b : c)$ satisfies

$$\frac{\partial f}{\partial x}(a, b, c) = \frac{\partial f}{\partial y}(a, b, c) = \frac{\partial f}{\partial z}(a, b, c) = 0,$$

then $p \in \mathrm{V}(f)$.

The notion of smooth curves and singular curves certainly extends beyond the study of conics. We will briefly discuss higher degree curves here. Throughout, we will see that *singular* corresponds to not having a well-defined tangent.

Exercise 1.10.12. Graph the curve

$$f(x, y) = x^3 + x^2 - y^2 = 0$$

in the real plane $\mathbb{R}^2$. What is happening at the origin $(0, 0)$? Find the singular points.

Exercise 1.10.13. Graph the curve

$$f(x, y) = x^3 - y^2 = 0$$

in the real plane $\mathbb{R}^2$. What is happening at the origin $(0, 0)$? Find the singular points.

For any two polynomials, $f_1(x, y)$ and $f_2(x, y)$, let $f(x, y) = f_1(x, y)f_2(x, y)$ be their product. We have

$$\mathrm{V}(f) = \mathrm{V}(f_1) \cup \mathrm{V}(f_2).$$

The picture of these curves is

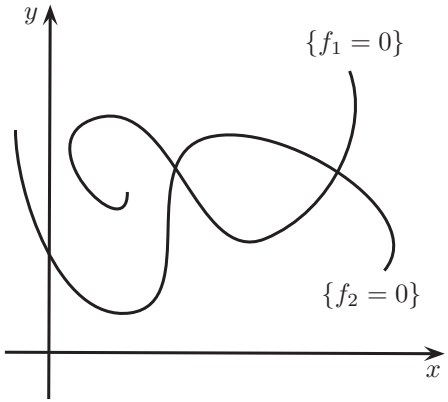


Figure 8. Curves $\mathrm{V}(f_1)$ and $\mathrm{V}(f_2)$.

From the picture, it seems that the curve $\mathrm{V}(f)$ should have singular points at the points of intersection of $\mathrm{V}(f_1)$ and $\mathrm{V}(f_2)$.

Exercise 1.10.14. Suppose that

$$f_1(a,b) = 0 \quad \text{and} \quad f_2(a,b) = 0$$

for a point $(a,b) \in \mathbb{C}^2$. Show that (a,b) is a singular point on $\mathrm{V}(f)$, where $f = f_1 f_2$.

While it is safe to say for higher degree curves and especially for higher dimensional algebraic geometric objects that "singularness" is far from understood, that is not the case for conics. A complete description is contained in the following theorem.

Theorem 1.10.15. All ellipses, hyperbolas, and parabolas are smooth curves. All conics that are crossing lines have exactly one singular point, namely the point of intersection of the two lines. Every point on a double line is singular.

We have seen specific examples for each of these. The proof of the theorem relies on the fact that under projective transformations there are three distinct classes of conics. We motivated the idea of

projective changes of coordinates as just the relabeling of coordinate systems. Surely how we label points on the plane should not affect the lack of a well-defined tangent line. Hence a projective change of coordinates should not affect whether or not a point is smooth or singular. The next series of exercises proves this.

Consider a projective change of coordinates from xyz-space to uvw-space given by

$$\begin{aligned} u &= a_{11}x + a_{12}y + a_{13}z \\ v &= a_{21}x + a_{22}y + a_{23}z \\ w &= a_{31}x + a_{32}y + a_{33}z, \end{aligned}$$

where $a_{ij} \in \mathbb{C}$ and

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \neq 0.$$

In $\mathbb{P}^2$, with homogeneous coordinates $(u : v : w)$, consider a curve $C = \{(u : v : w) : f(u, v, w) = 0\}$, where f is a homogeneous polynomial. The (inverse) change of coordinates above gives a map from polynomials in $(u : v : w)$ to polynomials in $(x : y : z)$ described by

$$\begin{aligned} f(u, v, w) \mapsto & f(a_{11}x + a_{12}y + a_{13}z, a_{21}x + a_{22}y + a_{23}z, \\ & a_{31}x + a_{32}y + a_{33}z) = \widetilde{f}(x, y, z). \end{aligned}$$

The curve C corresponds to the curve $\widetilde{C} = \{(x : y : z) : \widetilde{f}(x, y, z) = 0\}$.

Exercise 1.10.16. Consider the curve

$$C = \{(u : v : w) \in \mathbb{P}^2 : u^2 - v^2 - w^2 = 0\}.$$

Suppose we have the projective change of coordinates given by

$$\begin{aligned} u &= x + y \\ v &= x - y \\ w &= z. \end{aligned}$$

Show that C corresponds to the curve

$$\widetilde{C} = \{(x : y : z) \in \mathbb{P}^2 : 4xy - z^2 = 0\}.$$

In other words, if $f(u,v,w) = u^2 - v^2 - w^2$, then $\widetilde{f}(x,y,z) = 4xy - z^2$.

Exercise 1.10.17. Suppose we have the projective change of coordinates given by

$$\begin{aligned} u &= x + y \\ v &= x - y \\ w &= x + y + z. \end{aligned}$$

If $f(u,v,w) = u^2 + uw + v^2 + vw$, find $\widetilde{f}(x,y,z)$.

Exercise 1.10.18. For a general projective change of coordinates given by

$$\begin{aligned} u &= a_{11}x + a_{12}y + a_{13}z \\ v &= a_{21}x + a_{22}y + a_{23}z \\ w &= a_{31}x + a_{32}y + a_{33}z \end{aligned}$$

and a polynomial $f(u,v,w)$, describe how to find the corresponding $\widetilde{f}(x,y,z)$.

We now want to show, under a projective change of coordinates, that singular points go to singular points and smooth points go to smooth points.

Exercise 1.10.19. Let

$$\begin{aligned} u &= a_{11}x + a_{12}y + a_{13}z \\ v &= a_{21}x + a_{22}y + a_{23}z \\ w &= a_{31}x + a_{32}y + a_{33}z \end{aligned}$$

be a projective change of coordinates. Show that $(u_0 : v_0 : w_0)$ is a singular point of the curve $C = \{(u : v : w) : f(u,v,w) = 0\}$ if and only if the corresponding point $(x_0 : y_0 : z_0)$ is a singular point of the corresponding curve $\widetilde{C} = \{(x : y : z) : \widetilde{f}(x,y,z) = 0\}$. (This is an exercise in the multi-variable chain rule; most people are not comfortable with the chain rule without a lot of practice. Hence the value of this exercise.)

Exercise 1.10.20. Use the previous exercise to prove Theorem 1.10.15.

1.11. Conics via Linear Algebra

The goal of this section is to show how to interpret conics via linear algebra. The linear algebra of symmetric 3×3 matrices will lead to straightforward proofs that, under projective changes of coordinates, all ellipses, hyperbolas, and parabolas are equivalent; all crossing line conics are equivalent; and all double lines are equivalent.

1.11.1. Conics via 3×3 Symmetric Matrices. We start by showing how to represent conics with symmetric 3×3 matrices. Consider the second degree homogeneous polynomial

$$\begin{aligned} f(x,y,z) &= x^2 + 6xy + 5y^2 + 4xz + 8yz + 9z^2 \\ &= x^2 + (3xy + 3yx) + 5y^2 + (2xz + 2zx) \\ &\quad + (4yz + 4zy) + 9z^2 \\ &= \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 \\ 3 & 5 & 4 \\ 2 & 4 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \end{aligned}$$

By using seemingly silly tricks such as $6xy = 3xy + 3yx$, we have written our initial second degree polynomial in terms of the symmetric 3×3 matrix

$$\begin{pmatrix} 1 & 3 & 2 \\ 3 & 5 & 4 \\ 2 & 4 & 9 \end{pmatrix}.$$

There is nothing special about this particular second degree polynomial. We can write all homogeneous second degree polynomials $f(x, y, z)$ in terms of symmetric 3×3 matrices. (Recall that a matrix $A = (a_{ij})$ is symmetric if $a_{ij} = a_{ji}$ for all i and j. Since the transpose of A simply switches the row and column entries, $A^T = (a_{ji})$, A is symmetric if and only if $A = A^T$.)

Exercise 1.11.1. Write the following conics in the form

$$\begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0.$$

That is, find a symmetric matrix A for each quadratic equation.

(1) $x^2 + y^2 + z^2 = 0$
(2) $x^2 + y^2 - z^2 = 0$
(3) $x^2 - y^2 = 0$
(4) $x^2 + 2xy + y^2 + 3xz + z^2 = 0$

Symmetric matrices can be used to define second degree homogeneous polynomials with any number of variables.

Definition 1.11.1. A *quadratic form* is a homogeneous polynomial of degree two in any given number of variables. Given a symmetric $n \times n$ matrix A and $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, then $f(X) = X^T AX$ is a quadratic form.

Thus conics are defined by quadratic forms in three variables.

Exercise 1.11.2. Show that any conic

$$f(x, y, z) = ax^2 + bxy + cy^2 + dxz + eyz + hz^2$$

can be written as

$$\begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

where A is a symmetric 3×3 matrix.

1.11.2. Change of Variables via Matrices. We want to see that a projective change of coordinates has a quite natural linear algebra interpretation.

Suppose we have a projective change of coordinates

$$\begin{aligned} u &= a_{11}x + a_{12}y + a_{13}z \\ v &= a_{21}x + a_{22}y + a_{23}z \\ w &= a_{31}x + a_{32}y + a_{33}z. \end{aligned}$$

The matrix

$$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

that encodes the projective change of coordinates will be key.

Suppose $f(u, v, w)$ is a second degree homogeneous polynomial and let $\widetilde{f}(x, y, z)$ be the corresponding second degree homogeneous polynomial in the xyz-coordinate system. From the previous subsection, we know that there are two 3×3 symmetric matrices A and B such that

$$f(u,v,w) = \begin{pmatrix} u & v & w \end{pmatrix} A \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \quad \widetilde{f}(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} B \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

We want to find a relation between the three matrices M, A, and B.

Exercise 1.11.3. Let C be a 3×3 matrix and let X be a 3×1 matrix. Show that $(CX)^T = X^T C^T$.

Exercise 1.11.4. Let M be a projective change of coordinates

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = M \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

and suppose

$$f(u,v,w) = \begin{pmatrix} u & v & w \end{pmatrix} A \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \quad \widetilde{f}(x,y,z) = \begin{pmatrix} x & y & z \end{pmatrix} B \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Show that

$$B = M^T A M.$$

As a pedagogical aside, if we were following the format of earlier problems, before stating the above theorem, we would have given some concrete exercises illustrating the general principle. We have chosen not to do that here. In part, it is to allow readers to come up with their own concrete examples, if necessary. The other part is that this entire section's goal is not only to link linear algebra with conics but also to (not so secretly) force readers to review some linear algebra.

Recall the following definitions from linear algebra.

Definition 1.11.2. We say that two $n \times n$ matrices A and B are *similar*, $A \sim B$, if there is an invertible $n \times n$ matrix C such that

$$A = C^{-1}BC.$$

Definition 1.11.3. An $n \times n$ matrix C is *orthogonal* if $C^{-1} = C^T$.

Definition 1.11.4. A matrix A has an eigenvalue λ if $Av = \lambda v$ for some non-zero vector v. The vector v is called an *eigenvector* with associated *eigenvalue* λ.

Exercise 1.11.5. Given a 3×3 matrix A, show that A has exactly three eigenvalues, counting multiplicity. (For this problem, it is fine to find the proof in a linear algebra text. After looking it up, close the book and try to reproduce the proof on your own. Repeat as necessary until you get it. This is, of course, another attempt by the authors to coax the reader into reviewing linear algebra.)

Exercise 1.11.6.

(1) Let A and B be two symmetric matrices, neither of which has a zero eigenvalue. Show there is an invertible 3×3 matrix C such that

$$A = C^T BC.$$

(2) Let A and B be two symmetric matrices, each of which has exactly one zero eigenvalue (with the other two eigenvalues being non-zero). Show that there is an invertible 3×3 matrix C such that

$$A = C^T BC.$$

(3) Now let A and B be two symmetric matrices, each of which has a zero eigenvalue with multiplicity two (and hence the remaining eigenvalue must be non-zero). Show that there is an invertible 3×3 matrix C such that

$$A = C^T BC.$$

(Again, it is fine to look up this deep result in a linear algebra text. Just make sure that you can eventually reproduce it on your own.)

Exercise 1.11.7.

(1) Show that the 3×3 matrix associated to the ellipse $\mathrm{V}(x^2 + y^2 - z^2)$ has three non-zero eigenvalues.

(2) Show that the 3×3 matrix associated to the two crossing lines $\mathrm{V}(xy)$ has one zero eigenvalue and two non-zero eigenvalues.

(3) Finally, show that the 3×3 matrix associated to the double line $\mathrm{V}((x-y)^2)$ has a zero eigenvalue of multiplicity two and a non-zero eigenvalue.

Exercise 1.11.8. Based on the material of this section, give another proof that under projective changes of coordinates all ellipses, hyperbolas, and parabolas are the same, all crossing line conics are the same, and all double lines are the same.

1.11.3. Conics in $\mathbb{R}^2$. We have shown that all smooth conics can be viewed as the same in the complex projective plane $\mathbb{P}^2$. As we saw earlier, ellipses, hyperbolas, and parabolas are quite different in the real plane $\mathbb{R}^2$. There is a more linear-algebraic approach that captures these differences.

Let $f(x, y, z) = ax^2 + bxy + cy^2 + dxz + eyz + hz^2 = 0$, with $a, b, c, d, e, h \in \mathbb{R}$. Dehomogenize by setting $z = 1$, so that we are looking at the polynomial

$$f(x, y) = ax^2 + bxy + cy^2 + dx + ey + h,$$

which can be written as

$$f(x, y) = \begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} a & \frac{b}{2} & \frac{d}{2} \\ \frac{b}{2} & c & \frac{e}{2} \\ \frac{d}{2} & \frac{e}{2} & h \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

In $\mathbb{P}^2$, the coordinates x, y, and z all play the same role. That is no longer the case after setting $z = 1$. The second order term of f,

$$ax^2 + bxy + cy^2,$$

determines whether we have an ellipse, hyperbola, or parabola.

Exercise 1.11.9. Explain why we need to consider only the second order terms. [Hint: We have already answered this question earlier in this chapter.]

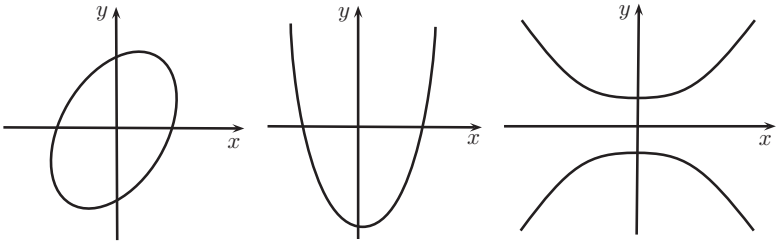


Figure 9. Three types of conics.

This suggests that the matrix

$$\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}$$

might be worth investigating.

Definition 1.11.5. The *discriminant* of a conic over $\mathbb{R}^2$ is

$$\Delta = -4 \det \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}.$$

Exercise 1.11.10. Find the discriminant of each of the following conics.

(1) $9x^2 + 4y^2 = 1$
(2) $9x^2 - 4y^2 = 1$
(3) $9x^2 - y = 0$

Exercise 1.11.11. Based on the previous exercise, describe the conic obtained if $\Delta = 0$, $\Delta < 0$, or $\Delta > 0$. State what the general result ought to be. (To rigorously prove it should take some time. In fact, if you have not seen this before, this type of problem will have to be spread out over a few days. We do not intend for you to spend all

of your time on this problem; no, we intend for you to work on it for thirty minutes to an hour, put it aside, and then come back to it.)

Exercise 1.11.12. Consider the equation $ax^2+bxy+cy^2=0$, where all coefficients are real numbers. Dehomogenize the equation by setting $y=1$. Solve the resulting quadratic equation for x. You should see a factor involving Δ in your solution. How does Δ relate to the discriminant used in the quadratic formula?

Exercise 1.11.13. The discriminant in the quadratic formula tells us how many (real) solutions a given quadratic equation in a single variable has. Classify a conic $\mathrm{V}(f(x,y))$ based on the number of solutions to its dehomogenized quadratic equation.

1.12. Duality

The first goal of this section is show that there is a duality between points and lines in the projective plane. The second goal of this section is to use duality to map any smooth curve in $\mathbb{P}^2$ to another curve called the dual curve in $\mathbb{P}^2$.

1.12.1. Duality in $\mathbb{P}^2$ between Points and Lines. Given a triple of points $a,b,c\in\mathbb{C}$, not all zero, we have a line

$$L=\{(x:y:z)\in\mathbb{P}^2 : ax+by+cz=0\}.$$

Exercise 1.12.1. Show that the line associated to $a_1=1, b_1=2, c_1=3$ is the same line as that associated to $a_2=-2, b_2=-4, c_2=-6$.

Exercise 1.12.2. Show that the line associated to a_1,b_1,c_1 is the same as the line associated to a_2,b_2,c_2 if and only if there is a nonzero constant $\lambda\in\mathbb{C}$ such that $a_1=\lambda a_2,\ b_1=\lambda b_2,\ c_1=\lambda c_2$.

Hence all representatives in the equivalence class for $(a:b:c)\in\mathbb{P}^2$ define the same line.

Exercise 1.12.3. Show that the set of all lines in $\mathbb{P}^2$ can be identified with $\mathbb{P}^2$ itself.

Even though the set of lines in $\mathbb{P}^2$ can be thought of as another $\mathbb{P}^2$, we want notation to be able to distinguish $\mathbb{P}^2$ as a set of points and $\mathbb{P}^2$ as the set of lines. Let $\mathbb{P}^2$ be our set of points and let $\widetilde{\mathbb{P}}^2$ denote the set of lines in $\mathbb{P}^2$. To help our notation, given $(a : b : c) \in \mathbb{P}^2$, let

$$L_{(a:b:c)} = \{(x : y : z) \in \mathbb{P}^2 : ax + by + cz = 0\}.$$

Then we define the map $\mathcal{D} : \widetilde{\mathbb{P}}^2 \to \mathbb{P}^2$ by

$$\mathcal{D}(L_{(a:b:c)}) = (a : b : c).$$

The $\mathcal{D}$ stands for *duality*.

Let us look for a minute at the equation of a line:

$$ax + by + cz = 0.$$

Though it is traditional to think of a, b, c as constants and x, y, z as variables, this is only a convention. Think briefly of x, y, z as fixed, and consider the set

$$M_{(x:y:z)} = \{(a : b : c) \in \widetilde{\mathbb{P}}^2 : ax + by + cz = 0\}.$$

Exercise 1.12.4. Explain in your own words why, given $(x_0 : y_0 : z_0) \in \mathbb{P}^2$, we can interpret $M_{(x_0:y_0:z_0)}$ as the set of all lines containing the point $(x_0 : y_0 : z_0)$.

We are beginning to see a duality between lines and points.

Let

$$\Sigma = \{((a : b : c), (x_0 : y_0 : z_0)) \in \widetilde{\mathbb{P}}^2 \times \mathbb{P}^2 : ax_0 + by_0 + cz_0 = 0\}.$$

There are two natural projection maps:

$$\pi_1 : \Sigma \to \widetilde{\mathbb{P}}^2$$

given by

$$\pi_1(((a : b : c), (x_0 : y_0 : z_0))) = (a : b : c)$$

and

$$\pi_2 : \Sigma \to \mathbb{P}^2$$

given by

$$\pi_2(((a : b : c), (x_0 : y_0 : z_0))) = (x_0 : y_0 : z_0).$$

Exercise 1.12.5. Show that both maps π_1 and π_2 are onto.

Exercise 1.12.6. Given a point $(a : b : c) \in \widetilde{\mathbb{P}}^2$, consider the set

$$\pi_1^{-1}(a : b : c) = \{((a : b : c), (x_0 : y_0 : z_0)) \in \Sigma\}.$$

Show that the set $\pi_2(\pi_1^{-1}(a : b : c))$ is identical to a set in $\mathbb{P}^2$ that we defined near the beginning of this section.

As evidence for a type of duality, show:

Exercise 1.12.7. Given a point $(x_0 : y_0 : z_0) \in \mathbb{P}^2$, consider the set

$$\pi_2^{-1}(x_0 : y_0 : z_0) = \{((a : b : c), (x_0 : y_0 : z_0)) \in \Sigma\}.$$

Show that the set $\pi_1(\pi_2^{-1}(x_0 : y_0 : z_0))$ is identical to a set in $\widetilde{\mathbb{P}}^2$ that we defined near the beginning of this section.

Exercise 1.12.8. Let $(1 : 2 : 3), (2 : 5 : 1) \in \widetilde{\mathbb{P}}^2$. Find

$$\pi_2(\pi_1^{-1}(1 : 2 : 3)) \cap \pi_2(\pi_1^{-1}(2 : 5 : 1)).$$

Explain why this is just a fancy way for finding the point of intersection of the two lines

$$x + 2y + 3z = 0$$

$$2x + 5y + z = 0.$$

As another piece of evidence for duality, consider:

Exercise 1.12.9. Let $(1 : 2 : 3), (2 : 5 : 1) \in \mathbb{P}^2$. Find

$$\pi_1(\pi_2^{-1}(1 : 2 : 3)) \cap \pi_1(\pi_2^{-1}(2 : 5 : 1)).$$

Explain that this is just a fancy way for finding the unique line containing the two points $(1 : 2 : 3), (2 : 5 : 1)$.

Principle 1.12.1. The *duality principle* for points and lines in the complex projective plane is that for any theorem for points and lines there is a corresponding theorem obtained by interchanging the words "points" and "lines".

Exercise 1.12.10. Use the duality principle to find the corresponding theorem to:

Theorem 1.12.11. Any two distinct points in $\mathbb{P}^2$ determine a unique line.

This duality extends to higher dimensional projective spaces. The following is a fairly open-ended exercise:

Exercise 1.12.12. Given $(x_0, y_0, z_0, w_0), (x_1, y_1, z_1, w_1) \in \mathbb{C}^4 - \{(0,0,0,0)\}$, define

$$(x_0, y_0, z_0, w_0) \sim (x_1, y_1, z_1, w_1)$$

if there exists a non-zero λ such that

$$x_0 = \lambda x_1, y_0 = \lambda y_1, z_0 = \lambda z_1, w_0 = \lambda w_1.$$

Define

$$\mathbb{P}^3 = \mathbb{C}^4 - \{(0,0,0,0)\}/\sim .$$

Show that the set of all planes in $\mathbb{P}^3$ can be identified with another copy of $\mathbb{P}^3$. Explain how the duality principle can be used to link the fact that three non-collinear points define a unique plane to the fact three planes with linearly independent normal vectors intersect in a unique point.

1.12.2. Dual Curves to Conics. Let $f(x, y, z)$ be a homogeneous polynomial and let

$$C = \{(x : y : z) \in \mathbb{P}^2 : f(x, y, z) = 0\}.$$

We know that the normal vector at a point $p = (x_0 : y_0 : z_0) \in C$ is

$$\nabla f(p) = \left(\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p)\right).$$

Further the tangent line at $p = (x_0 : y_0 : z_0) \in C$ is defined as

$$T_p(C) = \{(x : y : z) \in \mathbb{P}^2 : x\frac{\partial f}{\partial x}(p) + y\frac{\partial f}{\partial y}(p) + z\frac{\partial f}{\partial z}(p) = 0\}.$$

Recall from Section 1.10 that if f has degree n, then

$$nf(x, y, z) = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z}.$$

Exercise 1.12.13. Show for any $p = (x_0 : y_0 : z_0) \in C$, we have

$$\begin{aligned} T_p(C) \ = \{(x : y : z) \in \mathbb{P}^2 : (x - x_0)&\frac{\partial f}{\partial x}(p) \\ + \ (y - y_0)&\frac{\partial f}{\partial y}(p) + (z - z_0)\frac{\partial f}{\partial z}(p) = 0\}. \end{aligned}$$

Recall that $p \in C$ is smooth if the gradient

$$\nabla f(p) \neq (0,0,0).$$

Definition 1.12.1. For a smooth curve C, the *dual curve* $\widetilde{C}$ is the composition of the map, for $p \in C$,

$$p \mapsto T_p(C)$$

with the dual map

$$\mathcal{D} : \widetilde{\mathbb{P}}^2 \to \mathbb{P}^2$$

from the last subsection. We also denote this map by $\mathcal{D}$. Then

$$\mathcal{D}(p) = \left(\frac{\partial f}{\partial x}(p) : \frac{\partial f}{\partial y}(p) : \frac{\partial f}{\partial z}(p) \right).$$

To make sense of this, we, of course, need some examples.

Exercise 1.12.14. For $f(x,y,z) = x^2+y^2-z^2$, let $C = \mathrm{V}(f(x,y,z))$. Show for any $(x_0 : y_0 : z_0) \in C$ that

$$\mathcal{D}(x_0 : y_0 : z_0) = (2x_0 : 2y_0 : -2z_0).$$

Show that in this case the dual curve $\widetilde{C}$ is the same as the original C.

Exercise 1.12.15. Consider $f(x,y,z) = x^2 - yz = 0$. Then for any $(x_0 : y_0 : z_0) \in C$, where $C = \mathrm{V}(f)$, show that

$$\mathcal{D}(x_0, y_0, z_0) = (2x_0 : -z_0 : -y_0).$$

Show that the image is in $\mathbb{P}^2$ by showing that $(2x_0, -z_0, -y_0) \neq (0,0,0)$. Letting $(u : v : w) = (2x : -z : -y)$, show that $u^2 - 4vw = 0$ defines the dual curve $\widetilde{C}$. Note that here $\widetilde{C} \neq C$.

Exercise 1.12.16. For $C = \mathrm{V}(x^2 + 4y^2 - 9z^2)$, show that the dual curve is

$$\widetilde{C} = \{(x : y : z) \in \mathbb{P}^2 : x^2 + \frac{1}{4}y^2 - \frac{1}{9}z^2 = 0\}.$$

Exercise 1.12.17. For $C = \mathrm{V}(5x^2 + 2y^2 - 8z^2)$, find the dual curve.

Exercise 1.12.18. For a line $L = \{(x : y : z) \in \mathbb{P}^2 : ax + by + cz\}$, find the dual curve. Explain why calling this set the "dual curve" might seem strange.

Chapter 2

Cubic Curves and Elliptic Curves

The goal of this chapter is to provide an introduction to cubic curves (smooth cubic curves are also known as elliptic curves). Cubic curves have a far richer structure than that of conics. Many of the deepest questions in mathematics still involve questions about cubics. After a few preliminaries, we will show how each smooth cubic curve is a group, meaning that its points can be added together. This group structure provides a fascinating interplay between algebra, geometry, analysis, and topology. We will then see that there are many different cubics, even up to projective change of coordinates. In fact, we will see that there are a complex numbers' worth of different cubics. That is, we can parametrize cubics up to isomorphism by the complex numbers. This is in marked contrast to conics, since all smooth conics are the same up to projective change of coordinates. Next, we will see that, as surfaces, all smooth cubics are tori. Finally, we see how all cubics can be viewed as the quotient $\mathbb{C}/\Lambda$, where Λ is a lattice in $\mathbb{C}$.

2.1. Cubics in $\mathbb{C}^2$

The goal of this section is to begin the study of cubic curves by looking at some specific examples.

A cubic curve $\mathrm{V}(P)$ is simply the zero set of a degree three polynomial P. If P is in two variables, then $\mathrm{V}(P)$ will be a cubic in $\mathbb{C}^2$ while if P is homogeneous in three variables, then $\mathrm{V}(P)$ is a cubic in the projective plane $\mathbb{P}^2$.

Exercise 2.1.1. Sketch the following cubics in the real plane $\mathbb{R}^2$.

(1) $y^2 = x^3$
(2) $y^2 = x(x-1)^2$
(3) $y^2 = x(x-1)(x-2)$
(4) $y^2 = x(x^2+x+1)$

Of course, we are sketching these curves in the real plane only to get a feel for cubics.

Exercise 2.1.2. Consider the cubics in the above exercise.

(1) Give the homogeneous form for each cubic, which extends each of the above cubics to the complex projective plane $\mathbb{P}^2$.
(2) For each of the above cubics, dehomogenize by setting $x = 1$, and graph the resulting cubic in $\mathbb{R}^2$ with coordinates y and z.

Recall that a point $(a : b : c) \in \mathrm{V}(P)$ on a curve is *singular* if

$$\begin{aligned} \frac{\partial P}{\partial x}(a,b,c) &= 0, \\ \frac{\partial P}{\partial y}(a,b,c) &= 0, \\ \frac{\partial P}{\partial z}(a,b,c) &= 0. \end{aligned}$$

If a curve has a singular point, then we call the curve *singular*. If a curve has no singular points, we call it *smooth*.

Exercise 2.1.3. Show that the following cubics are singular.

(1) $\mathrm{V}(xyz)$
(2) $\mathrm{V}(x(x^2+y^2-z^2))$
(3) $\mathrm{V}(x^3)$

The only singular conics are unions of two lines or double lines. The above singular cubics are similar, in that they are all the zero sets of reducible polynomials $P(x, y, z)$. Unlike for conics, though, there are singular cubics that do not arise from reducible polynomials $P(x, y, z)$.

Exercise 2.1.4. Sketch the cubic $y^2 = x^3$ in the real plane $\mathbb{R}^2$. Show that the corresponding cubic $\mathrm{V}(x^3 - y^2z)$ in $\mathbb{P}^2$ has a singular point at $(0 : 0 : 1)$. Show that this is the only singular point on this cubic.

Exercise 2.1.5. Show that the polynomial $P(x, y, z) = x^3 - y^2z$ is irreducible, i.e., cannot be factored into two polynomials. (This is a fairly brute force high school algebra problem.)

2.2. Inflection Points

The goal of this section is to show that every smooth cubic curve must have exactly nine points of inflection.

2.2.1. Intuitions about Inflection Points. One of the strengths of algebraic geometry is the ability to move freely between the symbolic language of algebra and the visual capabilities of geometry. We would like to use this flexibility to convert what initially is a geometric problem into an algebraic one. While we can sometimes imagine what is happening geometrically, this will help us in situations that may be difficult to visualize.

We have seen that a line will intersect a smooth conic in two points. If the points are distinct, then the line will cut through the conic. However, there may be a line which has only one point in common with the conic, namely the tangent line. In this case, if we consider that the point of tangency is to be counted twice, then the line will intersect the conic in "two" points.

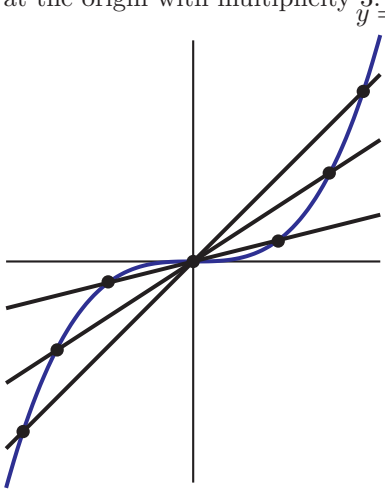
If we now consider a line intersecting a cubic, we have more points of intersection to consider. Intuitively, a line cannot cross a cubic in too many places. In fact, the Fundamental Theorem of Algebra shows that a line intersects a cubic in at most three points. As in the case

of conics, points may need to be counted more than once, but never more than three times.

If a line intersects a cubic in a single point (counted thrice), we call such a point a point of inflection or flex point. An *inflection point* of a curve $\mathrm{V}(P)$ is a nonsingular point $p \in \mathrm{V}(P)$ where the tangent line to the curve at p intersects $\mathrm{V}(P)$ with multiplicity 3 (or greater).

We will later define what it means for the tangent line at a point to intersect the curve with multiplicity 3 (or greater), but the idea can be illustrated with some examples.

(1) Consider the cubic curve $y = x^3$, that is, $\mathrm{V}(P)$ where $P(x, y) = x^3 - y$. Let the point p be the origin, and consider the line $y = \epsilon x$, where $\epsilon > 0$. This line intersects the curve in three distinct points no matter how small ϵ is, but as ϵ approaches zero, the three points of intersection coalesce into just one point. We say that the tangent line $y = 0$ intersects the cubic $y = x^3$ at the origin with multiplicity 3.

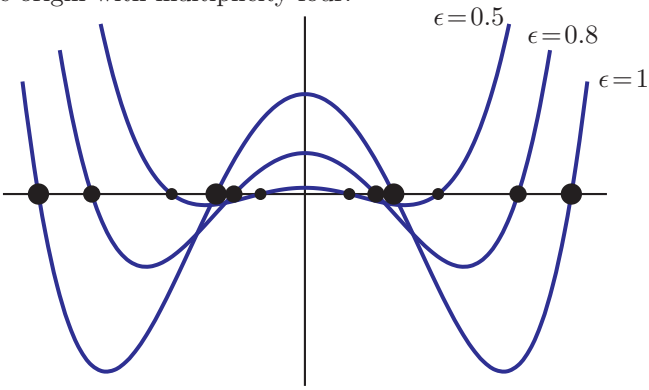


(2) If we look at the behavior of the quartic (fourth-degree) curve

$$y = (x - \epsilon)(x - \epsilon/3)(x + \epsilon/3)(x + \epsilon),$$

we see that the curve and the line $y = 0$ intersect at four points whenever $\epsilon > 0$. But as ϵ approaches zero, the four

points of intersection converge to one point, the origin. Here we say that the tangent line $y = 0$ intersects this curve at the origin with multiplicity four.



$$y = (x - \epsilon)(x - \epsilon/3)(x + \epsilon/3)(x + \epsilon)$$

(3) We will see later that the tangent line ℓ to a curve $\mathrm{V}(P)$ at a point p always intersects the curve with multiplicity at least 2.

2.2.2. Multiplicity of Roots. For a moment we will look at one-variable polynomials (which correspond to homogeneous two-variable polynomials).

Definition 2.2.1. Given a polynomial $P(x)$, a *root* or *zero* is a point a such that $P(a) = 0$.

Exercise 2.2.1. If $(x - a)$ divides $P(x)$, show that a is a root of $P(x)$.

Exercise 2.2.2. If a is a root of $P(x)$, show that $(x - a)$ divides $P(x)$. [Hint: Use the Division Algorithm for polynomials.]

Definition 2.2.2. Let a be a root of the polynomial $P(x)$. This root has *multiplicity* k if $(x - a)^k$ divides $P(x)$ but $(x - a)^{k+1}$ does not divide $P(x)$.

Exercise 2.2.3. Suppose that a is a root of multiplicity two for $P(x)$. Show there is a polynomial $g(x)$, with $g(a) \neq 0$, such that

$$P(x) = (x - a)^2 g(x)$$

Exercise 2.2.4. Suppose that a is a root of multiplicity two for $P(x)$. Show that $P(a) = 0$ and $P'(a) = 0$ but $P''(a) \neq 0$.

Exercise 2.2.5. Suppose that a is a root of multiplicity k for $P(x)$. Show there is a polynomial $g(x)$ such that

$$P(x) = (x - a)^k g(x)$$

with $g(a) \neq 0$.

Exercise 2.2.6. Suppose that a is a root of multiplicity k for $P(x)$. Show that $P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0$ but $P^{(k)}(a) \neq 0$.

The homogeneous version is the following.

Definition 2.2.3. Let $P(x, y)$ be a homogeneous polynomial. A *root* or *zero* is a point $(a : b) \in \mathbb{P}^1$ such that $P(a, b) = 0$. If $(a : b)$ is a root of $P(x, y)$, then $(bx - ay)$ divides $P(x, y)$. This root has *multiplicity* k if $(bx - ay)^k$ divides $P(x, y)$ but $(bx - ay)^{k+1}$ does not divide $P(x, y)$.

Exercise 2.2.7. Suppose that $(a : b)$ is a root of multiplicity two for $P(x, y)$. Show that

$$P(a, b) = \frac{\partial P}{\partial x}(a, b) = \frac{\partial P}{\partial y}(a, b) = 0,$$

but at least one of the second partials does not vanish at $(a : b)$.

Exercise 2.2.8. Suppose that $(a : b)$ is a root of multiplicity k for $P(x, y)$. Show that

$$P(a, b) = \frac{\partial P}{\partial x}(a, b) = \frac{\partial P}{\partial y}(a, b) = \cdots = \frac{\partial^{k-1} P}{\partial x^i \partial y^j}(a, b) = 0,$$

where $i + j = k - 1$ but

$$\frac{\partial^k P}{\partial x^i \partial y^j}(a, b) \neq 0,$$

for at least one pair $i + j = k$. This means that the first partials, second partials, etc. up to the $k - 1$ partials all vanish at $(a : b)$, but at least one of the k^{th} partials does not vanish at $(a : b)$.

Exercise 2.2.9. Suppose

$$P(a,b) = \frac{\partial P}{\partial x}(a,b) = \frac{\partial P}{\partial y}(a,b) = \cdots = \frac{\partial^{k-1} P}{\partial x^i \partial y^j}(a,b) = 0,$$

where $i + j = k - 1$ and

$$\frac{\partial^k P}{\partial x^i \partial y^j}(a,b) \neq 0,$$

for at least one pair $i+j = k$. Show that $(a : b)$ is a root of multiplicity k for $P(x, y)$.

2.2.3. Inflection Points. Let $P(x, y, z)$ be a homogeneous polynomial. We want to understand what it means for a line to intersect $\mathrm{V}(P)$ in a point with multiplicity three or more. Let

$$l(x, y, z) = ax + by + cz$$

be a linear polynomial and let $\ell = \mathrm{V}(l)$ be the corresponding line in $\mathbb{P}^2$. We are tacitly assuming that not all of a, b, c are zero. We might as well assume that $b \neq 0$. That is, by a projective change of coordinates we may assume that $b \neq 0$. We can multiply l by any nonzero constant and still have the same line, meaning that for $\lambda \neq 0$, we have $\mathrm{V}(l) = \mathrm{V}(\lambda l)$. So, we can assume that $b = -1$. The reason for the -1 is that we now know that all points on the line have the property that $y = ax + cz$.

Exercise 2.2.10. Let $(x_0 : y_0 : z_0) \in \mathrm{V}(P) \cap \mathrm{V}(l)$. Show that $(x_0 : z_0)$ is a root of the homogeneous two-variable polynomial $P(x, ax + cz, z)$ and that $y_0 = ax_0 + cz_0$.

Definition 2.2.4. The *intersection multiplicity* of $\mathrm{V}(P)$ and $\mathrm{V}(l)$ at $(x_0 : y_0 : z_0)$ is the multiplicity of the root $(x_0 : z_0)$ of $P(x, ax + cz, z)$.

Exercise 2.2.11. Let $P(x, y, z) = x^2 - yz$ and $l(x, y, z) = \lambda x - y$. Show that the intersection multiplicity of $\mathrm{V}(P)$ and $\mathrm{V}(l)$ at $(0 : 0 : 1)$ is one when $\lambda \neq 0$ and is two when $\lambda = 0$.

The key to the definition above is that, when $b = -1$, the system $x = x, y = ax + cz, z = z$ gives a parametrization of the line $\mathrm{V}(l)$, and the intersection multiplicity of $\mathrm{V}(P)$ and $\mathrm{V}(l)$ at $(x_0 : y_0 : z_0)$ is found by considering P evaluated as a function of these two parameters. The

next exercise proves that the intersection multiplicity is independent of the choice of parametrization used for the line $\mathrm{V}(l)$.

Exercise 2.2.12. Let $(x_0 : y_0 : z_0) \in \mathrm{V}(P) \cap \mathrm{V}(l)$. Let $x = a_1 s + b_1 t, y = a_2 s + b_2 t, z = a_3 s + b_3 t$ and $x = c_1 u + d_1 v, y = c_2 u + d_2 v, z = c_3 u + d_3 v$ be two parametrizations of the line $\mathrm{V}(l)$ such that $(x_0 : y_0 : z_0)$ corresponds to $(s_0 : t_0)$ and $(u_0 : v_0)$, respectively. Show that the multiplicity of the root $(s_0 : t_0)$ of $P(a_1 s + b_1 t, a_2 s + b_2 t, a_3 s + b_3 t)$ is equal to the multiplicity of the root $(u_0 : v_0)$ of $P(c_1 u + d_1 v, c_2 u + d_2 v, c_3 u + d_3 v)$. Conclude that our definition of the intersection multiplicity of $\mathrm{V}(P)$ and $\mathrm{V}(l)$ is independent of the parametrization used for the line $\mathrm{V}(l)$.

Exercise 2.2.13. Let $P(x, y, z) = x^2 + 2xy - yz + z^2$. Show that the intersection multiplicity of $\mathrm{V}(P)$ and any line ℓ at a point of intersection is at most two.

Exercise 2.2.14. Let $P(x, y, z)$ be an irreducible second degree homogeneous polynomial. Show that the intersection multiplicity of $\mathrm{V}(P)$ and any line ℓ at a point of intersection is at most two.

Exercise 2.2.15. Let $P(x, y, z) = x^2 + y^2 + 2xz - yz$.

(1) Find the tangent line $\ell = \mathrm{V}(l)$ to $\mathrm{V}(P)$ at $(-2 : 1 : 1)$.

(2) Show that the intersection multiplicity of $\mathrm{V}(P)$ and ℓ at $(-2 : 1 : 1)$ is two.

Exercise 2.2.16. Let $P(x, y, z) = x^3 - y^2 z + z^3$.

(1) Find the tangent line to $\mathrm{V}(P)$ at $(2 : 3 : 1)$ and show directly that the intersection multiplicity of $\mathrm{V}(P)$ and its tangent at $(2 : 3 : 1)$ is two.

(2) Find the tangent line to $\mathrm{V}(P)$ at $(0 : 1 : 1)$ and show directly that the intersection multiplicity of $\mathrm{V}(P)$ and its tangent at $(0 : 1 : 1)$ is three.

Exercise 2.2.17. Redo the previous two exercises using Exercise 2.2.9.

Exercise 2.2.18. Show for any nonsingular curve $\mathrm{V}(P) \subset \mathbb{P}^2$, the intersection multiplicity of $\mathrm{V}(P)$ and its tangent line ℓ at the point of tangency is at least two.

Exercise 2.2.19.

(1) Let $P(x, y, z)$ be an irreducible degree three homogeneous polynomial. Show that the intersection multiplicity of $\mathrm{V}(P)$ and any line ℓ at a point of intersection is at most three.

(2) Let $P(x, y, z)$ be an irreducible homogeneous polynomial of degree n. Show that the intersection multiplicity of $\mathrm{V}(P)$ and any line ℓ at a point of intersection is at most n.

Definition 2.2.5. Let $P(x, y, z)$ be an irreducible homogeneous polynomial of degree n. A nonsingular point $p \in \mathrm{V}(P) \subset \mathbb{P}^2$ is called a *point of inflection* or a *flex* of the curve $\mathrm{V}(P)$ if the tangent line to the curve at p intersects $\mathrm{V}(P)$ with multiplicity at least three.

Exercise 2.2.20. Let $P(x, y, z) = x^3 + yz^2$. Show that $(0 : 0 : 1)$ is an inflection point of $\mathrm{V}(P)$.

Exercise 2.2.21. Let $P(x, y, z) = x^3 + y^3 + z^3$ (the Fermat curve). Show that $(1 : -1 : 0)$ is an inflection point of $\mathrm{V}(P)$.

2.2.4. Hessians. We have just defined what it means for a point $p \in \mathrm{V}(P)$ to be a point of inflection. Checking to see whether a given point $p \in \mathrm{V}(P)$ is an inflection point can be tedious, but finding inflection points can be an extremely difficult task with our current tools. How did we know to check $(1 : -1 : 0)$ in Exercise 2.2.21? Since $\mathrm{V}(P)$ has an infinite number of points, it would be impossible to find the tangent at every point and to check the intersection multiplicity. Moreover, if these inflection points are related to the inflection points of calculus, where are the second derivatives? The Hessian curve will completely solve these difficulties. We will first define the Hessian curve, then determine how it can be used to find the points of inflection.

Definition 2.2.6. Let $P(x, y, z)$ be a homogeneous polynomial of degree n. The *Hessian* $H(P)$ is the polynomial

$$H(P)(x, y, z) = \det \begin{pmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{yx} & P_{yy} & P_{yz} \\ P_{zx} & P_{zy} & P_{zz} \end{pmatrix},$$

where

$$\begin{aligned} P_x &= \frac{\partial P}{\partial x} \\ P_{xx} &= \frac{\partial^2 P}{\partial x^2} \\ P_{yx} &= \frac{\partial^2 P}{\partial x \partial y}, \quad \text{etc.} \end{aligned}$$

The *Hessian curve* is $\mathrm{V}(H(P))$.

Exercise 2.2.22. Compute $H(P)$ for the following cubic polynomials.

(1) $P(x, y, z) = x^3 + yz^2$
(2) $P(x, y, z) = y^3 + z^3 + xy^2 - 3yz^2 + 3zy^2$
(3) $P(x, y, z) = x^3 + y^3 + z^3$

Exercise 2.2.23. Let $P(x, y, z)$ be an irreducible homogeneous polynomial of degree three. Show that $H(P)$ is also a third degree homogeneous polynomial.

We want to link the Hessian curve with inflection points.

Exercise 2.2.24. Let $P(x, y, z) = x^3 + y^3 + z^3$ (the Fermat curve). Show that $(1 : -1 : 0) \in \mathrm{V}(P) \cap \mathrm{V}(H(P))$.

Exercise 2.2.25. Let $P(x, y, z) = y^3 + z^3 + xy^2 - 3yz^2 + 3zy^2$. Show that $(-2 : 1 : 1) \in \mathrm{V}(P) \cap \mathrm{V}(H(P))$.

Exercise 2.2.26. Let $P(x, y, z) = x^3 + yz^2$. Show that $(0 : 0 : 1) \in \mathrm{V}(P) \cap \mathrm{V}(H(P))$.

These exercises suggest a link between inflection points of $\mathrm{V}(P)$ and points in $\mathrm{V}(P) \cap \mathrm{V}(H(P))$, but we need to be careful.

Exercise 2.2.27. Let $P(x, y, z) = x^3 + yz^2$.

(1) Show that $(0 : 1 : 0) \in \mathrm{V}(P) \cap \mathrm{V}(H(P))$.
(2) Explain why $(0 : 1 : 0)$ is not an inflection point of $\mathrm{V}(P)$.

We can now state the relationship we want.

Theorem 2.2.28. Let $P(x, y, z)$ be a homogeneous polynomial of degree d. If $\mathrm{V}(P)$ is smooth, then $p \in \mathrm{V}(P) \cap \mathrm{V}(H(P))$ if and only if p is a point of inflection of $\mathrm{V}(P)$.

We will prove this theorem through a series of exercises.[1] The first thing we need to show is that the vanishing of the Hessian $\mathrm{V}(H(P))$ is invariant under a projective change of coordinates.

Exercise 2.2.29. Consider the following projective change of coordinates

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Suppose that under the projective transformation A the polynomial $P(x, y, z)$ becomes the polynomial $Q(u, v, w)$.

(1) Show that the Hessian matrices of P and Q are related by

$$\begin{pmatrix} P_{xx} & P_{xy} & P_{xz} \\ P_{xy} & P_{yy} & P_{yz} \\ P_{xz} & P_{yz} & P_{zz} \end{pmatrix} = A^T \begin{pmatrix} Q_{uu} & Q_{uv} & Q_{uw} \\ Q_{uv} & Q_{vv} & Q_{vw} \\ Q_{uw} & Q_{vw} & Q_{ww} \end{pmatrix} A.$$

(2) Conclude that $H(P)(x, y, z) = 0$ if and only if $H(Q)(u, v, w) = 0$.

Next we need to show that inflection points are mapped to inflection points under a projective change of coordinates.

Exercise 2.2.30. Suppose p is a point of inflection of $\mathrm{V}(P)$, and that under a projective change of coordinates the polynomial P becomes the polynomial Q and $p \mapsto q$. Show that q is a point of inflection of $\mathrm{V}(Q)$.

In the next exercise, we will reduce the proof of Theorem 2.2.28 to the case where $p = (0 : 0 : 1) \in \mathrm{V}(P)$ and the tangent line to $\mathrm{V}(P)$ at p is $\ell = \mathrm{V}(y)$.

[1] The following exercises are based on the proof taken from C. G. Gibson's "Elementary Geometry of Algebraic Curves." [**Gib98**]

Exercise 2.2.31. Use Exercises 2.2.29 and 2.2.30 to explain why, in proving Theorem 2.2.28, it is enough to show that p is a point of inflection if and only if $H(P)(p) = 0$ in the case where $p = (0 : 0 : 1) \in \mathrm{V}(P)$ and the tangent line ℓ to $\mathrm{V}(P)$ at p is $y = 0$, i.e., $\ell = \mathrm{V}(y)$.

Thus we will assume that the point $p = (0 : 0 : 1) \in \mathrm{V}(P)$ and that the tangent line to $\mathrm{V}(P)$ at p is $y = 0$ from now until the end of Exercise 2.2.35.

Exercise 2.2.32. Explain why in the affine patch $z = 1$ the dehomogenized curve is

$$\lambda y + (ax^2 + bxy + cy^2) + \text{higher order terms},$$

where $\lambda \neq 0$. [Hint: We know that $p \in \mathrm{V}(P)$ and p is nonsingular.]

From this we can conclude that $P(x, y, z)$ is given by

(2.1)
$$P(x, y, z) = \lambda y z^{d-1} + (ax^2 + bxy + cy^2)z^{d-2} + \text{higher order terms}$$

where $d = \deg P$.

Exercise 2.2.33. Explain why the intersection of $\mathrm{V}(P)$ with the tangent $\mathrm{V}(y)$ at p corresponds to the root $(0 : 1)$ of the equation

$$P(x, 0, z) = ax^2 z^{d-2} + \text{higher order terms} = 0.$$

Exercise 2.2.34. Show that p is a point of inflection of $\mathrm{V}(P)$ if and only if $a = 0$. [Hint: For p to be an inflection point, what must the multiplicity of $(0 : 1)$ be in the equation in Exercise 2.2.33?]

We have now established that p is a point of inflection if and only if $a = 0$ in Equation (2.1). All that remains is to show that $p \in \mathrm{V}(H(P))$ if and only if $a = 0$.

Exercise 2.2.35.

(1) Show that

$$H(P)(p) = \det \begin{pmatrix} 2a & b & 0 \\ b & 2c & \lambda(d-1) \\ 0 & \lambda(d-1) & 0 \end{pmatrix}.$$

(2) Conclude that $p \in \mathrm{V}(H(P))$ if and only if $a = 0$.

This completes our proof of Theorem 2.2.28. In practice, we use the Hessian to locate inflection points even if $\mathrm{V}(P)$ is not smooth by finding the points of intersection of $\mathrm{V}(P)$ and $\mathrm{V}(H(P))$ and eliminating those that are singular on $\mathrm{V}(P)$.

Exercise 2.2.36. Let $P(x, y, z)$ be an irreducible second degree homogeneous polynomial. Using the Hessian curve, show that $\mathrm{V}(P)$ has no points of inflection.

We conclude this section with the following theorem, which we state without proof. Theorem 2.2.37 is a direct result of Bézout's theorem, which we will prove in Section 3.3.

Theorem 2.2.37. Two cubic curves in $\mathbb{P}^2$, with no common components, will intersect in exactly $3 \times 3 = 9$ points, counted up to intersection multiplicities. (We have not defined what is meant by intersection multiplicity; this is one of the goals of chapter three and is a bit subtle.)

Exercise 2.2.38. Use Exercise 2.2.23 and Theorems 2.2.28 and 2.2.37 to show that if $\mathrm{V}(P)$ is a smooth cubic curve, then $\mathrm{V}(P)$ has exactly nine inflection points.

Exercise 2.2.39. Find all nine points of inflection of the Fermat curve,

$$P(x, y, z) = x^3 + y^3 + z^3.$$

2.3. Group Law

The goal of this section is to illustrate that, as a consequence of their geometric structure, smooth cubic curves are abelian groups. While the group law can be stated algebraically, in this section we will develop it geometrically to see why it is important for the curve to have degree three.

2.3.1. Adding Points on Smooth Cubics. Let C denote a smooth cubic curve in the projective plane, $\mathbb{P}^2(\mathbb{C})$. We will develop a geometric method for adding points so that C is an abelian group under this operation. First, we define an abelian group.

Definition 2.3.1. A *group* is a set G equipped with a binary operation $\star$ satisfying the following axioms:

(G1) The binary operation is associative, i.e.,

$$g_1 \star (g_2 \star g_3) = (g_1 \star g_2) \star g_3$$

for all $g_1, g_2, g_3 \in G$.

(G2) There is an (unique) identity element $e \in G$ such that $e \star g = g = g \star e$ for all $g \in G$.

(G3) For each $g \in G$, there is an (unique) inverse element $g' \in G$ satisfying $g \star g' = e = g' \star g$.

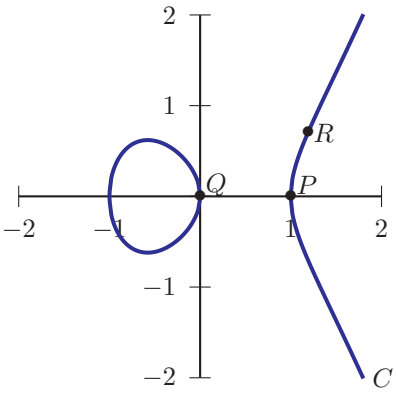
A group G is said to be an *abelian group* if, in addition, the binary operation $\star$ is commutative, i.e., $g_1 \star g_2 = g_2 \star g_1$ for all $g_1, g_2 \in G$.

For points P and Q on C, let $\ell(P, Q)$ denote the line in $\mathbb{P}^2$ through P and Q. In case P and Q are the same point, let $\ell(P, P)$ be the line tangent to C at P. (This is why we must assume the cubic curve C is smooth, in order to ensure there is a well-defined tangent line at every point.) In Section 2.2.3 we saw that the Fundamental Theorem of Algebra ensures there are exactly three points of intersection of $\ell(P, Q)$ with the cubic curve C, counting multiplicities. Let PQ denote this unique third point of intersection, so that the three points of intersection of C with $\ell(P, Q)$ are P, Q and PQ. If a line ℓ is tangent to C at P, then the multiplicity of P is at least two by Exercise 2.2.18. Therefore, if $P \neq Q$ and $\ell(P, Q)$ is tangent to C at P, then $PQ = P$, for P counted the second time is the third point of intersection of $\ell(P, Q)$ with C. The rule $(P, Q) \mapsto PQ$ gives a binary operation on C, which is called the *chord-tangent composition law*.

Exercise 2.3.1. Explain why the chord-tangent composition law is commutative, i.e., $PQ = QP$ for all points P, Q on C.

While this is a well-defined, commutative binary operation on C, the following exercises illustrate that the chord-tangent composition law lacks the properties required of a group law.

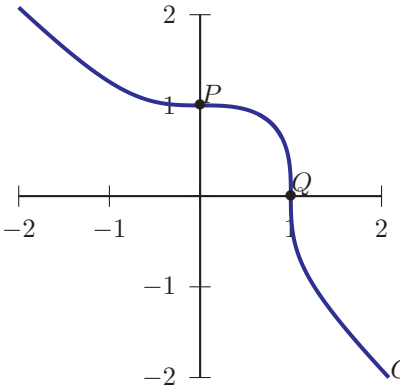
Exercise 2.3.2. Consider the cubic curve $C = \{(x, y) \in \mathbb{C}^2 : y^2 = x^3 - x\}$ and the points P, Q, R on C, as shown below. (Note that only the real part of C is shown.)



Using a straightedge, locate PQ and then $(PQ)R$ on the curve C. Now locate the point QR and the point $P(QR)$ on the curve C. Is it true that $P(QR) = (PQ)R$? That is, is the chord-tangent composition law associative for these points on C?

The preceding exercise demonstrates that the chord-tangent composition law is not associative. The next exercise illustrates that associativity is not the only group axiom that fails for the chord-tangent composition law.

Exercise 2.3.3. Consider the cubic curve $C = \{(x, y) \in \mathbb{C}^2 : x^3 + y^3 = 1\}$ and the points $P = (0, 1)$ and $Q = (1, 0)$ on C, as shown below. (Again, we note that only the real part is shown.)



(1) Using the equation of the cubic curve C and its Hessian, verify that P and Q are inflection points of C.

(2) Verify that $PP = P$. Conclude that if C has an identity element e, then $e = P$.

(3) Verify that $QQ = Q$. Conclude that if C has an identity element e, then $e = Q$.

(4) Conclude that C does not have an identity element for the chord-tangent composition law.

Therefore, the chord-tangent composition law will not serve as a binary operation for the group structure on C because it violates both axioms (G1) and (G2). However, we can find a way to make this work. By using the chord-tangent composition law twice in combination with a fixed inflection point, we will construct the group law on C in the next subsection.

2.3.2. Group Law with an Inflection Point. Let C denote a smooth cubic curve in the projective plane $\mathbb{P}^2(\mathbb{C})$. As was shown in Exercise 2.2.38, there are nine points of inflection (counting multiplicity) on C. These are the points of intersection of the cubic curve with its Hessian curve.

Select a point of inflection O on C. We define our binary operation, $+$, relative to this specific point O. For points P, Q on C, define $P + Q$ to be the unique third point of intersection of $\ell(O, PQ)$ with C, where PQ denotes the chord-tangent composition of P and Q, that is, $P + Q = O(PQ)$, using the chord-tangent composition law notation. We claim that with this binary operation $+$, C is an abelian group, and we call this operation addition, i.e., we can "add" points on C.

We will prove that for a given choice of inflection point, O, the cubic curve C with addition of points relative to O is an abelian group. Before we verify this claim, let's consider a specific example.

Consider the cubic curve $C = \mathrm{V}(x^3 - y^2z + z^3) \subset \mathbb{P}^2$, and the points $P_1 = (2 : 3 : 1)$, $P_2 = (0 : 1 : 1)$, $P_3 = (-1 : 0 : 1)$, $P_4 = (0 : -1 : 1)$, $P_5 = (2 : -3 : 1)$ on C. The figure shows C in the affine patch $z = 1$.

Exercise 2.3.4. Use the equations of the cubic curve C and its Hessian to verify that P_2 and P_4 are inflection points of C.

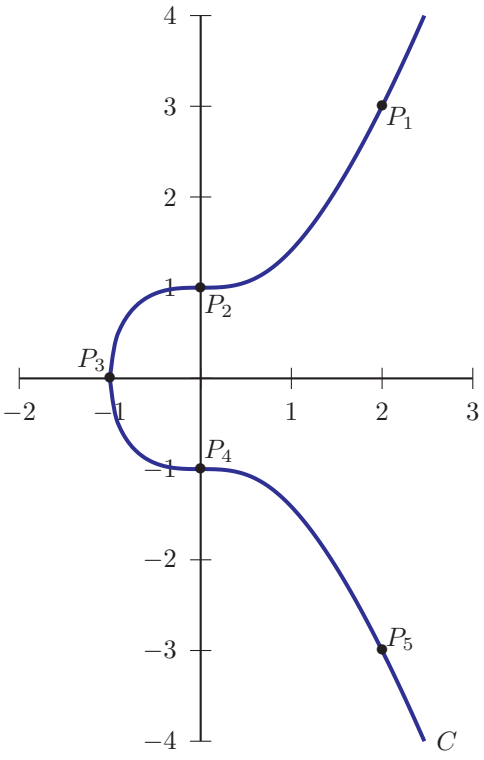


Figure 1. The cubic curve $C = \mathrm{V}(x^3 - y^2z + z^3)$ in the affine patch $z = 1$.

Exercise 2.3.5. Let $O = P_2$ be the specified inflection point so that $+$ is defined relative to P_2, i.e., $Q + R = P_2(QR)$ for points Q, R on C.

(1) Compute $P_1 + P_2$, $P_2 + P_2$, $P_3 + P_2$, $P_4 + P_2$, and $P_5 + P_2$.
(2) Explain why P_2 is the identity element for C.
(3) Find the inverses of P_1, P_2, P_3, P_4 and P_5 on C.
(4) Verify that $P_1 + (P_3 + P_4) = (P_1 + P_3) + P_4$. In general, addition of points on C is associative.

Exercise 2.3.6. Now let $O = P_4$ be the specified inflection point so that $+$ is defined relative to P_4, i.e., $Q + R = P_4(QR)$ for points Q, R on C.

(1) Compute $P_1 + P_2$, $P_2 + P_2$, $P_3 + P_2$, $P_4 + P_2$, and $P_5 + P_2$. [Hint: For $P_4 + P_2$ and $P_5 + P_2$ find the equations of the lines $\ell(P_4, P_2)$ and $\ell(P_5, P_2)$, respectively, to find the third points of intersection with C.] Are the answers the same as they were in Part (1) of Exercise 2.3.5? Is P_2 still the identity element for C?

(2) Now compute P_1+P_4, P_2+P_4, P_3+P_4, P_4+P_4, and P_5+P_4. Explain why P_4 is now the identity element for C.

(3) Using the fact that P_4 is now the identity element on C, find the inverses of P_1, P_2, P_3, P_4 and P_5 on C. [Hint: See the hint in Part (1).] Are these the same as the inverses found in Part (3) of Exercise 2.3.5?

Now we will prove that the cubic curve C with addition of points relative to a fixed inflection point O is an abelian group. First, we verify that the binary operation $+$ is commutative.

Exercise 2.3.7. Explain why $P + Q = Q + P$ for all points P, Q on C. This establishes that $+$ is a commutative binary operation on C.

In Exercises 2.3.5 and 2.3.6, the inflection point used to define the addition also served as the identity element for the curve $C = \mathrm{V}(x^3 - y^2z + z^3)$. In the exercise below, you will show this is true for any cubic curve.

Exercise 2.3.8. Let C be a smooth cubic curve and let O be one of its inflection points. Define addition $+$ of points on C relative to O. Show that $P + O = P$ for all points P on C and that there is no other point on C with this property. Thus O is the identity element for $+$ on C.

Thus $(C, O, +)$ satisfies group axiom (G2). Next, we verify that every point P on C has an inverse, so that C with $+$ also satisfies group axiom (G3).

Exercise 2.3.9. Let C be a smooth cubic curve and let O be one of its inflection points. Define addition $+$ of points on C relative to the identity O.

(1) Suppose that P, Q, R are collinear points on C. Show that $P + (Q + R) = O$ and $(P + Q) + R = O$.

(2) Let P be any point on C. Assume that P has an inverse element P^{-1} on C. Prove that the points P, P^{-1}, and O must be collinear.

(3) Use the results of Parts (1) and (2) to show that for any P on C there is an element P' on C satisfying $P + P' = P' + P = O$, i.e., every element P has an inverse P^{-1}. Then show this inverse is unique.

So far we have shown that $(C, O, +)$ has an identity, inverses, and is commutative. All that remains in order to prove that C is an abelian group is to show that $+$ is an associative operation. Establishing this fact is more involved than verifying the other axioms.

The following three exercises are based on [**Ful69**], pages 124–125. We will first develop some results regarding families of cubic curves.

Exercise 2.3.10. Start with two cubic curves, $C = \mathrm{V}(f)$ and $D = \mathrm{V}(g)$. By Theorem 2.2.37, there are exactly nine points of intersection, counting multiplicities, of C and D. Denote these points by $P_1, P_2, \ldots, P_9$.

(1) Let $\lambda, \mu \in \mathbb{C}$ be arbitrary constants. Show that $P_1, P_2, \ldots, P_9$ are points on the cubic curve defined by $\lambda f + \mu g = 0$.

(2) Let $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{C}$ be arbitrary constants. Show that $P_1, P_2, \ldots, P_9$ are the nine points of intersection of the cubic curves $C_1 = \mathrm{V}(\lambda_1 f + \mu_1 g)$ and $C_2 = \mathrm{V}(\lambda_2 f + \mu_2 g)$.

Let $F(x, y, z) = a_1x^3 + a_2x^2y + a_3x^2z + a_4xy^2 + a_5xyz + a_6xz^2 + a_7y^3 + a_8y^2z + a_9yz^2 + a_{10}z^3$ be a cubic whose coefficients, $a_1, a_2, \ldots, a_{10}$, are viewed as unknowns. Then, for any point $P = (x_0 : y_0 : z_0)$ in $\mathbb{P}^2$, the equation $F(P) = 0$ gives a linear equation in the unknown coefficients a_i. Explicitly, we obtain the linear equation

$$a_1x_0^3 + a_2x_0^2y_0 + a_3x_0^2z_0 + a_4x_0y_0^2 + a_5x_0y_0z_0 + a_6x_0z_0^2 + a_7y_0^3 + a_8y_0^2z_0 + a_9y_0z_0^2 + a_{10}z_0^3 = 0.$$

Recall that the coordinates of P are determined only up to nonzero scalar multiples. Since $F(x, y, z)$ is homogeneous of degree three, we have

$$F(\lambda x_0, \lambda y_0, \lambda z_0) = \lambda^3 F(x_0, y_0, z_0).$$

Therefore, the zero set of the equation in the ten unknowns

$$a_1, a_2, \ldots, a_{10}$$

is uniquely determined by P.

For k points $P_1, \ldots, P_k$ the system of equations

$$F(P_1) = F(P_2) = \cdots = F(P_k) = 0$$

is a system of k linear equations in the ten unknowns $a_1, a_2, \ldots, a_{10}$. The common solutions to this system are the coefficients of cubics through the k points.

In the next two exercises we will prove that eight "general" points impose eight conditions on the space of cubic polynomials. We state the following results, which are a direct consequence of Bézout's Theorem, which we will prove in Section 3.3.

(1) If a line and a cubic in $\mathbb{P}^2$ intersect in four points, then the cubic must be reducible and contain the line as a component.

(2) If a conic and a cubic intersect in seven points, then the cubic must be reducible and contain the conic as a component.

(3) Two conics meet in four points, counted up to intersection multiplicities.

Exercise 2.3.11. Consider eight distinct points $P_1, P_2, \ldots, P_8$ in $\mathbb{P}^2$, such that no four are collinear and no seven are on a single conic. Let F be a generic cubic polynomial with unknown coefficients $a_1, a_2, \ldots, a_{10}$. The system of simultaneous equations

$$F(P_1) = F(P_2) = \cdots = F(P_8) = 0$$

is a system of eight linear equations in the ten unknowns $a_1, a_2, \ldots, a_{10}$. Prove that the vector space of solutions to this linear system has dimension equal to 2 by considering each of the following cases.

(1) The eight points are in *general position*, which means that no three are collinear and no six are on a conic.

(2) Three of the points are collinear.

(3) Six of the points are on a conic.

Exercise 2.3.12. Show that there are two linearly independent cubics

$$F_1(x, y, z) \text{ and } F_2(x, y, z)$$

such that any cubic curve passing through the eight points

$$P_1, P_2, \ldots, P_8$$

has the form $\lambda F_1 + \mu F_2$. Conclude that for any collection of eight points with no four collinear and no seven on a conic, there is a *unique* ninth point P_9 such that *every* cubic curve passing through the eight given points must also pass through P_9.

In this next exercise, we prove the associativity of the newly defined addition $+$ of points on a smooth cubic curve.

Exercise 2.3.13. Let C be a smooth cubic curve in $\mathbb{P}^2$ and let P, Q, R be three points on C. We will show that $P + (Q + R) = (P + Q) + R$.

- Let $\mathrm{V}(l_1) = \ell(P, Q)$ and $S_1 = PQ$, so $\mathrm{V}(l_1) \cap C = \{P, Q, S_1\}$.
- Let $\mathrm{V}(l_2) = \ell(S_1, O)$ and $S_2 = OS_1 = P + Q$, so $\mathrm{V}(l_2) \cap C = \{S_1, O, S_2\}$.
- Let $\mathrm{V}(l_3) = \ell(S_2, R)$ and $S_3 = (P + Q)R$, so $\mathrm{V}(l_3) \cap C = \{S_2, R, S_3\}$.

Similarly:

- Let $\mathrm{V}(m_1) = \ell(Q, R)$ and $T_1 = QR$, so $\mathrm{V}(m_1) \cap C = \{Q, R, T_1\}$.
- Let $\mathrm{V}(m_2) = \ell(T_1, O)$ and $T_2 = OT_1 = Q + R$, so $\mathrm{V}(m_2) \cap C = \{T_1, O, T_2\}$.
- Let $\mathrm{V}(m_3) = \ell(T_2, P)$ and $T_3 = P(Q + R)$, so $\mathrm{V}(m_3) \cap C = \{T_2, P, T_3\}$.

(1) Notice that $C' = \mathrm{V}(l_1 m_2 l_3)$ is a cubic. Find $C' \cap C$.

(2) Likewise, $C'' = \mathrm{V}(m_1 l_2 m_3)$ is a cubic. Find $C'' \cap C$.

(3) Using Parts (1) and (2) together with Exercise 2.3.11, deduce that $(P + Q)R = P(Q + R)$.

(4) Explain why $(P+Q)R = P(Q+R)$ implies that $(P+Q)+R = P+(Q+R)$. Conclude that the addition of points on cubics is associative.

Therefore, a cubic curve C with a selected inflection point O determines a binary operation $+$ in such a way that $(C, O, +)$ is an abelian group under addition.[2]

Since $(C, O, +)$ is a group, it is natural to ask group theoretic questions about C, such as questions regarding the orders of its elements. First we define an integer multiple of a point and the order of a point.

Definition 2.3.2. Let $(C, O, +)$ be a smooth cubic curve and let $P \neq O$ be a point on the curve. For $n \in \mathbb{Z}$ we define $n \cdot P$ as follows:

- $0 \cdot P = O$ and $1 \cdot P = P$
- For $n \geq 2$, we have $n \cdot P = (n-1) \cdot P + P$
- For $n < 0$, we set $n \cdot P$ to be the inverse of $(-n) \cdot P$.

Definition 2.3.3. Let $(C, O, +)$ be a smooth cubic curve and let $P \neq O$ be a point on the curve. If there exists a positive integer n so that $n \cdot P = O$ and for $1 \leq m \leq n-1$ we have $m \cdot P \neq O$, then the point P has *order* n. If no such positive integer exists, then the point is said to have *infinite order*.

We can now examine points of finite order. In particular, we are interested here in points of order two and three. Many deep questions in mathematics are concerned with the computation of the order of various points on a cubic curve.

2.3.3. Points of Order Two and Three. Let C be a smooth cubic curve with $+$ defined relative to an inflection point O, the group identity. Let P be a point on C.

Exercise 2.3.14. Show that $2 \cdot P = O$ if and only if $\ell(O, P)$ is tangent to C at P.

[2] We defined addition on C relative to an inflection point, O, but we could define addition on C relative to any point O on C. See Husemöller, *Elliptic Curves* [**Hus87**], Theorem 1.2, for details.

Exercise 2.3.15. Show that if P and Q are two points on C of order two, then PQ, the third point of intersection of C with $\ell(P,Q)$, is also a point of order two on C.

Exercise 2.3.16. Let C be the cubic curve defined by $y^2z = x^3 - xz^2$.

(1) Show that $O = (0:1:0)$ is an inflection point.
(2) Graph C in the affine patch $z = 1$.
(3) Show that lines through $(0:1:0)$ correspond to vertical lines in the affine patch $z = 1$.
(4) Find three points of order two in the group $(C, O, +)$.

Let C be a smooth cubic curve with $+$ defined relative to the inflection point O.

Exercise 2.3.17. Let P be any inflection point on C. Show that $3 \cdot P = O$.

Exercise 2.3.18. Suppose P is point on C and $3 \cdot P = O$. Conclude that $PP = P$. From this, deduce that P is a point of inflection on C.

We have shown that the points of order 3 are inflection points, and an inflection point which is not O must have order 3. We will return to points of finite order in Section 2.4.3 after we have developed a more convenient way to express our smooth cubic curves.

2.4. Normal Forms of Cubics

The goal of this section[3] is to show that every smooth cubic is projectively equivalent to one of the form $y^2 = x^3 + Ax + B$, which is called the Weierstrass normal form, where the coefficients A and B are uniquely determined. We will also show that every smooth cubic is projectively equivalent to a curve of the form $y^2 = x(x-1)(x-\lambda)$, called the canonical form. For a given cubic, there are 6 possible values for this λ. We associate to each cubic a complex number called the j-invariant and show that we can parametrize all cubics by the complex numbers via their j-invariant.

[3]The development in this section follows the first two sections of chapter three of J. Silverman's *The Arithmetic of Elliptic Curves* [**Sil86**].

2.4.1. Weierstrass Normal Form. We will show that any smooth cubic curve can be transformed into the Weierstrass normal form $y^2 = x^3 + Ax + B$ under a projective change of coordinates. This will be accomplished using a sequence of several projective changes of coordinates, both in the general case and with a concrete example, the Fermat cubic $x^3 + y^3 - z^3 = 0$.

Let C be a smooth cubic curve in $\mathbb{P}^2$ given by the homogeneous equation $f(x, y, z) = 0$. Select an inflection point, $O = (a_0 : b_0 : c_0)$, on C and let ℓ denote the tangent line to C at O, where ℓ is defined by the linear equation $l(x, y, z) = 0$. Recall that we can projectively change coordinates with an invertible 3×3 matrix M

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = M \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

We choose M so that

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = M \begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix},$$

and ℓ is transformed to the line defined by $l_1(x_1, y_1, z_1) = z_1$, i.e., the inflection point O becomes $(0 : 1 : 0)$ and the tangent line ℓ becomes the line $z_1 = 0$ under the projective change of coordinates M. Recall that we carry out the computations of changing coordinates by using the inverse M^{-1} of M and replacing x, y, and z with expressions involving x_1, y_1, and z_1.

Exercise 2.4.1. Consider the smooth cubic curve C defined by $x^3 + y^3 - z^3 = 0$.

(1) Show that $O = (1 : 0 : 1)$ is an inflection point of C.

(2) Show that $x - z = 0$ is the equation of the tangent line to C at O.

(3) Find a 3×3 matrix M such that, under the change of variables

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = M^{-1} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix},$$

we have $(1 : 0 : 1) \mapsto (0 : 1 : 0)$ and $l(x, y, z) = x - z$ becoming $l_1(x_1, y_1, z_1) = z_1$.

(4) Find the equation $f_1(x_1, y_1, z_1) = 0$ for the curve C_1 that is associated to this projective change of coordinates.

Now we have transformed our original smooth cubic curve C into another smooth cubic curve C_1, which is projectively equivalent to C. Let's now work with the new curve C_1 that is defined by the equation $f_1(x_1, y_1, z_1) = 0$ in $\mathbb{P}^2$ with coordinates $(x_1 : y_1 : z_1)$.

Exercise 2.4.2.

(1) Explain why the homogeneous polynomial $f_1(x_1, y_1, z_1)$ can be expressed as

$$f_1(x_1, y_1, z_1) = \alpha x_1^3 + z_1 F(x_1, y_1, z_1),$$

where $\alpha \neq 0$ and $F(0, 1, 0) \neq 0$.

(2) Explain why the highest power of y_1 in the homogeneous polynomial $f_1(x_1, y_1, z_1)$ is two.

(3) Explain how by rescaling we can introduce new coordinates $(x_2 : y_2 : z_2)$ so that the coefficient of x_2^3 is 1 and the coefficient of $y_2^2 z_2$ is -1 in the new homogeneous polynomial $f_2(x_2, y_2, z_2) = 0$.

We can now rearrange the equation $f_2(x_2, y_2, z_2) = 0$ to be of the form

$$y_2^2 z_2 + a_1 x_2 y_2 z_2 + a_3 y_2 z_2^2 = x_2^3 + a_2 x_2^2 z_2 + a_4 x_2 z_2^2 + a_6 z_2^3. \tag{2.2}$$

Exercise 2.4.3. Use the Fermat curve defined in Exercise 2.4.1 for the following.

(1) Show that

$$M^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

is the desired matrix that solves Part (3) of Exercise 2.4.1.

(2) Find the homogeneous polynomial $f_1(x_1, y_1, z_1)$ that corresponds to this projective change of coordinates.

(3) Verify that f_1 is of the form $f_1(x_1, y_1, z_1) = \alpha x_1^3 + z_1 F(x_1, y_1, z_1)$, where $\alpha \neq 0$ and $F(0, 1, 0) \neq 0$.

(4) Rescale, if necessary, so that the coefficient of x_2 is 1 and the coefficient of $y_2^2 z_2$ is -1.

(5) Rearrange $f_2(x_2, y_2, z_2) = 0$ to be in the form of Equation (2.2).

Let's now work in the affine patch $z_2 = 1$, that is, in the affine (x_2, y_2)-plane, and consider the nonhomogeneous form of Equation (2.2),

$$y_2^2 + a_1 x_2 y_2 + a_3 y_2 = x_2^3 + a_2 x_2^2 + a_4 x_2 + a_6, \tag{2.3}$$

keeping in mind that there is an extra point at infinity. We can treat the left-hand side of Equation (2.3) as a quadratic expression in y_2. This means we can complete the square to remove some of the terms.

Consider the following concrete examples.

Exercise 2.4.4.

(1) Complete the square on the left-hand side of the following equation.

$$y^2 + 2y = 8x^3 + x - 1$$

(2) Find an affine change of coordinates so that $y^2 + 2y = 8x^3 + x - 1$ becomes $v^2 = f(u)$.

Exercise 2.4.5.

(1) Complete the square (with respect to y) on the left-hand side of the following equation.

$$y^2 + 4xy + 2y = x^3 + x - 3.$$

(2) Find an affine change of coordinates such that $y^2 + 2y = 8x^3 + x - 1$ becomes $v^2 = f(u)$.

Now we do this in general.

Exercise 2.4.6. Complete the square on the left-hand side of Equation (2.3) and verify that the affine change of coordinates

$$x_3 = x_2$$
$$y_3 = a_1 x_2 + 2y_2 + a_3$$

gives the new equation

$$y_3^2 = 4x_3^3 + (a_1^2 + 4a_2)x_3^2 + 2(a_1a_3 + 2a_4)x_3 + (a_3^2 + 4a_6). \tag{2.4}$$

To simplify notation, we introduce the following:

$$b_2 = a_1^2 + 4a_2,$$
$$b_4 = a_1a_3 + 2a_4,$$
$$b_6 = a_3^2 + 4a_6,$$

so that Equation (2.4) becomes

$$y_3^2 = 4x_3^3 + b_2x_3^2 + 2b_4x_3 + b_6. \tag{2.5}$$

We are now ready to make the final change of coordinates to achieve the Weierstrass normal form. Our goal is to scale the coefficient of x_3^3 to 1 and to eliminate the x_3^2 term.[4]

Consider the following concrete examples.

Exercise 2.4.7.

(1) Suppose we have the equation

$$y^2 = x^3 + 6x^2 - 2x + 5.$$

Show that the affine change of coordinates

$$u = x + 2$$
$$v = y$$

eliminates the quadratic term on the right-hand side.

(2) Suppose we have the equation

$$y^2 = 4x^3 + 12x^2 + 4x - 6.$$

Show that the affine change of coordinates

$$u = 36x + 36$$
$$v = 108y$$

[4]This change of coordinates is similar to completion of the square, but with cubics. This was first used by Cardano in *Ars Magna* (in 1545) to achieve a general solution to the cubic equation $x^3 + \alpha x^2 + \beta x + \gamma = 0$. Cardano needed to eliminate the x^2 term then, as we do now. Since the coefficient of the cubic term in his equation is already one, he simply made the substitution $u = x - \alpha/3$.

eliminates the quadratic term and rescales the coefficient of the cubic term to one on the right-hand side.

Exercise 2.4.8. Verify that the affine change of coordinates

$$u = 36x_3 + 3b_2$$
$$v = 108y_3$$

gives the Weierstrass normal form

$$v^2 = u^3 - 27(b_2^2 - 24b_4)u - 54(b_2^3 + 36b_2b_4 - 216b_6).$$

Again we can introduce the following to simplify notation:

$$c_4 = b_2^2 - 24b_4$$
$$c_6 = -b_2^3 + 36b_2b_4 - 216b_6.$$

Then we have the following for our Weierstrass normal form.

$$v^2 = u^3 - 27c_4u - 54c_6. \tag{2.6}$$

Let's collect all of the coefficient substitutions that we have made, recalling that the a_i's are the coefficients from Equation (2.3):

$$b_2 = a_1^2 + 4a_2$$
$$b_4 = 2a_4 + a_1a_3$$
$$b_6 = a_3^2 + 4a_6$$
$$c_4 = b_2^2 - 24b_4$$
$$c_6 = -b_2^3 + 36b_2b_4 - 216b_6.$$

For upcoming computations it is convenient to introduce the following as well.

$$b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2$$
$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$$
$$j = \frac{c_4^3}{\Delta}.$$

Exercise 2.4.9. Show the following relationships hold:

(1) $4b_8 = b_2b_6 - b_4^2$
(2) $1728\Delta = c_4^3 - c_6^2$

(3) $j = \dfrac{1728c_4^3}{c_4^3 - c_6^2}$.

These are simply brute-force computations.

The number Δ is called the discriminant of the cubic curve, since it is related to the discriminant of the cubic polynomial in x on the right-hand side of Equation (2.5). The number j is called the *j-invariant* of the cubic curve. We will see its significance soon.

Exercise 2.4.10. Follow the procedure outlined above to write the following cubics in Weierstrass normal form and calculate their j-invariants.

(1) $y^2 + 2y = 8x^3 + x - 1$

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

To avoid even more cumbersome notation, let's "reset" our variables. Consider the Weierstrass normal form of a smooth cubic C:

$$y^2 = x^3 - 27c_4x - 54c_6. \tag{2.7}$$

Notice that with the specific example $x^3 + y^3 - z^3 = 0$ in $\mathbb{P}^2$ in Exercises 2.4.1 and 2.4.3, we chose the initial change of coordinates, so that the chosen inflection point becomes $(0 : 1 : 0)$ with tangent line given by $z = 0$. This is not a unique transformation. Suppose we had chosen a different transformation. That is, suppose instead of having the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

we obtained the equation

$$y^2 + a_1'xy + a_3'y = x^3 + a_2'x^2 + a_4'x + a_6'.$$

How different would our Weierstrass normal form have been?

Exercise 2.4.11. Show that the only (affine) transformation that takes

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

to

$$v^2 + a_1'uv + a_3'v = u^3 + a_2'u^2 + a_4'u + a_6'$$

is given by

$$\begin{aligned} x &= \alpha^2 u + r \\ y &= \alpha^2 s u + \alpha^3 v + t, \end{aligned}$$

with $\alpha, r, s, t \in \mathbb{C}$ and $\alpha \neq 0$. [Hint: Start with the projective transformation, which is also affine,

$$\begin{aligned} x &= a_{11}u + a_{12}v + a_{13}w \\ y &= a_{21}u + a_{22}v + a_{23}w \\ z &= w \end{aligned}$$

and show that the only way to satisfy the condition in this exercise is for the specific a_{ij} to have the form above.]

Using this change of coordinates, we can compute the following relationships[5] between equivalent cubic curves with coefficients a_i in Equation (2.2) with coordinates $(x : y : z)$ and coefficients a'_i with coordinates $(u : v : w)$.

$$\begin{aligned} \alpha a'_1 &= a_1 + 2s \\ \alpha^2 a'_2 &= a_2 - s a_1 + 3r - s^2 \\ \alpha^3 a'_3 &= a_3 + r a_1 + 2t \\ \alpha^4 a'_4 &= a_4 - s a_3 + 2r a_2 - (t + rs) a_1 + 3r^2 - 2st \\ \alpha^6 a'_6 &= a_6 + r a_4 - t a_3 + r^2 a_2 - rt a_1 + r^3 - t^2 \end{aligned}$$

$$\begin{aligned} \alpha^2 b'_2 &= b_2 + 12r \\ \alpha^4 b'_4 &= b_4 + r b_2 + 6r^2 \\ \alpha^6 b'_6 &= b_6 + 2r b_4 + r^2 b_2 + 4r^3 \\ \alpha^6 b'_8 &= b_8 + 3r b_6 + 3r^2 b_4 + r^3 b_2 + 3r^4 \end{aligned}$$

$$\begin{aligned} \alpha^4 c'_4 &= c_4 \\ \alpha^6 c'_6 &= c_6 \\ \alpha^{12} \Delta' &= \Delta \\ j' &= j. \end{aligned}$$

[5] This is Table 1.2 in Silverman [**Sil86**].

Notice that if two smooth cubic plane curves are projectively equivalent, then the value j for each is the same, which is why we call this number the j-invariant. Let C and C' be two cubic plane curves, written in Weierstrass normal form

$$C : y^2 = x^3 + Ax + B$$
$$C' : y^2 = x^3 + A'x + B'.$$

Exercise 2.4.12. Suppose C and C' have the same j-invariant.

(1) Show that this implies

$$\frac{A^3}{4A^3 + 27B^2} = \frac{A'^3}{4A'^3 + 27B'^2}.$$

(2) Show that from the previous part we have $A^3B'^2 = A'^3B^2$.

In the next exercises we construct the transformations that send C to C'. We need to consider three cases: $A = 0$, $B = 0$, $AB \neq 0$.

Exercise 2.4.13. Suppose $A = 0$.

(1) Show that if $A = 0$, then $B \neq 0$. [Hint: Recall that C is smooth.]

(2) What is j if $A = 0$?

(3) Explain why $B' \neq 0$.

(4) Show that the following change of coordinates takes C to C'.

$$x = (B/B')^{1/3}u$$
$$y = (B/B')^{1/2}v.$$

Exercise 2.4.14. Suppose $B = 0$.

(1) What is j if $B = 0$?

(2) Explain why $A' \neq 0$.

(3) Show that the following change of coordinates takes C to C'.

$$x = (A/A')^{1/2}u$$
$$y = (A/A')^{3/4}v.$$

Exercise 2.4.15. Suppose $AB \neq 0$. Find a change of coordinates that takes C to C'. [Hint: See the two previous problems.]

We can summarize the preceding discussion with the following theorem.

Theorem 2.4.16. Two smooth cubic curves are projectively equivalent if and only if their j-invariants are equal.

The following exercises yield a characterization of smooth cubics via the j-invariant.

Exercise 2.4.17. Let γ be any complex number except 0 or 1728, and consider the cubic curve C defined by

$$y^2 + xy = x^3 - \frac{36}{\gamma - 1728}x - \frac{1}{\gamma - 1728}.$$

Compute j for this cubic.

There are natural, but technical, reasons for the appearance of the seemingly random number 1728.

Exercise 2.4.18. Compute j for the following cubics.

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

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

Exercise 2.4.19. Use Theorem 2.4.16 and Exercises 2.4.10 and 2.4.18 to show that $\mathrm{V}(x^3 + xz^2 - y^2z)$ and $\mathrm{V}(8x^3 + xz^2 - y^2z - 2yz^2 - z^3)$ are projectively equivalent.

Exercises 2.4.17 and 2.4.18 establish the following theorem.

Theorem 2.4.20. If γ is any complex number, then there exists a plane cubic curve whose j-invariant is γ.

2.4.2. Canonical Form. As we have just seen the Weierstrass normal form is very useful and provides a nice way to characterize smooth plane cubics. Another form that is equally useful is the canonical form of the cubic. Consider Equation 2.5 from above

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6.$$

Exercise 2.4.21. Rewrite Equation (2.5) in (x_1, y_1) using the change of coordinates

$$x = x_1$$
$$y = 2y_1.$$

The change of coordinates in Exercise 2.4.21 scales the cubic coefficient on the right-hand side to one. Now we can factor the resulting equation to obtain

$$y_1^2 = (x_1 - e_1)(x_1 - e_2)(x_1 - e_3). \tag{2.8}$$

Exercise 2.4.22. Show that e_1, e_2, e_3 are distinct. [Hint: Recall that the cubic curve $\mathrm{V}((x - e_1z)(x - e_2z)(x - e_3z) - y^2z)$ is smooth.]

Consider the following example.

Exercise 2.4.23. In Exercise 2.4.10 we found the Weierstrass normal form of $y^2+2y = 8x^3+x-1$ to be $y^2 = x^3+\frac{1}{2}x$. Factor the right-hand side to find values for e_1, e_2, and e_3.

Now we do this in general.

Exercise 2.4.24. Rewrite Equation (2.8) in (x_2, y_2) using the change of coordinates

$$x_1 = (e_2 - e_1)x_2 + e_1$$
$$y_1 = (e_2 - e_1)^{3/2}y_2.$$

Exercise 2.4.25. Show that if we make the substitution

$$\lambda = \frac{e_3 - e_1}{e_2 - e_1} \tag{2.9}$$

in the equation we found in Exercise 2.4.24, we get

$$y_2^2 = x_2(x_2 - 1)(x_2 - \lambda).$$

We say a smooth cubic is in canonical form if it is

$$y^2 = x(x - 1)(x - \lambda). \tag{2.10}$$

Exercise 2.4.26. Find an affine transformation that puts $y^2 + 2y = 8x^3 + x - 1$ in canonical form. What is λ?

We digress for a moment here. By now we have become comfortable working in $\mathbb{P}^2$ and in various affine patches. We have seen that the context often determines when it is most advantageous to work in an affine patch. We usually work in the affine xy-plane, i.e., the $z = 1$ patch, but we need to be sure that we are not missing anything that happens "at infinity."

Exercise 2.4.27. Let $C \subset \mathbb{P}^2$ be the smooth cubic defined by the homogeneous equation $y^2z = x(x - z)(x - \lambda z)$. Show that the only point at infinity $(x_1 : y_1 : 0)$ on C is the point $(0 : 1 : 0)$. (We will see the significance of the point $(0 : 1 : 0)$ in Section 2.5.)

In Equation (2.8) we factored the right-hand side and called the roots e_1, e_2, and e_3, but these labels are just labels. We could just as easily have written e_2, e_3, and e_1. In other words, we should get the same cubic curve no matter how we permuted the e_i's. There are $3! = 6$ distinct permutations of the set $\{e_1, e_2, e_3\}$, so we expect that there would be six equivalent ways to express our cubic in canonical form. Recall that we defined λ as a ratio in Equation (2.9). Changing the roles of e_2 and e_3 would give $1/\lambda$ rather than λ. The two cubics

$$y^2 = x(x - 1)(x - \lambda)$$

and

$$y^2 = x(x - 1)(x - 1/\lambda)$$

should still be equivalent.

Exercise 2.4.28. Suppose we have the following canonical cubic

$$y^2 = x(x - 1)(x - \lambda),$$

where λ corresponds to the order e_1, e_2, e_3 of the roots in (2.8). Show that the other five arrangements of $\{e_1, e_2, e_3\}$ yield the following values in place of λ.

$$\frac{1}{\lambda} \qquad 1 - \lambda \qquad \frac{1}{1 - \lambda} \qquad \frac{\lambda - 1}{\lambda} \qquad \frac{\lambda}{\lambda - 1}$$

As we have seen, the value of λ in a canonical form of C is almost uniquely determined by C. The correspondence between complex numbers $\lambda \neq 0, 1$ and smooth cubic curves C is a six-to-one correspondence, where if λ is a complex number assigned to C, then all of

the complex numbers in Exercise (2.4.28) are assigned to C. Though λ is not uniquely determined, the j-invariant, as we would expect, is unique.

Exercise 2.4.29. Show that if a smooth cubic curve C has an equation in canonical form

$$y^2 = x(x-1)(x-\lambda),$$

then its j-invariant is

$$j = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}.$$

[Hint: Write the equation $y^2 = x(x-1)(x-\lambda)$ in Weirstrass normal form and use Exercise 2.4.9 to compute j.]

Exercise 2.4.30. Use the λ found in Exercise 2.4.26 to compute the j-invariant of $y^2 + 2y = 8x^3 + x - 1$. [Hint: Use the expression in Exercise 2.4.29.] Check that this agrees with the computation of j in Exercise 2.4.10.

Exercise 2.4.31. Show that the j-invariant of a smooth cubic curve C can be written as

$$2^7 \left[\sum_{i=1}^{6} \mu_i^2 - 3 \right],$$

where the μ_i range over the six values $\lambda, 1/\lambda, \ldots$ from Exercise 2.4.28.

Exercise 2.4.31 demonstrates that the value of the j-invariant, while expressed in terms of a particular choice of λ associated to C, is independent of which λ corresponding to C we select. When we combine Exercise 2.4.31 and Theorem 2.4.16 we see that, as we would expect, the six values in Exercise 2.4.28 really do give the same smooth cubic.

2.4.3. An Application: Points of Finite Order. As we have seen it is often convenient to express a smooth cubic in canonical form. For our final application in this section we will prove that there are exactly three points of order two on a smooth cubic. We showed, in Exercise 2.3.15, that if we have two points P and Q of order two, then there is a third point PQ also of order two, but we are not assured of the existence of the two points P and Q or that

there is not another point R, of order two, not collinear with P and Q. Exercise 2.3.16 suggests there are exactly three such points and now we will prove this in general. Recall that in Exercise 2.3.14, we showed that a point $P \in C$ has order two if and only if the tangent to C at P passes through the identity element O.

Exercise 2.4.32. Let $C = \mathrm{V}(x(x-1)(x-\lambda)-y^2)$ be a smooth cubic curve with $+$ defined relative to the inflection point $O = (0:1:0)$.

(1) Homogenize Equation 2.10 and find the equation of the tangent line $\mathrm{V}(l)$ to C at the point $P = (x_0 : y_0 : z_0)$.

(2) Show that $(0:1:0) \in \mathrm{V}(l)$ if and only if either $z_0 = 0$ or $y_0 = 0$.

(3) Show that O is the only point in C with $z_0 = 0$.

(4) Show that $(0:0:1)$, $(1:0:1)$, and $(\lambda:0:1)$ are the only points in C with $y_0 = 0$.

(5) Conclude that there are exactly three points of order two on C.

We have just shown that any cubic C has exactly three points of order two. In fact, we have found these points explicitly, but we can say even more.

Exercise 2.4.33.

(1) Show that the points of order two on C, together with $O = (0:1:0)$, form a subgroup of C.

(2) Show that this subgroup is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

We showed, in Exercises 2.3.17 and 2.3.18, that a point $P \in C$ satisfies $3P = O$ if and only if P is an inflection point. By Exercise 2.2.38 there are exactly nine inflection points on C, but O has order one. Thus there are eight points of order three on C.

In general, there are n^2 points on C whose order divides n. Hence there are twelve points of order four on C, as there will be sixteen whose order divides four, but four of these are already counted among the three points of order two and O.

2.5. The Group Law for a Smooth Cubic in Canonical Form

The goal of this section is to reformulate the group law for a smooth cubic in canonical form $y^2 = x(x-1)(x-\lambda)$. By doing so, we will see that the group law for cubics is valid not only over $\mathbb{C}$, but over fields of positive characteristic[6] and non-algebraically closed fields, too.

We have already shown that the set of points of a smooth cubic curve C forms a group under the binary operation $+$ we defined in Section 2.3. In what follows we will use the canonical form developed in Section 2.4 to determine the (affine) coordinates of the point $P+Q$ given coordinates of P and Q. We will use the point at infinity $(0:1:0)$ as our identity O on C. When we work in the affine patch $z=1$, we will see that the line $\ell(O, PQ)$ that we use to determine $P+Q$ will correspond to the vertical line through PQ.

2.5.1. The Identity, Addition, and Inverses. First, we need to establish that $O \in C$ and that any vertical line in the affine xy-plane does indeed pass through O.

Exercise 2.5.1. Consider the cubic curve C in homogeneous canonical form given by $y^2z = x(x-z)(x+z)$, i.e., $C = \mathrm{V}(x^3 - xz^2 - y^2z)$.

(1) Show that the point at infinity $(0:1:0)$ is on C.

(2) Show that $(0:1:0) \in \mathrm{V}(H(x^3 - xz^2 - y^2z))$, the Hessian curve of C, and conclude that $O = (0:1:0)$ is an inflection point.

(3) Show that every vertical line in the affine xy-plane meets C at $(0:1:0)$.

(4) Sketch the graph of the real affine part of C, $y^2 = x^3 - x$.

(5) Let P and Q be two points on the real affine curve. Explain geometrically that if the line $\ell(P, Q)$ through P and Q intersects C a third time at the point $PQ = (a, b)$, then $P + Q = (a, -b)$.

[6] We would need to modify our calculations from the previous sections for fields of characteristic two or three.

(6) Now suppose that $R = (a : b : 1)$ is a point on C. Show that the line $\ell(O, R)$ is given by the equation $x - az = 0$, which is the vertical line $x = a$ in the xy-plane.

Exercise 2.5.2. Let $\lambda \neq 0, 1$ be a complex number and consider the cubic curve C in homogeneous canonical form given by $y^2z = x(x - z)(x - \lambda z)$, i.e., $C = \mathrm{V}(x(x - z)(x - \lambda z) - y^2z)$.

(1) Show that the point at infinity, $(0 : 1 : 0)$ is on C.

(2) Show that $(0 : 1 : 0) \in \mathrm{V}(H(x(x - z)(x - \lambda z) - y^2z))$, the Hessian curve of C, and conclude that $O = (0 : 1 : 0)$ is an inflection point.

(3) Show that every vertical line in the affine xy-plane meets C at O.

(4) Suppose that $P = (a : b : 1)$ is a point on C. Show that the line $\ell(O, P)$ is given by the equation $x - az = 0$, which is the vertical line $x = a$ in the xy-plane.

Now we have established that if $C = \mathrm{V}(x(x - z)(x - \lambda z) - y^2z)$ is given in canonical form, then $(0 : 1 : 0)$ is an inflection point, so henceforth we let $O = (0 : 1 : 0)$ be our identity element and define $+$ relative to it. Before we develop an algebraic expression for the coordinates of $P + Q$, we first consider the coordinates of P^{-1}, the inverse of the point P. Recall that if $P \in C$ then the inverse P^{-1} of P is the third point of intersection of C and $\ell(O, P)$.

Exercise 2.5.3. First, we want to work in the affine patch $z = 1$, so we dehomogenize our cubic equation $y^2 = x(x - 1)(x - \lambda)$. Let $P = (x_1, y_1)$ be a point in the xy-plane on C with $y_1 \neq 0$.

(1) Find the linear equation that defines $\ell(O, P)$.

(2) Find the point $P' = (x_2, y_2)$ that is the third point of intersection of $\ell(O, P)$ and C in the xy-plane.

(3) Show that $P + P' = O$. Conclude that $P' = P^{-1}$.

Therefore, if $P = (x_1 : y_1 : 1)$ is a point on C, the additive inverse of P is the point $P^{-1} = (x_1 : -y_1 : 1)$ on C. Notice in Exercise 2.5.3 we assumed $y_1 \neq 0$ for our point P. Now we see what the inverse of a point on the x-axis in the affine xy-plane is.

Exercise 2.5.4. Let $P = (x_1, 0)$ be a point in the xy-plane on C defined by $y^2 = x(x-1)(x-\lambda)$.

(1) Show that $2P = O$, so that $P = P^{-1}$.

(2) Show that this agrees with Exercise 2.3.14, that is, show that the tangent to C at $P = (x_1, y_1)$ in the xy-plane is a vertical line if and only if $y_1 = 0$.

2.5.2. The Group Law. Our goal in this section is to obtain an algebraic formula for the sum of two points on a cubic in canonical form.

Exercise 2.5.5. Consider the cubic curve $C = \mathrm{V}(x^3 - xz^2 + z^3 - y^2z)$ and the points $P_1 = (1:1:1)$, $P_2 = (0:1:1)$, $P_3 = (-1:1:1)$, $P_4 = (-1:-1:1)$, $P_5 = (0:-1:1)$, $P_6 = (1:-1:1)$ on C. Figure 2 shows C in the affine $z = 1$ patch.

(1) Use a straightedge and the figure below to find $P_1 + P_2$, $P_1 + P_3$, $P_1 + P_4$, and $P_3 + P_4$ geometrically. [Hint: $O = (0:1:0)$, the point at infinity, is the identity and we use the vertical line through P_iP_j to find $P_i + P_j$.]

(2) Find the coordinates of $P_1 + P_2$, $P_1 + P_3$, $P_1 + P_4$, and $P_3 + P_4$. [Hint: Use the equation of the line through P_i and P_j to find the coordinates of the point P_iP_j. Now find the coordinates of $P_i + P_j$ using the equation of the vertical line through P_iP_j.]

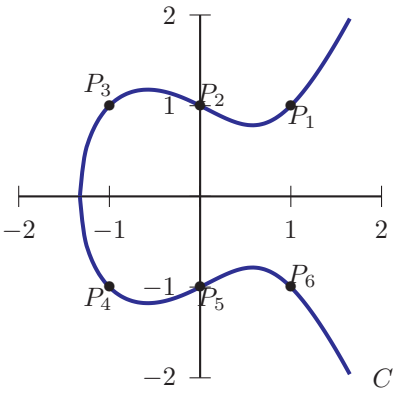


Figure 2. C in the affine xy-plane.

Exercise 2.5.6. Let C be the affine cubic curve defined by the equation $y^2 = x^3 + x^2 - 2x$. Let P denote the point $(-1/2, -3\sqrt{2}/4)$ and Q denote the point $(0, 0)$.

(1) Write the defining equation of C in canonical form and verify that P and Q are on C.

(2) Find the equation of $\ell(P, Q)$, the line through P and Q.

(3) Find the coordinates of the point PQ on C, that is, the coordinates of the third point of intersection of C and $\ell(P, Q)$.

(4) Let O denote the inflection point $(0 : 1 : 0)$ and find the coordinates of the point $P + Q$ on C using O as the identity element.

(5) Find the coordinates of $2P$ on C.

(6) Find the coordinates of the point P^{-1} on C.

(7) Show that $2Q = O$. [Hint: Show that the tangent to C at Q passes through O and invoke Exercise 2.3.14.]

(8) Find the coordinates of all three points of order two on C.

Now we carry out these computations in a more general setting to derive an expression for the coordinates of $P + Q$. Let $C = \mathrm{V}(x(x - z)(x - \lambda z) - y^2 z)$ be a smooth cubic curve. Dehomogenize the defining equation $x(x - z)(x - \lambda z) - y^2 z = 0$ to get the affine equation $y^2 = f(x)$, where $f(x) = x(x - 1)(x - \lambda)$.

Exercise 2.5.7. Suppose $P = (x_1 : y_1 : 1)$ and $Q = (x_2 : y_2 : 1)$ are two points on C, with $Q \neq P^{-1}$ (that is, $x_1 \neq x_2$), and let $y = \alpha x + \beta$ be the equation of the line $\ell(P, Q)$ through P and Q.

(1) Suppose $P \neq Q$. Express α in terms of x_1, x_2, y_1, y_2.

(2) Suppose $P = Q$ (in which case $\ell(P, Q)$ is the tangent line to C at P). Use implicit differentiation to express α in terms of x_1, y_1.

(3) Substitute $\alpha x + \beta$ for y in the equation $y^2 = f(x)$ to get a new equation in terms of x only. Write the resulting equation of x in the form $x^3 + Bx^2 + Cx + D = 0$.

(4) If $P + Q$ has coordinates $(x_3 : y_3 : 1)$, explain why $x^3 + Bx^2 + Cx + D$ must factor as $(x - x_1)(x - x_2)(x - x_3)$.

(5) By equating coefficients of x^2 in Parts (3) and (4), conclude that

$$x_3 = -x_1 - x_2 + \alpha^2 + \lambda + 1,$$

where α is the slope of the line $\ell(P, Q)$.

(6) We now have an expression for the x-coordinate of $P + Q$. Use this to conclude that

$$PQ = \left(-x_1 - x_2 + \alpha^2 + \lambda + 1 : y_1 + \alpha(x_3 - x_1) : 1\right),$$

where α is the slope of $\ell(P, Q)$ and therefore

$$P + Q = \left(-x_1 - x_2 + \alpha^2 + \lambda + 1 : -\left(y_1 + \alpha(x_3 - x_1)\right) : 1\right).$$

[Hint: Use the relationship between the y-coordinates of PQ and $P + Q$ along with the fact that (x_1, y_1) lies on the line defined by $y = \alpha x + \beta$.]

Therefore, if $P = (x_1 : y_1 : 1), Q = (x_2 : y_2 : 1)$ are points on the curve $C = \mathrm{V}(x(x - 1)(x - \lambda) - y^2)$, then $P + Q$ has coordinates $(x_3 : y_3 : 1)$ given by

$$\begin{aligned} x_3 &= \begin{cases} -x_1 - x_2 + \lambda + 1 + \left(\dfrac{y_2 - y_1}{x_2 - x_1}\right)^2 & \text{if } P \neq Q \\ -2x_1 + \lambda + 1 + \left(\dfrac{f'(x_1)}{2y_1}\right)^2 & \text{if } P = Q \end{cases} \\ y_3 &= -\left(y_1 + \alpha(x_3 - x_1)\right). \end{aligned}$$

Exercise 2.5.8. Verify the results in Exercise 2.5.6 using the above formula.

We may perform a similar sequence of calculations for a cubic in general form. Let C be the cubic curve defined by $y^2z = ax^3 + bx^2z + cxz^2 + dz^3$, where $a, b, c, d \in \mathbb{C}$. Dehomogenize this defining equation to get the affine equation $y^2 = f(x)$, where $f(x) = ax^3 + bx^2 + cx + d$ and f has distinct roots.

Exercise 2.5.9. Suppose $P = (x_1 : y_1 : 1)$ and $Q = (x_2 : y_2 : 1)$ are two points on C, with $Q \neq P^{-1}$, and let $y = \alpha x + \beta$ be the equation of line $\ell(P, Q)$ through the points P and Q.

(1) Suppose $P \neq Q$. Express α in terms of x_1, x_2, y_1, y_2.

(2) Suppose $P = Q$ (in which case $\ell(P, Q)$ is the tangent line to C at P). Use implicit differentiation to express α in terms of x_1, y_1.

(3) Substitute $\alpha x + \beta$ for y in the equation $y^2 = f(x)$ to get a new equation in terms of x only. Write the resulting equation in the form $ax^3 + Bx^2 + Cx + D = 0$.

(4) If $P + Q$ has coordinates $P + Q = (x_3 : y_3 : 1)$, explain why $ax^3 + Bx^2 + Cx + D$ must factor as $a(x - x_1)(x - x_2)(x - x_3)$.

(5) By equating coefficients of x^2, conclude that

$$x_3 = -x_1 - x_2 + \frac{\alpha^2 - b}{a},$$

where α is the slope of the line $\ell(P, Q)$.

(6) We now have an expression for the x-coordinate of $P + Q$. Use this to conclude that

$$PQ = \left(-x_1 - x_2 - \frac{b}{a} + \frac{1}{a}\alpha^2 : y_1 + \alpha(x_3 - x_1) : 1 \right),$$

where α is the slope of $\ell(P, Q)$ and use this to find $P + Q$.

Therefore, if $P = (x_1 : y_1 : 1), Q = (x_2 : y_2 : 1)$ are points on the curve $C = \mathrm{V}(ax^3 + bx^2 + cx + d - y^2)$, then $P + Q$ has coordinates $(x_3 : y_3 : 1)$ given by

$$\begin{aligned} x_3 &= \begin{cases} -x_1 - x_2 - \dfrac{b}{a} + \dfrac{1}{a}\left(\dfrac{y_2 - y_1}{x_2 - x_1}\right)^2 & \text{if } P \neq Q \\ -2x_1 - \dfrac{b}{a} + \dfrac{1}{a}\left(\dfrac{f'(x_1)}{2y_1}\right)^2 & \text{if } P = Q \end{cases} \\ y_3 &= -(y_1 + \alpha(x_3 - x_1)). \end{aligned}$$

2.5.3. Rational Points on Cubics. Of particular importance in number theory and the theory of elliptic curves is the following property of the group law.

Definition 2.5.1. Let $y^2 = f(x)$ be an affine equation of a smooth cubic curve, where $f(x)$ is a polynomial with rational coefficients. A point $P = (x, y)$ on the curve is a *rational point* if $x, y \in \mathbb{Q}$.

Once we have a rational point, a natural follow-up would be to ask how many rational points exist on a given curve. We first note the following property of rational points.

Exercise 2.5.10. Let $y^2 = f(x)$ be an affine equation of a smooth cubic curve, where $f(x)$ is a degree three polynomial with rational coefficients. Suppose P and Q are rational points on this curve, so that $P, Q \in \mathbb{Q}^2$ and $Q \neq P^{-1}$. Prove that $P + Q$ is also a rational point.

This shows that the rational points on a cubic form a subgroup. The study of the structure of this subgroup leads to some of the most significant open questions in number theory, including the Birch and Swinnerton-Dyer Conjecture.

2.5.4. Cubics over Other Fields. Another important consequence of our algebraic formulation for the group law is that the operations involved are independent of the field of definition. With this addition law, we can define the group law for cubic curves not only over $\mathbb{C}$, but also over $\mathbb{R}$, $\mathbb{Q}$, and even over finite fields. However, there is one subtlety, namely some of the calculations need to be modified if the characteristic of the field is equal to 2.

Exercise 2.5.11. This is inspired by [**AG06**], pages 105–109. Let C be the cubic curve given by $y^2 = x^3 + 1$.

(1) Show that $(0, 4)$ and $(2, 3)$ are points of C over $\mathbb{Z}_5$, the field of order five.

(2) Use the formulas for addition above to compute $(0, 4)+(2, 3)$.

(3) Find all of the points on C that are defined over $\mathbb{Z}_5$.

2.6. Cross-Ratios and the j-Invariant

We have seen that the j-invariant uniquely determines a cubic curve up to a projective change of coordinates. In this section, we will develop another way to understand the j-invariant. We start by showing that any three points in $\mathbb{P}^1$ can be sent to any other three points under a projective change of coordinates. It is critical, though, to understand that it is not possible for four points to be sent to an arbitrary collection of four other points. It is here that the cross-ratio appears. The key is that two ordered sets of four points are projectively equivalent if and only if they have the same cross-ratio. The cross-ratio will then return us to the j-invariant for a cubic curve.

2.6.1. Projective Changes of Coordinates for $\mathbb{P}^1$. Given any three points $(x_1 : y_1), (x_2 : y_2), (x_3 : y_3) \in \mathbb{P}^1$, we want to find a projective change of coordinates $T : \mathbb{P}^1 \to \mathbb{P}^1$ such that

$$\begin{aligned} T(x_1 : y_1) &= (1 : 0) \\ T(x_2 : y_2) &= (0 : 1) \\ T(x_3 : y_3) &= (1 : 1). \end{aligned}$$

We will see that not only does such a map exist, but that it is unique.

We first have to define what we mean by a projective change of coordinates for $\mathbb{P}^1$. In Section 1.5, we gave a definition for projective changes of coordinates for $\mathbb{P}^2$. The definition for $\mathbb{P}^1$ is similar, namely that a projective change of coordinates is given by

$$\begin{aligned} u &= ax + by \\ v &= cx + dy, \end{aligned}$$

where $ad - bc \neq 0$. We write this as

$$T(x : y) = (ax + by : cx + dy).$$

Now, we could write $(x : y) \in \mathbb{P}^1$ as a column vector

$$\begin{pmatrix} x \\ y \end{pmatrix}.$$

If we let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then we can think of $T(x : y) = (ax + by : cx + dy)$ in terms of the matrix multiplication

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

In $\mathbb{P}^1$, we have that $(x : y) = (\lambda x : \lambda y)$ for any constant $\lambda \neq 0$. This suggests the following:

Exercise 2.6.1. Show that the matrices

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 6 & 4 \\ 2 & 8 \end{pmatrix} = 2 \cdot A$$

give rise to the same projective change of coordinates of $\mathbb{P}^1$.

Exercise 2.6.2. Show that the matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix},$$

for any $\lambda \neq 0$, give rise to the same change of coordinates of $\mathbb{P}^1$.

This means that the projective changes of coordinates

$$(x : y) \to (ax + by : cx + dy)$$

and

$$(x : y) \to (\lambda ax + \lambda by : \lambda cx + \lambda dy)$$

are the same.

The projective change of coordinates T such that

$$T(x_1 : y_1) = (1 : 0),\ T(x_2 : y_2) = (0 : 1),\ T(x_3 : y_3) = (1 : 1)$$

is

$$T(x : y) = ((x_2y - y_2x)(x_1y_3 - x_3y_1) : (x_1y - y_1x)(x_2y_3 - x_3y_2)).$$

(It should not be at all clear how this T was created.)

Exercise 2.6.3. Let

$$(x_1 : y_1) = (1 : 2),\ (x_2 : y_2) = (3 : 4),\ (x_3 : y_3) = (6 : 5).$$

Show that

(1) $T(x : y) = (28x - 21y : 18x - 9y)$

(2) $T(1 : 2) = (1 : 0),\ T(3 : 4) = (0 : 1),\ T(6 : 5) = (1 : 1).$

These problems give no hint as to how anyone could have known how to create T; the goal of these last problems was to show that this T actually works.

Now we want to start looking at uniqueness questions.

Exercise 2.6.4. Let $T(x : y) = (ax + by : cx + dy)$ be a projective change of coordinates such that $T(1 : 0) = (1 : 0),\ T(0 : 1) = (0 : 1),\ T(1 : 1) = (1 : 1)$. Show that

$$a = d \neq 0$$

and that

$$b = c = 0.$$

Explain why T must be the same as the projective change of coordinates given by $T(x : y) = (x : y)$.

Part of showing uniqueness will be in finding an easy-to-use formula for the inverse of our map T.

Exercise 2.6.5. Let $T(x : y) = (ax + by : cx + dy)$ be a projective change of coordinates and let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be its associated matrix. Let

$$B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Show that

$$A \cdot B = \det(A) I,$$

where I is the 2×2 identity matrix.

Exercise 2.6.6. Let $(x_1 : y_1),\ (x_2 : y_2),\ (x_3 : y_3) \in \mathbb{P}^1$ be three distinct points. Let T_1 and T_2 be two projective changes of coordinates such that

$$T_1(x_1 : y_1) = (1 : 0),\ T_1(x_2 : y_2) = (0 : 1),\ T_1(x_3 : y_3) = (1 : 1),$$

and

$$T_2(x_1 : y_1) = (1 : 0),\ T_2(x_2 : y_2) = (0 : 1),\ T_2(x_3 : y_3) = (1 : 1).$$

Show that $T_1 \circ T_2^{-1}$ is a projective change of coordinates such that

$$\begin{aligned} T_1 \circ T_2^{-1}(1 : 0) &= (1 : 0), \\ T_1 \circ T_2^{-1}(0 : 1) &= (0 : 1), \\ T_1 \circ T_2^{-1}(1 : 1) &= (1 : 1). \end{aligned}$$

Show that T_1 and T_2 must be the same projective change of coordinates.

Thus our desired map T is unique.

Exercise 2.6.7. Mathematicians will say that any three points in $\mathbb{P}^1$ can be sent to any other three points, but any fourth point's image must be fixed. Using the results of this section, explain what this means. (This problem is not so much a typical math exercise but is instead an exercise in exposition.)

We have seen that the map

$$T(x : y) = ((x_2y - y_2x)(x_1y_3 - x_3y_1) : (x_1y - y_1x)(x_2y_3 - x_3y_2))$$

works, but we have not explained how it was derived. We just have to find a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

such that

$$A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ A \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ A \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Solving for the coefficients of A is now just a (somewhat brutal) exercise in algebra that yields the map T.

2.6.2. Working in $\mathbb{C}$. Algebraic geometers like to work in projective space $\mathbb{P}^n$. Other mathematicians prefer to work in affine space, such as $\mathbb{C}^n$, allowing for points to go off to infinity. In this subsection we interpret the projective change of coordinates $T : \mathbb{P}^1 \to \mathbb{P}^1$ from the previous section as a map $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$.

Given three points x_1, x_2 and x_3 in $\mathbb{C}$, we want to find a map $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ such that

$$\begin{aligned} T(x_1) &= \infty \\ T(x_2) &= 0 \\ T(x_3) &= 1. \end{aligned}$$

For now, set

$$T(x) = \frac{(x_2 - x)(x_1 - x_3)}{(x_1 - x)(x_2 - x_3)}.$$

The next three exercises are in parallel with those in the previous subsection.

Exercise 2.6.8. Let $x_1 = 1/2$, $x_2 = 3/4$, and $x_3 = 6/5$. (Note that these correspond to the dehomogenization of the three points $(x_1 : y_1) = (1 : 2)$, $(x_2 : y_2) = (3 : 4)$, $(x_3 : y_3) = (6 : 5)$ in the previous subsection's first problem.) Show that

(1) $T(x) = \dfrac{28x - 21}{18x - 9}$

(2) $T(1/2) = \infty$, $T(3/4) = 0$, $T(6/5) = 1$.

The next exercise will link the map $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ with the map $T : \mathbb{P}^1 \to \mathbb{P}^1$. Recall in $\mathbb{P}^1$ that

$$(x : y) = \left(\frac{x}{y} : 1\right),$$

provided that $y \neq 0$. By a slight abuse of notation, we can think of dehomogenizing as just setting all of the y's equal to one.

Exercise 2.6.9. Show that the map $T : \mathbb{P}^1 \to \mathbb{P}^1$ given by

$$T(x : y) = (ax + by : cx + dy)$$

will correspond to a map $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ given by

$$T(x) = \frac{ax + b}{cx + d}.$$

Exercise 2.6.10. Show that the map $T : \mathbb{P}^1 \to \mathbb{P}^1$ given by

$$T(x : y) = ((x_2y - y_2x)(x_1y_3 - x_3y_1) : (x_1y - y_1x)(x_2y_3 - x_3y_2))$$

will correspond to the map $T : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$ given by

$$T(x) = \frac{(x_2 - x)(x_1 - x_3)}{(x_1 - x)(x_2 - x_3)}.$$

Here the dehomogenization is the map achieved by setting $y = 1$.

2.6.3. Cross-Ratio: A Projective Invariant. We introduce the fundamental invariant for points on the projective line $\mathbb{P}^1$, the cross-ratio of four points.

Suppose we are given some points in $\mathbb{P}^1$. We can label these points in many ways, by choosing different coordinate systems. This is the same as studying the points under projective changes of coordinates. We would like to associate to our points something (for us, a number) that will not change, no matter how we write the points. We call such numbers *invariants*.

If we start with three points in $\mathbb{P}^1$, no such invariant can exist, since any three points can be sent to any other three points. But we cannot send any four points to any other four points. This means that any collection of four points has some sort of intrinsic geometry.

Definition 2.6.1. The *cross-ratio* of the four distinct points p_1, p_2, p_3, p_4 is

$$[p_1, p_2, p_3, p_4] = \frac{(x_2y_4 - y_2x_4)(x_1y_3 - x_3y_1)}{(x_1y_4 - y_1x_4)(x_2y_3 - x_3y_2)},$$

where $p_i = (x_i : y_i)$.

We need to show that this number does not change under a projective change of coordinates.

Exercise 2.6.11. Let

$$p_1 = (1 : 2),\ p_2 = (3 : 1),\ p_3 = (1 : 1),\ p_4 = (5 : 6).$$

(1) Calculate the cross-ratio $[p_1, p_2, p_3, p_4]$.

(2) Let $T : \mathbb{P}^1 \to \mathbb{P}^1$ be

$$T(x : y) = (3x + 2y : 2x + y).$$

Find $T(p_1),\ T(p_2),\ T(p_3),\ T(p_4)$.

(3) Show

$$[T(p_1), T(p_2), T(p_3), T(p_4)] = [p_1, p_2, p_3, p_4].$$

Exercise 2.6.12. Let $p_1 = (x_1 : y_1)$, $p_2 = (x_2 : y_2)$, $p_3 = (x_3 : y_3)$, $p_4 = (x_4 : y_4)$ be any collection of four distinct points in $\mathbb{P}^1$ and let $T(x, y) = (ax + by : cx + dy)$ be any projective change of coordinates. Show

$$[T(p_1), T(p_2), T(p_3), T(p_4)] = [p_1, p_2, p_3, p_4].$$

(This is a long exercise in algebra, but at the end, there should be satisfaction at seeing everything being equal.)

The above cross-ratio depends, though, on how we ordered our four points p_1, p_2, p_3, p_4. If we change the order, the cross-ratio might change.

Exercise 2.6.13. Let p_1, p_2, p_3, p_4 be any four distinct points in $\mathbb{P}^1$. Show

$$[p_1, p_2, p_3, p_4] = \frac{1}{[p_2, p_1, p_3, p_4]}.$$

Exercise 2.6.14. Let $p_1 = (x_1 : y_1)$, $p_2 = (x_2 : y_2)$, $p_3 = (x_3 : y_3)$, $p_4 = (x_4 : y_4)$ such that $[p_1, p_2, p_3, p_4] \neq \pm 1$. Show that there is no projective change of coordinates $T(x : y) = (ax + by : cx + dy)$ such that T interchanges p_1 with p_2 but leaves p_3 and p_4 alone. In other words, show there is no T such that

$$T(p_1) = p_2,\ T(p_2) = p_1,\ T(p_3) = p_3,\ T(p_4) = p_4.$$

Exercise 2.6.15. Let $p_1 = (x_1 : y_1)$, $p_2 = (x_2 : y_2)$, $p_3 = (x_3 : y_3)$, $p_4 = (x_4 : y_4)$ be any collection of four distinct points in $\mathbb{P}^1$. Show that

$$[p_2, p_1, p_4, p_3] = [p_1, p_2, p_3, p_4].$$

Exercise 2.6.16. Using the notation from the previous problem, find two other permutations of the points p_1, p_2, p_3, p_4 so that the cross-ratio does not change.

Let

$$[p_1, p_2, p_3, p_4] = \lambda.$$

We have shown that there are four permutations of the $p_1,\ p_2,\ p_3,\ p_4$ (including the identity) that do not change the cross-ratio, but we have also shown

$$[p_2, p_1, p_3, p_4] = \frac{1}{\lambda}.$$

Exercise 2.6.17. Using the above notation, find permutations of $p_1,\ p_2,\ p_3,\ p_4$ so that all of the following cross-ratios occur:

$$\lambda,\ \frac{1}{\lambda},\ \frac{1}{1-\lambda},\ 1-\lambda,\ \frac{\lambda}{\lambda-1},\ \frac{\lambda-1}{\lambda}.$$

Exercise 2.6.18. Given any four distinct points $p_1,\ p_2,\ p_3,\ p_4$ in $\mathbb{P}^1$, show that the j-invariant of the cross-ratio does not change under any reordering of the four points or under any projective change of coordinates. (This is why we are justified in using the term "invariant" in the name j-invariant.)

Thus given a smooth cubic curve, we can put the curve into Weierstrass normal form and associate to this curve a single number j. A natural question is if two different curves could have the same j-invariant. The next exercises will show that this is not possible. (We have already done this in Section 2.4, but now we will do this in the context of the cross-ratio.)

Exercise 2.6.19. Suppose that

$$j(\lambda) = 2^8 \frac{(\lambda^2-\lambda+1)^3}{\lambda^2(\lambda-1)^2} = a$$

for some constant a.

(1) Show that any solution μ of the equation

$$2^8(\lambda^2-\lambda+1)^3 - a\lambda^2(\lambda-1)^2 = 0$$

has the property that

$$j(\mu) = a.$$

(2) Show that the above equation has only six solutions.

(3) Show that if λ is a solution, then the other five solutions are $\frac{1}{\lambda},\ \frac{1}{1-\lambda},\ 1-\lambda,\ \frac{\lambda}{\lambda-1},\ \frac{\lambda-1}{\lambda}$.

(4) Show that if we have two curves $zy^2 = x(x-z)(x-\lambda z)$ and $zy^2 = x(x-z)(x-\mu z)$ with

$$j(\lambda) = j(\mu),$$

then there is a projective change of coordinates taking the first curve to the second.

2.7. Torus as $\mathbb{C}/\Lambda$

In the last chapter, we showed that all smooth conics are topologically spheres. Our long-term goal for the rest of this chapter is to show that all smooth cubics are topologically tori. This will take some work. In this section we will take the first step, which is to realize a torus as a quotient group $\mathbb{C}/\Lambda$, where Λ is a lattice.

2.7.1. Quotient Groups. Since we want to show that a torus is a quotient group, we will begin this section with some background material from abstract algebra to make clear what a quotient group is.

Given an abelian group G with binary operation $+$, a subset S of G is said to be a *subgroup* if S is itself a group using the operation $+$. Given a known group G, a way to generate examples of groups is to look at all of its subgroups. Another way of generating examples is to "collapse" a subgroup N of the group G into the identity element of a new "quotient group" G/N.

Before we make this notion precise, we need to introduce partitions and their connection to equivalence relations.

Definition 2.7.1. Given a nonempty set A, we say that a collection P of subsets of A is a *partition* of A if P consists of nonempty, pairwise disjoint sets whose union is A. This means that if

$$P = \{P_\alpha\}_{\alpha \in I},$$

where I is an indexing set, then the elements of P satisfy the following two conditions:

(1) $P_\alpha \cap P_\beta = \emptyset$, for all $\alpha \neq \beta \in I$;

(2) $A = \bigcup_{\alpha \in I} P_\alpha$.

Exercise 2.7.1. Let A be a nonempty set.

(1) Let $\sim$ be an equivalence relation on the set A. Show that the set of equivalence classes of $\sim$ is a partition of A.

(2) Suppose P is a partition of A. Show that the relation $\sim$, defined by $x \sim y$ if and only if x and y belong to the same element of P, is an equivalence relation.

The previous exercise shows a natural correspondence between partitions and equivalence relations.

Let G be an abelian group with binary operation $+$, and let H be a subgroup of G. Define the relation $\sim$ on G by $x \sim y$ if $x - y \in H$.

Exercise 2.7.2. Show that $\sim$ is an equivalence relation.

Exercise 2.7.3. Let $G = \mathbb{Z}$ with the binary operation $+$. Let $H = 3\mathbb{Z}$, the multiples of 3, which is a subgroup of G. Show that $1 \sim 4$, $2 \sim 5$ and $2 \sim -1$.

This equivalence relation determines a partition of G. Denote the equivalence class of $x \in G$ by $x + H$. This is called a *coset* of H in G. Denote the set of all cosets by G/H.

Exercise 2.7.4. Let $G = \mathbb{Z}$ and $H = 3\mathbb{Z}$, a subgroup of G. Find all of the cosets of H in G.

Define an operation $+$ on G/H by

$$(x + H) + (y + H) = (x + y) + H.$$

(Note that $x + H$, $y + H$, $(x + y) + H$ are all sets.)

Exercise 2.7.5. Suppose $x \sim x'$ and $y \sim y'$ for elements $x, x', y, y' \in G$. Show that $(x + y) + H = (x' + y') + H$. (This demonstrates that the operation on G/H is well-defined.)

Under this operation, G/H will be a group, which is called the *quotient group* of G by H.

Exercise 2.7.6. Find the quotient group of $G = \mathbb{Z}$ by $H = 3\mathbb{Z}$.

In the discussion above, we have produced some ways of generating examples of groups: finding subgroups and taking quotients. How do we compare groups? One way of doing this is to look for maps between groups that preserve group structure.

Definition 2.7.2. Suppose $(G, +_G)$ and $(H, +_H)$ are two groups. A map $\varphi : G \to H$ is said to be a *homomorphism* if $\varphi(x +_G y) = \varphi(x) +_H \varphi(y)$ for all $x, y \in G$. If a homomorphism is bijective, we call it an *isomorphism* and say that the groups G and H are *isomorphic*. We denote this by $G \cong H$.

If two groups are isomorphic, they are essentially "the same." If there is a homomorphism between two groups there is still a nice relationship between G and H.

Exercise 2.7.7. Let $\varphi : G \to H$ be a homomorphism between abelian groups, and let e be the identity element of H. Let $\ker(\varphi) := \{g \in G : \varphi(g) = e\}$. (We call $\ker(\varphi)$ the *kernel* of φ.)

(1) Show that $\ker(\varphi)$ is a subgroup of G.

(2) Show that if $\varphi : G \to H$ is onto, then the quotient group $G/\ker(\varphi)$ is isomorphic to H.

If the groups are not abelian, the formation of quotient groups is much more involved.

2.7.2. The Torus. In order to understand some of the geometry of a torus, we need to determine how a torus is formed. We will begin by using a little group theory to realize a circle, S^1, as the quotient group $\mathbb{R}/\mathbb{Z}$.

Exercise 2.7.8.

(1) Show that $\mathbb{R}$ is an abelian group under addition.

(2) Show that $\mathbb{Z}$ is a subgroup of $\mathbb{R}$.

Exercise 2.7.9. Define a relation on $\mathbb{R}$ by $x \sim y$ if and only if $x - y \in \mathbb{Z}$.

(1) Verify that $\sim$ is an equivalence relation.

(2) Let $[x]$ denote the equivalence class of x, that is, $[x] = \{y \in \mathbb{R} : x \sim y\}$. Find the following equivalence classes: $[0]$, $[\frac{1}{2}]$, and $[\sqrt{2}]$.

(3) The equivalence relation $\sim$ gives a partition of $\mathbb{R}$. Explain how this partition $\mathbb{R}/\mathbb{Z}$ is the realization of a circle. [Hint: Explain how progressing from 0 to 1 is the same as going around a circle once.]

We can also use Exercise 2.7.7 to give an isomorphism between $\mathbb{R}/\mathbb{Z}$ and the circle. Let S^1 denote the unit circle centered at the origin in $\mathbb{R}^2$. As we have already seen $\mathbb{R}^2$ is in one-to-one correspondence with $\mathbb{C}$, so we can regard S^1 as the set $S^1 = \{x \in \mathbb{C} : |x| = 1\}$. Recall that any complex number has a polar representation $x = r(\cos\theta + i\sin\theta)$, so we can express S^1 as $S^1 = \{\cos\theta + i\sin\theta : \theta \in \mathbb{R}\} \subset \mathbb{C}$.

Exercise 2.7.10. Show that S^1 is a group under (complex) multiplication.

Exercise 2.7.11. Define a map $\phi : \mathbb{R} \to S^1$ by $\phi(\theta) = \cos 2\pi\theta + i\sin 2\pi\theta$.

(1) Show that ϕ is onto.

(2) Show that ϕ is a homomorphism, i.e., show that $\phi(\alpha+\beta) = \phi(\alpha)\phi(\beta)$ for all $\alpha, \beta \in \mathbb{R}$.

(3) Find $\ker\phi$ and conclude that $\mathbb{R}/\mathbb{Z} \cong S^1$.

We now want to extend the ideas in the previous exercises to the complex plane. Let ω_1 and ω_2 be complex numbers such that $\dfrac{\omega_1}{\omega_2}$ is not purely real. Let the integer lattice Λ be defined as

$$\Lambda = \{m\omega_1 + n\omega_2 : m, n \in \mathbb{Z}\}.$$

We will call the parallelogram formed by joining 0, ω_1, $\omega_1 + \omega_2$, ω_2, and 0 in succession the fundamental period-parallelogram. We will realize a torus as a quotient group $\mathbb{C}/\Lambda$.

Exercise 2.7.12.

(1) Sketch the lattice generated by $\omega_1 = 1$ and $\omega_2 = i$. [Hint: Sketch the fundamental period-parallelogram of this lattice.]

(2) Sketch the lattice generated by $\omega_1 = 1 + i$ and $\omega_2 = i$.

Exercise 2.7.13.

(1) Show that $\mathbb{C}$ is an abelian group under addition.

(2) Show that Λ is a subgroup of $\mathbb{C}$.

Exercise 2.7.14. Define a relation on $\mathbb{C}$ by $x \sim y$ if and only if $x - y \in \Lambda$. Show that $\sim$ is an equivalence relation.

Since $\sim$ is an equivalence relation, it is natural to ask about the quotient group $\mathbb{C}/\Lambda$.

Exercise 2.7.15. Let $\Lambda \subset \mathbb{C}$ be the integer lattice generated by $\{\omega_1 = 1, \omega_2 = i\}$ and let $a, b \in \mathbb{R}$.

(1) Find all points in $\mathbb{C}$ equivalent to $\frac{1}{2} + \frac{1}{2}i$.

(2) Find all points in $\mathbb{C}$ equivalent to $\frac{1}{3} + \frac{1}{4}i$.

(3) Show that $a \sim a + i$.

(4) Show that $bi \sim 1 + bi$.

Exercise 2.7.16. Sketch a sequence of diagrams to show that $\mathbb{C}/\Lambda$ is a torus. [Hint: Construct a torus using $\omega_1 = 1$ and $\omega_2 = i$ by identifying the horizontal and vertical sides of the fundamental period-parallelogram as in the previous problem. Now repeat with any lattice.]

Exercise 2.7.17. Let $\Lambda \subset \mathbb{C}$ be the integer lattice generated by $\{\omega_1 = 1, \omega_2 = i\}$.

(1) Sketch a vertical segment in the fundamental period-parallelogram and illustrate to what this corresponds on our torus. Sketch a horizontal line in the fundamental period-parallelogram and illustrate to what this corresponds on our torus.

(2) Show that $\frac{1}{4} + i \in \mathbb{C}/\Lambda$ has order 4 and write all of the elements of $\langle \frac{1}{4} + i \rangle$.

(3) Represent the fact that $\frac{1}{4} + i$ has order 4 geometrically on the fundamental period-parallelogram by sketching a line in $\mathbb{C}$ that has slope $\frac{1}{4}$ and considering its image in $\mathbb{C}/\Lambda$.

(4) Sketch the paths traced by these segments on the torus. What do you notice about this path on the torus?

(5) Pick any element $\alpha \in \mathbb{C}/\Lambda$ and show that if α has finite order, then the path on the torus represented by the line through 0 and α is a closed path.

(6) Suppose an element α has infinite order. What can you say about the slope of the line through 0 and α? Illustrate this phenomenon on the fundamental period-parallelogram in $\mathbb{C}$ and on the torus.

2.8. Mapping $\mathbb{C}/\Lambda$ to a Cubic

The goal of this section is to construct a map from a torus $\mathbb{C}/\Lambda$ to a cubic curve.

In this section we assume some knowledge about complex variables and analysis.

Our goal is to construct a map from the quotient group $\mathbb{C}/\Lambda$ to $\mathbb{C}^2$ whose image is the zero locus of a nonsingular cubic polynomial. In order to do this we will use the Weierstrass $\wp$-function $\wp : \mathbb{C}/\Lambda \to \mathbb{C}$ defined by

$$\wp(x) = \frac{1}{x^2} + \sum_{\substack{m,n\in\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{(x - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2}.$$

Then our map $\mathbb{C}/\Lambda \to \mathbb{C}^2$ will be given by the map $x \mapsto (\wp(x), \wp'(x))$, and the smooth cubic will be defined by the differential equation

$$[\wp'(x)]^2 = 4[\wp(x)]^3 + A\wp(x) + B,$$

where A, B are constants depending on the lattice.

At this point it is not at all clear how we arrived at the function $\wp$. We begin by considering the minimal properties that are essential

for a map $\mathbb{C}/\Lambda \to \mathbb{C}$. We will then show that $\wp$ has these properties and gives us our desired cubic.

Exercise 2.8.1. Let $f : \mathbb{C} \to \mathbb{C}$ be a function. Show that $x + \Lambda \mapsto f(x)$ is a well-defined function from $\mathbb{C}/\Lambda$ to $\mathbb{C}$ if and only if f is doubly-periodic, that is,

$$f(x + \omega_1) = f(x) \quad \text{and} \quad f(x + \omega_2) = f(x),$$

for all x in the domain of f.

To define the function f, we need to consider only what happens on the fundamental period-parallelogram. We would like our function f to be as nice as possible. For example, we would like our function f to be analytic on its fundamental period-parallelogram, i.e., f equals its Taylor series, $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Unfortunately, this will not work.

Exercise 2.8.2. Show that if a doubly-periodic function f is analytic on its fundamental period-parallelogram, then f is constant. [Hint: Use Liouville's Theorem.]

We see then that f cannot be analytic on its entire fundamental period parallelogram. The next hope is that f is analytic except with a single pole at 0, and hence at the other lattice points by double periodicity. Furthermore, we hope that the pole at 0 is not too bad. We can do this, but 0 will be a pole of order two, as the next two exercises illustrate.

Recall that a function $f(x)$ on $\mathbb{C}$ has a *pole* of order k at a point $\alpha \in \mathbb{C}$ if near α we can write the function as

$$\begin{aligned} f(x) = &\frac{b_k}{(x-\alpha)^k} + \frac{b_{k-1}}{(x-\alpha)^{k-1}} + \cdots + \frac{b_1}{x-\alpha} \\ &+ a_0 + a_1(x-\alpha) + a_2(x-\alpha)^2 + \cdots, \end{aligned}$$

where $b_1, \ldots b_k, a_0, a_1, \ldots$ are all complex numbers and with $b_k \neq 0$. Thus

$$f(x) = \frac{3}{(x-2i)^2} + \frac{1}{(x-7)} + 4 + 5x$$

has a pole of order two at $2i$, a pole of order one at 7 and no other poles.

It is inconvenient to integrate over these parallelograms if the singularities are on the boundaries, but we can translate the vertices, without rotating, so that the singularities are in the interior. The translated parallelograms will be called *cells*.

Exercise 2.8.3. Show that the sum of the residues of f at its poles in any cell is zero.

Exercise 2.8.4. Show that if f has a single pole at 0 in its fundamental period-parallelogram, not including the other vertices, then 0 must be a pole of order at least two.

We have now established that a candidate for our function could have the form

$$f(x) = \frac{a_{-2}}{x^2} + a_0 + a_1 x + a_2 x^2 + \dots$$

Exercise 2.8.5. Show that if

$$f(x) = \frac{a_{-2}}{x^2} + a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

is doubly-periodic, then f is an even function, i.e., $a_1 = a_3 = \dots = 0$. [Hint: Consider the function $f(x) - f(-x)$.]

We can change coordinates to eliminate a_0 so that f is now of the form

$$f(x) = \frac{a_{-2}}{x^2} + a_2 x^2 + a_4 x^4 + \dots$$

Now we are ready to introduce the Weierstrass $\wp$-function,

$$\wp(x) = \frac{1}{x^2} + \sum_{\substack{m,n\in\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{(x - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2}. \tag{2.11}$$

A series $\sum_{n=0}^{\infty} a_n$ is *absolutely convergent* whenever $\sum_{n=0}^{\infty} |a_n| < \infty$.

A series of functions $\sum f_n$ is *uniformly convergent* with limit f if, for all $\epsilon > 0$, there exists a natural number N such that for all x in the domain and all $n \geq N$, $|\sum_{k=1}^{n} f_k(x) - f(x)| < \epsilon$.

Exercise 2.8.6. Show that $\wp(x)$ converges uniformly and absolutely except near its poles. Conclude that $\wp(x)$ is analytic on the complex plane except at the lattice points $\Lambda = \{m\omega_1 + n\omega_2\}$.

Since $\wp(x)$ converges uniformly and absolutely, we can differentiate term-by-term to find $\wp'(x)$, and the order of summation does not affect the value of the function, so we can rearrange the terms.

Exercise 2.8.7. Find $\wp'(x)$ and show that $\wp'(x)$ is doubly-periodic.

Exercise 2.8.8. Show that $\wp(x)$ is doubly-periodic. [Hint: Consider the functions $F_i(x) = \wp(x + \omega_i) - \wp(x)$ for $i = 1, 2$.]

Consider the function $F(x) = \wp(x) - x^{-2}$.

Exercise 2.8.9. Show that F is analytic in a neighborhood of 0.

Exercise 2.8.10. Find the Taylor series expansion of F at 0.

Exercise 2.8.11. From Exercise 2.8.5 we know that $\wp(x)$ is even, so F is also even. Show that the odd powers of x vanish in the Taylor expansion of F at 0.

Exercise 2.8.12. Now we can rewrite $\wp(x) = x^{-2} + F(x)$. Find the coefficients of x^2 and x^4 in this expression for $\wp(x)$.

Exercise 2.8.13. Let

$$g_2 = 60 \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^4}$$

and

$$g_3 = 140 \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^6}.$$

Find the coefficients of x^2 and x^4 in $\wp(x)$ in terms of g_2 and g_3.

Exercise 2.8.14. Find the coefficients of x and x^3 in $\wp'(x)$ in terms of g_2 and g_3.

We will now establish a cubic relationship between $\wp(x)$ and $\wp'(x)$. In the previous exercises we found the following expressions for $\wp(x)$ and $\wp'(x)$:

$$\begin{aligned} \wp(x) &= \frac{1}{x^2} + \frac{1}{20} g_2 x^2 + \frac{1}{28} g_3 x^4 + O(x^6) \\ \wp'(x) &= -\frac{2}{x^3} + \frac{1}{10} g_2 x + \frac{1}{7} g_3 x^3 + O(x^5). \end{aligned}$$

Exercise 2.8.15. Compute $\wp(x)^3$ and $\wp'(x)^2$, and consider only terms up to first order, that is, find f and g such that $\wp(x)^3 = f(x) + O(x^2)$ and $\wp'(x)^2 = g(x) + O(x^2)$.

Exercise 2.8.16. Show that $\wp'(x)^2 = 4\wp(x)^3 - g_2\wp(x) - g_3$.

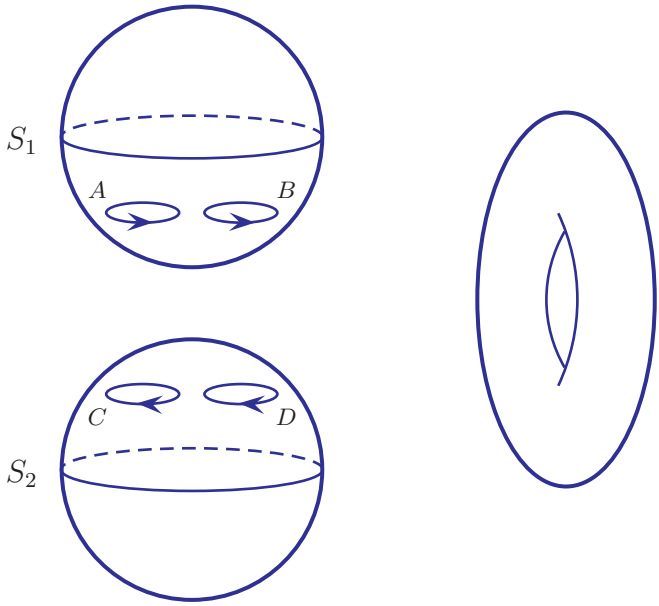
Thus, starting with the lattice Λ, we have a map to a cubic curve.

2.9. Cubics as Tori

The goal of this section is to show that a smooth cubic curve in $\mathbb{P}^2(\mathbb{C})$ is topologically a torus. We will be sketching the argument, as opposed to providing rigorous proofs.

In the following exercise, we first see how to obtain a torus from two spheres. Let A and B be two disjoint circles on a sphere. Let S_1 be the sphere with the interiors of A and B removed. Similarly, define S_2 as a different sphere with the interiors of disjoint circles C and D removed.

Exercise 2.9.1. Draw a sequence of diagrams to show that if we attach S_1 and S_2 by identifying the circle A to the circle C and the circle B to circle D, we obtain a torus. (Note we are working topologically, where we can deform objects, but not tear them.)



Now we return to cubics. Our goal is to show that a cubic in canonical form

$$y^2 = x(x-1)(x-\lambda)$$

can be realized as two spheres attached along two discs, as in the above problem and hence is topologically a torus. Since any smooth cubic curve can be put into canonical form, we will have shown that all smooth cubic curves are topologically tori.

The heart of the construction lies in the nature of the square root function, $\sqrt{z} = z^{1/2}$. As it stands, the square root function is not well-defined, but in fact has two possible values. For example, the square root of 4 is either 2 or -2. In high school, if you are just taking square roots of positive real numbers, the convention is usually to say that the square root is the positive value. Such a convention is more difficult to make over the complex numbers. Before starting the series of exercises to contruct a torus from a cubic, note that there is one and only one complex number for which the square root is unambiguous, namely $\sqrt{0} = 0$. Thus for any nonzero complex number w there are

two possible values for $\sqrt{w}$ but only one value for $\sqrt{0}$. This will be critical in a moment.

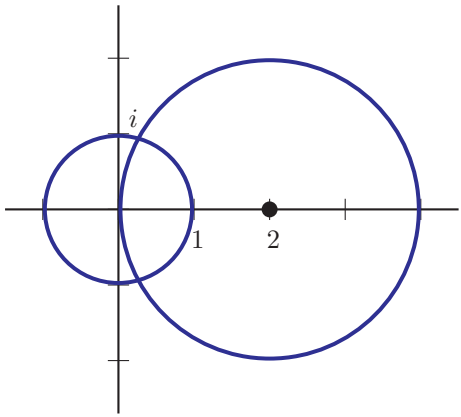
Exercise 2.9.2. Let $T : [0, 2\pi] \to \mathbb{C}$ be defined by $T(\theta) = e^{i\theta}$ and let $f : \mathbb{C} \to \mathbb{C}$ be defined by $f(x) = \sqrt{x}$.

(1) Show that $T([0, 2\pi])$ is a unit circle in $\mathbb{C}$.

(2) Show that $f \circ T([0, 2\pi])$ is a half circle.

This problem shows how there can be no easy sign convention for $\sqrt{w}$. We have $f \circ T(0) = 1$, while $f \circ T(2\pi) = -1$, even though $T(0) = T(2\pi) = 1$. Since f is the square root function, this means that the square root of 1 must be both 1 and -1. This is problematic at best. Even if we start with $\sqrt{1} = 1$, going once around the circle we must now have $\sqrt{1} = -1$.This forces the square root function to be two-valued and hence not a real function at all.

To remedy this situation, we do the following. Let $p(x)$ be a polynomial. If we go around a circle, the $\sqrt{p(x)}$ will change sign, provided that $p(x)$ has a zero in the interior. For example, $\sqrt{(x-2)}$ will change sign when we go around a circle of radius two centered at 2 but will not change sign if we go around a circle of radius one centered at the origin.



Exercise 2.9.3. Now let $T : [0, 2\pi] \to \mathbb{C}$ be defined by $T(\theta) = 2e^{i\theta}$ and let $f : \mathbb{C} \to \mathbb{C}$ be defined by $f(x) = \sqrt{x(x-1)}$.

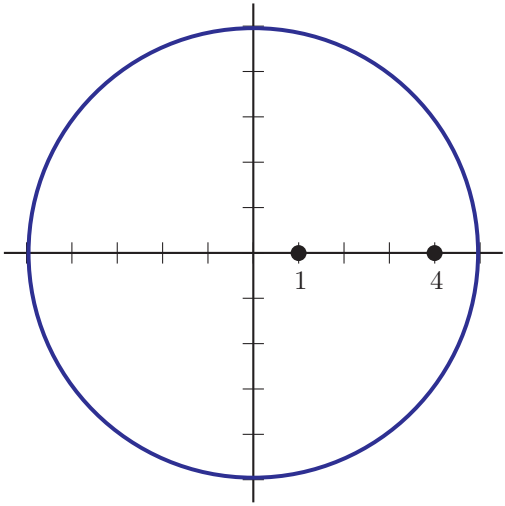
(1) Show that $T([0, 2\pi])$ is a circle of radius 2 in $\mathbb{C}$.

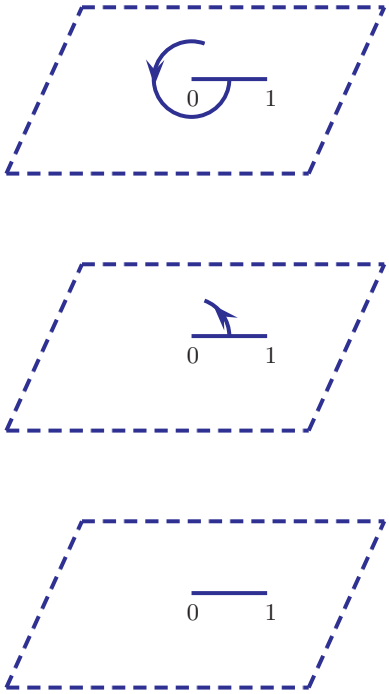
(2) Show that $f \circ T(0) = f \circ T(2\pi)$.

(3) Show that $f \circ T([0, 2\pi])$ is a closed curve in $\mathbb{C} - [0, 1]$.

Exercise 2.9.4. Using the notation from the previous problem, sketch an intuitive argument for $f(x) = \sqrt{x(x-1)}$ being well-defined on $\mathbb{C} - [0, 1]$ in two ways: (i) by setting $\sqrt{2(2-1)} = +\sqrt{2}$, and then (ii) by setting $\sqrt{2(2-1)} = -\sqrt{2}$. This construction establishes a two-sheeted cover of $\mathbb{C} - [0, 1]$.

This suggests that a function of the form $\sqrt{p(x)}$ will not change sign if $p(x)$ has two zeros in the interior. Thus we will assume, for example, that $\sqrt{(x-1)(x-4)}$ will not change sign along a circle of radius 5 centered at the origin.





Exercise 2.9.5. Let $T : [0, 2\pi] \to \mathbb{C}$ be defined by $T(\theta) = \frac{1}{2}e^{i(\theta+\pi/2)}$ and let $f : \mathbb{C} \to \mathbb{C}$ be defined by $f(x) = \sqrt{x(x-1)}$.

(1) Show that $T([0, 2\pi])$ is the circle of radius $\frac{1}{2}$, with center 0, starting at the point $\frac{1}{2}i$, in the counterclockwise direction.

(2) Show that $f \circ T(0)$ and $f \circ T(2\pi)$ give different values and that these exist on each of the two sheets.

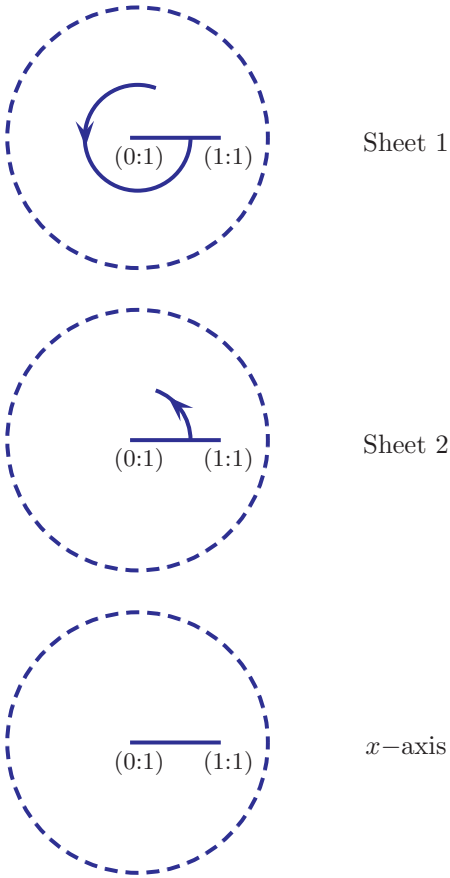
(3) Justify intuitively why $f \circ T([0, 2\pi])$ can be viewed as illustrated where Sheet 1 corresponds to $\sqrt{2}$ and Sheet 2 corresponds to $-\sqrt{2}$, as in the previous problem.

We now see how a particular conic can be viewed, topologically, as a sphere (a fact that we already showed in Chapter 1).

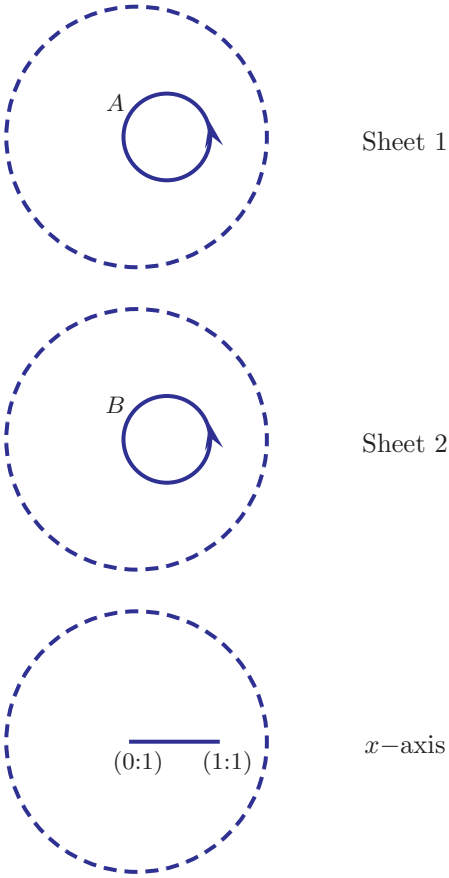
Exercise 2.9.6. Consider $\mathrm{V}(y^2 - x(x-z))$ in $\mathbb{P}^2$. Now instead of considering two $\mathbb{C}$ sheets, we include the point at infinity, so we have two $\mathbb{P}^1$ sheets, i.e., our two sheets are now spheres rather than planes.

(1) Show that for each $(x : z) \in \mathbb{P}^1$ there are two possible values for y, except at $(0 : 1)$ and $(1 : 1)$.

(2) Let γ be the straight real line from $(0 : 1)$ to $(1 : 1)$ where x/z is a positive real. Consider the following figure in which the bottom sphere corresponds to $\mathbb{P}^1 - \gamma$. Show that sitting over this projective line are two sheets, each of which is a copy of $\mathbb{P}^1 - \gamma$.



(3) Replace the segments in $[(0:1),(1:1)]$ in Sheets 1 and 2 with circles A and B. Draw a sequence of diagrams to show that if we attach circle A in Sheet 1 to circle B in Sheet 2, then we obtain a sphere.



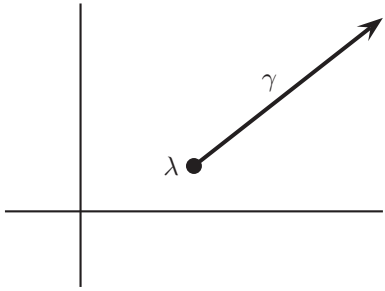
(4) Conclude that $\mathrm{V}(y^2 - x(x-z)) \subset \mathbb{P}^2$ is a sphere.

Exercise 2.9.7. Now consider $f : \mathbb{C} \to \mathbb{C}$ defined by

$$f(x) = \sqrt{x(x-1)(x-\lambda)}.$$

(1) Show that f is a 2-to-1 cover of the x-axis except at $x = 0$, $x = 1$, and $x = \lambda$.

(2) Let γ be a straight line in $\mathbb{C}$ from λ to infinity as in :



Via pictures, intuitively argue that $x(x-1)(x-\lambda)$ can contain either no zeros or two zeros in the interior of any circle in $\mathbb{C} - [0,1] - \gamma$.

(3) Sketch an intuitive argument for $f(x) = \sqrt{x(x-1)(x-\lambda)}$ being well-defined on $\mathbb{C} - [0,1] - \gamma$ in two ways: by fixing one of the values $f(x) = \sqrt{x(x-1)(x-\lambda)}$ at a point on $\mathbb{C} - [0,1] - \gamma$ and then by fixing the opposite value. This construction establishes a 2-sheeted cover of $\mathbb{C} - [0,1] - \gamma$.

(4) Homogenize $y^2 = x(x-1)(x-\lambda)$ to show that we now have a two-to-one cover of $\mathbb{P}^1$ except at $(0:1)$, $(1:1)$, $(\lambda:1)$, and $(1:0)$, where each of the two sheets is itself a $\mathbb{P}^1$.

(5) Use the earlier exercises to draw a sequence of diagrams illustrating how $y^2 = x(x-z)(x-\lambda z)$ in $\mathbb{P}^2$ is a torus.

Thus topologically all cubics are the same, while algebraically they can be quite different.

Chapter 3

Higher Degree Curves

The goal of this chapter is to explore higher degree curves in $\mathbb{P}^2$. There are seven parts. In the first, we define what is meant by an irreducible curve and its degree. We next show how curves in $\mathbb{P}^2$ can be thought of as real surfaces, similar to our observations for conics (Section 1.7) and cubics (Section 2.9). In the third part, we develop Bézout's Theorem, which tells us the number of points of intersection of two curves. We then introduce the ring of regular functions and the function field of a curve. In the fifth and sixth sections, we develop Riemann-Roch, an amazing theorem that links functions on the curve, the degree of the curve, and the genus (the number of holes) of the curve into one formula. In the last section, we consider singular points on a curve and develop methods for resolving them.

3.1. Higher Degree Polynomials and Curves

The goals of this section are to define what it means for a curve to be irreducible and to define the degree of a curve.

In Chapter 1 we dealt with conics, which are the zero sets of second degree polynomials. In Chapter 2 we looked at cubics, which are the zero sets of third degree polynomials. It is certainly natural to consider zero sets of higher degree polynomials.

By now, we know that it is most natural to work in the complex projective plane, $\mathbb{P}^2$, which means in turn that we want our zero sets to be the zero sets of homogeneous polynomials. Suppose that $P(x, y, z) \in \mathbb{C}[x, y, z]$ is a homogeneous polynomial. As before, we denote this polynomial's zero set by

$$\mathrm{V}(P) = \{(a : b : c) \in \mathbb{P}^2 : P(a, b, c) = 0\}.$$

Exercise 3.1.1. Let $P(x, y, z) = (x+y+z)(x^2+y^2-z^2)$. Show that $\mathrm{V}(P)$ is the union of the two curves $\mathrm{V}(x+y+z)$ and $\mathrm{V}(x^2+y^2-z^2)$.

Thus, if we want to understand $\mathrm{V}(P)$, we should start by looking at its two components: $\mathrm{V}(x+y+z)$ and $\mathrm{V}(x^2+y^2-z^2)$. In many ways, this might remind us of working with prime factorization of numbers. If we understand these building blocks—those numbers that cannot be broken into a product of two smaller numbers—then we start to understand the numbers formed when they are strung together.

Exercise 3.1.2. Let $P(x, y, z) = (x + y + z)^2$. Show that $\mathrm{V}(P) = \mathrm{V}(x + y + z)$.

Both $(x+y+z)(x^2+y^2-z^2)$ and $(x+y+z)^2$ are *reducible*, meaning that both can be factored. When we are considering a factorization, we do not consider trivial factorizations, such as $P = 1 \cdot P$. We would prefer, for now, to restrict our attention to curves that are the zero sets of irreducible homogeneous polynomials.

Definition 3.1.1. If the defining polynomial P cannot be factored, we say the curve $\mathrm{V}(P)$ is *irreducible*.

For the rest of this chapter, all polynomials used to define curves will be irreducible unless otherwise indicated.

Definition 3.1.2. The *degree* of the curve $\mathrm{V}(P)$ is the degree of its defining polynomial $P(x, y, z)$.

The degree of a curve is the most basic number associated to a curve that is invariant under change of coordinates. The following is an example of this phenomenon.

Exercise 3.1.3. Let $P(x, y, z) = x^3 + y^3 - z^3$. Then $\mathrm{V}(P)$ is a degree three curve. Consider the projective change of coordinates

$$\begin{aligned} x &= u - w \\ y &= iv \\ z &= u + v. \end{aligned}$$

Find the polynomial $\widetilde{P}(u, v, w)$ whose zero set $\mathrm{V}(\widetilde{P})$ maps to $\mathrm{V}(P)$. Show that $\mathrm{V}(\widetilde{P})$ also has degree three.

3.2. Higher Degree Curves as Surfaces

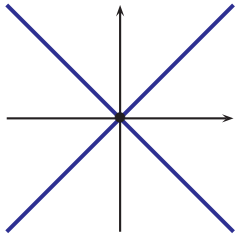
The goal of this section is to generalize our work in Sections 1.7 (Chapter 1) and 2.9 (Chapter 2), where we realized smooth conics and cubics over $\mathbb{C}$ as topological surfaces over $\mathbb{R}$.

3.2.1. Topology of a Curve. Suppose $f(x, y, z)$ is a homogeneous polynomial, so $\mathrm{V}(f)$ is a curve in $\mathbb{P}^2$. Recall that the degree of $\mathrm{V}(f)$ is, by definition, the degree of the homogeneous polynomial f. We will see that this algebraic invariant of the curve is closely linked to the topology of the curve viewed as a surface over $\mathbb{R}$. Specifically, it is related to the "genus" of the curve, which counts the number of holes in the surface.

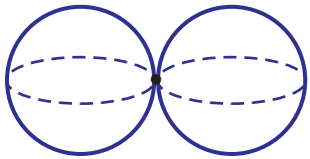
Before we proceed to higher degree curves, we return to our previous experience with conics and cubics.

Exercise 3.2.1. Consider the conics defined by the homogeneous equation $x^2 - y^2 = \lambda z^2$, where λ is a parameter. Sketch affine patches of these in the chart $z = 1$ for $\lambda = 4, 1, 0.25$.

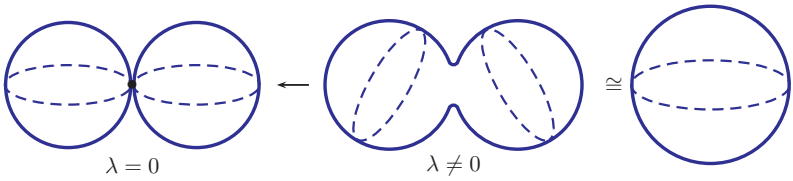
As $\lambda \to 0$, we get $x^2 - y^2 = 0$, or $(x - y)(x + y) = 0$. In $\mathbb{R}^2$, this looks like



but this picture isn't accurate over $\mathbb{C}$ in $\mathbb{P}^2$. Instead, topologically the picture looks like "kissing spheres":

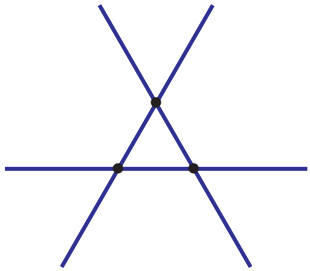


Thus, the topological version of the original equation, $x^2 - y^2 = \lambda z^2$, should be found by perturbing the kissing spheres a little to account for $\lambda \neq 0$:

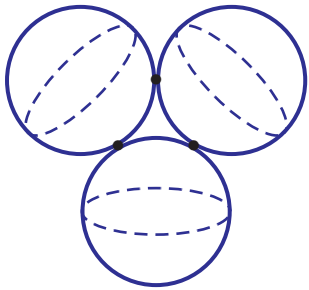


Therefore, by mildly perturbing the specialized, non-smooth conic, we find that topologically a smooth conic (those in this exercise for which $\lambda \neq 0$) is a sphere with no holes, which agrees with our work in Section 1.7.

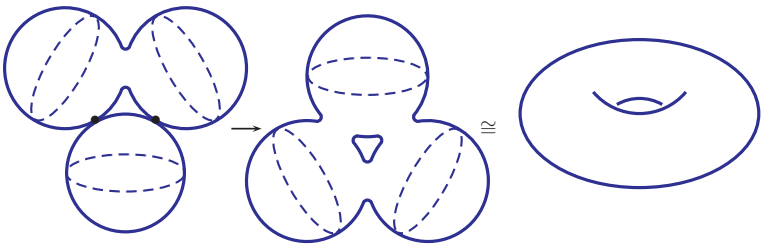
Following this same reasoning, we find another proof that a smooth cubic must be a torus when realized as a surface over $\mathbb{R}$. We begin with the highly degenerate cubic, $f(x, y, z) = (a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z)(a_3x + b_3y + c_3z)$. In the real affine chart $z = 1$, the picture looks like



Again, our picture isn't valid over $\mathbb{C}$ in $\mathbb{P}^2$. Instead, the correct topological picture is that of three spheres meeting at three points, as shown.



Perturbing the top two spheres slightly, we find they join into the topological equivalent of a single sphere, but that this new figure is joined to the third sphere at two points of contact. Perturbing each of these points of intersection independently of one another, we obtain a single surface with a hole through the middle as depicted in the sequence of figures below.



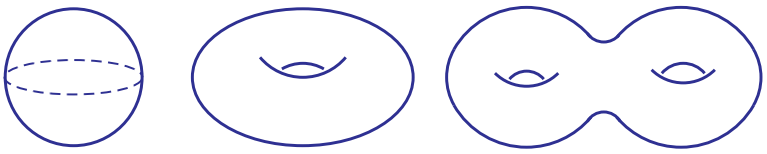
Thus a smooth cubic over $\mathbb{C}$ is topologically equivalent to a torus (a sphere with a hole through it) as a surface over $\mathbb{R}$. Note that this agrees with our results in Section 2.9.

Exercise 3.2.2. Mimic the arguments illustrated above to describe the real surface corresponding to a smooth quartic (fourth degree) curve over $\mathbb{C}$ in $\mathbb{P}^2$. Start with a highly degenerate quartic (the product of four pairwise non-parallel lines), draw the corresponding four spheres, and deform this surface by merging touching spheres two at a time. How many holes will the resulting figure possess?

Now do the same for a smooth quintic (fifth degree) curve. How many holes must it have?

3.2.2. Genus of a Curve. The number of holes in the real surfaces corresponding to smooth conics, cubics, quartics, and quintics is a topological invariant of these curves. That is, every smooth conic is topologically equivalent to a real sphere with no holes. Every smooth cubic is topologically equivalent to a real torus (a sphere with exactly one hole through it), every smooth quartic is equivalent to a sphere with three holes, and every smooth quintic to a sphere with six holes. Therefore, all smooth conics are topologically equivalent to one another, all smooth cubics are topologically equivalent, and so on, and each equivalence class is completely determined by the number of holes in the associated real surface.

Definition 3.2.1. Let $\mathrm{V}(P)$ be a smooth, irreducible curve in $\mathbb{P}^2(\mathbb{C})$. The number of holes in the corresponding real surface is called the *topological genus* of the curve $\mathrm{V}(P)$.



topological genus = 0 | topological genus = 1 | topological genus = 2

Presently, this notion of genus makes sense only when we are thinking of our complex curves as surfaces over the reals. We see that there is a connection between the genus g and the degree d of a

curve. That is, all smooth curves of degree d have the same genus, so we now wish to find a formula expressing the genus as a function of the degree.

Exercise 3.2.3. Find a quadratic function in d, the degree of a smooth curve, that agrees with the topological genus of curves of degrees $d = 2, 3, 4$ found earlier. Now use this formula to compute the genus of a smooth quintic (fifth degree) curve. Does it match your answer to the last exercise?

Definition 3.2.2. Let $\mathrm{V}(P)$ be a smooth curve of degree d. The number $\dfrac{(d-1)(d-2)}{2}$ is the *arithmetic genus* of the curve, which is an algebraic invariant of $\mathrm{V}(P)$.

Exercise 3.2.4. Argue by induction on d, the degree, that the topological genus agrees with the arithmetic genus for smooth curves, or in other words that

$$g = \frac{(d-1)(d-2)}{2}.$$

It is a theorem that the topological genus and the arithmetic genus do agree with one another whenever both are defined and make sense. However, the arithmetic version is independent of base field and enables us to exploit the genus of curves even over finite fields in positive characteristic.

3.3. Bézout's Theorem

The goal of this section is to develop the needed definitions that allow the statement and proof of Bézout's Theorem, which says that in $\mathbb{P}^2$ a curve of degree n will intersect a curve of degree m in exactly nm points, provided the points of intersection are "counted correctly."

3.3.1. Intuition behind Bézout's Theorem. We look at how many points a straight line will intersect a conic in $\mathbb{P}^2$. Both the need to work in the complex projective plane $\mathbb{P}^2$ and the need to define intersection numbers correctly will become apparent.

Exercise 3.3.1. Show that the line $\mathrm{V}(x-y)$ will intersect the circle $\mathrm{V}(x^2+y^2-1)$ in two points in the real plane, $\mathbb{R}^2$.

Exercise 3.3.2. Show that the line $\mathrm{V}(x - y + 10)$ will not intersect $\mathrm{V}(x^2 + y^2 - 1)$ in $\mathbb{R}^2$ but will intersect $\mathrm{V}(x^2 + y^2 - 1)$ in two points in $\mathbb{C}^2$.

The last exercise demonstrates our need to work over the complex numbers. Now we must demonstrate the need for projective space.

Exercise 3.3.3. Show that the two lines $\mathrm{V}(x-y+2)$ and $\mathrm{V}(x-y+3)$ do not intersect in $\mathbb{C}^2$. Homogenize both polynomials and show that they now intersect at a point at infinity in $\mathbb{P}^2$.

Exercise 3.3.4. Show that $\mathrm{V}(y - \lambda)$ will intersect $\mathrm{V}(x^2 + y^2 - 1)$ in two points in $\mathbb{C}^2$, unless $\lambda = \pm 1$. Show that $\mathrm{V}(y - 1)$ and $\mathrm{V}(y + 1)$ are tangent lines to the circle $\mathrm{V}(x^2+y^2-1)$ at their respective points of intersection. Explain why it is reasonable to say that $\mathrm{V}(y - 1)$ intersects the circle $\mathrm{V}(x^2+y^2-1)$ in one point with multiplicity two.

Exercise 3.3.5. Show that the line $\mathrm{V}(y-\lambda x)$ will intersect the curve $\mathrm{V}(y - x^3)$ in three points in $\mathbb{C}^2$, unless $\lambda = 0$. Letting $\lambda = 0$, show that $\mathrm{V}(y)$ will intersect the curve $\mathrm{V}(y - x^3)$ in only one point in $\mathbb{C}^2$. Explain why we might say that $\mathrm{V}(y)$ intersects $\mathrm{V}(y-x^3)$ in one point with multiplicity three.

Exercise 3.3.6. Show that there are no points in $\mathbb{C}^2$ in the intersection of $\mathrm{V}(xy - 1)$ with $\mathrm{V}(y)$. Homogenize both equations $xy = 1$ and $y = 0$. Show that there is a point of intersection at infinity. Explain why we might say that $\mathrm{V}(xy - 1)$ will intersect $\mathrm{V}(y)$ in one point at infinity with multiplicity two.

3.3.2. Fundamental Theorem of Algebra. Polynomials have roots. Much of the purpose of high school algebra is the exploration of this fact. The need for complex numbers stems from our desire to have all possible roots for polynomials.

In this subsection we briefly review the Fundamental Theorem of Algebra. The exercises will lead us to the realization that such a generalization requires a precise definition of the multiplicity of a point of intersection.

Consider a polynomial $f(x)$ with real coefficients. Of course, the number of real roots of f is less than or equal to the degree of f, with

equality in the case that f can be written as a product of distinct linear factors over $\mathbb{R}$.

Exercise 3.3.7. Give examples of second degree polynomials in $\mathbb{R}[x]$ that have zero, one, and two distinct real roots, respectively.

Exercise 3.3.8. Find the complex roots of your first example.

The moral of the preceding exercises is that by considering complex roots, and defining multiplicity appropriately, we can make a uniform statement about the number of roots of a polynomial.

Definition 3.3.1. Let $f(x)$ be a polynomial in $\mathbb{C}[x]$. If $f(x) = (x - a)^m g(x)$, $m > 0$, such that $(x - a)$ does not divide $g(x)$, then we say that the *multiplicity of the root* a of $f(x)$ is m.

Theorem 3.3.9 (Fundamental Theorem of Algebra)**.** If $f(x)$ is a polynomial of degree d in $\mathbb{C}[x]$, then

$$f(x) = c(x - a_1)^{m_1}(x - a_2)^{m_2} \cdots (x - a_r)^{m_r},$$

where c is a nonzero constant, each a_i is a complex root of multiplicity m_i and $\sum_{i=1}^{r} m_i = d$.

Another way of stating this theorem is that the graph of $y = f(x)$ in $\mathbb{C}^2$ intersects the complex line $y = 0$ in d points, counted with multiplicity. A natural generalization of this would be to consider the intersection of a curve defined by $f(x, y) = 0$, where f is a degree d polynomial in $\mathbb{C}[x, y]$, and a line defined by $ax + by + c = 0$.

Exercise 3.3.10. Let $f(x, y) = x^2 - y^2 - 1$ and $g(x, y) = x$. Sketch $\mathrm{V}(f)$ and $\mathrm{V}(g)$ in $\mathbb{R}^2$. Do they intersect? Find $\mathrm{V}(f) \cap \mathrm{V}(g)$ in $\mathbb{C}^2$.

Exercise 3.3.11. Let $g(x, y) = ax + by + c$, $b \neq 0$, in $\mathbb{C}[x, y]$. Let $f(x, y) = \sum_i a_i x^{r_i} y^{s_i}$ be any polynomial of degree d in $\mathbb{C}[x, y]$. Show that the number of points in $\mathrm{V}(f) \cap \mathrm{V}(g)$ is d, if the points are counted with an appropriate notion of multiplicity. [Hint: Substitute $y = \dfrac{-ax - c}{b}$ into $f = 0$, so that $f = 0$ becomes a polynomial equation of degree d in the single variable x. Apply the Fundamental Theorem of Algebra.]

What about the intersection of two curves? To answer this question we will need a more general definition of multiplicity—one that is inspired by the previous exercise, and for the most uniform statement we will need to consider curves in the complex projective plane.

3.3.3. Intersection Multiplicity. The goal of this section is to understand Bézout's Theorem on the number of points in the intersection of two plane curves. The statement of Bézout's Theorem requires the definition of the intersection multiplicity of a point p in the intersection of two plane curves, which is the goal of this subsection. Here we will present an axiomatic development for intersection multiplicity, but we will eventually show that calculating the intersection multiplicity can be reduced to applying the Fundamental Theorem of Algebra for a one-variable polynomial.

The following theorem establishes the existence of a well-behaved intersection multiplicity. We will not prove this theorem. (The proof is certainly not beyond the scope of this text; it would take a number of pages and problems to actually prove it, though.) The statement of this theorem and our treatment of it closely follows that of Kirwan [**Kir92**] and, to a lesser extent, Fulton [**Ful69**].

Theorem 3.3.12 (Intersection Multiplicity)**.** Given polynomials f and g in $\mathbb{C}[x, y]$ and a point p in $\mathbb{C}^2$, there is a uniquely defined number $I(p, \mathrm{V}(f) \cap \mathrm{V}(g))$ such that the following axioms are satisfied.

(1) $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) \in \mathbb{Z}_{\geq 0}$, unless p lies on a common component of $\mathrm{V}(f)$ and $\mathrm{V}(g)$, in which case $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = \infty$.

(2) $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = 0$ if and only if $p \notin \mathrm{V}(f) \cap \mathrm{V}(g)$.

(3) Two distinct lines meet with intersection number one at their common point of intersection.

(4) $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = I(p, \mathrm{V}(g) \cap \mathrm{V}(f))$.

(5) $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = \sum r_i s_j I(p, \mathrm{V}(f_i) \cap \mathrm{V}(g_j))$ when $f = \prod f_i^{r_i}$ and $g = \prod g_j^{s_j}$.

(6) $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = I(p, \mathrm{V}(f) \cap \mathrm{V}(g + af))$ for all $a \in \mathbb{C}[x, y]$.

Definition 3.3.2. The number $I(p, \mathrm{V}(f) \cap \mathrm{V}(g))$ is the *intersection multiplicity* of f and g at p.

We can easily extend this definition to curves in $\mathbb{P}^2$, by dehomogenizing the projective curves, making them into curves in $\mathbb{C}^2$ containing the point in question.

Exercise 3.3.13. Use the above axioms to show that for $p = (0, 0)$,

$$I(p, \mathrm{V}(x^2) \cap \mathrm{V}(y)) = 2.$$

Sketch $\mathrm{V}(x^2)$ and $\mathrm{V}(y)$.

Exercise 3.3.14. Show for $p = (0, 0)$,

$$I(p, \mathrm{V}(x^2 - y) \cap \mathrm{V}(y)) = 2.$$

Sketch $\mathrm{V}(x^2 - y)$ and $\mathrm{V}(y)$.

Exercise 3.3.15. Show for $p = (0, 0)$,

$$I(p, \mathrm{V}(y^2 - x^2 - x^3) \cap \mathrm{V}(x)) = 2.$$

Sketch $\mathrm{V}(y^2 - x^2 - x^3)$ and $\mathrm{V}(x)$.

Finally, we want to see how the intersection multiplicity varies, or more accurately doesn't vary, under change of coordinates.

Exercise 3.3.16. Show that for any polynomials f and g in $\mathbb{C}[x, y]$ and a point p in $\mathbb{C}^2$, for any affine change of coordinates T we have

$$I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = I(T(p), \mathrm{V}(T^{-1}f) \cap \mathrm{V}(T^{-1}g)).$$

[Hint: This problem is actually not that hard. Its solution involves little or no calculations.]

3.3.4. Multiplicity of a Curve at a Point. In this subsection we give the definition of the multiplicity of a point on a curve. The first step in this direction is to generalize the idea of multiplicity of a root. Besides being of independent interest, the multiplicity of a point on a curve will help us in easily calculating intersection multiplicities.

In order to have a rigorous definition for the multiplicity of a point on a curve $\mathrm{V}(P)$, we will need to review multivariable Taylor series expansions.

Exercise 3.3.17. Show that

$$P(x,y) = 5 - 8x + 5x^2 - x^3 - 2y + y^2$$

is equal to

$$(y-1)^2 - (x-2)^2 - (x-2)^3$$

by directly expanding the second polynomial. Now, starting with $P(x,y) = 5 - 8x + 5x^2 - x^3 - 2y + y^2$, calculate its Taylor series expansion at the point $(2,1)$:

$$\sum_{n,m=0}^{\infty} \frac{1}{n!m!} \frac{\partial^{n+m} P}{\partial x^n \partial y^m}(2,1)(x-2)^n(y-1)^m,$$

which is

$$P(2,1) + \frac{\partial P}{\partial x}(2,1)(x-2) + \frac{\partial P}{\partial y}(2,1)(y-1) + \frac{1}{2}\frac{\partial^2 P}{\partial x^2}(2,1)(x-2)^2 + \dots.$$

Definition 3.3.3. Let f be a nonhomogeneous polynomial (in any number of variables) and let p be a point in the set $\mathrm{V}(f)$. The *multiplicity of f at p*, denoted $m_p f$, is the degree of the lowest degree nonzero term of the Taylor series expansion of f at p.

Notice that if $p \notin \mathrm{V}(f)$, then $m_p f = 0$, since the lowest degree nonzero term of the Taylor expansion of f at p is $f(p) \neq 0$, which has degree zero. If $p \in \mathrm{V}(f)$, then $f(p) = 0$, so $m_p f$ must be at least one.

Exercise 3.3.18. Let f be a nonhomogeneous polynomial (in any number of variables) of degree n.

(1) Show that $m_p f = 1$ if and only if p is a nonsingular point. Hence, $m_p(f) = 1$ for every point $p \in \mathrm{V}(f)$ if and only if $\mathrm{V}(f)$ is nonsingular.

(2) Show that $m_p f \leq n$ for all $p \in \mathrm{V}(f)$. Hence, $1 \leq m_p f \leq n$ for all $p \in \mathrm{V}(f)$.

Exercise 3.3.19. Let $f(x,y) = xy$. What is the multiplicity of f at the origin? Let $p = (0,1)$, and calculate $m_p f$.

Exercise 3.3.20. Let $f(x,y) = x^2 + xy - 1$. Calculate the multiplicity of f at $p = (1,0)$.

We are interested in curves in the complex projective plane, $\mathbb{P}^2$, and hence in zero sets of homogeneous polynomials. By considering appropriate affine patches, we can apply our definition in this case.

Exercise 3.3.21. Consider the homogeneous polynomial

$$P(x, y, z) = zy^2 - (x - z)^3.$$

We want to show that the point $(1 : 0 : 1) \in \mathrm{V}(P)$ has multiplicity two, no matter how P is dehomogenized. Show when we dehomogenize by setting $z = 1$, that the point $x = 1, y = 0$ has multiplicity two for $P(x, y, 1)$. Then show when we dehomogenize by setting $x = 1$, that the point $y = 0, z = 1$ has multiplicity two for $P(1, y, z)$.

Exercise 3.3.22. Let $(a : b : c) \in \mathrm{V}(f)$. Show that the multiplicity of f at the point $(a : b : c)$ remains the same no matter how we dehomogenize. (This is quite a long problem to work out in full detail.)

The following theorem links intersection multiplicity of two plane curves at a point p with the multiplicity of the point for each of the curves.

Theorem 3.3.23. Given polynomials f and g in $\mathbb{C}[x, y]$ and a point p in $\mathbb{C}^2$, we have

$$I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) \geq m_p(f) \cdot m_p(g),$$

with equality if and only if $\mathrm{V}(f)$ and $\mathrm{V}(g)$ have no common tangent at p.

(We will not give a proof of this result; once we know how to compute intersection multiplicity from resultants, which is the goal of the next few sections, we could indeed show that it is true.)

3.3.5. Statement of Bézout's Theorem. Bézout's Theorem tells us how many points are in the intersection of two plane curves. We start with some examples.

Exercise 3.3.24. Let $f(x, y) = x^2 + y^2 - 1$. Give examples of real polynomials

$$g(x, y) = ax + by + c$$

such that

$$\mathrm{V}(x^2+y^2-1)\cap\mathrm{V}(ax+by+c)$$

in $\mathbb{R}^2$ has zero, one or two points, respectively. Now consider the intersections $\mathrm{V}(f)\cap\mathrm{V}(g)$ in $\mathbb{C}^2$. In each of your three examples, find these points of intersection, calculate their multiplicities, and verify that $\sum_p I(p,\mathrm{V}(f)\cap\mathrm{V}(g)) = (\deg f)(\deg g)$. [Hint: To help with the intersection multiplicity, use Theorem 3.3.23.]

Exercise 3.3.25. Let

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

and

$$g = x^2-y^2-1.$$

Find all points of intersection of the curves $\mathrm{V}(f)$ and $\mathrm{V}(g)$. For each point of intersection p, send p to $(0,0)$ via a change of coordinates T. Find $I(p,\mathrm{V}(f)\cap\mathrm{V}(g))$ by calculating $I((0,0),T(\mathrm{V}(f))\cap T(\mathrm{V}(g)))$. Verify that

$$\sum_p I(p,\mathrm{V}(f)\cap\mathrm{V}(g)) = (\deg f)(\deg g).$$

Exercise 3.3.26. Let

$$f = y - x(x-2)(x+2)$$

and

$$g = y - x.$$

Find all points of intersection of the curves $\mathrm{V}(f)$ and $\mathrm{V}(g)$. At each point of intersection p, show that the curves have distinct tangent lines. Using Theorem 3.3.23 from the previous section, find $I(p,\mathrm{V}(f)\cap\mathrm{V}(g))$. Verify that

$$\sum_p I(p,\mathrm{V}(f)\cap\mathrm{V}(g)) = (\deg f)(\deg g).$$

The previous exercises may have led you to conjecture that if f and g are any polynomials, then $\sum_p I(p,\mathrm{V}(f)\cap\mathrm{V}(g)) = (\deg f)(\deg g)$. This is not true for all curves $\mathrm{V}(f)$ and $\mathrm{V}(g)$ in $\mathbb{C}^2$ (though it will be true in $\mathbb{P}^2$), as the next exercise illustrates.

Exercise 3.3.27. Let $f = y - x^2$ and $g = x$. Verify that the origin is the only point of $\mathrm{V}(f) \cap \mathrm{V}(g)$ in $\mathbb{C}^2$ and that $I((0,0), \mathrm{V}(f) \cap \mathrm{V}(g)) = 1$.

We would like the number of points of intersection of two curves to be the product of their degrees. Unfortunately, we have seen in the previous exercise that this is not true in the affine plane. The corresponding curves in the projective plane, however, will always intersect in the "correct" number of points. This is Bézout's Theorem, which we will prove later in this section.

Theorem 3.3.28 (Bézout's Theorem)**.** Let f and g be homogeneous polynomials in $\mathbb{C}[x, y, z]$ with no common factors, and let $\mathrm{V}(f)$ and $\mathrm{V}(g)$ be the corresponding curves in $\mathbb{P}^2(\mathbb{C})$. Then

$$\sum_{p \in \mathrm{V}(f) \cap \mathrm{V}(g)} I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = (\deg f)(\deg g).$$

Exercise 3.3.29. Homogenize the polynomials in Exercise 3.3.27, and find the two points of $\mathrm{V}(f) \cap \mathrm{V}(g)$ in $\mathbb{P}^2(\mathbb{C})$. Check that Bézout's Theorem holds.

Exercise 3.3.30. Let $f = x^2 - y^2 - 1$ and $g = x - y$. Sketch $\mathrm{V}(f)$ and $\mathrm{V}(g)$ in $\mathbb{R}^2$. Homogenize f and g and verify Bézout's Theorem in this case. Describe the relationship between the points of intersection in $\mathbb{P}^2(\mathbb{C})$ and the sketch in $\mathbb{R}^2$. Repeat this exercise with $g = x + y$.

Exercise 3.3.31. Confirm that the curves defined by $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$ do not intersect in $\mathbb{C}^2$. Homogenize these equations and confirm Bézout's Theorem in this case. Would a sketch of the circles in $\mathbb{R}^2$ give you any insight into the intersections in $\mathbb{P}^2(\mathbb{C})$?

3.3.6. Resultants. The goal of this subsection is to define the resultant of two polynomials, which will be the main tool in our proof of Bézout's Theorem.

While the Fundamental Theorem of Algebra tells us that a one-variable polynomial of degree d has exactly d roots, counting multiplicities, it gives us no means for actually finding these roots. Similarly, what if we want to know if two one-variable polynomials have a common root? The most naive method would be to find the roots for each of the polynomials and see if any of the roots are the same. In

practice, though, this method is quite difficult to implement, since we have no easy way for finding these roots. The resultant is a totally different approach for determining if the polynomials share a root. The resultant is the determinant of a matrix; this determinant will be zero precisely when the two polynomials have a common root.

Definition 3.3.4. The *resultant* $\mathrm{Res}(f,g)$ of two polynomials

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

and

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$$

is defined to be the determinant of the $(m+n)\times(m+n)$ matrix

$$\begin{pmatrix} a_n & a_{n-1} & \cdots & a_0 & 0 & 0 & \cdots & 0 \\ 0 & a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & 0 & 0 & \cdots & 0 \\ 0 & b_m & b_{m-1} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & b_{m-1} & \cdots & \cdots & \cdots & b_0 \end{pmatrix},$$

where there are m rows of a's and n rows of b's.

For example, if $f(x) = a_2x^2 + a_1x + a_0$ and $g(x) = b_4x^4 + b_3x^3 + b_2x^2 + b_1x + b_0$, then

$$\mathrm{Res}(f,g) = \det \begin{pmatrix} a_2 & a_1 & a_0 & 0 & 0 & 0 \\ 0 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & 0 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & 0 & a_2 & a_1 & a_0 \\ b_4 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & b_4 & b_3 & b_2 & b_1 & b_0 \end{pmatrix}.$$

An important property of resultants is that $f(x)$ and $g(x)$ have a common root if and only if $\mathrm{Res}(f,g) = 0$. The following three exercises will illustrate this property. You will then prove this result.

Exercise 3.3.32. Let $f(x) = x^2 - 1$ and $g(x) = x^2 + x - 2$.

(1) Find the roots of f and g and show that they share a root.

(2) Show that $\mathrm{Res}(f,g) = 0$.

Exercise 3.3.33. Let $f(x) = x^2 - 1$ and $g(x) = x^2 - 4$.

(1) Find the roots of f and g and show that they have no roots in common.

(2) Show that $\mathrm{Res}(f,g) \neq 0$.

Exercise 3.3.34.

(1) Let $f(x) = x - r$ and $g(x) = x - s$. Find $\mathrm{Res}(f,g)$. Verify that $\mathrm{Res}(f,g) = 0$ if and only if $r = s$.

(2) Let $f(x) = x - r$ and $g(x) = (x - s_1)(x - s_2)$. Find $\mathrm{Res}(f,g)$. Verify that $\mathrm{Res}(f,g) = 0$ if and only if $r = s_1$ or $r = s_2$.

Exercise 3.3.35. For a degree two polynomial $f(x) = a_2x^2 + a_1x + a_0 = a_2(x - r_1)(x - r_2)$, we have

$$\begin{aligned} \frac{a_1}{a_2} &= -(r_1 + r_2) \\ \frac{a_0}{a_2} &= r_1 r_2. \end{aligned}$$

Use these relations between the coefficients and roots to show that if

$$\begin{aligned} f(x) &= a_2x^2 + a_1x + a_0 = a_2(x - r_1)(x - r_2) \\ g(x) &= b_2x^2 + b_1x + b_0 = b_2(x - s_1)(x - s_2), \end{aligned}$$

then

$$\mathrm{Res}(f,g) = a_2^2 b_2^2 (r_1 - s_1)(r_1 - s_2)(r_2 - s_1)(r_2 - s_2).$$

Exercise 3.3.36. Let $f(x,y) = x^2 + y^2 - 2$ and $g(x,y) = x^2 - xy + y^2 + y - 2$.

(1) Treating f and g as polynomials in x, compute

$$R(y) = \mathrm{Res}(f,g;x) = \det \begin{pmatrix} 1 & 0 & y^2 - 2 & 0 \\ 0 & 1 & 0 & y^2 - 2 \\ 1 & -y & y^2 + y - 2 & 0 \\ 0 & 1 & -y & y^2 + y - 2 \end{pmatrix}$$

(2) Set $R(y) = 0$ and solve for y to find the projections on the y-axis of the points of intersection of $\mathrm{V}(f)$ and $\mathrm{V}(g)$.

Exercise 3.3.37. The two lines $\mathrm{V}(x-y)$ and $\mathrm{V}(x-y+2)$ are parallel in the affine plane, but intersect at $(1:1:0)$ in $\mathbb{P}^2$. Treating $f(x,y,z) = x-y$ and $g(x,y,z) = x-y+2z$ as one-variable polynomials in x, show that $\mathrm{Res}(x-y, x-y+2z; x) = 0$ when $z = 0$.

Exercise 3.3.38. Let $f(x,y) = 4x-3y$ and $g(x,y) = x^2+y^2-25$. Use the resultant $\mathrm{Res}(f,g;x)$ to find the points of intersection of $\mathrm{V}(f)$ and $\mathrm{V}(g)$.

Exercise 3.3.39. Let $f(x) = ax^2+bx+c$.

(1) Find $\mathrm{Res}(f, f')$.

(2) Under what conditions will $\mathrm{Res}(f, f') = 0$?

We are now ready to prove that two polynomials f and g have a common root if and only if $\mathrm{Res}(f,g) = 0$.

Exercise 3.3.40. Show that if r is a common root of f and g, then the vector $\mathbf{x} = \begin{pmatrix} r^{m+n-1} \\ r^{m+n-2} \\ \vdots \\ r \\ 1 \end{pmatrix}$ is in the null space of the resultant matrix of f and g, and thus $\mathrm{Res}(f,g) = 0$.

Exercise 3.3.41 (from Kirwan, *Complex Algebraic Curves* [**Kir92**], Lemma 3.3, p. 67)**.** Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ and $g(x) = b_mx^m + b_{m-1}x^{m-1} + \cdots + b_1x + b_0$, with $a_n \neq 0$ and $b_m \neq 0$.

(1) Prove that f and g have a common root $x = r$ if and only if there exists a polynomial $p(x)$ of degree $m-1$ and a polynomial $q(x)$ of degree $n-1$ such that $p(x)f(x) = q(x)g(x)$.

(2) Write $p(x) = \alpha_{m-1}x^{m-1} + \cdots + \alpha_1x + \alpha_0$ and $q(x) = \beta_{n-1}x^{n-1} + \cdots + \beta_1x + \beta_0$. By comparing coefficients, show that the polynomial equation $p(x)f(x) = q(x)g(x)$ corresponds to the linear system in the α_i and β_j

$$\begin{aligned} \alpha_{m-1}a_n &= \beta_{n-1}b_m \\ \alpha_{m-1}a_{n-1} + \alpha_{m-2}a_n &= \beta_{n-1}b_{m-1} + \beta_{n-2}b_m \\ &\vdots \\ \alpha_0a_0 &= \beta_0b_0. \end{aligned}$$

(3) Prove that this system of equations has a nonzero solution

$$(\alpha_{m-1}, \alpha_{m-2}, \ldots, \alpha_0, \beta_{n-1}, \beta_{n-2}, \ldots, \beta_0)$$

if and only if $\operatorname{Res}(f, g) = 0$.

(This solution is quite a bit longer than most of the other problems.)

We have shown for $f(x) = x - r$ and $g(x) = x - s$ that

$$\operatorname{Res}(f, g) = r - s$$

and for $f(x) = x - r$ and $g(x) = (x - s_1)(x - s_2)$ that

$$\operatorname{Res}(f, g) = (r - s_1)(r - s_2).$$

We want to show in the next series of exercises that if

$$f(x) = a_n(x - r_1) \cdots (x - r_n)$$

and

$$g(x) = b_m(x - s_1) \cdots (x - s_m),$$

then

$$\operatorname{Res}(f, g) = a_n^m b_m^n \prod_{i=1, j=1}^{i=n, j=m} (r_i - s_j).$$

One technical point first: if $a_n \neq 0$, then the roots of

$$f(x) = a_n x^n + \cdots + a_0$$

are the same as the roots of

$$\tilde{f}(x) = x^n + \left(\frac{a_{n-1}}{a_n}\right) x_{n-1} + \cdots + \frac{a_0}{a_n}.$$

Exercise 3.3.42. If

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1} x^{n-1} \cdots + a_0 \\ g(x) &= b_m x^m + \cdots + b_0 \\ \tilde{f}(x) &= x^n + \left(\frac{a_{n-1}}{a_n}\right) x^{n-1} + \cdots + \frac{a_0}{a_n} \\ \tilde{g}(x) &= x^m + \left(\frac{b_{m-1}}{b_m}\right) x^{m-1} + \cdots + \frac{b_0}{b_m}, \end{aligned}$$

then show

$$\operatorname{Res}(f, g) = a_n^m b_m^n \operatorname{Res}(\tilde{f}, \tilde{g}).$$

We will use this to show that if

$$f(x) = (x - r_1) \cdots (x - r_n)$$

and

$$g(x) = (x - s_1) \cdots (x - s_m),$$

then $\text{Res}(f, g) = \prod(r_i - s_j)$.

First we recall the relationship between the roots of a polynomial and its coefficients.

Exercise 3.3.43. Let r_1, r_2 and r_3 be the three roots of the cubic

$$f(x) = x^3 + a_2x^2 + a_1x + a_0.$$

Show that

$$\begin{aligned} a_2 &= -(r_1 + r_2 + r_3) \\ a_1 &= r_1r_2 + r_1r_3 + r_2r_3 \\ a_0 &= -r_1r_2r_3. \end{aligned}$$

Exercise 3.3.44. Let $r_1, \ldots, r_n$ be the roots of

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

Show that

$$\begin{aligned} a_{n-1} &= -(r_1 + \cdots + r_n) \\ a_{n-2} &= r_1r_2 + r_1r_3 + \cdots + r_1r_n + r_2r_3 + \cdots + r_{n-1}r_n \\ &\vdots \\ a_{n-k} &= (-1)^k \sum_{i_1 < \cdots < i_k} r_{i_1} \cdots r_{i_k} \\ &\vdots \\ a_0 &= (-1)^n r_1 \cdots r_n. \end{aligned}$$

(If the polynomial has a multiple root, say of degree k, we list this root k times. For example, the roots of $(x - 3)^2(x - 5)$ are listed as $3, 3, 5$.)

We can describe a polynomial $f(x)$ either via its coefficients

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

or by its roots

$$f(x) = (x - r_1) \cdots (x - r_n).$$

We say that the n-tuple $(a_1, \ldots, a_{n-1})$ is in the coefficient space while the roots $(r_1, \ldots, r_n)$ are in the root space. The above exercise shows that there is an easy map

$$\text{Root space} \to \text{Coefficient Space}.$$

Much of the difficulty in algebra and algebraic geometry lies with the inverse map

$$\text{Coefficient Space} \to \text{Root Space}.$$

A large part of high school algebra is the development of the quadratic equation, which is simply the map from the coefficient space of a second degree polynomial to its two roots:

$$(a_1, a_2) \to \left(\frac{-a_1 + \sqrt{a_1^2 - 4a_2}}{2}, \frac{-a_1 - \sqrt{a_1^2 - 4a_2}}{2} \right).$$

Back to resultants: Given polynomials $f(x)$ and $g(x)$, the resultant is the determinant of a matrix whose entries are coefficients of f and g and is hence a polynomial of the coefficients. Since the coefficients are in turn polynomials of the roots of f and g, we can write the resultant as a polynomial in the roots $r_1, \ldots, r_n$ of f and the roots $s_1, \ldots s_m$ of g, namely as

$$\text{Res}(r_1, \ldots r_n, s_1, \ldots, s_m).$$

Exercise 3.3.45. Show that $(r_i - s_j)$ divides $\text{Res}(r_1, \ldots r_n, s_1, \ldots, s_m)$.

Exercise 3.3.46. If f has degree n and g has degree m, show that

$$\prod (r_i - s_j)$$

has degree m as a polynomial in r_i and degree n as a polynomial in s_j.

Exercise 3.3.47. For

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

with roots $r_1, \ldots, r_n$ and for

$$g(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$$

with roots $s_1, \ldots, s_m$, show that $\mathrm{Res}(r_1, \ldots r_n, s_1, \ldots, s_m)$ has degree m as a polynomial in r_i and degree n as a polynomial in s_j.

Exercise 3.3.48. With the notation of the previous exercise, show that there is a nonzero constant λ such that

$$\mathrm{Res}(f, g) = \lambda \prod (r_i - s_j).$$

The goal of the next two exercises is to show that this constant λ equals 1.

Exercise 3.3.49. For

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

and

$$g(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0,$$

show that the highest power of a_0 in the resultant is m with leading coefficient $(-1)^{nm}$.

Exercise 3.3.50. Show that the highest power of $r_1, \cdots, r_n$ in $\prod(r_i - s_j)$ is m, with leading coefficient one. Conclude for

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

and

$$g(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$$

that

$$\mathrm{Res}(f, g) = \prod (r_i - s_j).$$

Exercise 3.3.51. For

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

with roots $r_1, \ldots, r_n$ and $a_n \neq 0$ and for

$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$$

with roots $s_1, \ldots, s_m$ and $b_m \neq 0$, show that

$$\mathrm{Res}(f, g) = a_n^m b_m^n \prod_{i=1,j=1}^{i=n,j=m} (r_i - s_j).$$

Exercise 3.3.52. Suppose $f(x) = f_1(x)f_2(x)$ and $g(x)$ is any other polynomial. Show that

$$\mathrm{Res}(f, g) = \mathrm{Res}(f_1, g) \cdot \mathrm{Res}(f_2, g).$$

3.3.7. Proof of Bézout's Theorem. Now we are ready to outline a proof of Bézout's Theorem. Full details can be found in Cox, Little, and O'Shea's *Ideals, Varieties, and Algorithms* [**CLO07**], Chapter 8, Section 7.

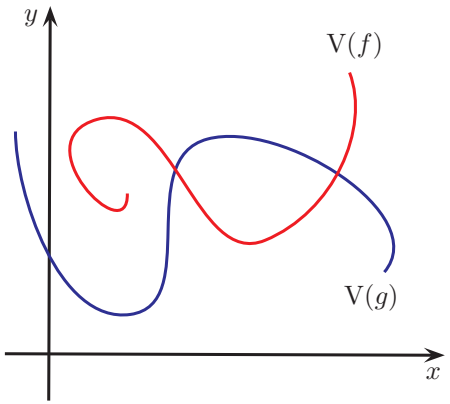
We start by linking resultants to intersection points of curves.

Exercise 3.3.53. Let $f(x, y, z) = 3x + y + 2z$ and $g(x, y, z) = x + 5z$. Show that $\mathrm{Res}(f, g; z)$ is a homogeneous polynomial in x and y of degree 1.

Exercise 3.3.54. Letting $f(x, y, z) = 3x + y + 2z$ and $g(x, y, z) = x + 5z$, show that $\mathrm{V}(f) \cap \mathrm{V}(g)$ contains no point with $y = 0$.

Exercise 3.3.55. Use the notation from the previous two exercises. After dehomogenizing by setting $y = 1$, sketch $\mathrm{V}(f)$ and $\mathrm{V}(g)$ in the xz-plane. Find the point of intersection $(a : 1 : b)$ of $\mathrm{V}(f) \cap \mathrm{V}(g)$. Show that $(a : 1)$ is the root of $\mathrm{Res}(f, g; z)$. Hence the zero of the resultant is the projection of the point of intersection of $\mathrm{V}(f) \cap \mathrm{V}(g)$.

This suggests the following approach for understanding intersections of curves. Consider two curves $\mathrm{V}(f(x, y, z))$ and $\mathrm{V}(g(x, y, z))$:



The resultant $\mathrm{Res}(f, g; z)$, calculated with respect to the variable z, is a homogeneous polynomial in the variables x and y. The zeros of this polynomial will be the projections of the points in the intersection of $\mathrm{V}(f) \cap \mathrm{V}(g)$ along the z-axis to the projective line $\mathbb{P}^1$ with homogeneous coordinates x, y. This will allow us to translate questions about points of intersections of plane curves to questions about

the roots of one-variable (or two-variable homogeneous) polynomials. The next few exercises show how we do this. We will then use the resultant to determine the intersection multiplicity, which in turn will provide us with the tools to prove Bézout's Theorem.

Exercise 3.3.56. Let $f(x, y, z) = x^2 + y^2 + z^2$ and $g(x, y, z) = 2x + 3y - z$. Show that $\text{Res}(f, g; z)$ is a homogeneous polynomial of degree 2.

Exercise 3.3.57. Let $f(x, y, z) = x^2 + xz + z^2$ and $g(x, y, z) = x^2 + y^2 + z^2$. Show that $\text{Res}(f, g; z)$ is a homogeneous polynomial of degree 4.

The next exercise is a generalization of the previous two exercises.

Exercise 3.3.58. (Cox, Little, and O'Shea [**CLO07**], Lemma 5, p. 425). Let $f, g \in \mathbb{C}[x, y, z]$ be homogeneous polynomials of degrees n and m, respectively. If $f(0, 0, 1)$ and $g(0, 0, 1)$ are nonzero, then $\text{Res}(f, g; z)$ is homogeneous of degree mn in x and y.

Exercise 3.3.59. Let $f(x, y) = x^2 - 8xy + 15y^2$. Show that $\text{V}(f) = \{(3 : 1), (5 : 1)\}$ and that $f(x, y) = (x - 3y)(x - 5y)$.

Exercise 3.3.60. Let $f(x, y) = x^2 + y^2$. Show that $\text{V}(f) = \{(i : 1), (-i : 1)\}$ and that $f(x, y) = (x + iy)(x - iy)$.

Exercise 3.3.61. Let $f(x, y) = x^3 - 5x^2y - 14xy^2$. Show that $\text{V}(f) = \{(0 : 1), (7 : 1), (-2 : 1)\}$ and that $f(x, y) = x(x + 2y)(x - 7y)$.

The previous exercises are special cases of the general result presented next.

Exercise 3.3.62. ([**CLO07**], Lemma 6, p. 427). Let $f \in \mathbb{C}[x, y]$ be homogeneous, and let $\text{V}(f) = \{(r_1 : s_1), \ldots, (r_t : s_t)\}$. Show that

$$f = c(s_1x - r_1y)^{m_1} \cdots (s_tx - r_ty)^{m_t},$$

where c is a nonzero constant. (This is actually not that hard.)

We now link intersection multiplicity with resultants. The method is a bit cumbersome. The most important part of the following is when our two curves contain no common components. Earlier, we defined intersection multiplicity axiomatically. We will now define

intersection multiplicity in terms of the resultant and show that the two definitions are equivalent.

Definition 3.3.5. Let $\mathrm{V}(f)$ and $\mathrm{V}(g)$ be curves in $\mathbb{P}^2(\mathbb{C})$ with no common components. Choose homogeneous coordinates for $\mathbb{P}^2(\mathbb{C})$ so that the point $(0:0:1)$ is not in $\mathrm{V}(f)$ or $\mathrm{V}(g)$ and is not collinear with any two points of $\mathrm{V}(f) \cap \mathrm{V}(g)$. (What follows will be independent of this choice of coordinates, though this is not obvious.) For

$$p = (u : v : w) \in \mathrm{V}(f) \cap \mathrm{V}(g),$$

define

$$I(p, \mathrm{V}(f) \cap \mathrm{V}(g))$$

to be the exponent of $(vx - uy)$ in the factorization of $\mathrm{Res}(f, g; z)$, while if

$$p \notin \mathrm{V}(f) \cap \mathrm{V}(g),$$

define

$$I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = 0.$$

If $\mathrm{V}(f)$ and $\mathrm{V}(g)$ are curves in $\mathbb{P}^2(\mathbb{C})$ with a common component and if p is an element of this common component, define

$$I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = \infty.$$

Finally, if $\mathrm{V}(f)$ and $\mathrm{V}(g)$ share a common component but p is an element not on this common component, factor out the common component and then compute the intersection multiplicity, as in the first case.

The condition that $(0:0:1)$ is not in $\mathrm{V}(f)$ or $\mathrm{V}(g)$ is needed to guarantee that

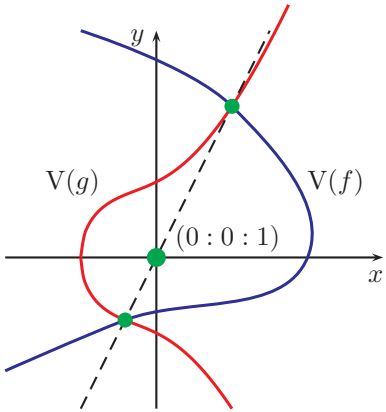
$$f(x, y, z) = az^n + \text{lower order terms with respect to } z,$$

with $a \neq 0$, and

$$g(x, y, z) = bz^m + \text{lower order terms with respect to } z,$$

with $b \neq 0$. Thus f and g, thought of as functions of z, will be of degree n and m respectively.

The condition that $(0:0:1)$ is not collinear with any two points of $\mathrm{V}(f) \cap \mathrm{V}(g)$ is to guarantee that two distinct points in the intersection will not project to the same point in the line $\mathbb{P}^1$ whose coordinates are x and y. Thus we do not want something like:



(In general, if $(0:0:1)$ is on both curves, we can do a simple change of coordinates to reduce finding the intersection points to the above.)

We want to show that this definition satisfies the axioms for the intersection multiplicity given earlier.

Exercise 3.3.63. Show that $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) \in \mathbb{Z}_{\geq 0}$, unless p lies on a common component of $\mathrm{V}(f)$ and $\mathrm{V}(g)$, in which case $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = \infty$.

Exercise 3.3.64. Show that $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = 0$ if and only if $p \notin \mathrm{V}(f) \cap \mathrm{V}(g)$.

Exercise 3.3.65. Show that two distinct lines meet with intersection number one at their common point of intersection.

Exercise 3.3.66. Show that $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = I(p, \mathrm{V}(g) \cap \mathrm{V}(f))$.

Exercise 3.3.67. Show that $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = \sum r_i s_i I(p, \mathrm{V}(f_i) \cap \mathrm{V}(g_i))$ when $f = \prod f_i^{r_i}$ and $g = \prod g_i^{s_i}$.

Exercise 3.3.68. Show that $I(p, \mathrm{V}(f) \cap \mathrm{V}(g)) = I(p, \mathrm{V}(f) \cap \mathrm{V}(g + af))$ for all homogeneous polynomials $a \in \mathbb{C}[x, y, z]$ of degree $m - n$, where g has degree m and f has degree n.

Thus the axiomatic and the resultant definitions of intersection multiplicity are equivalent.

Exercise 3.3.69. Deduce Bézout's Theorem from Exercises 3.3.58, 3.3.62, and 3.3.68.

3.4. The Ring of Regular Functions and Function Fields

The goal of this section is to associate a ring and a field to any curve. Both will encode algebraic information about the curve and play a critical role throughout algebraic geometry.

3.4.1. The Affine Case. We want to understand functions defined on a curve.

Exercise 3.4.1. Let $P(x,y) = x^2 + xy + 1$. Consider the two polynomials

$$f_1(x,y) = x^2 \quad \text{and} \quad f_2(x,y) = 2x^2 + xy + 1.$$

Find a point $(a,b) \in \mathbb{C}^2$ such that

$$f_1(a,b) \neq f_2(a,b).$$

Now show that if $(a,b) \in \mathrm{V}(P)$, then

$$f_1(a,b) = f_2(a,b).$$

To some extent, we would like to say that the polynomials f_1 and f_2 are the same as far as points on the curve $\mathrm{V}(P)$ are concerned. Why is it in the above exercise that $f_1(a,b) = f_2(a,b)$ for any point $(a,b) \in \mathrm{V}(P)$? The key is to look at $f_2(x,y) - f_1(x,y)$.

Definition 3.4.1. Let $\mathrm{V}(P)$ be an irreducible curve. Let $f(x,y)$ and $g(x,y)$ be two polynomials. We say that

$$f(x,y) \sim g(x,y)$$

if $P(x,y)$ divides $f(x,y) - g(x,y)$.

Exercise 3.4.2. Show that $\sim$ defines an equivalence relation on polynomials.

Definition 3.4.2. Let $V = \mathrm{V}(P)$ be an irreducible curve. The *ring of regular functions* on V is the set of all polynomials $f(x, y)$ modulo the equivalence relation $\sim$. Denote this ring by $\mathcal{O}(V)$. (We will also denote this by $\mathcal{O}_V$.)

You should be worried that we are calling $\mathcal{O}(V)$ a ring without proof. We shall remedy that situation now.

Exercise 3.4.3. We want to show that addition and multiplication are well-defined on $\mathcal{O}(V)$. Suppose that

$$f_1(x, y) \sim f_2(x, y) \quad \text{and} \quad g_1(x, y) \sim g_2(x, y).$$

Show that

$$f_1(x, y) + g_1(x, y) \sim f_2(x, y) + g_2(x, y),$$

which means that addition is well-defined in $\mathcal{O}(V)$. Also show

$$f_1(x, y)g_1(x, y) \sim f_2(x, y)g_2(x, y),$$

which means that multiplication is well-defined in $\mathcal{O}(V)$.

Hence for any curve V, we have the ring $\mathcal{O}(V)$ of regular functions defined on V. (Once we know the operations are well-defined, checking the ring axioms is straightforward and left as an exercise for the interested reader.)

Exercise 3.4.4. Suppose $\mathrm{V}(P)$ is an irreducible curve. Let f and g be two polynomials. Show that if $fg \sim 0$, then either $f \sim 0$ or $g \sim 0$. Conclude that the ring of regular functions on an irreducible curve is an integral domain.

There is also a field of functions associated to $\mathrm{V}(P)$. Morally this field will simply be all of the fractions formed by the polynomials in $\mathcal{O}(V)$.

Definition 3.4.3. Let the function field, $\mathcal{K}(V)$, for the curve $V = \mathrm{V}(P)$ be all rational functions

$$\frac{f(x, y)}{g(x, y)}$$

where

(1) P does not divide g (which is a way of guaranteeing that g, the denominator, is not identically zero on the curve V), and
(2) $\dfrac{f_1(x,y)}{g_1(x,y)}$ is identified with $\dfrac{f_2(x,y)}{g_2(x,y)}$ if P divides $f_1g_2 - f_2g_1$.

Exercise 3.4.5. Show that $\sim$ is an equivalence relation.

We want $\mathcal{K}(V)$ to mimic the rational numbers. Recall that the rational numbers $\mathbb{Q}$ are all the fractions

$$\frac{a}{b}$$

such that $a, b \in \mathbb{Z}$, $b \neq 0$ and $\frac{a}{b}$ is equal to $\frac{c}{d}$ if $ad - bc = 0$.

Now, you should be concerned that we are calling $\mathcal{K}(V)$ a field. We will define addition and multiplication on $\mathcal{K}(V)$ using the rational numbers as a guide.

Definition 3.4.4. On $\mathcal{K}(V)$, define addition and multiplication by

$$\frac{f(x,y)}{g(x,y)} + \frac{h(x,y)}{k(x,y)} = \frac{f(x,y)k(x,y) + g(x,y)h(x,y)}{g(x,y)k(x,y)}$$

and

$$\frac{f(x,y)}{g(x,y)} \cdot \frac{h(x,y)}{k(x,y)} = \frac{f(x,y)h(x,y)}{g(x,y)k(x,y)}.$$

In the next exercise, we will verify that these operations are well-defined.

Exercise 3.4.6. Suppose

$$\frac{f_1}{g_1} = \frac{f_2}{g_2} \text{ and } \frac{h_1}{k_1} = \frac{h_2}{k_2}$$

in $\mathcal{K}(V)$. Show that $\frac{f_1}{g_1} + \frac{h_1}{k_1}$ can be identified with $\frac{f_2}{g_2} + \frac{h_2}{k_2}$ in $\mathcal{K}(V)$. Similarly, show that $\frac{f_1}{g_1} \cdot \frac{h_1}{k_1}$ can be identified with $\frac{f_2}{g_2} \cdot \frac{h_2}{k_2}$ in $\mathcal{K}(V)$.

Exercise 3.4.7. Show that $\mathcal{K}(V)$ is a field. (This is an exercise in abstract algebra; its goal is not only to show that $\mathcal{K}(V)$ is a field but also to provide the reader with an incentive to review what a field is.)

3.4.2. The Projective Case. We have seen that the natural space for the study of curves is not $\mathbb{C}^2$ but the projective plane $\mathbb{P}^2$. The corresponding functions will have to come from homogeneous polynomials. To a large extent, after we add words about homogeneity, this subsection will be a reworking of the previous subsection.

Exercise 3.4.8. Let $P(x, y, z) = x^2 + xy + z^2$. Consider the two polynomials

$$f_1(x, y, z) = x^2 \quad \text{and} \quad f_2(x, y, z) = 2x^2 + xy + z^2.$$

Find a point $(a : b : c) \in \mathbb{P}^2$ such that

$$f_1(a, b, c) \neq f_2(a, b, c).$$

Now show that if $(a : b : c) \in \mathrm{V}(P)$, then

$$f_1(a, b, c) = f_2(a, b, c).$$

In the above exercise, why is $f_1(a, b, c) = f_2(a, b, c)$ for any point $(a : b : c) \in \mathrm{V}(P)$? Again, the key is to look at $f_2(x, y, z) - f_1(x, y, z)$. In the affine case, we used the analogous equivalence relation to define the ring of regular functions on the curve $\mathrm{V}(P)$. However, if we wish to use homogeneous polynomials to define functions on a projective curve, we encounter problems.

Exercise 3.4.9. Let $V = \mathrm{V}(P)$ be an irreducible curve in $\mathbb{P}^2$, let $(a : b : c) \in V$ be a point on this curve, and let $f(x, y, z)$ be a homogeneous polynomial of degree d.

(1) Show that $f(\lambda a, \lambda b, \lambda c) = \lambda^d f(a, b, c)$ for all $\lambda \neq 0$.
(2) Conclude that if $f(a, b, c) \neq 0$ and $d > 0$, then f is not a well-defined function from V to $\mathbb{C}$.

Therefore, the only regular functions on a projective curve are constants. Luckily we have a projective analogue to the function field.

Definition 3.4.5. Let the function field $\mathcal{K}(V)$ for the curve $V = \mathrm{V}(P)$, where $P(x, y, z)$ is a homogeneous polynomial, be the set of all rational functions

$$\frac{f(x, y, z)}{g(x, y, z)}$$

where

(1) both f and g are homogeneous of the same degree,

(2) P does not divide g (which is a way of guaranteeing that g, the denominator, is not identically zero on the curve V), and

(3) $\dfrac{f_1(x,y,z)}{g_1(x,y,z)}$ is identified with $\dfrac{f_2(x,y,z)}{g_2(x,y,z)}$ if P divides $f_1g_2 - f_2g_1$. We denote this identification by setting

$$\frac{f_1(x,y,z)}{g_1(x,y,z)} \sim \frac{f_2(x,y,z)}{g_2(x,y,z)}.$$

As before, we want $\mathcal{K}(V)$ to mimic the rational numbers.

Definition 3.4.6. On $\mathcal{K}(V)$, define addition and multiplication by

$$\frac{f(x,y,z)}{g(x,y,z)} + \frac{h(x,y,z)}{k(x,y,z)} = \frac{f(x,y,z)k(x,y,z) + g(x,y,z)h(x,y,z)}{g(x,y,z)k(x,y,z)}$$

and

$$\frac{f(x,y,z)}{g(x,y,z)} \cdot \frac{h(x,y,z)}{k(x,y,z)} = \frac{f(x,y,z)h(x,y,z)}{g(x,y,z)k(x,y,z)},$$

when f, g, h and k are all homogeneous, f and g have the same degree and h and k have the same degree.

We now want to link the equivalence relation for the projective case with the equivalence relation for the affine case.

In fact, we will show that this $\mathcal{K}(V)$ is isomorphic, in some sense, to the function field of the previous subsection (which is why we are using the same notation for both). For now, we will specify the $\mathcal{K}(V)$ of the projective curve as $\mathcal{K}_{\mathbb{P}}(V)$ and the $\mathcal{K}(V)$ of the affine curve as $\mathcal{K}_{\mathbb{A}}(V)$.

Define

$$T : \mathcal{K}_{\mathbb{P}}(V) \to \mathcal{K}_{\mathbb{A}}(V)$$

by setting

$$T\left(\frac{f(x,y,z)}{g(x,y,z)}\right) = \frac{f(x,y,1)}{g(x,y,1)}.$$

We first show that T is well-defined.

Exercise 3.4.10. Let $f(x,y,z)$ and $g(x,y,z)$ be two homogeneous polynomials of the same degree such that $f(x,y,z) \sim g(x,y,z)$ with

respect to the homogeneous polynomial $P(x, y, z)$. Show that $f(x, y, 1)$ $\sim g(x, y, 1)$ with respect to the dehomogenized polynomial $P(x, y, 1)$.

Exercise 3.4.11. Let $f_1(x, y, z), f_2(x, y, z), g_1(x, y, z)$ and $g_2(x, y, z)$ be homogeneous polynomials of the same degree such that $f_1(x, y, z)$ $\sim f_2(x, y, z)$ and $g_1(x, y, z) \sim g_2(x, y, z)$ with respect to the homogeneous polynomial $P(x, y, z)$. Show that in $\mathcal{K}_{\mathbb{A}}(V)$ we have

$$T\left(\frac{f_1(x, y, z)}{g_1(x, y, z)}\right) \sim T\left(\frac{f_2(x, y, z)}{g_2(x, y, z)}\right).$$

Hence T indeed maps the field $\mathcal{K}_{\mathbb{P}}(V)$ to the field $\mathcal{K}_{\mathbb{A}}(V)$. Next we want to show that T is a field homomorphism, which is the point of the next two exercises.

Exercise 3.4.12. Let $f(x, y, z)$ and $g(x, y, z)$ be two homogeneous polynomials of the same degree, and let $h(x, y, z)$ and $k(x, y, z)$ be two other homogeneous polynomials of the same degree. Show that

$$T\left(\frac{f(x, y, z)}{g(x, y, z)} + \frac{h(x, y, z)}{k(x, y, z)}\right) = T\left(\frac{f(x, y, z)}{g(x, y, z)}\right) + T\left(\frac{h(x, y, z)}{k(x, y, z)}\right).$$

Exercise 3.4.13. Let $f(x, y, z)$ and $g(x, y, z)$ be two homogeneous polynomials of the same degree, and let $h(x, y, z)$ and $k(x, y, z)$ be two other homogeneous polynomials of the same degree. Show that

$$T\left(\frac{f(x, y, z)}{g(x, y, z)} \cdot \frac{h(x, y, z)}{k(x, y, z)}\right) = T\left(\frac{f(x, y, z)}{g(x, y, z)}\right) \cdot T\left(\frac{h(x, y, z)}{k(x, y, z)}\right).$$

To show that T is one-to-one, we will use the fact that it is equivalent to show that the only element mapping to zero is zero itself.

Exercise 3.4.14. Suppose $f(x, y, z)$ and $g(x, y, z)$ are two homogeneous polynomials of the same degree such that

$$T\left(\frac{f(x, y, z)}{g(x, y, z)}\right) = 0$$

in $\mathcal{K}_{\mathbb{A}}(V)$. Show that

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

in $\mathcal{K}_{\mathbb{P}}(V)$.

To finish the proof that T is an isomorphism, we must show that T is onto.

Exercise 3.4.15. Given two polynomials $f(x, y)$ and $g(x, y)$, find two homogeneous polynomials $F(x, y, z)$ and $G(x, y, z)$ of the same degree such that

$$T\left(\frac{F(x,y,z)}{G(x,y,z)}\right) = \frac{f(x,y)}{g(x,y)}.$$

3.5. Divisors

The goal of this section is to define divisors on a curve. To each divisor, there is a naturally associated vector space of rational functions. The dimension of these vector spaces is essential for understanding the Riemann-Roch Theorem, a result that links the algebraic and topological properties of a curve. We will discuss and prove Riemann-Roch in the next section.

3.5.1. Intuition behind Riemann-Roch. Here is a fairly simple question. Let $C = \mathrm{V}(P)$ be a curve in $\mathbb{P}^2$. Choose some point p on the curve. Is there a rational function with a pole (an infinity) of order one exactly at the point p and no other poles? Recall that a rational function in $\mathcal{K}(C)$ can be written as

$$F(x, y, z) = \frac{f(x, y, z)}{g(x, y, z)},$$

where f and g are homogeneous polynomials of the same degree with the additional property that g is not identically zero on $\mathrm{V}(P)$. The poles of F on the curve $\mathrm{V}(P)$ occur when the denominator of F is zero. Thus we must look at the set of intersection points

$$\mathrm{V}(g) \cap \mathrm{V}(P).$$

By Bézout's Theorem, there should be $\deg(g) \cdot \deg(P)$ points of intersection. Unless P has degree one, F cannot have a single isolated pole of order one on C.

There is a subtlety that we need to consider. It could be that the number of intersection points in $\mathrm{V}(g) \cap \mathrm{V}(P)$ is greater than one but that at all of these points, besides our chosen point p, the numerator f has the same zeros, canceling those from the denominator. The heart of Riemann-Roch is showing that this does not happen. The

Riemann-Roch Theorem will give us information about what type of elements in $\mathcal{K}(C)$ can exist with prescribed poles on $C = \mathrm{V}(P)$.

We now want to see that the projective line $\mathbb{P}^1$ has a particularly well-behaved function field.

Exercise 3.5.1. If x and y are homogeneous coordinates for $\mathbb{P}^1$, show that the rational function

$$F(x, y) = \frac{x}{y}$$

has a single zero at $(0 : 1)$ and a single pole at $(1 : 0)$.

Exercise 3.5.2. For $\mathbb{P}^1$, find a rational function with a single zero at $(1 : -1)$ and a single pole at $(1 : 0)$.

Exercise 3.5.3. For $\mathbb{P}^1$, find a rational function with zeros at $(1 : -1)$ and at $(0 : 1)$ and a double pole at $(1 : 0)$.

Exercise 3.5.4. For $\mathbb{P}^1$, find a rational function with zeros at $(1 : -1)$ and $(0 : 1)$ and poles at $(1 : 0)$ and $(1 : 1)$.

Exercise 3.5.5. For $\mathbb{P}^1$, show that there cannot be a rational function with zeros at $(1 : -1)$ and at $(0 : 1)$ and a single pole at $(1 : 0)$ with no other poles.

3.5.2. Divisors.

Definition 3.5.1. A *divisor* on a curve $C = \mathrm{V}(P)$ is a formal finite linear combination of points on C with integer coefficients, $D = n_1p_1 + n_2p_2 + \cdots + n_kp_k$. The sum $\sum_{i=1}^{k} n_i$ of the coefficients is called the *degree* of D. When each $n_i \geq 0$ we say that D is *effective*.

Given two divisors D_1 and D_2 on $\mathrm{V}(P)$, we say

$$D_1 \leq D_2$$

if and only if $D_2 - D_1$ is effective. This defines a partial ordering on the set of all divisors on $\mathrm{V}(P)$.

One reason that divisors are natural tools to study a curve is their link with rational functions. Consider a nonzero function F in the function field $\mathcal{K}(C)$ of the curve $C = \mathrm{V}(P)$. Associate to F the divisor

$$\mathrm{div}(F) = \sum n_i p_i,$$

where the sum is taken over all zeros and poles of F on $\mathrm{V}(P)$, n_i is the multiplicity of the zero at p_i, and $-n_j$ is the order of the pole at p_j.

Definition 3.5.2. Any divisor that can be written as $\mathrm{div}(w)$ for a function $w \in \mathcal{K}(C)$ is called a *principal divisor* on $C = \mathrm{V}(P)$.

Note that for the plane curve $C = \mathrm{V}(P)$ defined by $P(x, y, z) = 0$, any $w \in \mathcal{K}(C)$ can be written as $w = \frac{f(x,y,z)}{g(x,y,z)}$, where f and g are homogeneous polynomials of the same degree.

Exercise 3.5.6. Let x and y be homogeneous coordinates on $\mathbb{P}^1$ and let $w = \frac{x}{y}$. Write the divisor $\mathrm{div}(w)$ as a formal sum of points.

Exercise 3.5.7. Let x, y, z be homogeneous coordinates on $\mathbb{P}^2$. For the cubic curve $\mathrm{V}(y^2z - x^3 - xz^2)$, write the divisor $\mathrm{div}(\frac{y}{z})$ as a formal sum of points.

Exercise 3.5.8. Let x, y, z be homogeneous coordinates on $\mathbb{P}^2$. For the cubic curve $\mathrm{V}(y^2z - x^3 - xz^2)$, show that the divisor $D = 2(0 : 0 : 1) - 2(0 : 1 : 0)$ is principal.

Exercise 3.5.9. Show that a principal divisor has degree zero.

Exercise 3.5.10. Prove that the set of all divisors on a curve $\mathrm{V}(P)$ forms an abelian group under addition and that the subset of principal divisors is a subgroup.

3.5.3. Vector Space $L(D)$ Associated to a Divisor. To any divisor on a curve C we want to associate a vector space that is a subspace of the function field $\mathcal{K}(C)$. The dimension of this vector space will be critical for the Riemann-Roch Theorem.

Definition 3.5.3. For a divisor D on a curve C, define $L(D)$ to be

$$L(D) = \{F \in \mathcal{K}(C) : F = 0 \text{ or } \mathrm{div}(F) + D \geq 0\}.$$

Thus for $D = \sum n_p p$, we have $F \in L(D)$ when F has a pole of order at most n_p for points p with $n_p > 0$ and F has a zero of multiplicity at least $-n_p$ at points p with $n_p < 0$.

Exercise 3.5.11. Consider the curve $\mathbb{P}^1$. Let $D = (1 : 0) + (0 : 1)$. Show that

$$\frac{(x - y)(x + y)}{xy} \in L(D).$$

Exercise 3.5.12. Consider the curve $\mathbb{P}^1$. Let $D = (1:0) + (0:1)$. Show that

$$\frac{(x-y)(x+y)}{xy} \in L(kD)$$

for any positive integer k.

Exercise 3.5.13. Continuing with the previous exercise, show that

$$\frac{xy}{(x-y)(x+y)} \notin L(D).$$

Exercise 3.5.14. Let $D = (1:0:1) + (-1:0:1)$ be a divisor on $\mathrm{V}(x^2+y^2-z^2)$. Show that

$$\frac{x}{y} \in L(D),$$

but that

$$\frac{y}{x} \notin L(D).$$

Exercise 3.5.15. Let D be a divisor on $\mathrm{V}(P)$. Show that $L(D)$ is a complex vector space.

Exercise 3.5.16. For a smooth curve $\mathrm{V}(P)$, find $L(0)$.

Exercise 3.5.17. Find $L(D)$ for the divisor $D = (0:1)$ on $\mathbb{P}^1$.

Exercise 3.5.18. Prove if $\deg(D) < 0$, then $L(D) = \{0\}$, the trivial space.

Exercise 3.5.19. Prove if $D_1 \leq D_2$, then $L(D_1) \subseteq L(D_2)$.

In the next subsection, we will see that the dimension of $L(D)$ is finite.

3.5.4. $L(D+p)$ versus $L(D)$. We write $l(D)$ for the dimension of $L(D)$ as a vector space over $\mathbb{C}$. In the next section we will be discussing the Riemann-Roch Theorem, which gives sharp statements linking the dimension of the vector space $L(D)$ with the degree of D and the genus of the curve C. We will start the proof here, by showing:

Theorem 3.5.20. Let D be a divisor on a curve C and let $p \in C$ be any point on the curve. Then

$$l(D+p) \leq l(D) + 1.$$

By Exercise 3.5.19, we know that $l(D) \leq l(D+p)$. Thus the above theorem is stating that by adding a single point to a divisor, we can increase the dimension of the corresponding vector space by at most one.

Exercise 3.5.21. Let $D = \sum n_p p$ be a divisor on the curve $\mathrm{V}(P)$. Use this theorem, together with the result of Exercise 3.5.18, to prove that $l(D)$ is finite.

The proof of Theorem 3.5.20 uses some basic linear algebra.

Exercise 3.5.22. Let V be a complex vector space. Let

$$T : V \to \mathbb{C}$$

be a linear transformation. Recall that the kernel of T is

$$\ker(T) = \{v \in V : T(v) = 0\}.$$

Show that $\ker(T)$ is a subspace of V.

Exercise 3.5.23. Using the above notation, show that

$$\dim(\ker(T)) \leq \dim(V) \leq \dim(\ker(T)) + 1.$$

(This problem will require you to look up various facts about linear transformations and dimensions.)

For the next few exercises, assume that D is a divisor on a curve C and $p \in C$ is a point on the curve.

Exercise 3.5.24. Suppose there is a linear transformation

$$T : L(D+p) \to \mathbb{C}$$

such that

$$\ker(T) = L(D).$$

Show that

$$l(D+p) \leq l(D) + 1.$$

Thus to prove the theorem it suffices to construct such a linear transformation. Let $D = \sum n_q q$, where each $n_q \in \mathbb{Z}$, the q are points on C and all but a finite number of the coefficients, n_q, are zero. We call the integer n_q the *multiplicity* of the point q for the divisor D.

Exercise 3.5.25. Show that the multiplicity of the point p for the divisor $D + p$ is exactly one more than the multiplicity of p for the divisor D.

Exercise 3.5.26. Let $p = (0 : 1 : 1) \in \mathrm{V}(x^2 + y^2 - z^2)$. Set $D = 2p + (1 : 0 : 1)$. Let $F \in L(D)$. Even though $F(x, y, z)$ can have a pole at the point p, show that the function $x^2F(x, y, z)$ cannot have a pole at p. Show if p is a zero of the function $x^2F(x, y, z)$, then $F \in L(D - p)$.

Exercise 3.5.27. Use the same notation from the previous exercise. Define a map

$$T : L(D) \to \mathbb{C}$$

as follows. Dehomogenize by setting $z = 1$. Set $T(F)$ to be the number obtained by plugging in $(0, 1)$ to the function $x^2F(x, y, 1)$. Show that

$$T\left(\frac{(2y - z)(2y + z)}{x^2}\right) = 3.$$

Exercise 3.5.28. Use the notation from the previous exercises. Show that

$$T : L(D) \to \mathbb{C}$$

is a linear transformation with kernel $L(D - p)$.

We need to make a few choices about our curve C and our point p. By choosing coordinates correctly, we can assume that $p = (0 : y : 1)$. We choose a line that goes through the point p and is not tangent to the curve C. By rotating our coordinates, if necessary, we can assume this line is $\mathrm{V}(x)$.

Exercise 3.5.29. Let n be the multiplicity of the point p for the divisor $D+p$. For any $F \in L(D+p)$, show that the function $x^nF(x, y, 1)$ does not have a pole at p.

Exercise 3.5.30. Using the notation from the previous exercise, show that $x^nF(x, y, 1)$ has a zero at p means that $F \in L(D)$.

Exercise 3.5.31. Let n be the multiplicity of the point p for the divisor $D + p$. Define

$$T : L(D + p) \to \mathbb{C}$$

by setting $T(F)$ to be the number obtained by plugging in $(0, y)$ to the function $x^n F(x, y, 1)$. Show that T is a linear transformation with kernel $L(D)$.

Thus we have shown that

$$l(D) \leq l(D + p) \leq l(D) + 1.$$

3.5.5. Linear Equivalence of Divisors. Recall that a divisor D on a curve C is called principal if it is of the form $\mathrm{div}(w)$ for some $w \in \mathcal{K}(C)$.

Definition 3.5.4. Two divisors, D_1 and D_2, are *linearly equivalent*, written as $D_1 \equiv D_2$, if $D_1 - D_2$ is principal.

Exercise 3.5.32. Prove that linear equivalence is an equivalence relation on the set of all divisors on $\mathrm{V}(P)$.

Exercise 3.5.33. Prove for any two points p and q in $\mathbb{P}^1$, $p \equiv q$.

Exercise 3.5.34. For any fixed point p, prove that any divisor on $\mathbb{P}^1$ is linearly equivalent to mp for some integer m.

Exercise 3.5.35. Prove if $D_1 \equiv D_2$, then $L(D_1) \cong L(D_2)$ as vector spaces over $\mathbb{C}$.

3.5.6. Hyperplane Divisors. In general, calculating $l(D)$ is difficult and is, in part, one of the goals of Riemann-Roch. There is a special class of divisors on any curve, called the hyperplane divisors, for which we can explicitly calculate this dimension.

We have defined divisors on a curve C as finite formal sums of points on C. Given any homogeneous polynomial $f(x, y, z)$, we can associate a divisor on C by setting $\mathrm{div}(f) = \sum n_p p$, where the sum is taken over all points p in $\mathrm{V}(f) \cap C$ and n_p is the intersection multiplicity. We frequently write $\mathrm{div}(f)$ as $\mathrm{V}(f) \cap C$. We now look at an important case where $f(x, y, z)$ is linear.

Exercise 3.5.36. Consider the curve $\mathrm{V}(x^2 + y^2 - z^2)$. Determine the divisor

$$D_1 = \mathrm{V}(x - y) \cap \mathrm{V}(x^2 + y^2 - z^2)$$

and the divisor

$$D_2 = \mathrm{V}(x) \cap \mathrm{V}(x^2 + y^2 - z^2).$$

Show that $D_1 \equiv D_2$.

Exercise 3.5.37. Using the notation from the previous exercise, let D_3 be the divisor on

$$\mathrm{V}(x^4 + 2y^4 - x^3z + z^4) \cap \mathrm{V}(x^2 + y^2 - z^2).$$

Show that $D_3 \equiv 4D_1$. [Hint: Do not explicitly calculate the divisor D_3.]

Exercise 3.5.38. Using the notation from the previous problems, let $f(x, y, z)$ be a homogeneous polynomial of degree 3. Show that

$$\frac{f(x, y, z)}{(x - y)^3} \in L(3D_1).$$

Exercise 3.5.39. Using the notation from the previous problems, let $f(x, y, z)$ be a homogeneous polynomial of degree k. Show that

$$\frac{f(x, y, z)}{(x - y)^k} \in L(kD_1).$$

Definition 3.5.5. Let $C = \mathrm{V}(P)$ be a plane curve defined by a homogeneous polynomial $P(x, y, z)$ of degree d. Define a *hyperplane divisor* H on C to be the divisor of zeros of a linear function $\ell(x, y, z)$ in $\mathbb{C}[x, y, z]$, meaning that

$$H = \mathrm{V}(\ell) \cap \mathrm{V}(P).$$

Exercise 3.5.40. Suppose that H and H' are hyperplane divisors on a curve C. Prove that $H \equiv H'$.

Exercise 3.5.41. Using the same notation from the previous problem, show for any homogeneous polynomial $f(x, y, z)$ of degree m in $\mathbb{C}[x, y, z]$,

$$\frac{f(x, y, z)}{\ell^m} \in L(mH).$$

Now we start calculating $l(mH) = \dim L(mH)$, for any hyperplane divisor H.

We know from the above exercise that $L(mH)$ contains elements of the form $\frac{f(x,y,z)}{\ell^m}$. In fact, every element in $L(mH)$ can be written in this form. To prove this we use the following

Theorem 3.5.42 (Noether's AF+BG Theorem)**.** ([**Ful69**], p. 61). Let $F(x, y, z)$ and $G(x, y, z)$ be homogeneous polynomials defining plane curves that have no common component. Assume that F defines a nonsingular curve. Let $U(x, y, z)$ be a homogeneous polynomial that satisfies the following condition: suppose for every point p in the intersection $\mathrm{V}(F) \cap \mathrm{V}(G)$, there is the following inequality of intersection multiplicities:

$$I(p, \mathrm{V}(F) \cap \mathrm{V}(U)) \geq I(p, \mathrm{V}(F) \cap \mathrm{V}(G)).$$

Then there are homogeneous polynomials A and B such that $U = AF + BG$.

We have restricted our statement to the case where $C = \mathrm{V}(F)$ is nonsingular; we can state this more generally, although the condition on intersections becomes more complicated.

Exercise 3.5.43. In the case of the theorem, what are the degrees of the polynomials A and B?

Exercise 3.5.44. Let $F(x, y, z) = x$ and $G(x, y, z) = y$. Show that any polynomial U vanishing at $(0 : 0 : 1)$ satisfies the hypothesis of the theorem, and thus there are A and B such that $U = AF + BG$.

Exercise 3.5.45. Let $F(x, y, z) = xz+yz+xy$ and $G(x, y, z) = xz+yz-xy$. Show that the polynomial $U = x^2(y+z)+y^2(x+z)+z^2(x+y)$ satisfies the hypothesis of the theorem, and find A and B such that $U = AF + BG$.

We now use this theorem to determine the form of the general element in $L(mH)$ in the following steps.

Exercise 3.5.46. Let $U \in L(mH)$. Show that U can be written as $U = \frac{u}{v}$, where u and v are homogeneous polynomials of the same degree in $\mathbb{C}[x, y, z]$ and $\mathrm{div}(v) \leq \mathrm{div}(u) + \mathrm{div}(\ell^m)$.

Exercise 3.5.47. Let $C = \mathrm{V}(F)$ and let $U = \frac{u}{v} \in L(mH)$, where u and v are homogeneous polynomials of the same degree in $\mathbb{C}[x, y, z]$. Show for all $p \in \mathrm{V}(F) \cap \mathrm{V}(v)$,

$$I(p, \mathrm{V}(F) \cap \mathrm{V}(\ell^m)) \geq I(p, \mathrm{V}(F) \cap \mathrm{V}(v)).$$

Exercise 3.5.48. Under the assumptions of the previous exercise, use Noether's Theorem to conclude there exist A and B with $u\ell^m = AF + Bv$. Show that this implies $U = \frac{B}{\ell^m}$ in $\mathcal{K}(C)$.

Thus the vector space $L(mH)$ consists of all functions in $\mathcal{K}(C)$ of the form $\frac{f}{\ell^m}$ for homogeneous polynomials f of degree m. To find the dimension of $L(mH)$, we need to find the dimension of the vector space of all possible numerators f. The key will be that P cannot divide f.

Exercise 3.5.49. Let $\mathbb{C}_m[x, y, z]$ denote the set of all homogeneous polynomials of degree m together with the zero polynomial. Show that if $f, g \in \mathbb{C}_m[x, y, z]$ and if $\lambda, \mu \in \mathbb{C}$, then

$$\lambda f + \mu g \in \mathbb{C}_m[x, y, z].$$

Conclude that $\mathbb{C}_m[x, y, z]$ is a vector space over $\mathbb{C}$.

Exercise 3.5.50. Show that $\dim \mathbb{C}_1[x, y, z] = 3$ and a basis is

$$\{x, y, z\}.$$

Exercise 3.5.51. Show that $\dim \mathbb{C}_2[x, y, z] = 6$ and a basis is

$$\{x^2, xy, xz, y^2, yz, z^2\}.$$

Exercise 3.5.52. Show that

$$\dim \mathbb{C}_m[x, y, z] = \binom{m+2}{m}.$$

(By definition

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

This number, called "n choose k", is the number of ways of choosing k items from n things, where order does not matter.)

Exercise 3.5.53. Let $P(x, y, z)$ be a homogeneous polynomial of degree d. In the vector space $\mathbb{C}_m[x, y, z]$, let

$$W = \{f(x, y, z) \in \mathbb{C}_m[x, y, z] : P \text{ divides } f\}.$$

If $f, g \in W$ and if $\lambda, \mu \in \mathbb{C}$, show

$$\lambda f + \mu g \in W.$$

Conclude that W is a vector subspace of $\mathbb{C}_m[x, y, z]$.

Exercise 3.5.54. With the notation of the previous problem, show that the vector space W is isomorphic to the vector space $\mathbb{C}_{m-d}[x,y,z]$. (Recall that this means you must find a linear map $T : \mathbb{C}_{m-d}[x,y,z] \to W$ that is one-to-one and onto.) Conclude that

$$\dim(W) = \dim \mathbb{C}_{m-d}[x,y,z].$$

Exercise 3.5.55. Show that

$$\begin{aligned} l(mH) &= \dim \mathbb{C}_m[x,y,z] - \dim \mathbb{C}_{m-d}[x,y,z] \\ &= \frac{(m+1)(m+2)}{2} - \frac{(m-d+1)(m-d+2)}{2}. \end{aligned}$$

Exercise 3.5.56. Let ℓ be a linear function and let H be the corresponding hyperplane divisor on $\mathrm{V}(P)$, where $P(x,y,z)$ is homogeneous of degree d. Show that $\deg(H) = d$ and in general, that $\deg(mH) = md$. [Hint: Think Bézout.]

Exercise 3.5.57. Use the degree-genus formula $g = \dfrac{(d-1)(d-2)}{2}$ to show that

$$l(mH) = md - g + 1.$$

3.6. The Riemann-Roch Theorem

We will show Riemann's Theorem as our first step toward the true goal of this section, the Riemann-Roch Theorem. Riemann-Roch is a fundamental result that links the algebraic and topological properties of a curve.

3.6.1. Riemann's Theorem. Throughout this section, let $C = \mathrm{V}(P)$ be a smooth plane curve of degree d and genus g.

Theorem 3.6.1 (Riemann's Theorem)**.** If D is a divisor on a plane curve C of genus g, then

$$l(D) \geq \deg D - g + 1.$$

(The Riemann-Roch Theorem, our eventual goal, finds the explicit term that is needed to change the above inequality into an equality.)

Exercise 3.6.2. Show that for any hyperplane divisor H and any positive integer m, we have

$$l(mH) = \deg(mH) - g + 1.$$

Following notation used in Fulton's *Algebraic Curves* [**Ful69**], set

$$S(D) = \deg D + 1 - l(D).$$

Exercise 3.6.3. Suppose that for all divisors D we have

$$S(D) \leq g.$$

Show that this implies Riemann's Theorem is true.

Thus we want to show that $S(D) \leq g$, for any divisor D.

Exercise 3.6.4. Show for any hyperplane divisor H that $S(mH) = g$ for all positive integers m.

Exercise 3.6.5. Let $D_1 \leq D_2$. Show that $l(D_1) \leq l(D_2)$.

Exercise 3.6.6. Recall for any divisor D and point p on the curve C that $l(D) \leq l(D+p) \leq l(D)+1$. Show that

$$S(D+p) \geq S(D).$$

Exercise 3.6.7. Suppose that $D_1 \equiv D_2$ for two divisors on the curve C. Show that

$$S(D_1) = S(D_2).$$

Exercise 3.6.8. Let $f(x,y,z) \in \mathcal{O}(C)$ be a homogeneous polynomial of degree m. Let D be the divisor

$$\mathrm{V}(f) \cap C$$

and let H be a hyperplane divisor on C. Show that $D \equiv mH$ and that $\deg(D) = md$.

Exercise 3.6.9. Let $p = (a : b : c) \in C$. Suppose that a and b are not both zero. (This is not a big restriction on the point.) Let

$$f(x,y,z) = ay - bx.$$

Define the divisor

$$D = \mathrm{V}(f) \cap \mathrm{V}(P).$$

Show that $p \leq D$.

Exercise 3.6.10. Let $p_1 = (a_1 : b_1 : c_1)$ and $p_2 = (a_2 : b_2 : c_2)$ be points on C. Suppose that a_1 and b_1 are not both zero and similarly for a_2 and b_2. Let

$$f(x, y, z) = (a_1 y - b_1 x)(a_2 y - b_2 x)$$

and define the divisor

$$D = \mathrm{V}(f) \cap C.$$

Show that $p_1 + p_2 \leq D$.

Exercise 3.6.11. Let $p_1, p_2, \ldots, p_k \in C$. Find a polynomial f such that if

$$D = \mathrm{V}(f) \cap C,$$

then $p_1 + \cdots + p_k \leq D$.

Exercise 3.6.12. Let H be a hyperplane divisor on C. Using the divisor D from the previous problem, show that there is a positive integer m such that $D \equiv mH$.

Exercise 3.6.13. Let $D = \sum n_i p_i$ be an effective divisor on C. Let n be any positive integer. Prove that there is an $m \geq n$ and points $q_1, \ldots, q_k$ on C such that $D + \sum q_i \equiv mH$.

Exercise 3.6.14. Let $D = \sum n_i p_i$ be a divisor on C. Show that there are points $q_1, \ldots, q_k$ on C, which need not be distinct, such that $D + q_1 + \cdots + q_k$ is an effective divisor.

Exercise 3.6.15. Let $D = \sum n_i p_i$ be a divisor on C (not necessarily effective). Let n be a positive integer. Prove that there exists an integer m, $m \geq n$, and points $q_1, \ldots, q_k$ on C such that $D + \sum q_i \equiv mH$.

Exercise 3.6.16. Let D be a divisor on a curve C and let H be any hyperplane divisor. Show that there is a positive integer m so that

$$S(D) \leq S(mH).$$

Exercise 3.6.17. Prove Riemann's Theorem.

3.6.2. Differentials. In calculus we learn that the slope of the graph $y = f(x)$ is given by the derivative $\frac{dy}{dx}$ at each point where it is defined. For a curve defined implicitly, say by an equation $P(x, y) = 0$, using implicit differentiation we compute $\frac{dy}{dx} = -(\frac{\partial P}{\partial x})/(\frac{\partial P}{\partial y})$. Using this as motivation, we define the *differential* of the function $P(x, y)$ to be

$$dP = \frac{\partial P}{\partial x}dx + \frac{\partial P}{\partial y}dy.$$

More generally, a *differential form* on $\mathbb{C}^2$ is a sum of terms gdf, for functions $f, g \in \mathcal{K}(\mathbb{C}^2)$. (Recall that this means f and g are ratios of polynomials in two variables.) Of course we have the usual rules from calculus:

$$\begin{aligned} d(f+g) &= df + dg \\ d(cf) &= cdf \\ d(fg) &= gdf + fdg \end{aligned}$$

for $c \in \mathbb{C}$ and $f, g \in \mathcal{K}(\mathbb{C}^2)$.

Exercise 3.6.18.

(1) Find the differential of $f(x, y) = x^2 + y^2 - 1$.

(2) Use your answer from Part (1) to find the slope of the circle $f(x, y) = 0$ at a point (x, y).

(3) For which points on the circle is this slope undefined?

Exercise 3.6.19.

(1) Find the differential of $f(x, y) = x^3 + x - y^2$.

(2) Use your answer from Part (1) to find the slope of the curve $f(x, y) = 0$ at a point (x, y).

(3) For which points on the curve is this slope undefined?

Exercise 3.6.20. Prove that the set of all differential forms on $\mathbb{C}^2$ is a vector space over $\mathcal{K}(\mathbb{C}^2)$ with basis $\{dx, dy\}$.

To define differentials on an affine curve $P(x, y) = 0$ in $\mathbb{C}^2$, we use the relation $dP = \frac{\partial P}{\partial x}dx + \frac{\partial P}{\partial y}dy = 0$. As in calculus this gives the slope $-\frac{\partial P/\partial x}{\partial P/\partial y}$ of the curve when $\frac{\partial P}{\partial y} \neq 0$. We can also use this expression to express dy in the form $g(x, y)dx$ for a function $g \in \mathcal{K}(\mathbb{C}^2)$ (namely, $g = -\frac{\partial P/\partial x}{\partial P/\partial y}$, the slope of our curve).

Suppose that $f \in \mathcal{K}(C)$ is determined by some $F(x,y) \in \mathcal{K}(\mathbb{C}^2)$ restricted to C. We wish to define the differential df to be dF restricted to C. This appears to depend on the choice of $F(x,y)$, which is only well-defined up to the addition of terms of the form $G(x,y)P(x,y)$ for $G(x,y) \in \mathcal{K}(\mathbb{C}^2)$. Yet $d(GP) = G(x,y)\ dP + P(x,y)\ dG$, and we know that $P(x,y) = dP = 0$ on C. Thus any $F+GP \in \mathcal{K}(\mathbb{C}^2)$ that represents $f \in \mathcal{K}(C)$ has $d(F+GP) = dF$ when restricted to C, so taking df to be the restriction of dF is well-defined. With this established, we may define differentials on an affine curve $C = \mathrm{V}(P)$ to be sums of terms of the form $g\ df$ for $g, f \in \mathcal{K}(C)$.

Exercise 3.6.21. Prove that the set of all differential forms on a nonsingular curve $C = \mathrm{V}(P)$ in $\mathbb{C}^2$ is a vector space over $\mathcal{K}(C)$.

Exercise 3.6.22. Prove that the vector space of differentials on a nonsingular curve $C = \mathrm{V}(P)$ in $\mathbb{C}^2$ has dimension one over $\mathcal{K}(C)$.

3.6.3. Local Coordinates. To extend our definition of differential forms to projective curves $C = \mathrm{V}(P)$ in $\mathbb{P}^2$, we will consider the affine pieces of C obtained by dehomogenizing the defining polynomial $P(x,y,z)$. We can cover $\mathbb{P}^2$ by three affine coordinate charts, that is, three copies of $\mathbb{C}^2$, as follows. The bijective map

$$\varphi : \mathbb{P}^2 - \mathrm{V}(z) \to \mathbb{C}^2$$

defined by $\varphi(x:y:z) = \left(\frac{x}{z}, \frac{y}{z}\right)$ assigns coordinates

$$r = \frac{x}{z}, s = \frac{y}{z}$$

for all points $(x:y:z)$ with $z \neq 0$. Similarly we can set

$$t = \frac{x}{y}, u = \frac{z}{y}$$

for all $(x:y:z)$ with $y \neq 0$, and

$$v = \frac{y}{x}, w = \frac{z}{x}$$

when $x \neq 0$. (These three coordinate systems give a more careful way to "dehomogenize" polynomials in $\mathbb{P}^2$, compared to simply setting one coordinate equal to 1, as we did in Chapter 1.)

Exercise 3.6.23. Verify that the map $\varphi : \mathbb{P}^2 - \mathrm{V}(z) \to \mathbb{C}^2$ is a bijection.

Exercise 3.6.24. Use the above coordinates in each of the three affine charts on $\mathbb{P}^2$.

(1) Find coordinates for the point $(-1 : 2 : 3)$ in each of the three coordinate charts.

(2) Find all points in $\mathbb{P}^2$ that cannot be represented in rs-affine plane.

(3) Find the points in $\mathbb{P}^2$ that are not in either rs- or tu-affine planes.

Exercise 3.6.25. In this exercise you will find the change of coordinates functions between coordinate charts.

(1) Write the affine coordinates r and s as functions of t and u.

(2) Write the affine coordinates r and s as functions of v and w.

(3) Write the affine coordinates v and w as functions of t and u.

Now let C be a smooth curve defined by the vanishing of a homogeneous polynomial $P(x, y, z)$. We work locally by considering an affine part of the curve in an affine chart. Let $p = (a : b : c) \in C$ and assume $c \neq 0$, so p is a point of the affine curve defined by $P(\frac{x}{z}, \frac{y}{z}, 1) = P(r, s) = 0$ in $\mathbb{C}^2$. Since C is smooth, $\frac{\partial P}{\partial r} \neq 0$ or $\frac{\partial P}{\partial s} \neq 0$ at $(r, s) = \left(\frac{a}{c}, \frac{b}{c}\right)$.

The differential of P is

$$dP = \frac{\partial P}{\partial r} dr + \frac{\partial P}{\partial s} ds = 0.$$

Thus when $\frac{\partial P}{\partial r} \neq 0$ we can write

$$dr = \frac{-\frac{\partial P}{\partial s}}{\frac{\partial P}{\partial r}} ds.$$

Therefore we can write any differential form as $f(r, s)ds$, for some rational function $f(r, s)$, at all points where $\frac{\partial P}{\partial r} \neq 0$. We call s a *local coordinate* at these points.

Similarly, when $\frac{\partial P}{\partial s} \neq 0$ we may write

$$ds = \frac{-\frac{\partial P}{\partial r}}{\frac{\partial P}{\partial s}} dr.$$

We say that r is a local coordinate at points where $\frac{\partial P}{\partial s}$ is nonzero.

Exercise 3.6.26. Let $C = \mathrm{V}(x^2 - yz)$.

(1) Show that this curve is covered by the rs- and tu-charts, that is, every point $p \in C$ can be written in at least one of these coordinate systems.

(2) Show that r is a local coordinate at all points $p = (a : b : c) \in C$ with $c \neq 0$.

(3) Show that t is a local coordinate if $c = 0$.

Exercise 3.6.27. Let $C = \mathrm{V}(x^3 - y^2 z - xz^2)$.

(1) Show that every point $p \in C$ can be written in either rs- or tu-coordinates.

(2) Show that r is a local coordinate at all points $p = (a : b : c) \in C$ with $b, c \neq 0$.

(3) Find all points on C with $b = 0$ or $c = 0$ and determine a local coordinate at each point.

In the next two exercises we use local coordinates to write differential forms on affine curves. As the derivative provides local (that is, in a small neighborhood of a point) information about a curve, it makes sense to use this approach for differentials.

Exercise 3.6.28. Consider the curve $\mathrm{V}(x^2 - y)$ in $\mathbb{C}^2$.

(1) Show that x is a coordinate at all points on this curve.

(2) Show that any differential form can be written as $h(x, y)dx$ for some rational function $h(x, y)$.

Exercise 3.6.29. Consider the curve $\mathrm{V}(x^2 + y^2 - 1)$ in $\mathbb{C}^2$.

(1) Show that x is a local coordinate at all points (a, b) with $b \neq 0$.

(2) Write the differential dy in the form $f(x, y)\, dx$, where f is a rational function.

We now extend our definition of differential forms to curves in the projective plane. With the previous notation we have three affine pieces of a curve, corresponding to the $(r,s) = (\frac{x}{z}, \frac{y}{z})$, $(t,u) = (\frac{x}{y}, \frac{z}{y})$, and $(v,w) = (\frac{y}{x}, \frac{z}{x})$ coordinate charts. For an affine piece, say in the rs-coordinate system, we can write a differential form as $h(r,s)\,dr$ (or $h(r,s)\,ds$) for a rational function h. Using the changes of coordinates between the three affine charts we can translate this form to each set of coordinates. Thus a differential form on C is a collection of differential forms on each affine piece of C, such that these pieces "match" under our changes of coordinates.

Exercise 3.6.30. Let C be the curve $\mathrm{V}(x^2 - yz)$ in $\mathbb{P}^2$, which dehomogenizes to $r^2 - s = 0$ in the rs-affine chart.

(1) Show that the differential form ds can be written as $2r\,dr$.

(2) Use an appropriate change of coordinates to write ds in the form $f(t,u)\,dt$.

(3) Use an appropriate change of coordinates to write ds in the form $g(v,w)\,dw$.

Exercise 3.6.31. Let C be the curve $\mathrm{V}(x^2 + y^2 - z^2)$ in $\mathbb{P}^2$.

(1) Write the differential form dr as $f(r,s)\,ds$ for a rational function $f(r,s)$.

(2) Use an appropriate change of coordinates to write dr in the form $g(v,w)\,dw$.

3.6.4. The Canonical Divisor. We next define the divisor associated to a differential form on a smooth projective curve C in $\mathbb{P}^2$. For any differential form ω, we want to determine a divisor $\mathrm{div}(\omega) = \sum n_p p$, a finite sum of points $p \in C$ with integer coefficients n_p.

Throughout this section we use the notation:

$$r = \frac{x}{z},\ s = \frac{y}{z},$$

$$t = \frac{x}{y},\ u = \frac{z}{y},$$

$$v = \frac{y}{x},\ w = \frac{z}{x}.$$

To define the divisor of ω, let $p = (a : b : c)$ be any point on C and assume $c \neq 0$. By dehomogenizing we can consider p as a point on the affine piece of C given by $P(\frac{x}{z}, \frac{y}{z}, 1) = 0$ in $\mathbb{C}^2$. As C is nonsingular, at least one of

$$\frac{\partial P}{\partial x}, \frac{\partial P}{\partial y}, \frac{\partial P}{\partial z}$$

is nonzero at $(a : b : c)$. Moreover, as $c \neq 0$, either $\frac{\partial P(r,s)}{\partial r} \neq 0$ or $\frac{\partial P(r,s)}{\partial s} \neq 0$ at $(r, s) = (\frac{a}{c}, \frac{b}{c})$. Assume $\frac{\partial P(r,s)}{\partial s} \neq 0$; then we have r as local coordinate at p. Thus we can write

$$\omega = h(r, s)\, dr$$

near p. We define the *order* n_p of $\mathrm{div}(\omega)$ at p to be the order of the divisor of the rational function $h(r, s)$ at p.

As a first example, let C be the curve $\mathrm{V}(x^2 - yz)$, and let $\omega = ds$. In Exercise 3.6.26 we showed that r is a local coordinate for all points $p = (a : b : c)$ on C with $c \neq 0$. In Exercise 3.6.30 we determined how to transform ω among the different affine charts and we showed that ω is of the form $2rdr$ for all points with $c \neq 0$.

We now use these expressions to compute the divisor of ω.

Exercise 3.6.32.

(1) Show that $2r$ has a simple zero at $(0 : 0 : 1)$ and thus the divisor of $2r$, and hence the divisor of ω in the rs-chart, is $(0 : 0 : 1)$.

(2) In Exercise 3.6.26 we showed that t is a local coordinate for C at $(0 : 1 : 0)$. Show that ω has a pole of order 3 at $(0 : 1 : 0)$.

(3) Conclude that the divisor of ω is $(0 : 0 : 1) - 3(0 : 1 : 0)$.

The above computation for the divisor of $\omega = ds$ depended on our choice of local coordinates. This divisor should be independent of the local coordinates chosen.

Exercise 3.6.33. Suppose x and y are both local coordinates at a point p on a smooth curve C. We will show that the divisor of a differential form ω does not depend on the choice of local coordinates.

(1) Suppose x and y are coordinates in the same affine patch of $\mathbb{P}^2$. We can describe C near p by $P(x,y)=0$. Use this to prove the divisors of dx and dy both have order zero at p. [Hint: Since $P(x,y)=0$, we have $0=dP=P_x dx+P_y dy$.]

(2) Suppose x and y are coordinates in two different affine patches. We can describe C near p by either $P(x,u)=0$ or $P(y,v)=0$, where (x,u) and (y,v) are two affine charts of $\mathbb{P}^2$. Use that x and u can each be written as rational functions in y and v to show that the divisors of dx and dy both have order zero at p. [Hint: This is an exercise in the chain rule. Let $f(y,v)$ be the rational function with $x=f(y,v)$. Then

$$\begin{aligned} 1 &= \frac{dx}{dx} \\ &= \frac{df}{dx} \\ &= f_y\frac{dy}{dx}+f_v\frac{dv}{dx} \\ &= f_y y_x + f_v v_x. \end{aligned}$$

Also,

$$dx = f_y dy + f_v dv,$$

which can be written as a rational function times dy, since $dv=\dfrac{-P_y}{P_v}dy$. You will also need to show and use

$$-\frac{P_y}{P_v}=\frac{v_x}{y_x}.]$$

(3) Let $\omega = h(x,u)\,dx$ in the xu-chart and $\omega = g(y,v)\,dy$ in the yv-chart. Show that the divisors of both g and h have the same order at p. [Hint: This is actually not that hard. Recall Exercise 3.3.16.]

(4) Conclude that the divisor of ω does not depend on the choice of local coordinates.

The previous exercise shows that the divisor of a differential form does not depend on the choice of local coordinates. Thus we can make the following definition.

Definition 3.6.1. The *canonical divisor* K_C on a curve C is the divisor associated to any differential form ω on C.

It is key that the linear equivalence class of the divisor K_C does not depend on the choice of differential forms, which the next exercise shows.

Exercise 3.6.34. Assume C is a nonsingular curve.

(1) Let $f, g \in \mathcal{K}(C)$. Show that $\operatorname{div} f\, dg \equiv \operatorname{div} dg$.

(2) Let ω_1, ω_2 be two differential forms on C. Show that

$$\operatorname{div}\omega_1 \equiv \operatorname{div}\omega_2.$$

[Hint: In an earlier section we showed that the vector space of differential forms over the field of rational functions is one-dimensional.]

Exercise 3.6.35. To compute the canonical divisor of the projective line $\mathbb{P}^1$, write $(x : y)$ for coordinates of $\mathbb{P}^1$, with affine charts $u = \frac{x}{y}$ and $v = \frac{y}{x}$.

(1) Show that the divisor of du is equal to $-2(1 : 0)$.

(2) Show that the divisor of dv is equal to $-2(0 : 1)$.

(3) Prove that the divisors of the two differential forms du and dv are linearly equivalent.

Exercise 3.6.36. Let $C = \mathrm{V}(x^2 - yz)$. We have already seen that

$$\operatorname{div} ds = (0 : 0 : 1) - 3(0 : 1 : 0).$$

The goal of this exercise is to show that

$$\operatorname{div} dr \equiv \operatorname{div} ds.$$

(1) Compute the divisor of the differential form dr.

(2) Prove that the divisors of the two differential forms dr and ds are linearly equivalent and of degree -2.

Exercise 3.6.37. Let C be the curve defined by $P(x, y, z) = x^2 + y^2 - z^2 = 0$. We will compute the divisor of the differential form $\omega = dr$.

(1) For points $p = (a : b : c) \in C$ with $c = 0$, show that $w = \frac{z}{x}$ is a local coordinate. Use that $r = \frac{1}{w}$ to write dr as $h(v, w)dw$. Show that there are two points on C with $w = 0$ and that $h(v, w)$ has a pole of order two at each.

(2) For points $p = (a : b : c) \in C$ with $c \neq 0$ and $\frac{\partial P}{\partial y} \neq 0$, show that r is a local coordinate. Conclude that the divisor of ω has no zeros or poles when $z \neq 0$, $\frac{\partial P}{\partial y} \neq 0$.

(3) For points $p = (a : b : c) \in C$ with $c \neq 0$ and $\frac{\partial P}{\partial y} = 0$, show that $\frac{\partial P}{\partial x} \neq 0$ and therefore $a \neq 0$. Conclude that $s = \frac{y}{z}$ is a local coordinate at these points. Use $r^2 + s^2 = 1$ to write $dr = h(r, s)ds$ and show that $h(r, s)$ has zeros of multiplicity one at each of these points.

(4) Conclude that $\operatorname{div} \omega$ is a divisor of degree -2.

In the previous exercises we found that the divisor of a differential form on a conic has degree -2. For a general smooth curve we have the following relation between genus and degree of K_C.

Theorem 3.6.38. The degree of a canonical divisor on a nonsingular curve C of genus g is $2g - 2$.

We outline a proof of this theorem in the following exercises.

Exercise 3.6.39. Let C be a nonsingular curve defined by a homogeneous polynomial $P(x, y, z)$ of degree n.

(1) Show that by changing coordinates if necessary we may assume $(1 : 0 : 0) \notin C$.

(2) Show that the curve C is covered by two copies of $\mathbb{C}^2$, $\{(a : b : c) : c \neq 0\}$ and $\{(a : b : c) : b \neq 0\}$. Conclude that at every point of C we may use either the coordinates (r, s), where $r = \frac{x}{z}, s = \frac{y}{z}$, or (t, u), where $t = \frac{x}{y}, u = \frac{z}{y}$.

(3) Let $P_1(r, s) = P(r, s, 1)$ and $P_2(t, u) = P(t, 1, u)$ be the dehomogenized polynomials defining C in the two coordinate

systems. Prove that

$$\begin{aligned}\frac{\partial P_1}{\partial r} &= \frac{\partial P}{\partial x}(r,s,1)\\ \frac{\partial P_1}{\partial s} &= \frac{\partial P}{\partial y}(r,s,1)\\ \frac{\partial P_2}{\partial t} &= \frac{\partial P}{\partial x}(t,1,u)\\ \frac{\partial P_2}{\partial u} &= \frac{\partial P}{\partial z}(t,1,u).\end{aligned}$$

(4) Explain why $(1:0:0) \notin C$ implies that $\frac{\partial P_1}{\partial r}$ has degree $n-1$.

(5) Show that by changing coordinates if necessary we may assume if $p = (a:b:c) \in C$ with $\frac{\partial P}{\partial x}(a,b,c) = 0$, then $c \neq 0$.

We will find the degree of K_C by computing the divisor of the differential form $\omega = ds$, where $s = \frac{y}{z}$. By the previous exercise we may assume $(1:0:0) \notin C$ and if $p = (a:b:c) \in C$ with $\frac{\partial P}{\partial x}(a,b,c) = 0$, then $c \neq 0$.

Exercise 3.6.40. First consider points $(a:b:c)$ on the curve with $c \neq 0$ and $\frac{\partial P}{\partial x} \neq 0$. Show that s is a local coordinate and ω has no zeros or poles at these points.

Exercise 3.6.41. Next we determine $\mathrm{div}\omega$ at points $(a:b:c)$ with $c \neq 0$ and $\frac{\partial P}{\partial x} = 0$.

(1) Show that we must have $\frac{\partial P}{\partial y} \neq 0$ at these points, and that r is a local coordinate.

(2) Use $P(r,s,1) = 0$ to write $\omega = ds$ in the form $f(r,s)\,dr$.

(3) Compute the degree of $\mathrm{div}\omega$ at these points by determining the order of the zeros or poles of $f(r,s)$.

Exercise 3.6.42. Now we determine $\mathrm{div}\omega$ at points $(a:b:c)$ with $c = 0$. By our choice of coordinates, we are assuming that $(a:b:c) \in \mathrm{V}(P)$ with $c = 0$ can happen only if $\frac{\partial P}{\partial x} \neq 0$.

(1) Show that u is a local coordinate.

(2) Write $\omega = ds$ in the form $g(t,u)du$.

(3) Compute the degree of divω at these points by determining the order of the zeros or poles of $g(t, u)$.

(4) Conclude that divω has degree $n(n-1) - 2n = n(n-3)$. Use Exercise 3.2.4 to show that this is equal to $2g-2$, where g is the genus of C.

Exercise 3.6.43. Let $C = V(xy + xz + yz)$.

(1) Find a change of coordinates to transform C to an equivalent curve C' such that $(1:0:0) \notin C'$ and such that if $p = (a : b : c) \in C$ with $c = 0$, then $\frac{\partial P}{\partial x}(a, b, c) \neq 0$.

(2) Compute the canonical divisor class of C' by computing the divisor of $\omega = ds$.

3.6.5. The Space $L(K_C - D)$. We will now see the important role that the canonical divisor plays in the Riemann-Roch Theorem. We proved previously Riemann's Theorem,

$$l(D) \geq \deg D - g + 1$$

for any divisor D on a smooth curve C of genus g. We now improve this result by determining the value of $l(D) - (\deg D - g + 1)$. We will show that for all D on C, this difference is equal to the dimension of the space $L(K_C - D)$.

We have seen for any point $p \in C$, $l(D) \leq l(D+p) \leq l(D)+1$, that is, $L(D)$ is either equal to $L(D + p)$ or is a subspace of codimension one. Applying this to the divisor $K_C - D$, we have either $l(K_C - D) = l(K_C - D - p)$ or $l(K_C - D) = l(K_C - D - p) + 1$.

For our next result we need an important consequence of the Residue Theorem: there is no differential form on C with a simple (order one) pole at one point and no other poles.

Exercise 3.6.44. We will show if $L(D) \subsetneq L(D + p)$ then $L(K_C - D - p) = L(K_C - D)$.

(1) Assume $L(D) \subsetneq L(D+p)$ and $L(K_C - D - p) \subsetneq L(K_C - D)$. Show that this implies the existence of functions $f, g \in \mathcal{K}(C)$ with div$f + D + p \geq 0$ and div$g + K_C - D \geq 0$, such that these relations are equalities at p.

(2) Let ω be a differential form on C so that $\mathrm{div}\omega \equiv K_C$. Show that $\mathrm{div} fg\omega + p \geq 0$ and thus the form $fg\omega$ has a simple pole at p.

(3) Explain why this contradicts the Residue Theorem.

(4) Show that this result is equivalent to the inequality $l(D + p) - l(D) + l(K_C - D) - l(K_C - D - p) \leq 1$.

Exercise 3.6.45. Let $q_1, \ldots, q_k$ be points on the curve C. Use the previous exercise and induction to show

$$l(D + \sum_1^k q_i) - l(D) + l(K_C - D) - l(K_C - D - \sum_1^k q_i) \leq k.$$

The next problem has nothing to do with the previous one, but does use critically that $l(D) = 0$ if D has negative degree. It will be the last step that we need before proving Riemann-Roch in the next section.

Exercise 3.6.46. Prove there exists a positive integer n such that $l(K_C - nH) = 0$, where H is a hyperplane divisor.

3.6.6. Riemann-Roch Theorem. We have previously seen Riemann's Theorem: for a divisor D on a smooth plane curve C of genus g, $l(D) \geq \deg D - g + 1$. This result provides a bound for the dimension of the space of functions on C with poles bounded by the divisor D. A remarkable fact is that we can explicitly calculate the error term in this inequality; that is, we can improve this result in the Riemann-Roch Theorem:

Theorem 3.6.47 (Riemann-Roch)**.** Let C be a smooth plane curve of genus g with canonical divisor K_C. If D is a divisor on C, then

$$l(D) - l(K_C - D) = \deg D - g + 1.$$

This theorem allows us to explicitly calculate the dimensions of spaces of functions on our curve C in terms of the genus of C and the degree of the bounding divisor D. As before we will prove this for smooth curves in the plane, but in fact the result also holds for singular curves. The Riemann-Roch Theorem can also be generalized to higher dimensional varieties. In the next several exercises we complete the proof.

Exercise 3.6.48. Let n be a positive integer with $l(K_C - nH) = 0$; use Exercise 3.6.15 to show there exists $m > n$ and $q_1, \ldots, q_k \in C$ with $D + \sum_1^k q_i \equiv mH$. Show that the degree of D is $m \deg C - k$.

Exercise 3.6.49. Using the notation from the previous exercise and Exercise 3.6.45, show that

$$l(mH) - l(D) + l(K_C - D) \leq k.$$

Exercise 3.6.50. Using the notation from the previous exercise and that

$$l(mH) = m \deg(C) - g + 1$$

(Exercise 3.6.2), show that

$$l(D) - l(K_C - D) \geq \deg D - g + 1.$$

Exercise 3.6.51. Show that

$$\deg(D) - g + 1 \geq l(D) - l(K_C - D).$$

[Hint: Think of $K_C - D$ as the divisor.]

Exercise 3.6.52. Prove the Riemann-Roch Theorem: show that

$$l(D) - l(K_C - D) = \deg D - g + 1.$$

Exercise 3.6.53. Use the Riemann-Roch Theorem to prove for a divisor D with $\deg D > 0$ on an elliptic curve, $l(D) = \deg D$.

Exercise 3.6.54. For a smooth curve C prove that the genus g is equal to the dimension of the vector space $L(K_C)$.

Exercise 3.6.55. Suppose D is a divisor of degree $2g-2$ with $l(D) = g$. Prove that D is linearly equivalent to the canonical divisor.

3.6.7. Associativity of the Group Law for a Cubic. As an application of Riemann-Roch, we will provide a more conceptual proof of associativity for the group law on a cubic curve. Starting with a smooth cubic curve C, we must show, given any three points $P, Q, R \in C$, that

$$(P + Q) + R = P + (Q + R).$$

Most of the following exercises, which were inspired by Theorem 6.39 of Kirwan [**Kir92**], depend on the material in Chapter 2. We start, though, by explaining how we will use the Riemann-Roch Theorem.

Exercise 3.6.56. Let T be a point on the smooth cubic curve C. Show that $L(T)$ is one-dimensional and conclude that the only rational functions in $L(T)$ are constant functions.

Exercise 3.6.57. Let S and T be two points on the smooth cubic curve C. Suppose there is a rational function f such that

$$\mathrm{div} f + T = S.$$

Show that $S = T$.

Let

$$S = (P + Q) + R, \quad T = P + (Q + R).$$

Here the $+$ refers to the group law addition, not the divisor addition. Our goal is to show that $S = T$.

Let

$$A = P + Q, \quad B = Q + R.$$

Again, the addition is the group law addition. Let O denote the identity element of the smooth cubic curve C.

Exercise 3.6.58. Show there exists a linear function $l_1(x, y, z)$ such that

$$\mathrm{V}(l_1) \cap C = \{P, Q, -A\}.$$

Here $-A$ refers to the inverse of A with respect to the group law of the cubic.

Exercise 3.6.59. Show there exists a linear function $l_2(x, y, z)$ such that

$$\mathrm{V}(l_2) \cap C = \{A, O, -A\}.$$

Exercise 3.6.60. Find a rational function ϕ such that

$$\mathrm{div}\phi = P + Q - A - O.$$

Here the addition is the addition for divisors.

Exercise 3.6.61. Show there exists a linear function $l_3(x, y, z)$ such that

$$\mathrm{V}(l_3) \cap C = \{A, R, -S\}.$$

Here $-S$ refers to the inverse of S with respect to the group law of the cubic.

Exercise 3.6.62. Show there exists a linear function $l_4(x, y, z)$ such that

$$\mathrm{V}(l_4) \cap C = \{S, O, -S\}.$$

Exercise 3.6.63. Find a rational function ψ such that

$$\mathrm{div}\psi = A + R - S - O.$$

Here the addition is the addition for divisors.

Exercise 3.6.64. Show that

$$\mathrm{div}\, \psi\phi = P + Q + R - S - 2O.$$

Here the addition is the addition for divisors.

Exercise 3.6.65. Following the outline of the last six exercises, find a rational function μ so that

$$\mathrm{div}\, \mu = P + Q + R - T - 2O.$$

Here the addition is the addition for divisors.

Exercise 3.6.66. Show that $\frac{\mu}{\psi\phi}$ is a rational function such that

$$\mathrm{div}\, \frac{\mu}{\psi\phi} + T = S.$$

Exercise 3.6.67. Put these exercises together to prove that the group law for cubics is associative.

3.7. Blowing Up

In this section we will study a classical technique to take any singular curve X and construct a new curve Y that is closely related to X and is smooth. This technique will be called blowing up.

We begin by describing the blow-up of the plane $\mathbb{C}^2$ at the origin. We will later see how this will apply to resolving singularities of curves. Let

$$\pi : \mathbb{C}^2 \times \mathbb{P}^1 \longrightarrow \mathbb{C}^2$$

be the projection

$$((x, y), (u : v)) \mapsto (x, y).$$

Let

$$\tilde{Y} = \{((x,y),(x:y)) \ : \ \text{at least one of } x \text{ or } y \text{ is nonzero}\} \subset \mathbb{C}^2 \times \mathbb{P}^1.$$

Set

$$Y = \tilde{Y} \cup \pi^{-1}((0,0)).$$

Exercise 3.7.1. Verify that $\pi^{-1}((0,0))$ can be identified with $\mathbb{P}^1$. Show that the restriction of π to $\tilde{Y}$ is a bijection between $\tilde{Y}$ and $\mathbb{C}^2 - (0,0)$. (Neither of these are deep.)

The set Y, along with the projection $\pi : Y \longrightarrow \mathbb{C}^2$, is called the *blow-up* of $\mathbb{C}^2$ at the point $(0,0)$. (For the rest of this section, the map π will refer to the restriction projection $\pi : Y \longrightarrow \mathbb{C}^2$.)

We look at the blow-up a bit more carefully. We can describe $\tilde{Y}$ as

$$\begin{aligned} \tilde{Y} &= \{((x,y),(x:y)) \ : \ \text{at least one of } x \text{ or } y \text{ is nonzero}\} \subset \mathbb{C}^2 \times \mathbb{P}^1 \\ &= \{((x,y),(u:v)) \in \mathbb{C}^2 \times \mathbb{P}^1 : xv = yu, (x,y) \neq (0,0)\}. \end{aligned}$$

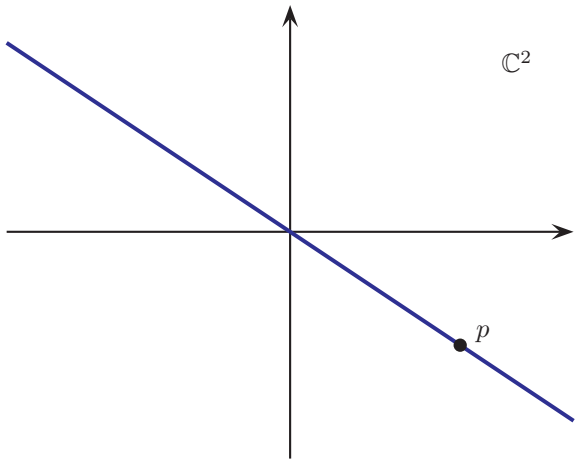
Then Y is simply

$$Y = \{((x,y),(u:v)) \in \mathbb{C}^2 \times \mathbb{P}^1 : xv = yu\}.$$

Recall that the projective line $\mathbb{P}^1$ can be thought of as all lines in $\mathbb{C}^2$ containing the origin. Thus Y is the following set:

$$\{(\text{points } p \text{ in } \mathbb{C}^2) \times (\text{lines } \ell \text{ through } (0,0)) : p \in \ell\}.$$

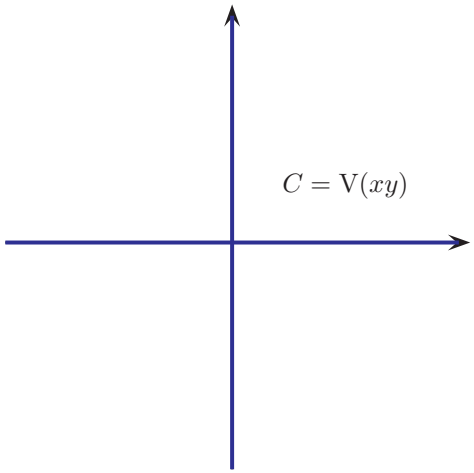
The above exercise is simply a restatement that through every point p in $\mathbb{C}^2 - (0,0)$ there is a unique line through that point and the origin.



More generally, if C is a curve in $\mathbb{C}^2$ that passes through the origin, then there is a bijection between $C-(0,0)$ and the set $\pi^{-1}(C-(0,0))$ in Y. The blow-up of C at the origin, denoted $Bl_{(0,0)}C$, is the closure of $\pi^{-1}(C-(0,0))$ in Y along with the restricted projection map.

Intuitively, $\pi^{-1}(C-(0,0))$ resembles a punctured copy of C in $\mathbb{C}^2 \times \mathbb{P}^1$, and there is an obvious way to complete this punctured curve. If the origin is a smooth point of C, then the blow-up at the origin is simply a copy of C. If the origin is a singular point, then the blow-up contains information about how the tangents to C behave near the origin.

We want to look carefully at an example. Consider $C = \mathrm{V}(xy)$ in $\mathbb{C}^2$. Here we are interested in the zero locus of $xy = 0$,



or, in other words, the union of the x-axis (when $y = 0$) and the y-axis (when $x = 0$). We will show in two ways that the blow-up of C has two points over the origin $(0, 0)$, namely $((0, 0), (1 : 0))$ corresponding to the x-axis and $((0, 0), (0 : 1))$ corresponding to the y-axis.

We know that π is a bijection away from the origin. We have

$$\pi^{-1}(C - (0, 0)) = \{(x, y) \times (x : y) : xy = 0, (x, y) \neq (0, 0)\}.$$

We know that

$$C = \mathrm{V}(xy) = \mathrm{V}(x) \cup \mathrm{V}(y).$$

We will show that there is one point over the origin on the blow-up of $\mathrm{V}(x)$ and one point (a different point) over the origin on the blow-up of $\mathrm{V}(y)$.

We have

$$\begin{aligned} \pi^{-1}(\mathrm{V}(x) - (0, 0)) &= \{((x, y), (0 : y)) : 0 = x, (x, y) \neq (0, 0)\} \\ &= \{((0, y), (0 : y)) : y \neq 0\} \\ &= \{((0, y), (0 : 1)) : y \neq 0\}. \end{aligned}$$

Then as $y \to 0$, we have

$$((0, y), (0 : 1)) \to ((0, 0), (0 : 1)),$$

a single point as desired, corresponding to the y-axis.

Similarly, we have

$$\begin{aligned}\pi^{-1}(\mathrm{V}(y)-(0,0)) &= \{((x,y),(x:0)) : y=0,(x,y)\neq(0,0)\}\\ &= \{((x,0),(x:0)) : x\neq 0\}\\ &= \{((x,0),(1:0)) : x\neq 0\}.\end{aligned}$$

Then as $x \to 0$, we have

$$((x,0),(1:0)) \to ((0,0),(1:0)),$$

a single, different point, again as desired, corresponding to the x-axis.

Now for a slightly different way of thinking of the blow-up. The projective line can be covered by two copies of $\mathbb{C}$, namely by $(u:1)$ and $(1:v)$. For any point $(u:v) \in \mathbb{P}^1$, at least one of u or v cannot be zero. If $u \neq 0$, then we have

$$(u:v) = (1:v/u)$$

while if $v \neq 0$, we have

$$(u:v) = (u/v:1).$$

In either case, we can assume that $u = 1$ or that $v = 1$.

Start with $u = 1$. We can identify $\{((x,y),(1:v))\}$ with $\mathbb{C}^3$ having coordinates x, y, v. Then the blow-up of $\mathrm{V}(xy)$ is given by the equations

$$\begin{aligned}xy &= 0\\ y &= xv\\ (x,y) &= (0,0).\end{aligned}$$

Plugging xv for y into the top equation, we have

$$x^2 v = 0.$$

Since $x \neq 0$, we can divide through by x to get

$$v = 0.$$

Then we can describe points of our curve as $((x, xv), (1 : 0)) = ((x, 0), (1 : 0))$. Therefore, as $x \to 0$, we have

$$((x, 0), (1 : 0)) \to ((0, 0), (1 : 0)).$$

Now let $v = 1$. We can identify $\{((x, y), (u : 1))\}$ with $\mathbb{C}^3$ having coordinates x, y, u. Then the blow-up of $\mathrm{V}(xy)$ is given by

$$\begin{aligned} xy &= 0 \\ yu &= x \\ (x, y) &= (0, 0). \end{aligned}$$

Plugging yu for x into the top equation, we have

$$y^2 u = 0.$$

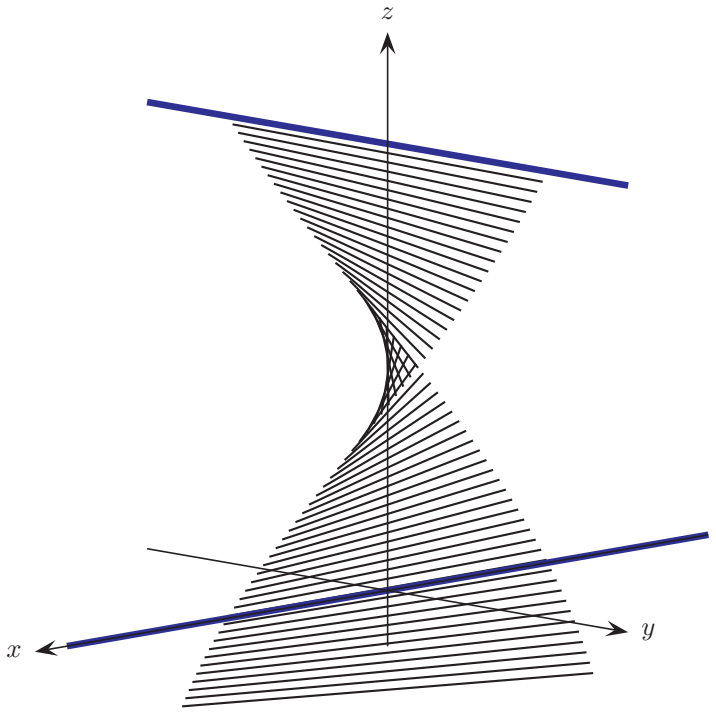
Since $y \neq 0$, we can divide through by y to get

$$u = 0.$$

Then points on our curve have the form $((yu, y), (0 : 1)) = ((0, y), (0 : 1))$. Thus as $y \to 0$, we have

$$((0, y), (0 : 1)) \to ((0, 0), (0 : 1)).$$

In either case, the blow-up looks like



Each of these techniques will be needed for the following problems.

Exercise 3.7.2. Let $C = \mathrm{V}(y - x^2)$ in $\mathbb{C}^2$. Show that this curve is smooth. Sketch this curve in $\mathbb{C}^2$. Sketch a picture of $Bl_{(0,0)}C$. Show that the blow-up projects bijectively to C.

Exercise 3.7.3. Let $C = \mathrm{V}(x^2 - y^2)$ in $\mathbb{C}^2$. Show that this curve has a singular point at the origin. Sketch this curve in $\mathbb{C}^2$. Blow up C at the origin, showing that there are two points over the origin, and then sketch a picture of the blow-up.

Exercise 3.7.4. Let $C = \mathrm{V}(y^2 - x^3 + x^2)$. Show that this curve has a singular point at the origin. Sketch this curve in $\mathbb{C}^2$. Blow up C at the origin, and sketch a picture of the blow up. Show that there are two points in the blow-up over the origin.

Exercise 3.7.5. Let $C = \mathrm{V}(y^2 - x^3)$. Show that this curve has a singular point at the origin. Sketch this curve in $\mathbb{C}^2$. Blow up C at the origin, and sketch a picture of the blow-up. Show that there is only one point over the origin.

Exercise 3.7.6. Let $C = \mathrm{V}((x - y)(x + y)(x + 2y))$ be a curve in $\mathbb{C}^2$. Show that this curve has a singular point at the origin. Sketch this curve in $\mathbb{C}^2$. Blow up C at the origin, and sketch a picture of the blow-up. Show that there are three points over the origin.

The previous exercises should convey the idea that if the original curve is singular at the origin, then the blow-up seems to be less singular at its points over the origin. We currently can't express precisely what this means, since our definition of singularity applies only to curves in the plane, and the blow-up does not lie in a plane. Algebraic ideas developed in Chapter 4 will allow us to make this idea precise.

Of course, there is nothing special about the origin in affine space, and we could just as easily blow up curves at any other point. Also, the definition of blowing up can easily be extended to curves in projective spaces. A significant part of current algebraic geometry involves resolving singularities of more complicated algebraic varieties.

Chapter 4

Affine Varieties

The goal of this chapter is to use abstract algebra to describe the geometry of curves, surfaces, and more general geometric objects called varieties. By considering the set of points where a polynomial vanishes, we will see there is a correspondence between the algebra of polynomials and the geometry of points on a variety. This chapter is devoted to understanding this correspondence. Here tools from abstract algebra, especially commutative ring theory, will become key. You will need to know some basic facts about rings and ideals, which can be found in most undergraduate abstract algebra texts.

4.1. Zero Sets of Polynomials

We begin the study of affine varieties with some examples of zero sets of polynomials.

The natural ambient space for affine varieties is affine space.

Definition 4.1.1. For a field k, the *affine n-space over k* is the set

$$\mathbb{A}^n(k) = \{(a_1, a_2, \ldots, a_n) : a_i \in k \quad \text{for} \quad i = 1, \ldots, n\}.$$

We write simply $\mathbb{A}^n$ when the field k is understood.

For example, $\mathbb{A}^2(\mathbb{R})$ is the familiar Euclidean plane $\mathbb{R}^2$ from calculus, and $\mathbb{A}^1(\mathbb{C})$ is the complex line. We are interested in subsets of $\mathbb{A}^n$ that are the zero sets of a collection of polynomials over k.

Recall that $k[x_1, x_2, \ldots, x_n]$ is the commutative ring of all polynomials in the variables $x_1, x_2, \ldots, x_n$ with coefficients in the field k. Frequently for us, our field will be the complex numbers $\mathbb{C}$, with the field of the real numbers $\mathbb{R}$ being our second most common field.

The goal of this chapter is to explore the link between the zero sets of polynomials in $\mathbb{A}^n$ and the ideals of $k[x_1, x_2, \ldots, x_n]$.

Exercise 4.1.1. Describe or sketch the zero set of each polynomial over $\mathbb{C}$.

(1) $x^2 + 1$

(2) $y - x^2$

Exercise 4.1.2. Show that the zero set of $x^2 + y^2 - 1$ in $\mathbb{C}^2$ is unbounded, in contrast with the zero set of $x^2 + y^2 - 1$ in $\mathbb{R}^2$.

Exercise 4.1.3. Find a set of polynomials $\{P_1, \ldots, P_n\}$, all of whose coefficients are real numbers, whose common zero set is the given set.

(1) $\{(3, y) : y \in \mathbb{R}\}$ in $\mathbb{R}^2$

(2) $\{(1, 2)\}$ in $\mathbb{R}^2$

(3) $\{(1, 2), (0, 5)\}$ in $\mathbb{R}^2$

(4) Generalize the method from Part (3) to any finite set of points in $\mathbb{R}^2$.

Exercise 4.1.4. Find a set of polynomials $\{P_1, \ldots, P_n\}$, all of whose coefficients are complex numbers, whose common zero set is the given set.

(1) $\{(3 + 2i, -i)\}$ in $\mathbb{C}^2$

(2) $\{(3 + 2i, -i), (0, 1 - 4i)\}$ in $\mathbb{C}^2$

(3) Generalize the method from Part (2) to any finite set of points in $\mathbb{C}^2$.

Exercise 4.1.5.

(1) Is every finite subset of $\mathbb{C}^2$ the zero set of a collection of polynomials in $\mathbb{C}[x, y]$? Prove or find a counterexample.

(2) Is there an infinite subset of $\mathbb{C}^2$ that is the common zero set of a finite collection of polynomials in $\mathbb{C}[x, y]$?

(3) Find an infinite set of points in $\mathbb{C}$ that is not the common zero set of a finite collection of polynomials in $\mathbb{C}[x]$.

(4) Is there any infinite set of points in $\mathbb{C}$, besides $\mathbb{C}$ itself, that is the common zero set of a finite collection of polynomials in $\mathbb{C}[x]$?

Much of the reason that modern algebraic geometry heavily influences not only geometry but also number theory is that we can allow our coefficients to be in any field, even those for which no geometry is immediately apparent. As we saw in earlier chapters, our fields do not even need to be infinite!

Exercise 4.1.6. Find the zero set of each polynomial in $\mathbb{A}^1(\mathbb{Z}_3)$.

(1) $x^2 + 2$

(2) $x^2 - 2$

Exercise 4.1.7. Find the zero set of each polynomial in $\mathbb{A}^2(\mathbb{Z}_5)$.

(1) $y - x^2$

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

(3) $xy - 3y - x^2 + 3x$

Exercise 4.1.8.

(1) Show that if k is an infinite field, and $P \in k[x_1, \ldots, x_n]$ is a polynomial whose zero set is $\mathbb{A}^n(k)$, then $P = 0$. [Hint: Use induction on n.]

(2) Is there any finite field for which this result holds?

4.2. Algebraic Sets and Ideals

In this section we explore the relationship between algebraic sets in $\mathbb{A}^n$ and ideals in $k[x_1, \ldots, x_n]$.

4.2.1. Algebraic Sets. The zero sets of polynomials in affine space are called algebraic sets.

Definition 4.2.1. Let $S \subseteq k[x_1, \ldots, x_n]$ be a set of polynomials. The *algebraic set* defined by S is

$$\mathrm{V}(S) = \{(a_1, a_2, \ldots, a_n) \in \mathbb{A}^n(k) : P(a_1, a_2, \ldots, a_n) = 0 \quad \text{for all} \quad P \in S\}.$$

Exercise 4.2.1. Sketch the algebraic sets.

(1) $\mathrm{V}(x^3 - 1)$ in $\mathbb{A}^1(\mathbb{C})$
(2) $\mathrm{V}((y - x^2)(y^2 - x))$ in $\mathbb{A}^2(\mathbb{R})$
(3) $\mathrm{V}(y - x^2, y^2 - x)$ in $\mathbb{A}^2(\mathbb{R})$
(4) $\mathrm{V}(y^2 - x^3 + x)$ in $\mathbb{A}^2(\mathbb{R})$
(5) $\mathrm{V}(x - 2y + 3z)$ in $\mathbb{A}^3(\mathbb{R})$
(6) $\mathrm{V}(z - 3, z - x^2 - y^2)$ in $\mathbb{A}^3(\mathbb{R})$
(7) $\mathrm{V}(xy - z^2 y) = \mathrm{V}(y(x - z^2))$ in $\mathbb{A}^3(\mathbb{R})$
(8) $\mathrm{V}(y - x + x^2)$ in $\mathbb{A}^2(\mathbb{Z}_3)$

Exercise 4.2.2. Algebraic Sets in $\mathbb{R}^n$ and $\mathbb{C}^n$:

(1) Show that for any $a \in \mathbb{R}$, the singleton $\{a\}$ is an algebraic set.
(2) Show that any finite collection of numbers $\{a_1, a_2, \ldots, a_k\}$ in $\mathbb{R}$ is an algebraic set.
(3) Show that a circle in $\mathbb{R}^2$ is an algebraic set.
(4) Show that the set $\{(-1/\sqrt{2}, 1/\sqrt{2}), (1/\sqrt{2}, -1/\sqrt{2})\} \subset \mathbb{R}^2$ is an algebraic set.
(5) Show that any line in $\mathbb{R}^3$ is an algebraic set.
(6) Show that the positive numbers are not an algebraic set in $\mathbb{R}$.
(7) Show that the region inside the unit circle $|z| < 1$ in $\mathbb{C}$ is not an algebraic set.
(8) Give an example of a nonconstant polynomial P in $\mathbb{R}[x, y]$ such that the algebraic set $X = \{(x, y) \in \mathbb{R}^2 : P(x, y) = 0\}$ is the empty set.

(9) Is there a nonconstant polynomial P in $\mathbb{C}[x, y]$ such that the algebraic set $X = \{(x, y) \in \mathbb{C}^2 | P(x, y) = 0\}$ is the empty set? Explain why or why not.

(10) Suppose $X_1 = \{(x, y) \in \mathbb{C}^2 | x + y = 0\}$ and $X_2 = \{(x, y) \in \mathbb{C}^2 | x - y = 0\}$. Find a polynomial $Q \in \mathbb{C}[x, y]$ such that $X_1 \cup X_2 = \{(x, y) \in \mathbb{C}^2 | Q(x, y) = 0\}$.

(11) Suppose $X_1 = \{(x_1, x_2, \ldots, x_n) \in \mathbb{C}^n | P_1(x_1, x_2, \ldots, x_n) = 0\}$ and $X_2 = \{(x_1, x_2, \ldots, x_n) \in \mathbb{C}^n | P_2(x_1, x_2, \ldots, x_n) = 0\}$. Give a single polynomial Q such that

$$X_1 \cup X_2 = \{(x_1, x_2, \ldots, x_n) \in \mathbb{C}^n | Q(x_1, x_2, \ldots, x_n) = 0\}.$$

Exercise 4.2.3.

(1) Is every finite subset of $\mathbb{A}^2(\mathbb{R})$ an algebraic set?

(2) Is every finite subset of $\mathbb{A}^2(\mathbb{C})$ an algebraic set?

Exercise 4.2.4. Show that the set $\{(x, y) \in \mathbb{A}^2(\mathbb{R}) : 0 \leq x \leq 1, y = 0\}$ is not an algebraic set. [Hint: Any one-variable polynomial, which is not the zero polynomial, can have only a finite number of roots.]

Exercise 4.2.5. Show that both the empty set and $\mathbb{A}^n(k)$ are algebraic sets in $\mathbb{A}^n(k)$.

Exercise 4.2.6. Show that if $X = \mathrm{V}(f_1, \ldots, f_s)$ and $W = \mathrm{V}(g_1, \ldots, g_t)$ are algebraic sets in $\mathbb{A}^n(k)$, then $X \cup W$ and $X \cap W$ are algebraic sets in $\mathbb{A}^n(k)$.

4.2.2. Zero Sets via $\mathrm{V}(I)$. We next will see how to define algebraic sets using ideals in the polynomial ring.

Exercise 4.2.7. Let $f(x, y), g(x, y) \in \mathbb{C}[x, y]$. Show that

$$\mathrm{V}(f, g) = \mathrm{V}(f - g, f + g).$$

Exercise 4.2.8. Show that $\mathrm{V}(x + y, x - y, 2x + y^2, x + xy + y^3, y + x^2 y) = \mathrm{V}(x, y)$.

Thus the polynomials that define a zero set are far from being unique. But there is an algebraic object that comes close to being uniquely defined by a zero set.

The following exercise is key to algebraic geometry.

Exercise 4.2.9. Let I be the ideal in $k[x_1, \ldots, x_n]$ generated by a set $S \subset k[x_1, \ldots, x_n]$. Show that $\mathrm{V}(S) = \mathrm{V}(I)$. Thus every algebraic set is defined by an ideal.

While it is not quite true that the set $\mathrm{V}(I)$ uniquely determines the ideal I, we will soon see how to restrict our class of ideals so that the associated ideal will be unique.

Exercise 4.2.10. For $f(x_1, \ldots, x_n), g(x_1, \ldots, x_n) \in \mathbb{C}[x_1, \ldots, x_n]$, let I be the ideal generated by f and g and let J be the ideal generated by f alone.

(1) Show that $J \subset I$.

(2) Show that $\mathrm{V}(I) \subset \mathrm{V}(J)$.

Exercise 4.2.11. Show that if I and J are ideals in $k[x_1, \ldots, x_n]$ with $I \subset J$, then $\mathrm{V}(I) \supset \mathrm{V}(J)$.

Exercise 4.2.12. You may find Exercise 4.2.6 useful here.

(1) Show that an arbitrary intersection of algebraic sets is an algebraic set.

(2) Show that a finite union of algebraic sets is an algebraic set.

We will see in Section 4.11 that the algebraic sets can be used to help define a topology.

4.2.3. Ideals Associated to Zero Sets. We have seen that the set of polynomials that define a zero set is not unique. While an ideal uniquely determines an algebraic set, the converse is not true.

Definition 4.2.2. Let V be a set of points in $\mathbb{A}^n(k)$. The *ideal* of V is given by

$$I(\mathrm{V}) = \{P \in k[x_1, \ldots, x_n] : P(a_1, \ldots, a_n) = 0 \quad \text{for all} \quad (a_1, \ldots, a_n) \in V\}.$$

We will be most interested when V is an algebraic set.

Exercise 4.2.13. Show that $I(V)$ is an ideal in the ring $k[x_1, \ldots, x_n]$.

Exercise 4.2.14. Let X be a set of points in $\mathbb{A}^n(\mathbb{C})$.

(1) Show that $X \subseteq \mathrm{V}(I(X))$.

(2) Find a set X with $X \neq \mathrm{V}(I(X))$.

(3) Show that if X is a algebraic set, then $X = \mathrm{V}(I(X))$.

Exercise 4.2.15. Let I be an ideal in $k[x_1, \ldots, x_n]$.

(1) Show that $I \subseteq I(\mathrm{V}(I))$.

(2) Find an ideal I with $I \neq I(\mathrm{V}(I))$.

(3) Show that if I is the ideal of an algebraic set, then $I = I(\mathrm{V}(I))$.

It looks as if there is a correspondence between algebraic sets and some ideals.

Definition 4.2.3. Let I be an ideal in $k[x_1, \ldots, x_n]$. The *radical of* I is defined as

$$\mathrm{Rad}(I) = \{P \in k[x_1, \ldots, x_n] : P^m \in I \text{ for some } m > 0\}.$$

An ideal I is called a *radical ideal* if $I = \mathrm{Rad}(I)$.

Exercise 4.2.16. Let $f(x, y) = (x^2 - y + 3)^2 \in \mathbb{C}[x, y]$. Show that the ideal I generated by f is not radical. Find $\mathrm{Rad}(I)$.

Exercise 4.2.17. Let I be an ideal in $k[x_1, \ldots, x_n]$. Show that $\mathrm{Rad}(I)$ is an ideal.

Exercise 4.2.18. Let X be a set of points in $\mathbb{A}^n(k)$. Show that $I(X)$ is a radical ideal.

Exercise 4.2.19. Show that $\mathrm{Rad}(I) \subset I(\mathrm{V}(I))$ for any ideal I in $k[x_1, \ldots, x_n]$.

Exercise 4.2.20. Let I be an ideal in $k[x_1, \ldots, x_n]$. Show $\mathrm{V}(I) = \mathrm{V}(\mathrm{Rad}(I))$.

Exercise 4.2.21. Let X and W be algebraic sets in $\mathbb{A}^n(k)$. Show that $X \subset W$ if and only if $I(X) \supset I(W)$. Conclude that $X = W$ if and only if $I(X) = I(W)$.

4.3. Hilbert Basis Theorem

The goal of this section is prove the Hilbert Basis Theorem, which has as a consequence that every ideal in $k[x_1, \ldots, x_n]$ is finitely generated.

How many polynomials are needed to define an algebraic set $V \subset \mathbb{A}^n$? Is there a finite number of polynomials $f_1, f_2, \ldots, f_m$ such that

$$V = \{a \in \mathbb{A}^n : f_i(a) = 0, \forall 1 \leq i \leq m\},$$

or are there times that we would always need an infinite number of defining polynomials?

Exercise 4.3.1. Let $\mathrm{V}(x^2 + y^2 - 1)$. Show that $I(\mathrm{V})$ contains an infinite number of elements.

We know that there are an infinite number of possible defining polynomials, but do we need all of them to define V? In the above exercise, all we need is the single $x^2 + y^2 - 1$ to define the entire algebraic set. If there are times when we need an infinite number of defining polynomials, then algebraic geometry would be extremely hard. Luckily, the Hilbert Basis Theorem has as its core that we need only a finite set of polynomials to generate any ideal. The rest of this section will be pure algebra.

A (commutative) ring R is said to be *Noetherian* if every ideal I in R is finitely generated. (Recall that all rings considered in this book are commutative.)

Exercise 4.3.2. Show that every field and principal ideal domain (PID) is Noetherian. (Recall that a ring is a PID if whenever $x \cdot y = 0$, then $x = 0$ or $y = 0$, and every nontrivial ideal is generated by a single element.)

Exercise 4.3.3. Let R be a ring. Prove that the following three conditions are equivalent:

(1) R is Noetherian.

(2) Every ascending chain $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$ of ideals in R is stationary, i.e., there exists N such that for all $n \geq$

N, $I_n = I_N$. This is called the ascending chain condition (ACC).

(3) Every nonempty set of ideals in R has a maximal element (with inclusion being the ordering between ideals). This means that if we have a set of ideals $\{I_1, I_2, \ldots\}$, there must be at least one ideal in the set, say I_k, such that there is no I_n in the set with

$$I_k \subsetneq I_n.$$

(There can be more than one maximal element.)

In what follows, we guide the reader through a proof of the Hilbert Basis Theorem.

Theorem 4.3.4 (Hilbert Basis Theorem)**.** If R is a Noetherian ring, then $R[x]$ is also a Noetherian ring.

Sketch of proof. Let $I \subset R[x]$ be an ideal of $R[x]$. We show I is finitely generated.

Step 1. Let f_1 be a nonzero element of least degree in I.

Step 2. For $i > 1$, let f_i be an element of least degree in $I - \langle f_1, \ldots, f_{i-1} \rangle$, if possible.

Step 3. For each i, write $f_i = a_i x^{d_i} +$ lower order terms. That is, let a_i be the leading coefficient of f_i. Set $J = \langle a_1, a_2, \ldots \rangle$.

Step 4. Since R is Noetherian, $J = \langle a_1, \ldots, a_m \rangle$ for some m.

Exercise 4.3.5. Justify Step 4.

Step 5. Claim that $I = \langle f_1, \ldots, f_m \rangle$. If not, there is an f_{m+1}, and we can subtract off its leading term using elements of $\langle f_1, \ldots, f_m \rangle$ to get a contradiction.

Exercise 4.3.6. Fill in the details of Step 5.

Exercise 4.3.7. Show that if R is Noetherian, then $R[x_1, \ldots, x_n]$ is Noetherian.

The Hilbert Basis Theorem shows that the ideal of any variety is finitely generated.

Working over $\mathbb{R}$ we can show that every variety is defined by a single polynomial.

First an example

Exercise 4.3.8. Show in $\mathbb{A}^2(\mathbb{R})$ that

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

Since we are working over the real numbers, the only way for

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

is for $(y-x^2)^2 = 0$ and for $(x-y^2)^2 = 0$.

Exercise 4.3.9. Let $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_n]$ and $V = \mathrm{V}(f_1, \ldots, f_k) \subset \mathbb{A}^k(\mathbb{R})$. Show that there is a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ such that $\mathrm{V}(f_1, \ldots, f_k) = \mathrm{V}(f)$. Give an example to show that this fails over $\mathbb{C}$.

4.4. The Strong Nullstellensatz

The goal of this section is to start the proof of Hilbert's Nullstellensatz, which shows that there is a one-to-one correspondence between algebraic sets in $\mathbb{A}^n(k)$ and radical ideals when k is algebraically closed. In this section we will prove the Strong Nullstellensatz, under the assumption of the Weak Nullstellensatz, which we prove in the next section.

We saw in Exercise 4.2.20 that given any ideal $I \subset \mathbb{C}[x_1, x_2, \ldots, x_n]$,

$$\mathrm{V}(I) = \mathrm{V}(\mathrm{Rad}(I)).$$

Could there be some other ideal $J \subset k[x_1, x_2, \ldots, x_n]$, with $\mathrm{V}(J) = \mathrm{V}(I)$ but $\mathrm{Rad}(I) \neq \mathrm{Rad}(J)$? The punch line for this section is that this is impossible when we work over an algebraically closed field.

The goal is to prove the Strong Nullstellensatz. (We will primarily be following the proof shown by Arrondo [**Arr06**], with a nod to Fulton [**Ful69**].) This is one of the key results in algebraic geometry. To some extent, if it were not true, then the subject would probably not be studied. It provides a clean connection between points in affine space with radical ideals in polynomial rings.

Theorem 4.4.1 (Strong Nullstellensatz)**.** Let k be an algebraically closed field and let I be an ideal of the polynomial ring $k[x_1, \ldots, x_n]$. Then

$$I(\mathrm{V}(I)) = \mathrm{Rad}(I).$$

(By the way, the word *Nullstellensatz* is German for "Theorem of Zeros.")

By Exercise 4.2.19, it is always true that $\mathrm{Rad}(I) \subseteq I(\mathrm{V}(I))$. Thus, to prove the Strong Nullstellensatz, we must show $I(\mathrm{V}(I)) \subseteq \mathrm{Rad}(I)$. For this section, suppose that k is an algebraically closed field and I is an ideal in $k[x_1, \ldots, x_n]$ such that I is generated by the polynomials $f_1, \ldots, f_r$. Suppose that g is a polynomial such that $g \in I(V(I))$. We must find a positive integer N and polynomials $A_1, \ldots, A_r$ such that

$$g^N = A_1 f_1 + \cdots + A_r f_r,$$

for then $g \in \mathrm{Rad}(I)$, as desired.

We will show this in the next few exercises, under the following assumption, which we will prove in the next section:

Theorem 4.4.2 (Weak Nullstellensatz—Version 1)**.** Let k be an algebraically closed field and let I be a proper ideal of the polynomial ring $k[x_1, \ldots, x_n]$. Then

$$\mathrm{V}(I) \neq \emptyset,$$

i.e., there exists $(a_1, \ldots, a_n) \in k^n$ such that $f(a_1, \ldots, a_n) = 0$ for all $f \in I$.

Starting with our original ideal I and the polynomial $g \in I(V(I))$, define a new ideal J in the slightly larger polynomial ring $k[x_1, \ldots, x_n, x_{n+1}]$, by setting

$$J = \langle f_1, \ldots, f_r, x_{n+1}g - 1 \rangle.$$

Exercise 4.4.3. Suppose $g \in I(V(I))$, or in other words $g(a_1, \ldots, a_n) = 0$ for all $(a_1, \ldots, a_n) \in \mathrm{V}(\langle f_1, \ldots, f_r \rangle)$. For $J = \langle f_1, \ldots, f_r, x_{n+1}g - 1 \rangle$, show $\mathrm{V}(J) = \emptyset$.

Exercise 4.4.4. Assuming the Weak Nullstellensatz, show that J is not a proper ideal and hence that there exist $A_1, \ldots, A_r, B$ in $k[x_1, \ldots, x_n, x_{n+1}]$ such that

$$1 = A_1 f_1 + A_2 f_2 + \cdots + B(x_{n+1} g - 1).$$

Exercise 4.4.5. Let $x_{n+1} = \dfrac{1}{y}$. Show there exists $N > 0$ and polynomials $C_1, \ldots, C_r, D$ in $k[x_1, \ldots, x_n, y]$ with

$$y^N = C_1 f_1 + \cdots + C_r f_r + D(g - y)$$

by clearing denominators.

Exercise 4.4.6. Letting $y = g$ show that $g^N \in I$ and hence $g \in \mathrm{Rad}(I)$.

4.5. The Weak Nullstellensatz

Here we prove the Weak Nullstellensatz, which states that an algebraic set for any proper ideal cannot be empty. The use of resultants will be critical. We will then give another formulation of the Weak Nullstellensatz, which we will call Version 2.

From the last section recall the statement:

Theorem (Weak Nullstellensatz—Version 1)**.** Let k be an algebraically closed field and let I be a proper ideal of the polynomial ring $k[x_1, \ldots, x_n]$. Then $\mathrm{V}(I) \neq \emptyset$, i.e., there exists $(a_1, \ldots, a_n) \in k^n$ such that $f(a_1, \ldots, a_n) = 0$ for all $f \in I$.

Note that if I is not proper, the result is false, since if I is not proper, then I must be the entire polynomial ring $k[x_1, \ldots, x_n]$. In particular, the constant function 1 must then be an element of I. Since the constant function 1 has no zeros, this would mean that $\mathrm{V}(I) = \emptyset$.

The proof of the Weak Nullstellensatz will take some work. Our argument will use induction on n, the number of variables of our polynomial ring. In the one variable case, we start with a proper ideal I in $k[x_1]$. As in the last section we assume k is an algebraically closed field.

For one variable polynomials, the Euclidean Algorithm can be used to show I is principal, that is, I can be generated by a single polynomial.

Exercise 4.5.1. Prove the Weak Nullstellensatz for $n = 1$.

We now assume the Weak Nullstellensatz for ideals in $k[x_1, \ldots, x_m]$ for any $1 \leq m < n$. We must prove that every proper ideal I in $k[x_1, \ldots, x_n]$ has a non-empty zero set.

Exercise 4.5.2. Let I be a proper ideal of $k[x_1, \ldots, x_n]$ and let I' be the set of all polynomials in I that do not contain the variable x_n. Prove that I' is a proper ideal of $k[x_1, \ldots, x_{n-1}]$.

Thus I' is a proper ideal in the polynomial ring $k[x_1, \ldots, x_{n-1}]$. Under our induction hypothesis, we must have

$$\mathrm{V}(I') \neq \emptyset.$$

Let $(a_1, \ldots, a_{n-1}) \in \mathrm{V}(I')$ be one of these points. We want to show that there is at least one $a_n \in k$ so that

$$(a_1, \ldots, a_{n-1}, a_n) \in \mathrm{V}(I).$$

Set

$$J = \{f(a_1, \ldots, a_{n-1}, x_n) : f \in I\}$$

in $k[x_n]$. If this J is a proper ideal, by induction our desired a_n must exist.

Exercise 4.5.3. Show that J is an ideal in $k[x_n]$.

Exercise 4.5.4. Under the assumption that J is proper, prove the Weak Nullstellensatz.

Thus to complete the proof of the Weak Nullstellensatz we must show that J, our newly created ideal in $k[x_n]$, is a proper ideal. This will be the technical heart of the proof, and this is where we will use resultants. We first will show there exists a polynomial in our original ideal I that is monic in x_n. (By *monic*, we mean that the coefficient of the highest power for x_n is one.) Why we do so will be apparent in a moment. We will first consider a concrete example and then turn to the general case.

Exercise 4.5.5. Let $g(x_1, x_2, x_3, x_4) = x_1x_2 + x_3x_4$. Prove that there exist $\lambda_1, \lambda_2, \lambda_3$ such that the coefficient of x_4^2 in $g(x_1 + \lambda_1 x_4, x_2 + \lambda_2 x_4, x_3 + \lambda_3 x_4, x_4)$ is nonzero.

Exercise 4.5.6. Let $I \subset k[x_1, \ldots, x_4]$ be an ideal containing the polynomial $g(x_1, x_2, x_3, x_4) = x_1x_2 + x_3x_4$. Prove that there is a change of coordinates so that I contains a polynomial monic in the variable x_4.

Exercise 4.5.7. Let k be an infinite field and g be a nonconstant polynomial in $k[x_1, \ldots, x_n]$ (with $n \geq 2$). Prove that there exist $\lambda_1, \ldots, \lambda_{n-1}$ in k such that the coefficient of x_n^d in $g(x_1 + \lambda_1 x_n, \ldots, x_{n-1} + \lambda_{n-1}x_n, x_n)$ is nonzero, where d is the total degree of $g(x_1 + \lambda_1 x_n, \ldots, x_{n-1} + \lambda_{n-1}x_n, x_n)$.

Exercise 4.5.8. Let k be an infinite field and I be a proper ideal of $k[x_1, \ldots, x_n]$. Prove that there is a change of coordinates so that I contains a polynomial g that is monic in the variable x_n.

Thus we can always assume that I has at least one polynomial g that is monic in the variable x_n.

We now return to show that our J is a proper ideal in $k[x_n]$. We will assume J is not proper and use a resultant to obtain a contradiction.

Exercise 4.5.9. Assume J is not proper. Show there exists $f \in I$ such that

$$f(x_1, \ldots, x_n) = f_0(x_1, \ldots, x_{n-1}) + f_1(x_1, \ldots, x_{n-1})x_n + \cdots + f_d(x_1, \ldots, x_{n-1})x_n^d$$

with $f_0(a_1, \ldots, a_{n-1}) = 1, f_i(a_1, \ldots, a_{n-1}) = 0$ for $1 \leq i \leq d$. [Hint: If J is not proper, then $1 \in J$.]

We now use this expression of f as a polynomial in x_n. Recall Definition 3.3.4.

Fixing a monic $g \in I$, we can similarly write

$$g(x_1, \ldots, x_n) = g_0(x_1, \ldots, x_{n-1}) + g_1(x_1, \ldots, x_{n-1})x_n + \cdots + x_n^e.$$

Consider the resultant

$$R = Res(f, g; x_n) = \det \begin{pmatrix} f_0 & f_1 & \cdots & f_d & 0 & 0 & \cdots & 0 \\ 0 & f_0 & f_1 & \cdots & f_d & 0 & \cdots & 0 \\ & & & & & & & \\ 0 & 0 & \ddots & \ddots & \cdots & \ldots & \ldots & 0 \\ 0 & 0 & \cdots & f_0 & f_1 & \cdots & \cdots & f_d \\ g_0 & g_1 & \cdots & g_{e-1} & 1 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & \cdots & g_{e-1} & 1 & \cdots & 0 \\ & & & & & & & \\ 0 & 0 & \ddots & \ddots & \cdots & \ldots & \ldots & 0 \\ 0 & 0 & \cdots & g_0 & \cdots & \cdots & g_{e-1} & 1 \end{pmatrix}.$$

(Note: The above matrix is the transpose, and then a switching of the order of a few columns, of the matrix that we used in 3.3.4. Since this will not change where the determinant is zero, the above is still the resultant. This formatting will make the following proof a bit easier to follow.) We want to show $R \in I$.

Exercise 4.5.10. Replace the first column in the resultant matrix by

$$1\text{st column} + x_n \cdot 2\text{nd column} + \cdots + x_n^{d+e-1} \cdot \text{last column}.$$

Show that the first column becomes

$$\begin{bmatrix} f(x_1, \ldots, x_{n_1}, x_n) \\ x_n f(x_1, \ldots, x_{n_1}, x_n) \\ \vdots \\ x_n^{e-1} f(x_1, \ldots, x_{n_1}, x_n) \\ g(x_1, \ldots, x_{n_1}, x_n) \\ \vdots \\ x_n^{d-1} g(x_1, \ldots, x_{n_1}, x_n) \end{bmatrix}.$$

Exercise 4.5.11. By expanding along the first column of the matrix obtained from the column operation in the previous exercise, show $R \in I$.

Exercise 4.5.12. Show $R \in I'$.

Exercise 4.5.13. Show $R(a_1, \ldots, a_{n-1}) = 1$.

Exercise 4.5.14. Prove that the ideal J must be proper.

Thus we have proven the Weak Nullstellensatz, our final step in the proof that $I(\mathrm{V}(I)) = \mathrm{Rad}(I)$.

We next give an alternate version of the Weak Nullstellensatz that will be useful later.

Exercise 4.5.15 (Weak Nullstellensatz—Version 2)**.** Let k be an algebraically closed field. An ideal I in $k[x_1, \ldots, x_n]$ is maximal if and only if there are elements $a_i \in k$ such that I is the ideal generated by the elements $x_i - a_i$; that is $I = \langle x_1 - a_1, \ldots, x_n - a_n \rangle$.

In our last exercises of this section we see that the Nullstellensatz can fail when the field k is not algebraically closed.

Exercise 4.5.16. Let $I = \langle x^2 + 1 \rangle \in \mathbb{R}[x]$ and show that $I(\mathrm{V}(I)) \neq \mathrm{Rad}(I)$.

Exercise 4.5.17. Show that $I = \langle x^2 + y^2 \rangle$ and $J = \langle x, y \rangle$ are radical ideals in $\mathbb{R}[x, y]$ with $\mathrm{V}(I) = \mathrm{V}(J)$. This demonstrates that the correspondence between algebraic sets and radical ideals is not one-to-one over $\mathbb{R}$.

4.6. Points in Affine Space as Maximal Ideals

In this section we give a geometric interpretation of the Weak Nullstellensatz to establish a correspondence between points in affine space and maximal ideals in $k[x_1, \ldots, x_n]$.

Exercise 4.6.1. Show that for $a_1, a_2, \ldots, a_n \in k$, the ideal $I \subset k[x_1, \ldots, x_n]$ defined as

$$I = \langle x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n \rangle$$

is maximal. [Hint: Suppose J is an ideal with $I \subsetneq J$, and show that J contains 1.]

Exercise 4.6.2. Show that $I(\{(a_1, \ldots, a_n)\}) = \langle x_1 - a_1, \ldots, x_n - a_n \rangle$.

Exercise 4.6.3. Show that if an ideal $I \subset k[x_1, \ldots, x_n]$ is maximal, then $\mathrm{V}(I)$ is either a point or empty.

Note that we are not requiring k to be algebraically closed. This is why $\mathrm{V}(I)$ can be empty. When k is algebraically closed, the Weak Nullstellensatz shows that $\mathrm{V}(I) \neq \emptyset$.

Exercise 4.6.4. Find a maximal ideal $I \subset \mathbb{R}[x_1, \ldots, x_n]$ for which $\mathrm{V}(I) = \emptyset$.

Such a maximal ideal in $k[x_1, \ldots, x_n]$ cannot exist when k is algebraically closed, for all maximal ideals are then of the form $\langle x_1 - a_1, \ldots, x_n - a_n \rangle$ for some $a_1, \ldots, a_n \in k$ by the Weak Nullstellensatz (Theorem 4.5.15). However, in Exercises 4.6.2 and 4.6.3, we saw that such maximal ideals correspond to points in $\mathbb{A}^n(k)$. This proves the following important fact.

Theorem 4.6.5. In an algebraically closed field k, there is a one-to-one correspondence between points of $\mathbb{A}^n(k)$ and maximal ideals of $k[x_1, \ldots, x_n]$.

4.7. Affine Varieties and Prime Ideals

The goal of this section is to define affine varieties and to show that they correspond to prime ideals.

Throughout this section we assume that k is an algebraically closed field.

4.7.1. Irreducible Components. An algebraic set V is *reducible* if

$$V = V_1 \cup V_2,$$

where V_1 and V_2 are algebraic sets with $V_1 \subsetneq V$ and $V_2 \subsetneq V$. An algebraic set that is not reducible is said to be *irreducible*.

Definition 4.7.1. An *affine variety* is an irreducible algebraic set in $\mathbb{A}^n$, for some n.

Exercise 4.7.1. Show that $\mathbb{A}^1(\mathbb{C})$ is irreducible, so $\mathbb{A}^1(\mathbb{C})$ is an affine variety.

Exercise 4.7.2. Decide if the following algebraic sets in $\mathbb{A}^2$ are reducible or irreducible.

(1) $\mathrm{V}(x)$

(2) $\mathrm{V}(x + y)$

(3) $\mathrm{V}(xy)$

Exercise 4.7.3. Let $f \in k[x_1, \ldots, x_n]$ and set $V = \mathrm{V}(f)$. Show that if f factors as a product $f = gh$ of distinct nonconstant irreducible polynomials $g, h \in k[x_1, \ldots, x_n]$, then V is reducible.

4.7.2. Prime and Non-Prime Ideals. A proper ideal $I \subset R$ is a *prime ideal* in R if, whenever $ab \in I$ for $a, b \in R$, either $a \in I$ or $b \in I$ (or both). A proper ideal $I \subset R$ is a *maximal ideal* in R if $I \subsetneq J \subset R$ for some ideal J implies that $J = R$.

Exercise 4.7.4. We know that every ideal I in $\mathbb{Z}$ is of the form $I = \langle m \rangle$ for some $m \in \mathbb{Z}$.

(1) For what values of m is the ideal $I = \langle m \rangle$ a prime ideal in $\mathbb{Z}$?

(2) For what values of m is the ideal $I = \langle m \rangle$ a maximal ideal in $\mathbb{Z}$?

Exercise 4.7.5. Let I be an ideal in a ring R.

(1) Show that $I \subset R$ is a prime ideal if and only if R/I is an integral domain.

(2) Show that $I \subset R$ is a maximal ideal if and only if R/I is a field.

(3) Explain why every maximal ideal in R is prime.

Exercise 4.7.6. Let $f(x, y) = xy \in k[x, y]$. Show that the ideal $\langle f \rangle$ is not a prime ideal.

Exercise 4.7.7. Let $f \in k[x]$ be a nonconstant polynomial. Prove that f is an irreducible polynomial if and only if $\langle f \rangle$ is a prime ideal.

Exercise 4.7.8. Let $\varphi : R \to S$ be a ring homomorphism. Let $J \subset S$ be a prime ideal in S. Show that $\varphi^{-1}(J)$ is a prime ideal in R.

4.7.3. Varieties and Prime Ideals. We now reach the key results of this section.

Exercise 4.7.9. Let $V \subset \mathbb{A}^n$ be an algebraic set.

(1) Suppose that V is reducible, say $V = V_1 \cup V_2$ where V_1 and V_2 are algebraic sets with $V_1 \subsetneq V$ and $V_2 \subsetneq V$. Show that

there are polynomials $P_1 \in I(V_1)$ and $P_2 \in I(V_2)$ such that $P_1P_2 \in I(V)$ but $P_1, P_2 \notin I(V)$. Conclude that $I(V)$ is not a prime ideal.

(2) Prove that if $I(V)$ is not a prime ideal in $k[x_1, \ldots, x_n]$, then V is a reducible algebraic set.

Exercise 4.7.10. Let V be an algebraic set in $\mathbb{A}^n$. Prove that the following are equivalent:

(1) V is an affine variety.

(2) $I(V)$ is a prime ideal in $k[x_1, \ldots, x_n]$.

(3) The quotient ring $k[x_1, \ldots, x_n]/I(V)$ is an integral domain. (Note: This quotient ring is denoted by $\mathcal{O}_V$ and is called either the coordinate ring or the ring of regular functions.)

Exercise 4.7.11. Show that $\mathbb{A}^n$ is an irreducible algebraic set for every $n \geq 1$. Thus every affine space is an affine variety.

Exercise 4.7.12. Let $f \in k[x, y]$ be an irreducible polynomial. Show that $\mathrm{V}(f)$, which is a curve in $\mathbb{A}^2$, is an irreducible algebraic set.

Typically an algebraic set is not irreducible, but it can always be written as a finite union of irreducible algebraic sets, which leads us to the following exercise.

Exercise 4.7.13. Let V be an algebraic set. Assume that V cannot be written as the union of a finite number of irreducible algebraic sets. (This is based on Proposition I.1.5 and Corollary I.1.6 of [**Har77**].)

(1) Show that there is an infinite descending chain of algebraic sets

$$V \supset V_1 \supset V_2 \supset \cdots$$

in $\mathbb{A}^n$.

(2) Show that

$$I(V) \subset I(V_1) \subset I(V_2) \subset \cdots$$

is an infinite ascending chain of ideals in $k[x_1, \ldots, x_n]$.

(3) Use the fact that $k[x_1, \ldots, x_n]$ is Noetherian to develop a contradiction.

Conclude that every algebraic set in $\mathbb{A}^n$ can be written as a union of a finite number of irreducible algebraic sets in $\mathbb{A}^n$.

Exercise 4.7.14.

(1) Let V be an algebraic set in $\mathbb{A}^n$. Show that V can be written as a union of finitely many irreducible algebraic sets in $\mathbb{A}^n$, $V = V_1 \cup \cdots \cup V_k$, such that no V_i contains any V_j.

(2) Suppose that $V_1 \cup \cdots \cup V_k = W_1 \cup \cdots \cup W_\ell$, where the V_i, W_j are irreducible algebraic sets in $\mathbb{A}^n$ such that no V_i contains any V_j and no W_i contains any W_j if $i \neq j$. Show that $k = \ell$ and, after rearranging the order, $V_1 = W_1, \ldots, V_k = W_k$.

Therefore, every algebraic set in $\mathbb{A}^n$ can be expressed uniquely as the union of finitely many affine varieties, no one containing another.

4.8. Regular Functions and the Coordinate Ring

In this section we define the natural ring of polynomial functions on an algebraic set: $\mathcal{O}(V)$.

One of the themes in 20th-century mathematics is that it is not clear what is more important in geometry: the actual geometric point set or the space of functions defined on the geometric point set. We now look at functions defined on algebraic sets.

Definition 4.8.1. Let $V \subseteq \mathbb{A}^n(k)$ be an algebraic set. The *coordinate ring* of V is the quotient ring $\mathcal{O}(V) = k[x_1, \ldots, x_n]/I(V)$. The elements of the coordinate ring are the *regular functions* on V.

The elements of $\mathcal{O}(V)$ can be thought of as polynomial functions on V.

Given an algebraic set V, recall that by $I(V)$ we mean the vanishing ideal of V, i.e., the ideal in $k[x_1, \ldots, x_n]$ consisting of polynomials f that satisfy $f(p) = 0$ for all $p \in V$.

Exercise 4.8.1. Let $f(x, y) = x^2 + y^2 - 1 \in \mathbb{C}[x, y]$. Consider the two polynomials $g(x, y) = y, h(x, y) = x^2 + y^2 + y - 1$.

(1) Find a point $(a, b) \in \mathbb{A}^2(\mathbb{C})$ such that

$$g(a, b) \neq h(a, b).$$

(2) Show for any point $(a, b) \in \mathrm{V}(f)$ that

$$g(a, b) = h(a, b).$$

Thus g and h are different as functions on $\mathbb{A}^2(\mathbb{C})$ but should be viewed as equal on the algebraic set $\mathrm{V}(f)$.

Exercise 4.8.2. Let $f(x, y) = x^2 + y^2 - 1 \in \mathbb{C}[x, y]$. Suppose that $g, h \in \mathbb{C}[x, y]$ such that for all $(a, b) \in \mathrm{V}(f)$ we have $g(a, b) = h(a, b)$. Show that the polynomial $g(x, y) - h(x, y) \in \langle x^2 + y^2 - 1 \rangle$.

Exercise 4.8.3. Let $V \subseteq k^n$ be an algebraic set. Prove that there is a one-to-one correspondence from the set of all ideals of $k[x_1, \ldots, x_n]/I(V)$ onto the set of all ideals of $k[x_1, \ldots, x_n]$ containing $I(V)$.

Exercise 4.8.4. Let $V \subseteq \mathbb{A}^n(k)$ and $W \subseteq \mathbb{A}^m(k)$ be algebraic sets. A function $f : V \to W$ is a *polynomial map* if there exist $f_1, \ldots, f_m \in \mathcal{O}(V)$ so that for all $p \in V$ we have $f(p) = (f_1(p), \ldots, f_m(p))$.

(1) Let $f : V \to W$ be a polynomial map, and define $\phi : \mathcal{O}(W) \to \mathcal{O}(V)$ by $\phi(g) = g \circ f$. Show that ϕ is a k-algebra homomorphism. Thus you must show that $\phi(g + h) = \phi(g) + \phi(h)$, for all $g, h \in \mathcal{O}(W)$ and $\phi(ag) = a\phi(g)$ for all $a \in k$.

(2) Show that for each k-algebra homomorphism $\phi : \mathcal{O}(W) \to \mathcal{O}(V)$ there exists a polynomial map $f : V \to W$ such that $\phi(g) = g \circ f$, for all $g \in \mathcal{O}(W)$.

4.9. Subvarieties

The goal of this section is to define subvarieties of an affine variety and to examine some of their algebraic properties.

Definition 4.9.1. Let W be an algebraic variety that is properly contained in an algebraic variety $V \subset \mathbb{A}^n(k)$. Then W is a *subvariety* of V.

Exercise 4.9.1. Let $V = \{(x, y) : x - y = 0\} \subset \mathbb{A}^2(\mathbb{C})$. Show that the point $p = (1, 1)$ is a subvariety of V, while the point $q = (1, 2)$ is not a subvariety of V.

Exercise 4.9.2. From the previous problem, find $I(V)$, $I(p)$ and $I(q)$. Show that

$$I(V) \subset I(p)$$

and

$$I(V) \not\subset I(q).$$

Exercise 4.9.3. Let W be a subvariety of V. Show that

$$I(V) \subset I(W).$$

Exercise 4.9.4. Let V and W be two algebraic varieties in $\mathbb{A}^n(k)$. Suppose that

$$I(V) \subset I(W).$$

Show that W is a subvariety of V.

Thus we have an elegant diagram:

$$\begin{array}{ccc} W & \subset & V \\ \text{if} & \text{and} & \text{only if} \\ I(W) & \supset & I(V) \end{array}$$

We now want to explore the relation between the coordinate ring $\mathcal{O}(V)$ and the coordinate ring $\mathcal{O}(W)$ for any subvariety W of a variety V.

Exercise 4.9.5. Continue letting $V = \{(x, y) : x - y = 0\} \subset \mathbb{A}^2(\mathbb{C})$, with subvariety $p = (1, 1)$. Find a polynomial $f \in \mathbb{C}[x, y]$ that is not identically zero on points of V but is zero at p, meaning there is a point $q \in V$ with $f(q) \neq 0$ but $f(p) = 0$. Show that

$$\text{Rad}\,\langle f, I(V)\rangle = I(p).$$

[Hint: Choose f reasonably.]

We have to worry a little about notation. For a variety $V \subset \mathbb{A}^n(k)$, by definition $\mathcal{O}(V) = k[x_1, \ldots, x_n]/I(V)$. Then, given any $f \in k[x_1, \ldots, x_n]$, we can think of f as a function on V and hence as an element of $\mathcal{O}(V)$, but we must keep in mind that if we write

$f \in \mathcal{O}(V)$, then f is standing for the equivalence class $f + I(V)$, capturing that if f and $g \in k[x_1, \ldots, x_n]/I(V)$ agree on all points of V, then $f - g \in I(V)$ and hence $f + I(V) = g + I(V)$, representing the same function in $\mathcal{O}(V)$.

We have a ring theoretic exercise first.

Exercise 4.9.6. Let R be a commutative ring. Let $I \subset J$ be two ideals in R. Show that J/I is an ideal in the quotient ring R/I. Show that there is a natural onto map

$$R/I \to R/J$$

whose quotient is the ideal J/I.

Exercise 4.9.7. Continue letting $V = \{(x, y) : x - y = 0\} \subset \mathbb{A}^2(\mathbb{C})$, with subvariety $p = (1, 1)$. Explicitly check the above exercise for $R = \mathbb{C}[x, y], I = I(V)$ and $J = I(p)$.

For any type of subsets $W \subset V$, if $f : V \to k$, then there is the natural restriction map $f|_W : W \to k$, which just means for all $p \in W$ that we define

$$f|_W(p) = f(p).$$

Exercise 4.9.8. Let W be a subvariety of a variety $V \subset \mathbb{A}^n(k)$. Let $f \in \mathcal{O}(V)$. Show that the above restriction map sends f to an element of $\mathcal{O}(W)$ and that this restriction map is a ring homomorphism.

Exercise 4.9.9. Show that the kernel of this restriction map is $I(W)/I(V)$ in the ring $\mathcal{O}(V)$.

Exercise 4.9.10. Discuss why each subvariety W of V should correspond to an onto ring homomorphism from the coordinate ring $\mathcal{O}(V)$ to a commutative ring.

Thus there are three equivalent ways for thinking of subvarieties of an algebraic variety V:

(1) W as an algebraic variety properly contained in an algebraic variety V.

(2) A prime ideal J properly containing the prime ideal $I(V)$.

(3) A quotient ring of the ring $\mathcal{O}(V) = k[x_1, \ldots, x_n]/I(V)$.

4.10. Function Fields

The goal of this section is to associate not just a ring but also a field to an algebraic variety. This field plays a critical role throughout algebraic geometry.

Every algebraic variety V corresponds to a prime ideal $I \subset k[x_1, \ldots, x_n]$. This allowed us to define the ring of functions on V, namely the quotient ring $\mathcal{O}(V) = k[x_1, \ldots, x_n]/I$. Every integral domain sits inside a field, much like the integers sit inside the rational numbers. In fact, the integers can be used to define the rational numbers as the smallest field containing the integers. The goal of this subsection is to define the *function field* $\mathcal{K}_V$, which is the smallest field that contains the quotient ring $\mathcal{O}(V)$.

Definition 4.10.1. Given an algebraic variety V corresponding to a prime ideal $I \subset k[x_1, \ldots, x_n]$, the *function field* $\mathcal{K}_V$ is:

$$\mathcal{K}_V = \left\{ \frac{f}{g} : f, g \in \mathcal{O}(V), g \neq 0 \right\} / \sim$$

where

$$\frac{f_1}{g_1} \sim \frac{f_2}{g_2} \quad \text{if and only if} \quad f_1 g_2 - f_2 g_1 \in I.$$

Exercise 4.10.1. Show that $\sim$ is an equivalence relation.

So far, $\mathcal{K}_V$ is simply a set. To make it into a field, we need to define how to add and multiply its elements. Define addition to be:

$$\frac{e}{f} + \frac{g}{h} = \frac{eh + fg}{fh}$$

and multiplication to be

$$\frac{e}{f} \cdot \frac{g}{h} = \frac{eg}{fh}.$$

Exercise 4.10.2. Show that addition is well-defined, that is, if

$$\frac{e_1}{f_1} \sim \frac{e_2}{f_2} \quad \text{and} \quad \frac{g_1}{h_1} \sim \frac{g_2}{h_2},$$

then

$$\frac{e_1}{f_1} + \frac{g_1}{h_1} \sim \frac{e_2}{f_2} + \frac{g_2}{h_2}.$$

Exercise 4.10.3. Show that multiplication is well-defined, that is, if

$$\frac{e_1}{f_1} \sim \frac{e_2}{f_2} \quad \text{and} \quad \frac{g_1}{h_1} = \frac{g_2}{h_2},$$

then

$$\frac{e_1}{f_1} \cdot \frac{g_1}{h_1} \sim \frac{e_2}{f_2} \cdot \frac{g_2}{h_2}.$$

Under these definitions, $\mathcal{K}_V$ is indeed a field.

4.11. The Zariski Topology

The goal of this section is to show that there is an algebraically defined topology for any ring.

4.11.1. Topologies. The development of topology is one of the great success stories of early 20th-century mathematics. With a sharp definition for a topological space, once tricky notions such as "continuity" and "dimension" now have rigorous, meaningful definitions. As with most good abstractions, these definitions could be applied to situations far removed from what their founders intended. This is certainly the case in algebraic geometry.

We start with the definition of a topology on a set X.

Definition 4.11.1. A *topology* on the set X is given by specifying a collection $\mathcal{U}$ of subsets of X having the properties:

(1) Both the empty set and the entire set X are elements of the collection $\mathcal{U}$.

(2) The union of any subsets of X in $\mathcal{U}$ is also in $\mathcal{U}$. (It is critical that we allow even infinite unions.)

(3) The finite intersection of any subsets of X in $\mathcal{U}$ is also in $\mathcal{U}$. (Here is it critical that we allow only finite intersections.)

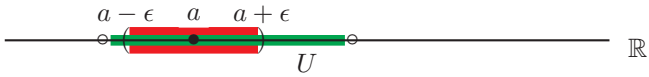
A set $U \in \mathcal{U}$ is said to be *open*. A set C is said to be *closed* if its complement $X - C$ is open.

Let us look at a few examples.

Start with the real numbers $\mathbb{R}$. We need to define what subsets will make up the collection $\mathcal{U}$.

Definition 4.11.2. A set $U \subset \mathbb{R}$ is a *standard open set in* $\mathbb{R}$ if for every $a \in U$, there exists an $\epsilon > 0$ such that

$$\{x \in \mathbb{R} : |x - a| < \epsilon\} \subset U.$$



Exercise 4.11.1. Let $a, b \in \mathbb{R}$ with $a < b$.

(1) Show that in $\mathbb{R}$, $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ is open.

(2) Show that in $\mathbb{R}$, $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ is closed.

(3) Show that in $\mathbb{R}$, $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ is neither open nor closed. (This type of set is often said to be half-open.)

Exercise 4.11.2. Show that the collection of standard open sets in $\mathbb{R}$ defines a topology on $\mathbb{R}$. This is called the *standard topology on* $\mathbb{R}$.

Let us now put a topology on $\mathbb{C}^n$.

Definition 4.11.3. A set $U \subset \mathbb{C}^n$ is a *standard open set in* $\mathbb{C}^n$ if for every $a \in U$, there exists an $\epsilon > 0$ such that

$$\{x \in \mathbb{C}^n : |x - a| < \epsilon\} \subset U.$$

(Note that $|x-a| = \sqrt{|x_1 - a_1|^2 + \cdots + |x_n - a_n|^2}$ for $a = (a_1, \ldots, a_n)$ and $x = (x_1, \ldots x_n)$.)

Thus a set U is open in $\mathbb{C}^n$ if any of its points can be made the center of a little open ball that lies entirely within U.

Exercise 4.11.3. Show that the collection of standard open sets in $\mathbb{C}^n$ defines a topology on $\mathbb{C}^n$. This is called the *standard topology on* $\mathbb{C}^n$.

Exercise 4.11.4. In $\mathbb{C}^2$, show that $\mathbb{C}^2 - \mathrm{V}(x^2 + y^2 - 1)$ is open in the standard topology.

Exercise 4.11.5. In $\mathbb{C}^2$, show that $\mathbb{C}^2 - \mathrm{V}(P)$ is open in the standard topology for any polynomial $P(x, y)$.

Exercise 4.11.6. In $\mathbb{C}^3$, show that $\mathbb{C}^3 - \mathrm{V}(x^2 + y^2 + z^2 - 1)$ is open in the standard topology.

Exercise 4.11.7. In $\mathbb{C}^n$, show that $\mathbb{C}^n - \mathrm{V}(P)$ is open for any polynomial $P(x_1, x_2, \ldots, x_n)$, so $\mathrm{V}(P)$ is closed in the standard topology on $\mathbb{C}^n$.

Exercise 4.11.8. In $\mathbb{C}^2$, show that $\{(x, y) \in \mathbb{C}^2 : |x|^2 + |y|^2 < 1\}$ is open in the standard topology. [Hint: Use the *Triangle Inequality*, which states that for any $z, w \in \mathbb{C}^n$, $|z + w| \leq |z| + |w|$.]

The standard topologies on $\mathbb{R}$ and $\mathbb{C}^n$ may be familiar to you. However, these are not the only topologies that can be defined on these sets. In the next exercise you will explore the finite complement topology on $\mathbb{R}$ and see that it is different than the standard topology.

Exercise 4.11.9. *Finite complement topology on X*: On $\mathbb{R}$ a set U is open if the complement of U is a finite collection of points, i.e., $U = X - \{p_1, \ldots, p_k\}$. X and $\emptyset$ are also considered to be open sets.

(1) Verify that any arbitrary union of open sets is again open.

(2) Verify that any finite intersection of open sets is open.

(3) Conclude that the open sets defined above form a topology on X. This is called the *finite complement topology on X*.

(4) Show that if a set U is open in the finite complement topology for $\mathbb{R}$, then it is open in the standard topology on $\mathbb{R}$.

(5) Give an example of an open set in the standard topology on $\mathbb{R}$ that is not open in the finite complement topology.

(6) Show that any two nonempty open sets in the finite complement topology on $\mathbb{R}$ must intersect. Is the same true in the standard topology on $\mathbb{R}$?

The definition of topology involves properties of open sets. When we work in the algebraic geometry world, most of our basic objects, such as varieties, are closed objects. We would like to be able to switch between open and closed sets from time to time. The next pair of exercises hold for any general topology.

Exercise 4.11.10. Let X be a set. Define for any set U in X its complement to be $U^c = X - U$. Show that

$$(U^c)^c = U.$$

Exercise 4.11.11. For subsets U_α, $\alpha \in A$, of a set X, let $C_\alpha = U_\alpha^c$.

(1) Show that
$$\bigcup_\alpha U_\alpha = X - \bigcap_\alpha C_\alpha.$$
(2) Show that
$$\bigcap_\alpha U_\alpha = X - \bigcup_\alpha C_\alpha.$$

The final part of Exercise 4.11.9 implies that the finite complement topology on $\mathbb{R}$ is not "Hausdorff" while the standard topology is. In a *Hausdorff topology* on a set X, for any pair of distinct points $p, q \in X$ you can find open sets U, V such that $p \in U$, $q \in V$ and $U \cap V = \emptyset$. That is, we can "separate" p and q in X with disjoint open sets. This is usually a desirable property in the study of topology, but it is not a property of the topology we use in algebraic geometry: the Zariski topology.

4.11.2. The Zariski Topology on $\mathbb{A}^n(k)$.

Definition 4.11.4. Let k be a field. A set $X \subset \mathbb{A}^n(k)$ is a *Zariski-closed* set if X is an algebraic set. A set U is *Zariski-open* if $U = \mathbb{A}^n(k) - X$ where X is an algebraic set.

Exercise 4.11.12.

(1) Use Exercises 4.2.5 and 4.2.12 to show that the collection of Zariski-open sets in $\mathbb{A}^n(k)$ is a topology. This is called the *Zariski topology on* $\mathbb{A}^n(k)$.
(2) Show that a finite collection of points in $\mathbb{A}^n(k)$ is a Zariski-closed set.

Exercise 4.11.13. In this exercise, we compare the Zariski and finite complement topologies.

(1) Show that if a set U is open in the finite complement topology on $\mathbb{A}^n(k)$, then it is open in the Zariski topology on $\mathbb{A}^n(k)$.
(2) Show that the finite complement topology on $\mathbb{R}$ is the same as the Zariski topology on $\mathbb{R}$. (Two topologies are the same

if and only if they are given by the same collection of open sets or, equivalently, the same collection of closed sets.)

(3) Show that the finite complement topology on $\mathbb{C}$ is the same as the Zariski topology on $\mathbb{C}$.

(4) Show that a circle in $\mathbb{R}^2$ is a Zariski-closed set. Conclude that the Zariski topology is not the same as the finite complement topology on $\mathbb{R}^2$.

Exercise 4.11.14. In this exercise we describe the Zariski topology on $\mathbb{C}^2$.

(1) Show that, in $\mathbb{C}^2$, the complement of a finite number of points and algebraic curves is Zariski-open. Note: By an "algebraic curve" in $\mathbb{C}^2$ we mean $\mathrm{V}(P)$ for some irreducible polynomial $P(x, y)$ of positive degree.

(2) Show that a non-empty Zariski-open set in $\mathbb{C}^2$ is the complement of a finite number of points and algebraic curves.

Exercise 4.11.15. Show geometrically that the Zariski topology on $\mathbb{C}^2$ is not Hausdorff.

4.12. Spec(R)

The goal of this section is to define a space and its Zariski topology associated to any ring R.

Let R be a commutative ring. In order to create a topological space, we first have to specify our set of points. We will see that our "points" will be the prime ideals in R. Recall that a proper ideal I in a ring R is *prime* if the following holds: whenever $f, g \in R$ with $fg \in I$, then $f \in I$ or $g \in I$. A proper ideal I of R is *maximal* if $I \subset J$ for some ideal J in R implies that either $J = I$ or $J = R$.

Definition 4.12.1. The *prime spectrum* or *spectrum* of a ring R is the collection of prime ideals in R, denoted by $\mathrm{Spec}(R)$.

For any ring R, the set on which we will define our topology is $\mathrm{Spec}(R)$. This definition might not appear to be the proper generalization of Theorem 4.6.5, where we learned that points of $\mathbb{A}^n(k)$

correspond to maximal ideals in $k[x_1, \ldots, x_n]$. However, it is the natural choice for the set of points we need, which will be further explained in Subsection 4.18.2 later in this chapter. In the meantime, we note that because maximal ideals are prime, the set $\mathrm{Spec}(R)$ includes all points from before and potentially more, as we explore in the following exercises.

Exercise 4.12.1. Describe the following sets.

(1) $\mathrm{Spec}(\mathbb{Z})$

(2) $\mathrm{Spec}(\mathbb{R})$

(3) $\mathrm{Spec}(k)$ for any field k

Exercise 4.12.2. Consider the polynomial ring $\mathbb{C}[x]$.

(1) Show that the ideal $\langle 0 \rangle$ is a prime ideal in $\mathbb{C}[x]$.

(2) Show that all prime ideals in $\mathbb{C}[x]$ are maximal ideals, except for the ideal $\langle 0 \rangle$.

(3) Show for each point $a \in \mathbb{C}$ there is a corresponding prime ideal.

(4) Explain why $\mathrm{Spec}(\mathbb{C}[x])$ can reasonably be identified with $\mathbb{C}$.

Exercise 4.12.3. Show that there are three types of points in $\mathrm{Spec}(\mathbb{R}[x])$:

i. The zero ideal $\langle 0 \rangle$,

ii. Ideals of the form $\langle x - a \rangle$ for a real number a,

iii. Ideals of the form $\langle x^2 + \beta x + \gamma \rangle$ for real numbers β, γ with $\beta^2 - 4\gamma < 0$.

Exercise 4.12.4. A curious property of "points" in $\mathrm{Spec}(R)$.

(1) Show that $\langle x - y \rangle$ is a prime ideal in $\mathbb{C}[x, y]$ and hence is a point in $\mathrm{Spec}(\mathbb{C}[x, y])$.

(2) For two fixed complex numbers a and b, show that $\langle x - a, y - b \rangle$ is a maximal ideal of $\mathbb{C}[x, y]$ and is hence also a point in $\mathrm{Spec}(\mathbb{C}[x, y])$.

(3) Show that for every $a \in \mathbb{C}$, $\langle x - a, y - a \rangle$ contains the ideal $\langle x - y \rangle$.

Thus, in Spec(R), some "points" can be contained in others. This suggests that not all points in Spec(R) are created equal. Returning to our motivation in Theorem 4.6.5, we make the following definition.

Definition 4.12.2. The *geometric points* in Spec(R) are the maximal ideals.

By Part (2) of Exercise 4.12.4, $\langle x-a, y-b\rangle$ is a maximal ideal in $\mathbb{C}[x,y]$, and hence a geometric point in Spec($\mathbb{C}[x,y]$). In general, by the Weak Nullstellensatz (Theorem 4.5.15), the maximal ideals in $\mathbb{C}[x_1,\ldots,x_n]$ are of the form $\langle x_1-a_1,\ldots,x_n-a_n\rangle$ for $(a_1,\ldots,a_n)\in \mathbb{C}^n$. Thus the set of geometric points of Spec($\mathbb{C}[x_1,\ldots,x_n]$) corresponds exactly to the set of points of $\mathbb{C}^n$. However, we should not confuse these two sets, for Spec($\mathbb{C}[x_1,\ldots,x_n]$) contains many points other than its geometric ones, as indicated in the previous exercises.

Now that we are better acquainted with our set of points, we are ready to define the topology.

Definition 4.12.3. Let $S\subseteq R$. Define the *Zariski closed set* given by S in Spec(R) to be

$$Z(S)=\{P\in \mathrm{Spec}(R) : P\supseteq S\}.$$

A subset U of Spec(R) is *Zariski open* if there is a set $S\subseteq R$ with

$$U=\mathrm{Spec}(R)-Z(S).$$

Exercise 4.12.5. Let R be a ring. For a subset S of R, recall that $\langle S\rangle$ denotes the ideal in R generated by S.

(1) For a set $S\subseteq R$, show that $Z(S)=Z(\langle S\rangle)$.

(2) Show that $Z(\{0\})=\mathrm{Spec}(R)$, and $Z(\{1\})=\emptyset$.

(3) For ideals $I,J\subset R$ with $I\subset J$, show that $Z(I)\supseteq Z(J)$.

Exercise 4.12.6. Show that a point I in Spec(R) is Zariski closed if and only if the ideal I is maximal in R.

Thus the geometric points of Spec(R) coincide with the points of Spec(R) that are Zariski closed. We could have defined a geometric point as a point of Spec(R) that is Zariski closed.

As in Section 4.2 we want to create a dictionary for going back and forth between Zariski closed sets in Spec(R) and ideals in R. We

have already described $Z(S)$, which assigns closed subsets of $\mathrm{Spec}(R)$ to ideals in R. Now we define the ideal associated to a subset of $\mathrm{Spec}(R)$.

Definition 4.12.4. For $X \subseteq \mathrm{Spec}(R)$, define the *ideal of* X to be

$$I(X) = \bigcap_{P \in X} P,$$

the intersection of all prime ideals of R that are in X.

Exercise 4.12.7. Let $X \subset \mathrm{Spec}(R)$. Show that $I(X)$ is a radical ideal.

Exercise 4.12.8. Let X and Y be subsets of $\mathrm{Spec}(R)$.

(1) Show that $X \subseteq Z(I(X))$.

(2) Show that if $X \subseteq Y$, then $I(Y) \subseteq I(X)$.

Exercise 4.12.9. Show that if X is a Zariski closed set in $\mathrm{Spec}(R)$, then $X = Z(I(X))$.

Definition 4.12.5. For a subset Y of $\mathrm{Spec}(R)$, the *Zariski closure* of Y in $\mathrm{Spec}(R)$ is $\overline{Y} = Z(I(Y))$.

Exercise 4.12.10. Compute the Zariski closure of the following sets.

(1) $\{\langle 2\rangle, \langle 3\rangle\}$ in $\mathrm{Spec}(\mathbb{Z})$

(2) $\{\langle 0\rangle\}$ in $\mathrm{Spec}(\mathbb{Z})$

(3) $\{\langle x - y\rangle\}$ in $\mathrm{Spec}(\mathbb{C}[x, y])$

This reinforces our previous result that the geometric points of $\mathrm{Spec}(R)$ coincide with the Zariski closed points. Part (2) is especially interesting and has a name:

Definition 4.12.6. A point of $\mathrm{Spec}(R)$ whose closure is the whole space is called a *generic point*.

As we have already noted, the nature of points in $\mathrm{Spec}(R)$ challenges our geometric intuition. Still, we have also seen that several of the results from Section 4.2 for our dictionary between closed sets and ideals in R continue to hold in $\mathrm{Spec}(R)$. Here is one more of these results, which will prove important in our proof that the collection of Zariski open sets defines a topology on $\mathrm{Spec}(R)$.

Exercise 4.12.11. Show that if X and Y are Zariski closed sets in $\mathrm{Spec}(R)$, then $X \cup Y = Z(I(X) \cap I(Y))$ and $X \cap Y = Z(I(X) + I(Y))$.

Exercise 4.12.12. Show that the Zariski closed sets are closed under arbitrary intersections.

Recall that a topology is a collection of open sets. We now convert the Zariski closed sets into open ones and prove that the set of Zariski open sets defines a topology on $\mathrm{Spec}(R)$.

Exercise 4.12.13. Let U_1 and U_2 be Zariski *open* sets in $\mathrm{Spec}(R)$. Show that $U_1 \cap U_2$ is a Zariski open set in $\mathrm{Spec}(R)$.

Exercise 4.12.14. Let $\{U_\alpha : \alpha \in A\}$ be an arbitrary collection of Zariski open sets in $\mathrm{Spec}(R)$. Show that $\bigcup_\alpha U_\alpha$ is a Zariski open set in $\mathrm{Spec}(R)$.

Exercise 4.12.15. Show that the collection of Zariski open sets forms a topology on $\mathrm{Spec}(R)$.

4.13. Points and Local Rings

The goal of this section is to show how to link points on an algebraic variety V with local rings $\mathcal{O}_V$, which are subrings of the function field $\mathcal{K}_V$.

We want to study what is going on around a point p in an algebraic variety. One approach would be to understand the behavior of the functions on V near p. If we just want to know what is going on at p, then what a function is doing far from p is irrelevant. The correct ring-theoretic concept will be that of a local ring.

We start with points in affine varieties $V \subset \mathbb{A}^n(k)$ and their local rings. We then see how to put this into a much more general language.

4.13.1. Points as Maximal Ideals in Affine Varieties. In Section 4.6 we proved that points in $\mathbb{A}^n(k)$ correspond to maximal ideals in $k[x_1, \ldots, x_n]$ and, conversely, that maximal ideals correspond to points when the field k is algebraically closed (Theorem 4.6.5). In the following exercises, we prove similar results for affine varieties

$V \subset \mathbb{A}^n(k)$. Throughout this subsection we assume that k is an algebraically closed field.

Exercise 4.13.1. Let $V = \mathrm{V}(x^2 + y^2 - 1) \subset \mathbb{A}^2(k)$. Let $p = (1, 0) \in V$. Define

$$\mathfrak{m}_p = \{f \in \mathcal{O}_V : f(p) = 0\}.$$

(1) Show that $\mathfrak{m}_p$ is an ideal in $\mathcal{O}_V$.

(2) Show that $\mathfrak{m}_p$ is in fact a maximal ideal in $\mathcal{O}_V$.

Exercise 4.13.2. Let $\mathfrak{m}$ be a maximal ideal in $\mathcal{O}_V$ for the variety $V = \mathrm{V}(x^2 + y^2 - 1)$ from the previous problem. Let

$$\mathrm{V}(\mathfrak{m}) = \{p \in V : \text{ for all } f \in \mathfrak{m},\ f(p) = 0\}.$$

Show that $\mathrm{V}(\mathfrak{m})$ must be a single point on V.

Exercise 4.13.3. Let $V \subset \mathbb{A}^n(k)$ be an algebraic variety. Let p be a point in V. Define

$$\mathfrak{m}_p = \{f \in \mathcal{O}_V : f(p) = 0\}.$$

Show that $\mathfrak{m}_p$ is a maximal ideal in $\mathcal{O}_V$.

Exercise 4.13.4. Let $\mathfrak{m}$ be a maximal ideal in $\mathcal{O}_V$, for $V \subset \mathbb{A}^n(k)$. Let

$$\mathrm{V}(\mathfrak{m}) = \{p \in V : \text{ for all } f \in \mathfrak{m},\ f(p) = 0\}.$$

Show that $\mathrm{V}(\mathfrak{m})$ must be a single point in V.

Thus points p define maximal ideals in the coordinate ring $\mathcal{O}_V$ and maximal ideals in $\mathcal{O}_V$ define points on V. This extends the results of Theorem 4.6.5 to affine varieties in general.

4.13.2. Local Ring at a Point. Let $V \subset \mathbb{A}^n(k)$ be an algebraic variety and let p be a point in V. We want to concentrate on the functions on V defined near p. Suppose there is a $g \in \mathcal{O}_V$ with $g(p) \neq 0$, say $g(p) = 1$. Then close to p, whatever that means, the function g looks a lot like the constant function 1. This means that we should be allowed to look at $1/g$, which is generally not allowed in $\mathcal{O}_V$ but is allowed in its function field $\mathcal{K}_V$.

Recall the construction of $\mathcal{K}_V$ from Section 4.10. By definition,

$$\mathcal{K}_V = \left\{ \frac{f}{g} : f, g \in \mathcal{O}_V, g \neq 0 \right\} / \left(\frac{f_1}{g_1} \sim \frac{f_2}{g_2} \right),$$

where $f_1/g_1 \sim f_2/g_2$ if $f_1g_2 - f_2g_1 \in I(V)$. Addition and multiplication were defined as usual for fractions,

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} = \frac{f_1g_2 + f_2g_1}{g_1g_2} \quad \text{and} \quad \frac{f_1}{g_1} \cdot \frac{f_2}{g_2} = \frac{f_1f_2}{g_1g_2},$$

both of which are well-defined. This set with these operations is then a field. We now define the local ring at p to be a subring of this field.

Definition 4.13.1. Let p be a point on an algebraic variety V. The *local ring* associated to p is

$$\mathcal{O}_p(V) = \left\{ \frac{f}{g} \in \mathcal{K}_V : \ g(p) \neq 0 \right\}.$$

Exercise 4.13.5. Let p be a point in an algebraic variety V. Prove that its local ring $\mathcal{O}_p(V)$ is a subring of the function field $\mathcal{K}_V$.

Exercise 4.13.6. Let $V = \mathrm{V}(x^2+y^2-1) \subset \mathbb{A}^2(k)$ and $p = (1,0) \in V$.

(1) Show for $f(x,y) = x \in \mathcal{O}_p(V)$ that there is an element $g \in \mathcal{O}_p(V)$ such that $f \cdot g = 1$ in $\mathcal{O}_p(V)$.

(2) Show for $f(x,y) = y \in \mathcal{O}_p(V)$ that there can exist no element $g \in \mathcal{O}_p(V)$ such that $f \cdot g = 1$ in $\mathcal{O}_p(V)$.

(3) Show that the ring $\mathcal{O}_p(V)$ cannot be a field.

Exercise 4.13.7. Let p be a point in an algebraic variety $V \subseteq \mathbb{A}^n(k)$ and let

$$\mathfrak{m}_p = \{f \in \mathcal{O}_p(V) : f(p) = 0\}.$$

(1) Suppose that $f \notin \mathfrak{m}_p$. Show that there exists an element $g \in \mathcal{O}_p(V)$ such that $f \cdot g = 1$ in $\mathcal{O}_p(V)$.

(2) Show that $\mathfrak{m}_p$ is the unique maximal ideal in the ring $\mathcal{O}_p(V)$.

4.13.3. Local Rings in Commutative Algebra. We now shift gears and make things quite a bit more abstract. Part of the power of algebraic geometry is that we can start with geometric insights and translate these into the language of ring theory, allowing us to think geometrically about rings for which there is little apparent geometry. This is not our emphasis in this book, but the following is included to give just a flavor of this.

In Exercise 4.13.7 we saw that the local ring at a point p in an affine variety V, $\mathcal{O}_p(V)$, has a unique maximal ideal, $\mathfrak{m}_p$. Inspired

by this, we make the following definition for commutative rings in general.

Definition 4.13.2. A *local ring* is a ring that has a unique maximal ideal.

Now we can talk about local rings quite generally. For example, every field is a local ring since the only proper ideal in a field is the zero ideal. However, as we have seen in Exercise 4.13.6, not every local ring is a field.

The rest of this section develops the method of localization for creating local rings from a given commutative ring R. This method is similar to the creation of $\mathcal{O}_p(V)$ above, where we create a new ring of "fractions" of elements from R with denominators from one of its subsets. In $\mathcal{O}_p(V)$, that set of denominators was $\{g \in \mathcal{O}_V : g(p) \neq 0\}$. In this case and in all of our other experiences with fractions, both addition and multiplication require that we multiply denominators and again have a valid denominator. This leads to the following definition.

Definition 4.13.3. A nonempty subset S of a ring R is said to be *multiplicatively closed* in R if, whenever $a, b \in S$, the product $ab \in S$.

Exercise 4.13.8.

(1) Show that $S = \{1, 3, 9, 27, \dots\} = \{3^k : k \geq 0\}$ is a multiplicatively closed set in $\mathbb{Z}$.

(2) Let R be a ring and let $a \neq 0$ be an element of R. Show that the set $S = \{a^k : k \geq 0\}$ is a multiplicatively closed set in R.

Exercise 4.13.9.

(1) Let $p \in \mathbb{Z}$ be a prime number. Show that the set $\mathbb{Z} - \langle p \rangle$ is multiplicatively closed.

(2) Let R be a ring and assume that $I \subset R$ is a maximal ideal in R. Show that $S = R - I$ is multiplicatively closed.

(3) Let R be a ring and $I \subset R$ be any ideal. Under what conditions on the ideal I will the subset $S = R - I$ be a multiplicatively closed subset of R? Prove your answer.

Let S be a multiplicatively closed set in R. Define an equivalence relation $\sim$ on the set $R \times S$ as follows:

$$(r, s) \sim (r', s') \iff \exists\, t \in S \text{ such that } t(s'r - sr') = 0.$$

Exercise 4.13.10. Show that $\sim$ is an equivalence relation on $R \times S$.

Exercise 4.13.11. Describe the equivalence relation $\sim$ on $R \times S$ if $0 \in S$.

Let $R_S = (R \times S)/\sim$ and let $[r, s]$ denote the equivalence class of (r, s) with respect to $\sim$. Define addition in R_S by

$$[r_1, s_1] \underset{S}{+} [r_2, s_2] = [r_1 s_2 + r_2 s_1, s_1 s_2]$$

and multiplication by

$$[r_1, s_1] \underset{S}{\cdot} [r_2, s_2] = [r_1 r_2, s_1 s_2].$$

Exercise 4.13.12. Show that $\underset{S}{+}$ and $\underset{S}{\cdot}$ are well-defined operations on R_S.

With a little work checking the axioms, one can show that R_S is a ring under the addition and multiplication defined above. This ring is called the *localization* of R at S.

Exercise 4.13.13. Let $S = \mathbb{Z} - \{0\}$. What is $\mathbb{Z}_S$? Is $\mathbb{Z}_S$ a local ring?

Exercise 4.13.14. Let $R = \mathbb{Z}$ and $S = \{2^k : k \geq 0\} = \{1, 2, 4, 8, \dots\}$.

(1) Show that S is multiplicatively closed in R.

(2) Show that, in $R_S = \mathbb{Z}_S$, addition and multiplication of $[a, 2^m]$ and $[b, 2^n]$ agrees with the addition and multiplication of the fractions $a/2^m$ and $b/2^n$ in $\mathbb{Q}$.

(3) Let $S' = \{2, 4, 8, \dots\} = \{2^k : k \geq 1\}$. Show that $R_{S'} \cong R_S$.

Exercise 4.13.15. Let R be a ring and $I \subset R$ be a prime ideal. Set $S = R - I$, which is a multiplicatively closed set in R, and consider the ring R_S.

(1) Show that R_S is a local ring. Describe its unique maximal ideal.

(2) Show that the proper ideals in R_S correspond to ideals J in R such that $J \subseteq I$.

Exercise 4.13.16. Let $f \neq 0$ be an element of an integral domain R. Let $S = \{f^m \mid m \geq 0\} = \{1, f, f^2, f^3, \dots\}$, which is a multiplicatively closed set in R by Part (2) of Exercise 4.13.8 and doesn't contain 0 since R is a domain.

(1) Show that the proper ideals in R_S correspond to ideals J in R such that $J \cap S = \emptyset$.

(2) Show that the prime ideals in R_S correspond to prime ideals in R that do not meet S.

(3) Conclude that $\operatorname{Spec}(R_S)$ coincides with the Zariski open set $\operatorname{Spec}(R) - Z(\{f\}) = \{P \in \operatorname{Spec}(R) : f \notin P\}$.

We conclude this section by showing that the method of localization developed above gives another way to create the local ring at a point in $\mathbb{A}^n(k)$.

Exercise 4.13.17. Let p be a point in $\mathbb{A}^n(k)$. Let $\mathfrak{m}_p = \{f \in k[x_1, \dots, x_n] : f(p) = 0\}$. By the Weak Nullstellensatz (Theorem 4.5.15), $\mathfrak{m}_p$ is a maximal ideal in $R = k[x_1, \dots, x_n]$. Prove that the localization of R at $S = R - \mathfrak{m}_p$ is isomorphic to $\mathcal{O}_p(\mathbb{A}^n)$.

4.14. Tangent Spaces

The goal of this section is to establish the equivalence among several different notions of the tangent space T_pV of a variety V at a point p.

4.14.1. Derivations. There are several equivalent notions of a tangent space in algebraic geometry. Before developing the algebraic idea of a tangent space we will consider the familiar tangent space as it is usually defined in a multivariable calculus course, but we want to be able to work over any field k, not just $\mathbb{R}$ and $\mathbb{C}$, so we need to generalize our idea of differentiation.

To motivate this new definition, let's consider the main properties of the derivative map. The derivative is linear, the derivative of a constant is zero, and the derivative obeys the Leibniz rule (the Product Rule). The derivative map is an example of a derivation.

Definition 4.14.1. A *derivation* is a map $L : R \to S$ from a k-algebra[1] R to a k-algebra S with the following properties:

(i) L is k-linear, i.e., $L(af+bg) = aL(f)+bL(g)$ for all $a, b \in k$ and $f, g \in R$,

(ii) L obeys the Leibniz rule, $L(fg) = gL(f) + fL(g)$ for all $f, g \in R$.

Exercise 4.14.1. Suppose R is a k-algebra. Show that if $L : R \to R$ is a derivation, then $L(a) = 0$ for all $a \in k$. [Hint: Show that $L(1) = 0$ and apply (i).]

Exercise 4.14.2. Verify that $\dfrac{d}{dx} : k[x] \to k[x]$ formally defined by

$$\begin{aligned}\frac{d}{dx}&[a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0] \\ &= na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1\end{aligned}$$

is a derivation.

4.14.2. First Definition. We will first give an extrinsic definition of the tangent space of an algebraic set at a point.

Definition 4.14.2. Let I be a prime ideal in $k[x_1, \ldots, x_n]$, $V = \mathrm{V}(I)$ the corresponding variety in $\mathbb{A}^n$, and $p = (p_1, p_2, \cdots, p_n) \in V$. The *tangent space* of V at p is the linear subspace

$$T_pV = \left\{ (x_1, x_2, \ldots, x_n) \in k^n : \sum_{i=1}^{n} (x_i - p_i)\frac{\partial f}{\partial x_i}(p) = 0, \right.$$
$$\left. \text{for all } f \in I \right\},$$

where $\dfrac{\partial}{\partial x_i}$ is the derivation defined formally by

$$\frac{\partial}{\partial x_i} x_j^m = \begin{cases} m x_j^{m-1} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

and imposing that it is k-linear and satisfies the Leibniz rule.

[1] A *k-algebra* is a k-vector space that also has a multiplication making it a ring.

If $k = \mathbb{C}$ or $\mathbb{R}$, then $\dfrac{\partial}{\partial x_i}$ can be regarded as the usual partial derivative.

Note that the above definition also makes sense for an algebraic set and will be used when we compute their tangent spaces. In the special case that V is a hypersurface, $V = \mathrm{V}(f)$ for $f \in k[x_1, \ldots, x_n]$, the tangent space of the hypersurface at p is simply

$$T_pV = \left\{ (x_1, x_2, \ldots, x_n) \in k^n : \sum_{i=1}^{n} (x_i - p_i) \frac{\partial f}{\partial x_i}(p) = 0 \right\}.$$

Exercise 4.14.3. In $\mathbb{R}^2$ let $f(x, y) = x^2 + y^2 - 1$ and consider the curve $C = \mathrm{V}(f)$. Let $p = (a, b)$ be a point on C.

(1) Find the normal direction to C at p.

(2) How is the normal direction to C at p related to the gradient of f at p, $\nabla f(p)$?

(3) Use Definition 4.14.2 to find T_pC.

(4) How is T_pC related to $\nabla f(p)$?

Exercise 4.14.4. Show that T_pV, as defined in Definition 4.14.2, is a vector space over k by identifying the vector $(x_1, \ldots, x_n)$ in T_pV with the vector $(x_1 - p_1, \ldots, x_n - p_n)$ in k^n.

Exercise 4.14.5. In this problem, let

$$\begin{aligned} z_1 &= x + iy \quad \in \mathbb{C}, \quad (x, y) \in \mathbb{R}, \\ z_2 &= u + iv \quad \in \mathbb{C}, \quad (u, v) \in \mathbb{R}. \end{aligned}$$

Suppose $V \subset \mathbb{C}^2$ is defined via $F(z_1, z_2) = z_1 - z_2^2 = 0$.

(1) Let $p_0 = (-1, i)$. Is $p_0 \in V$?

(2) Find the tangent line $h(z_1, z_2) = 0$ to V at p_0 using Definition 4.14.2.

(3) Show that V, viewed as a set $V_{\mathbb{R}} \subset \mathbb{R}^4$, is the intersection of two surfaces,

$$f(x, y, u, v) = 0 \quad \text{and} \quad g(x, y, u, v) = 0.$$

Find f and g explicitly. Intuitively, what is the real dimension of $V_{\mathbb{R}}$?

(4) Find the point $q_0 = (x_0, y_0, u_0, v_0) \in \mathbb{R}^4$ to which $p_0 = (-1, i) \in \mathbb{C}^2$ corresponds.

(5) Find two normal vectors in $\mathbb{R}^4$ to $V_{\mathbb{R}}$ at q_0 via $\vec{N_1} = \nabla f(q_0)$, $\vec{N_2} = \nabla g(q_0)$. The real tangent space $T_{\mathbb{R},q_0}$ to $V_{\mathbb{R}}$ at q_0 is the set of lines through q_0 perpendicular to both $\vec{N_1}$ and $\vec{N_2}$. Intuitively, what is the real dimension of $T_{\mathbb{R},q_0}$?

(6) In Part (2), you found the tangent line equation $h(z_1, z_2) = 0$ to V at p_0 in $\mathbb{C}^2$. Write the tangent line as a system of two equations in $\mathbb{R}^4$ using x, y, u, v. These equations correspond to two hyperplanes Π_1, Π_2 in $\mathbb{R}^4$. Let $T = \Pi_1 \cap \Pi_2$. Find two linearly independent vectors $\vec{D_1}, \vec{D_2} \in \mathbb{R}^4$ parallel to T. Show that $\vec{D_1} \perp \vec{N_1}, \vec{N_2}$ and $\vec{D_2} \perp \vec{N_1}, \vec{N_2}$. Is T the same as $T_{\mathbb{R},Q_0}$?

(7) Does this convince you that if C is a curve in $\mathbb{C}^2$ and $T_{\mathbb{C},p_0}$ is the tangent line to C at p_0, then $T_{\mathbb{C},p_0}$ is the usual geometric tangent space to C at p_0 when $\mathbb{C}^2$ is thought of as $\mathbb{R}^4$?

4.14.3. Second Definition. Next, we consider another definition of an affine tangent space. Recall the definition of the local ring of a variety V at p,

$$\mathcal{O}_p(V) = \left\{ \frac{f}{g} \in \mathcal{K}_V : g(p) \neq 0 \right\}.$$

This local ring captures the behavior of functions on V near p. That is, $\mathcal{O}_p(V)$ gives an algebraic description of V near p. On the other hand, the tangent space to V at p gives a geometric description of V near p. With our second definition of T_pV, we connect these descriptions, using derivations on $\mathcal{O}_p(V)$ to construct T_pV.

Definition 4.14.3. The *tangent space* of the variety V at a point p is the linear space

$$T_pV = \{L : \mathcal{O}_p(V) \to k \ : \ L \text{ is a derivation}\}.$$

We first justify our claim in the definition that this is a linear space.

Exercise 4.14.6. Show that T_pV, as defined in Definition 4.14.3, is a vector space over k.

For any point $p \in \mathbb{A}^n$, $T_p\mathbb{A}^n$ is the vector space $\text{span}\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\}$, where $\frac{\partial}{\partial x_i}$ are defined formally as before. When $V = \mathrm{V}(I) \subset \mathbb{A}^n$ is an affine variety, T_pV is the subspace of linear combinations of $\frac{\partial}{\partial x_i}$ that agree on I. In other words, T_pV consists of all derivations $L = \sum_{i=1}^n \alpha_i \frac{\partial}{\partial x_i}$ such that $L(f)(p) = 0$ for all $f \in I$. We verify this below.

Exercise 4.14.7. Show that $L = \displaystyle\sum_{i=1}^{n} \alpha_i \frac{\partial}{\partial x_i}$ defines a derivation $\mathcal{O}_p(V) \to k$ if and only if $L(f)(p) = 0$ whenever $f \in I$.

Now we are in a position to use this characterization of the tangent space to compute an example.

There is yet another description of the tangent space to V at p, which we explore in the next two exercises.

Exercise 4.14.8. In $\mathbb{A}^2(\mathbb{C})$, consider the complex curve $C = \mathrm{V}(x^2 + y^2 - 1)$. At a point $p = (a, b) \in C$, show that $\mathfrak{m}_p/\mathfrak{m}_p^2$ is a 1-dimensional vector space over $\mathbb{C}$. Relate this 1-dimensional vector space to the tangent line found in Exercise 4.14.3.

Exercise 4.14.9. Let $V \subset \mathbb{A}^n(k)$ be an algebraic variety and let $p = (p_1, \ldots, p_n)$ be a point in V. As in Definition 4.14.3, write

$$T_pV = \{L : \mathcal{O}_p(V) \to k : L \text{ is a derivation}\}$$

and let $L \in T_pV$ be given.

(1) Let $\mathfrak{m}_p = \{f \in \mathcal{O}_V : f(p) = 0\}$ be the maximal ideal in $\mathcal{O}_V$ corresponding to p. Show that for all $f, g \in \mathfrak{m}_p$, $L(fg)(p) = 0$.

(2) Show that L induces a k-linear map $L' : \mathfrak{m}_p/\mathfrak{m}_p^2 \to k$.

(3) Let $l : \mathfrak{m}_p/\mathfrak{m}_p^2 \to k$ be a k-linear map. Show that the function given by $D_l(f)(p) = l(f - f(p) + \mathfrak{m}_p^2)$ is a well-defined derivation $D_l : \mathcal{O}_p(V) \to k$.

(4) Show that the maps $L \mapsto L'$ in Part (2) and $l \mapsto D_l$ in Part (3) establish an isomorphism of vector spaces between T_pV and $\mathrm{Hom}_k(\mathfrak{m}_p/\mathfrak{m}_p^2, k)$, the space of k-linear maps from $\mathfrak{m}_p/\mathfrak{m}_p^2$ to k.

4.15. Dimension

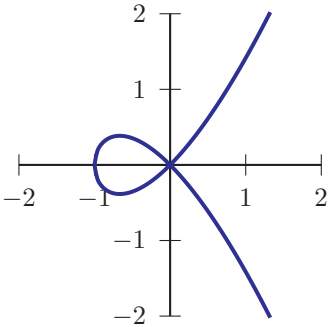
In this section we define the dimension of a variety using the tangent space at a point. We also have alternate definitions of dimension using the coordinate ring and the function field, as we will explore in several examples.

One may think of dimension of an affine variety V as the number of coordinates needed to describe V. The dimension will depend on our base field k, as we have seen in the first few chapters when we considered complex curves as surfaces over $\mathbb{R}$. Our first definition of dimension uses the tangent spaces studied in the previous section.

Definition 4.15.1. Let $V \subseteq \mathbb{A}^n$ be an irreducible variety. Then the *dimension* of V is the minimum dimension of T_pV over all points $p \in V$.

Here we define the dimension of T_pV to be its dimension as a vector space over the field k. For example, at any point p the tangent space $T_p\mathbb{A}^n$ is just $\mathbb{A}^n$, which is n-dimensional over k.

Exercise 4.15.1. Let $V = \mathrm{V}(x^3 + x^2 - y^2)$ be a curve in $\mathbb{A}^2$ whose graph in $\mathbb{R}^2$ is shown.



(1) Let $p = (-1, 0)$. Show that T_pV has dimension 1.

(2) Let $q = (0, 0)$. Show that T_qV has dimension 2.

We have defined smooth plane curves to be curves with well-defined tangent lines at each point, that is, curves $f(x, y) = 0$ where

at least one of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ is nonzero for every point on the curve. We next show that a smooth curve has dimension one.

Exercise 4.15.2. Let V be a smooth plane curve. Show that $\dim V = 1$.

In fact the dimension of any curve in $\mathbb{A}^2$ is one. More generally we can show that any hypersurface $\mathrm{V}(f)$ in $\mathbb{A}^n$ has dimension $n-1$.

Exercise 4.15.3. Let $V = \mathrm{V}(f)$ be an irreducible hypersurface in $\mathbb{A}^n$. Show that V has dimension $n-1$.

In fact the subset of points at which the dimension of the tangent space is greater than the dimension of the variety forms a closed subvariety. We will see in the next section that these are the *singular* points of V.

Definition 4.15.2. Let V be an affine variety. A point $p \in V$ is a *smooth* point of V if $\dim T_pV = \dim V$. Otherwise $\dim T_pV > \dim V$ and p is a *singular* point of V.

Exercise 4.15.4. In Definition 4.15.1, must V be an irreducible variety? Why?

The next exercise may help clarify.

Exercise 4.15.5. Let $V \subset \mathbb{A}^3$ be the variety defined by $\mathrm{V}(xz, yz)$.

(1) Show that V is reducible by showing that it is the union of a line and a plane.

(2) Find the dimension of the tangent space to the plane at a point $p \neq (0,0,0)$ on the plane.

(3) Find the dimension of the tangent space to the line at a point $p \neq (0,0,0)$ on the line.

We define the dimension of a algebraic set to be the maximum dimension of its components.

We have seen that varieties can be studied algebraically in terms of the spectrum of a ring. When we use this algebraic point of view we want a similar approach to compute dimension. In fact the dimension of a variety V can be defined in terms of prime ideals in the coordinate ring $\mathcal{O}(V)$. First we need some definitions from commutative algebra.

Definition 4.15.3. The *height* of a prime ideal $\mathcal{P}$ in a ring R is the length of the largest chain of prime ideals properly contained in $\mathcal{P}$, that is, the maximum n with

$$\mathcal{P}_0 \subsetneq \mathcal{P}_1 \subsetneq \cdots \subsetneq \mathcal{P}_{n-1} \subsetneq \mathcal{P}.$$

The *Krull dimension* of R is the maximum height over all prime ideals.

Exercise 4.15.6. Find a prime ideal in the polynomial ring $R = k[x_1, \ldots, x_n]$ of height n.

One can show this is the maximal height of any prime ideal in $k[x_1, \ldots, x_n]$; thus its Krull dimension is n and we have $\dim \mathbb{A}^n = n$.

As a third alternative we can define dimension using the field of functions of our irreducible variety V. For an extension field K of a field k, the *transcendence degree* of K over k is the maximum number of elements in K that form an algebraically independent set over k. The dimension of V is also equal to the transcendence degree of $\mathcal{K}_V$ over the base field k.

For example, the transcendence degree of $k(x)$, the rational functions in one variable, is one since $\{x\}$ is a maximal algebraically independent set; the transcendence degree of an algebraic extension field is 0. Thus a point in affine space, which has function field isomorphic to k, has dimension zero. The function field of the affine line $\mathbb{A}^1$ is $k(x)$, so $\mathbb{A}^1$ has dimension one. Similarly we can show that affine n-space has the expected dimension.

Exercise 4.15.7. Let $\mathcal{K}_{\mathbb{A}^n}$ be the function field of $\mathbb{A}^n$.

(1) Show that $\mathcal{K}_{\mathbb{A}^n} \cong k(x_1, \ldots, x_n)$.

(2) Show that $\{x_1, \ldots, x_n\}$ is a maximal set of algebraically independent elements over k.

(3) Conclude that the dimension of $\mathbb{A}^n$ is n.

We have seen that our three notions of dimension agree for affine n-space $\mathbb{A}^n$. In fact these definitions are equivalent for any irreducible affine variety. (See Theorems I.1.8 and I.5.1 in Hartshorne [**Har77**].)

Exercise 4.15.8. Check that our three notions of dimension agree for a smooth plane curve.

4.16. Arithmetic Surfaces

Our affine varieties have been in essence of the form $\mathrm{Spec}(k[x_1, \ldots, x_n]/I)$, for I a prime ideal, where k is an algebraically closed field. Arithmetic varieties occur when we replace the field k with the ring of integers $\mathbb{Z}$. In this section we will consider the ring $\mathbb{Z}[x]$ and see that $\mathrm{Spec}(\mathbb{Z}[x])$ is two dimensional.

In Section 4.15, we gave various ways for defining the dimension of a variety. For $\mathrm{Spec}(\mathbb{Z}[x])$, the definition that makes sense is the Krull dimension (Definition 4.15.3). Thus the dimension for us will be the number of terms of the longest chain of prime ideals $\mathcal{P}_0 \subsetneq \mathcal{P}_1 \subsetneq \cdots \subsetneq \mathcal{P}_n$ in $\mathbb{Z}[x]$.

We start with finding the dimension of $\mathrm{Spec}(\mathbb{Z})$:

Exercise 4.16.1. Show that the Krull dimension of $\mathrm{Spec}(\mathbb{Z})$ is one.

Next, we want to find the Krull dimension of $\mathrm{Spec}(k[x])$, for k a field:

Exercise 4.16.2. Show that the Krull dimension of $\mathrm{Spec}(k[x])$ is one.

Here is the rough idea for why $\mathrm{Spec}(\mathbb{Z}[x])$ should have dimension two and hence why $\mathrm{Spec}(\mathbb{Z}[x])$ should be called a surface. The $\mathbb{Z}$ part gives us one degree of freedom and the x part gives us another degree of freedom. Now to make this a bit more precise.

We first have to find the maximal ideals in $\mathbb{Z}[x]$.

Let $J = \langle p, f(x) \rangle$ be an ideal in $\mathbb{Z}[x]$ generated by a prime number p and a polynomial $f(x)$ whose reduction mod p is irreducible in $\mathbb{Z}_p$. Our goal is to show that every maximal ideal in $\mathbb{Z}[x]$ has this form.

Exercise 4.16.3. Show that $f(x) = x^2 + 4$ is irreducible mod 3.

Exercise 4.16.4. Given an ideal $J = \langle p, f(x) \rangle$, show that the quotient ring $\mathbb{Z}[x]/J$ is isomorphic to the quotient ring $\mathbb{Z}_p/\langle f(x) \rangle$. (This is actually not that hard of a problem.)

Exercise 4.16.5. Show that ideals $J = \langle p, f(x) \rangle$ for p prime and for polynomials $f(x)$ whose reduction mod p is irreducible in $\mathbb{Z}_p$, is a

maximal ideal in $\mathbb{Z}[x]$. (Use that an ideal I is maximal in a ring R if and only if R/I is a field.)

We now want to see why every maximal ideal in $\mathbb{Z}[x]$ is of the form $\langle p, f(x)\rangle$ for p prime and $f(x)$ a polynomial whose reduction mod p is irreducible in $\mathbb{Z}_p$.

Exercise 4.16.6. Let J be a prime ideal in the ring $\mathbb{Z}[x]$. Show that $J \cap \mathbb{Z}$ is a prime ideal in $\mathbb{Z}$, in which case either $J \cap \mathbb{Z} = \langle p\rangle$, the ideal in $\mathbb{Z}$ generated by a prime number p, or $J \cap \mathbb{Z} = \langle 0\rangle$.

Exercise 4.16.7. Show that in $\mathbb{Z}[x]$, the ideal $J_1 = \langle 3, x^2 + 8x + 5\rangle$ is equal to the ideal $J_2 = \langle 3, x^2 + 2x + 2\rangle$.

Exercise 4.16.8. Suppose two polynomials $f(x)$ and $g(x)$ in $\mathbb{Z}[x]$ are equal when reduced by mod p. Show that the ideals $\langle p, f(x)\rangle$ and $\langle p, g(x)\rangle$ are equal.

Exercise 4.16.9. Suppose J is a maximal ideal of the form $\langle p, f_1(x), \ldots, f_n(x)\rangle$, for some prime p. Show that there is a polynomial $f(x) \in \mathbb{Z}[x]$ that is irreducible mod p such that

$$J = \langle p, f(x)\rangle.$$

Exercise 4.16.10. Let J be a maximal ideal in $\mathbb{Z}[x]$. Show that $J = \langle p, f(x)\rangle$ for some prime number p and some polynomial $f(x)$ whose reduction mod p is irreducible in $\mathbb{Z}_p$.

Now to show that the height of a maximal ideal $J = \langle p, f(x)\rangle$ is two.

We have that

$$\langle 0\rangle \subsetneq \langle p\rangle \subsetneq \langle p, f(x)\rangle$$

is a chain of prime ideals, which means that the height of J is at least two. We also have another type of chain of prime ideals,

$$\langle 0\rangle \subsetneq \langle g(x)\rangle \subsetneq \langle p, f(x)\rangle,$$

where $g(t)$ equals $f(x)$ mod p. We have to show there is no chain of prime ideals of greater length.

Exercise 4.16.11. For a maximal ideal $J = \langle p, f(x)\rangle$, suppose we have a chain of prime ideals

$$\langle 0\rangle \subsetneq I \subsetneq \langle p, f(x)\rangle.$$

Show that either $I = \langle p \rangle$ or $I = \langle g(x) \rangle$, for some $g(x)$ equaling $f(x)$ mod p.

Exercise 4.16.12. Show that the ideal $J = \langle p, f(x) \rangle$ for p prime and $f(x)$ whose reduction mod p is irreducible in $\mathbb{Z}_p$, has height two.

4.17. Singular Points

A singularity of a variety is a point where the variety exhibits unusual behavior. In this section we will see two ways to find the singular points of a variety, either using the tangent space or computing the Jacobian matrix.

As we saw in the previous section, a singular point is where the dimension of the tangent space jumps and is larger than the dimension of the variety. In Section 1.10 we stated that a plane curve $\mathrm{V}(f(x,y))$ is singular at any point p where f, $\frac{\partial f}{\partial x}$, and $\frac{\partial f}{\partial y}$ vanish simultaneously.

We will first verify that these definitions coincide in the case of a plane curve.

Exercise 4.17.1. Let $V = \mathrm{V}(f(x,y)) \subset \mathbb{A}^2$ be an irreducible curve and let p be a point on V.

(1) Suppose at least one of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ is nonzero at p. Show that T_pV is one-dimensional and thus p is a smooth point of V.

(2) Suppose $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ both vanish at p. Show that T_pV has dimension two, meaning that p is a singular point of V.

(3) Show that the tangent space T_pV, as defined in Section 4.14, is equivalent to the tangent line of this curve as defined in Section 1.10.

Exercise 4.17.2. Determine the singular points of each curve.

(1) $\mathrm{V}(y^2 - x^3 + x^2)$

(2) $\mathrm{V}(y^2 - x^3)$

We have previously seen the nodal and cuspidal singular cubics of the last exercise. More generally we can determine when a cubic in normal form is singular.

Exercise 4.17.3. Let $f(x)$ be a polynomial and let $V = \mathrm{V}(y^2 - f(x))$. (In the case where f has degree three, V is the normal form of a cubic curve.) Show that V is singular at a point (x_0, y_0) if and only if $y_0 = 0$ and x_0 is a multiple root of $f(x)$.

The case of a hypersurface is similar to that of curves.

Exercise 4.17.4. Let V be the hypersurface $f(x_1, \ldots, x_n) = 0$ in $\mathbb{A}^n$ and let p be a point on V. Recall from Exercise 4.15.3 that V has dimension $n-1$.

(1) Suppose at least one of the $\frac{\partial f}{\partial x_i}$ is nonzero at p. Show that T_pV has dimension $n-1$. Conclude that p is a smooth point of V.

(2) Suppose $\frac{\partial f}{\partial x_i}(p) = 0$ for $i = 1, \ldots, n$. Show that $T_pV = \mathbb{A}^n$. Conclude that p is a singular point.

Exercise 4.17.5. Find all singular points of each surface in $\mathbb{A}^3$.

(1) $\mathrm{V}(x^2 + y^2 - z^2)$

(2) $\mathrm{V}(x^2 - y^2 z)$

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

Exercise 4.17.6. Let $V = \mathrm{V}(x^2 + y^2 + z^2 - 1, x - 1) \subset \mathbb{A}^3$.

(1) Show that V has dimension one, by visualizing V as the intersection of the surface $x^2 + y^2 + z^2 = 1$ and the plane $x = 1$.

(2) Show that the tangent space to V at $p = (1, 0, 0)$ is the plane $x - 1 = 0$. Thus T_pV has dimension two. Conclude that V is singular at p.

Exercise 4.17.7. Let $V = \mathrm{V}(fg)$ and let p be a point of intersection of the hypersurfaces $\mathrm{V}(f)$ and $\mathrm{V}(g)$. Show that p is a singular point of V.

We now present a second way to find the singular points of a variety V using a Jacobian matrix.

Definition 4.17.1. Let $\{f_1, f_2, \ldots, f_m\}$ be a generating set for $I(V)$, with each $f_i \in k[x_1, \ldots, x_n]$, where $V = \mathrm{V}(f_1, f_2, \ldots, f_m) \subset \mathbb{A}^n$. The *Jacobian* matrix for V at a point $p \in V$ is the $m \times n$ matrix $\left(\frac{\partial f_i}{\partial x_j}(p)\right)$.

This definition depends upon the set of generators for $I(V)$. We will see that the rank of the Jacobian matrix is independent of this choice; thus we can use the rank of the Jacobian to give an alternate definition of singularity.

Definition 4.17.2. Let V be a variety in $\mathbb{A}^n$ of dimension d. V is nonsingular at p if and only if the rank of the Jacobian matrix at p is equal to $n - d$.

Exercise 4.17.8. Let V be the curve $\mathrm{V}(x - yz, xz - y^2, y - z^2) \subset \mathbb{A}^3$. Show that the Jacobian matrix has rank two at every point $p \in V$. Conclude that V is a smooth curve.

Exercise 4.17.9. Compute the Jacobian matrix for $V = \mathrm{V}(x^2 + y^2 + z^2 - 1, x - 1) \subset \mathbb{A}^3$.

(1) Show the Jacobian has rank two when y or z is nonzero.

(2) Show the Jacobian has rank one when $y = z = 0$. Use this to determine the singular points of V.

Exercise 4.17.10. Let $V = \mathrm{V}(x + y + z, x - y + z) \subset \mathbb{A}^3$.

(1) Compute the Jacobian matrix for V and show that V is nonsingular everywhere.

(2) Show that $\{x + z, y\}$ is also a generating set for $I(V)$.

(3) Compute the Jacobian matrix using this alternate set of generators and show that it has the same rank as your matrix in Part (1).

Exercise 4.17.11. Let $V = \mathrm{V}(x^2 + y^2 - 1, x^2 + z^2 - 1) \subset \mathbb{A}^3$.

(1) Compute the Jacobian matrix for V and find all points where the rank is not equal to two.

(2) Show that $\left\{y^2 - z^2, 2x^2 + y^2 + z^2 - 2\right\}$ is also a generating set for $I(V)$.

(3) Compute the Jacobian matrix using this alternate set of generators and show that it has the same rank as your matrix in Part (1).

Exercise 4.17.12. For a hypersurface $\mathrm{V}(f) \subset \mathbb{A}^n$ the Jacobian matrix at p is

$$\left(\frac{\partial f}{\partial x_1}(p) \ \frac{\partial f}{\partial x_2}(p) \ \cdots \ \frac{\partial f}{\partial x_n}(p) \right).$$

Show that V is nonsingular at p if at least one of the $\frac{\partial f}{\partial x_i}(p)$ is nonzero. Thus for hypersurfaces this definition coincides with our previous one.

In the next exercise we will show more generally that our two definitions agree.

Exercise 4.17.13. Let p be a point of a d-dimensional variety $V \subset \mathbb{A}^n$ and let $\{f_1, f_2, \ldots, f_m\}$ be a generating set for $I(V)$.

(1) By identifying each point $q \in \mathbb{A}^n$ with the vector $q - p$, show that the Jacobian matrix $\left(\frac{\partial f_i}{\partial x_j}(p) \right)$ defines a linear transformation from $\mathbb{A}^n$ to $\mathbb{A}^m$.

(2) Show the kernel of this transformation is the tangent space T_pV.

(3) Use the Rank-Nullity Theorem to conclude that p is a nonsingular point of V if and only if the rank of the Jacobian matrix is equal to $n - d$.

It follows from this exercise that the rank of the Jacobian matrix at p is equal to $n - \dim T_pV$. Thus the rank is independent of the choice of generators for $I(V)$.

In each of our examples we have seen that the singular points form a proper subvariety of V. Our next exercises will show this is always true.

Exercise 4.17.14. Let V be an affine variety. Prove that the set of singular points of V is a closed subset of V. [Hint: Think of the minors of the Jacobian.]

Exercise 4.17.15. Let $V = \mathrm{V}(f) \subset \mathbb{A}^n$ be an irreducible hypersurface over k. Prove that the singularities of V are a proper subvariety. [Hint: Consult Exercise 4.15.3.]

The previous exercise shows that the singular points of a hypersurface are a proper subvariety. One can extend this result to all affine varieties using the fact that every irreducible variety is "equivalent" to a hypersurface. The type of equivalence we will use is a birational morphism, which will be defined in a later section.

4.18. Morphisms

The goal of this section is to define morphisms, which are a natural type of mapping between algebraic sets. We will then relate morphisms to the spectrum of a ring.

4.18.1. Definition of Morphism. The world of algebraic geometry is the world of polynomials[2]. For example, algebraic sets are defined as the set of common zeros of collections of polynomials. The morphisms, or mappings, between them should also be given by polynomials.

Suppose $X \subset \mathbb{A}^n(k)$ and $Y \subset \mathbb{A}^m(k)$ are algebraic sets. The *morphisms* between X and Y are the polynomial mappings:

$$\begin{aligned} \phi : X &\to Y \\ p &\mapsto (f_1(p), \ldots, f_m(p)) \end{aligned}$$

for some $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$.

The map ϕ induces a ring homomorphism

$$\begin{aligned} \mathcal{O}_Y &\to \mathcal{O}_X \\ f &\mapsto f \circ \phi. \end{aligned}$$

In terms of polynomials, suppose that

$$G(y_1, \ldots, y_m) \in I(Y) \subset k[y_1, \ldots, y_m].$$

Then we require that

$$G(f_1, f_2, \ldots, f_m) \in I(X).$$

Thus the map

$$\phi : X \to Y$$

[2]The first subsection is largely based on David Perkinson's lectures in the Undergraduate Summer School at the Park City Mathematics Institute in 2008.

induces a ring homomorphism

$$\phi^* : k[y_1, \ldots, y_m]/I(Y) = \mathcal{O}_Y \to \mathcal{O}_X = k[x_1, \ldots, x_n]/I(X).$$

Exercise 4.18.1. Let $X = \mathrm{V}(v - u^2)$, and let $Y = \mathrm{V}(z^2 - xy)$. We may think of X as a parabola and Y as a double cone. Define a morphism

$$\begin{array}{rcl} \phi : X & \to & Y \\ (u, v) & \mapsto & (1, v, u). \end{array}$$

Show that the image of ϕ is actually in Y.

Exercise 4.18.2. Keeping the same notation as in the above problem, show for the corresponding ring homomorphism

$$\phi^* : \mathcal{O}_Y = \mathbb{C}[x, y, z]/\langle z^2 - xy\rangle \to \mathcal{O}_X = \mathbb{C}[u, v]/\langle v - u^2\rangle,$$

we have

$$\phi^*(x^2 + xy + xz^3) = 1 + v + u^3.$$

Exercise 4.18.3. For each of the polynomial mappings $\phi : X \to Y$, describe the corresponding ring homomorphism $\phi^* : \mathcal{O}_Y \to \mathcal{O}_X$.

(1)
$$\begin{array}{rcl} \phi : \mathbb{A}^2(k) & \to & \mathbb{A}^3(k) \\ (x, y) & \mapsto & (y - x^2, xy, x^3 + 2y^2) \end{array}$$

(2) $X = \mathbb{A}^1(k)$ and $Y = \mathrm{V}(y - x^3, z - xy) \subset \mathbb{A}^3(k)$.

$$\begin{array}{rcl} \phi : X & \to & Y \\ t & \mapsto & (t, t^3, t^4) \end{array}$$

Exercise 4.18.4. For each of the ring homomorphisms $\sigma : \mathcal{O}_Y \to \mathcal{O}_X$, describe the corresponding morphism of algebraic sets, $X \to Y$. (This is actually not hard; in part, you should see that the answers are almost given to you. If that is not clear, you should go back and look at the definitions again.)

(1)
$$\begin{array}{rcl} \sigma : k[x, y] & \to & k[t] \\ x & \mapsto & t^2 - 1 \\ y & \mapsto & t(t^2 - 1) \end{array}$$

(2) $\sigma : k[s,t,u,v]/\langle s^2 - v, sv - tu\rangle \to k[x,y,z]/\langle xy - z^2\rangle$

$$\begin{aligned} s &\mapsto xy \\ t &\mapsto yz \\ u &\mapsto xz \\ v &\mapsto z^2 \end{aligned}$$

The morphism constructed here is a mapping of the saddle surface to a surface in $\mathbb{A}^4(k)$.

4.18.2. Spec and Morphisms. When you first read the definition of Spec(R) in Section 4.12, you probably asked yourself why we let the elements of Spec(R) be all prime ideals, rather than only consider those that are maximal. After all, in Sections 4.6 and 4.13 we established the correspondences between points in an affine variety and the maximal ideals in its coordinate ring, whereas prime ideals were shown to correspond to the irreducible subvarieties in Section 4.7. Thus our definition of "points" as prime ideals should be confusing. Additionally, our definition leads to complications such as points contained within other points, which results in our need to distinguish between geometric and generic points and those somehow in between. It is also the reason why we can only "reasonably identify" Spec($\mathbb{C}[x_1,\ldots,x_n]$) with $\mathbb{A}^n$ (and Proj($\mathbb{C}[x_0,x_1,\ldots,x_n]$) with $\mathbb{P}^n$ in Section 5.6) as $\mathbb{A}^n$ does not have non-closed points. Indeed, there are those who define the *maximal spectrum* of a ring R, denoted m-Spec(R) or Specm(R), to be the set of all maximal ideals of the ring R. So you may be wondering why we didn't.

In the following exercises, we will see one reason to prefer Spec over m-Spec, even though the former tends to challenge our geometric intuition regarding the nature of points.

Recall that whenever $\phi : X \to Y$ is a morphism of affine varieties, there is a corresponding ring homomorphism $\phi^* : \mathcal{O}_Y \to \mathcal{O}_X$ between their coordinate rings given by $f \mapsto f \circ \phi$. The converse is also true, as we now show, so that the set of morphisms from X to Y may be identified with the set of ring homomorphisms $\mathcal{O}_Y \to \mathcal{O}_X$.

Exercise 4.18.5. Let $X = \mathrm{V}(I) \subset \mathbb{A}^n$ and $Y = \mathrm{V}(J) \subset \mathbb{A}^m$ be affine varieties. Write $\mathcal{O}_X = k[x_1,\ldots,x_n]/I$ and $\mathcal{O}_Y = k[y_1,\ldots,y_m]/J$. Suppose $\varphi : \mathcal{O}_Y \to \mathcal{O}_X$ is a ring homomorphism.

(1) Let $g_i = \varphi(y_i + J) \in \mathcal{O}_X$ for $i = 1, \ldots, m$ and define $\psi : X \to \mathbb{A}^m$ by $\psi(p) = (g_1(p), \ldots, g_m(p))$. Show that $\psi(p) \in Y$ for all $p \in X$.

(2) Show that $\psi : X \to Y$ is a morphism of affine varieties.

(3) Show that, for all $f \in \mathcal{O}_Y$, $\varphi(f) = f \circ \psi$.

(4) Conclude that the set of morphisms $X \to Y$ is in a one-to-one correspondence with the set of ring homomorphisms $\mathcal{O}_Y \to \mathcal{O}_X$.

We want the same to be true for Spec. That is, for rings R and S, we want the set of morphisms $\mathrm{Spec}(R) \to \mathrm{Spec}(S)$ to be the same as the set of ring homomorphisms $S \to R$. This is where our need for prime rather than maximal ideals will arise. Before we show this, however, we want to give another view of how a morphism $X \to Y$ may be described in terms of the ring homomorphism $\mathcal{O}_Y \to \mathcal{O}_X$.

Exercise 4.18.6. Let $\phi : X \to Y$ be a morphism between affine varieties and let $\phi^* : \mathcal{O}_Y \to \mathcal{O}_X$ be the corrresponding ring homomorphism. For $p \in X$, let $\mathfrak{m}_p = \{f \in \mathcal{O}_X : f(p) = 0\}$ and $\mathfrak{m}_{\phi(p)} = \{g \in \mathcal{O}_Y : g(\phi(p)) = 0\}$. Prove that $\mathfrak{m}_{\phi(p)} = (\phi^*)^{-1}(\mathfrak{m}_p)$.

The points $p \in X$ and $\phi(p) \in Y$ are completely determined by their maximal ideals, $\mathfrak{m}_p$ and $\mathfrak{m}_{\phi(p)}$. That is, $\{p\} = \mathrm{V}(\mathfrak{m}_p)$ and $\{\phi(p)\} = \mathrm{V}(\mathfrak{m}_{\phi(p)})$, so the morphism ϕ may either be viewed as a function mapping points in X to points in Y or as a mapping from ideals in $\mathcal{O}_X$ to ideals in $\mathcal{O}_Y$ via $\mathfrak{m} \mapsto (\phi^*)^{-1}(\mathfrak{m})$.

This second view of the morphism ϕ is the one we will use to define morphisms $\mathrm{Spec}(R) \to \mathrm{Spec}(S)$. In order for our definition to agree with the definition for varieties above, it must be defined by mapping a "point" $P \in \mathrm{Spec}(R)$ to its preimage $\varphi^{-1}(P)$ under the corresponding ring homomorphism $\varphi : S \to R$. Before we can make this a definition, we need to confirm that $\varphi^{-1}(P) \in \mathrm{Spec}(S)$ whenever $P \in \mathrm{Spec}(R)$.

Exercise 4.18.7. Let $\varphi : S \to R$ be a ring homomorphism.

(1) Let $J \subset R$ be an ideal in R. Show that $\varphi^{-1}(J)$ is an ideal in S.

(2) Let $J \subset R$ be a prime ideal in R. Show that $\varphi^{-1}(J)$ is a prime ideal in S.

Based on these results, we can define morphisms between $\mathrm{Spec}(R)$ and $\mathrm{Spec}(S)$.

Definition 4.18.1. Let $\varphi : S \to R$ be a ring homomorphism. The corresponding *morphism* $\psi : \mathrm{Spec}(R) \to \mathrm{Spec}(S)$ is given by $\psi(P) = \varphi^{-1}(P)$ for all $P \in \mathrm{Spec}(R)$.

Exercise 4.18.8. Show that a morphism $\psi : \mathrm{Spec}(R) \to \mathrm{Spec}(S)$ is continuous in the Zariski topology. (Recall that a function $f : X \to Y$ between topological spaces is *continuous* if $f^{-1}(U)$ is open in X whenever U is open in Y.)

Now suppose that we were using m-Spec rather than Spec. Could we still define morphisms as we have above, which is the only way that is compatible with their definition as polynomial maps for affine varieties? This would require that the preimage of a maximal ideal under a ring homomorphism is again maximal. In the next exercises you will see that this does not always happen.

Exercise 4.18.9. Consider the ring homomorphism $\iota : \mathbb{Z} \to \mathbb{Q}$ given by inclusion, $\iota(n) = n$.

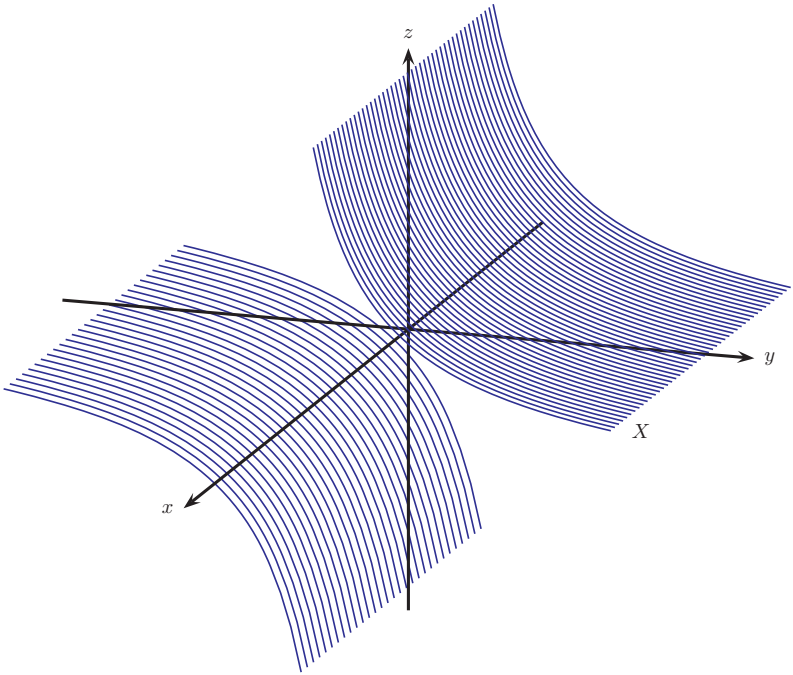
(1) What are the maximal ideals in $\mathbb{Z}$?

(2) Because $\mathbb{Q}$ is a field, $\langle 0 \rangle$ is its only maximal ideal. Find $\iota^{-1}(\langle 0 \rangle)$. Is it a maximal ideal in $\mathbb{Z}$?

We want to find an example that is algebraically similar to the above, but which has a more geometric flavor. In $\mathbb{C}^3$, with variables x, y and z, consider

$$X = \mathrm{V}(yz - 1).$$

This is a surface in $\mathbb{C}^3$ that is the product of $\mathbb{C}$ with the hyperbola $yz = 1$.



There is the natural projection from X to the affine plane $\mathbb{C}^2$, with coordinates x and y, given by

$$(x, y, z) \to (x, y).$$

The corresponding ring homomorphism is given by $T : \mathbb{C}[x, y] \to \mathbb{C}[x, y, z]/\langle yz - 1\rangle$, where $T(f(x, y)) = f(x, y)$.

Exercise 4.18.10. Consider the ring homomorphism

$$T : \mathbb{C}[x, y] \to \mathbb{C}[x, y, z]/\langle yz - 1\rangle$$

given by $T(f(x, y)) = f(x, y)$.

(1) Show that the ideal $\langle x\rangle$ of all multiples of x in $\mathbb{C}[x, y, z]/\langle yz - 1\rangle$ is a maximal ideal.

(2) Show that the inverse ideal $T^{-1}(\langle x\rangle)$ is the ideal $\langle x\rangle$ of all multiples of x in $\mathbb{C}[x, y]$.

(3) Show that the ideal $\langle x \rangle$ of all multiples of x in $\mathbb{C}[x, y]$ is not maximal.

Based on the results of the last exercises, if we were to use m-Spec rather than Spec, we would lose our ability to define morphisms as functions that send points to points. As we discussed in the preface, however, the approach to mathematics that has prevailed in algebraic geometry is linked to equivalence problems, and the solution of such problems not only requires the study of functions defined on the objects in question but also the functions between them. Thus morphisms play a critical role in algebraic geometry, as will the isomorphisms and rational maps that we encounter in the next sections, so their loss would be a great loss indeed.

4.19. Isomorphisms of Varieties

The goal of this section is to define a natural type of equivalence for algebraic sets.

4.19.1. Definition of Isomorphism. Let $V_1 = \mathrm{V}(I_1) \subset \mathbb{A}^n(k)$ and $V_2 = \mathrm{V}(I_2) \subset \mathbb{A}^m(k)$ be algebraic sets. We will assume in the following that each I_j is a radical ideal. As we have already seen, each ring $\mathcal{O}(V_i)$ is in a natural way the ring of (equivalence classes of) polynomial functions mapping V_i to k. We can then define a polynomial map (or morphism) $P : V_1 \to V_2$ by $P(x_1, \ldots, x_n) = (P_1(x_1, \ldots, x_n), \ldots, P_m(x_1, \ldots, x_n))$ where each $P_i \in \mathcal{O}(V_1)$.

Definition 4.19.1. A polynomial map $P : V_1 \to V_2$ is an *isomorphism* of algebraic sets if there exists a polynomial map $Q : V_2 \to V_1$ such that $Q \circ P = \mathrm{Id}|_{V_1}$ and $P \circ Q = \mathrm{Id}|_{V_2}$. Two algebraic sets are isomorphic if there exists an isomorphism between them, which we denote by $V_1 \cong V_2$.

Exercise 4.19.1. Let $V_1 = \mathrm{V}(x) \subset \mathbb{C}^2$ and $V_2 = \mathrm{V}(x + y) \subset \mathbb{C}^2$.

(1) Sketch V_1 and V_2 in $\mathbb{R}^2$.

(2) Find a one-to-one polynomial map $P(x, y) = (P_1(x, y), P_2(x, y))$ that maps V_1 onto V_2.

(3) Show $V_1 \cong V_2$ as varieties by finding an inverse polynomial map $Q(x, y)$ for the polynomial map $P(x, y)$ above. Verify that $Q \circ P = \mathrm{Id}\Big|_{V_1}$ and $P \circ Q = \mathrm{Id}\Big|_{V_2}$.

Exercise 4.19.2. Let $V_1 = \mathbb{C}$ and $V_2 = \mathrm{V}(x - y^2) \subset \mathbb{C}^2$ be algebraic sets.

(1) Sketch V_2 in $\mathbb{R}^2$.

(2) Find a one-to-one polynomial map P that maps V_1 onto V_2.

(3) Show $V_1 \cong V_2$ as algebraic sets by finding an inverse $Q(x, y)$ for the polynomial map $P(x)$ above. Verify that $Q \circ P = \mathrm{Id}\Big|_{V_1}$ and $P \circ Q = \mathrm{Id}\Big|_{V_2}$.

Exercise 4.19.3. Let $V_1 = \mathrm{V}(x^2 + y^2 - 1) \subset \mathbb{C}^2$ and $V_2 = \mathrm{V}(x^2 - y^2 - 1) \subset \mathbb{C}^2$ be varieties.

(1) Find a one-to-one polynomial map $P(x, y)$ that maps V_1 onto V_2.

(2) Show $V_1 \cong V_2$ as varieties by finding an inverse $Q(x, y)$ for the polynomial map $P(x, y)$ above. Verify that $Q \circ P = \mathrm{Id}\Big|_{V_1}$ and $P \circ Q = \mathrm{Id}\Big|_{V_2}$.

(3) Restricting to the real numbers, do you think $\mathrm{V}(x^2 + y^2 - 1) \subset \mathbb{R}^2$ and $\mathrm{V}(x^2 - y^2 - 1) \subset \mathbb{R}^2$ are isomorphic as varieties? Why or why not?

Exercise 4.19.4. Let k be any algebraically closed field. Let

$$V_1 = \mathrm{V}(x + y, z - 1) \subset \mathbb{A}^3(k) \quad \text{and } V_2 = \mathrm{V}(x - z^2, y + z) \subset \mathbb{A}^3(k)$$

be varieties.

(1) Find a polynomial map $P(x, y, z)$ that is a one-to-one and onto map from V_1 to V_2.

(2) Show $V_1 \cong V_2$ as varieties by finding an inverse $Q(x, y, z)$ for the polynomial map $P(x, y, z)$ above. Verify that $Q \circ P = \mathrm{Id}\Big|_{V_1}$ and $P \circ Q = \mathrm{Id}\Big|_{V_2}$.

4.19.2. Link to Ring Isomorphisms. Let's now consider the relationship between the coordinate rings $\mathcal{O}(V_1)$ and $\mathcal{O}(V_2)$ of two varieties. From the last section we know that there is a correspondence between polynomial maps

$$\phi : V_1 \to V_2$$

of varieties and ring homomorphisms

$$\phi^* : \mathcal{O}(V_2) \to \mathcal{O}(V_1)$$

of their coordinate rings, given by $\phi^*(f) = f \circ \phi$. We now want to show that two algebraic sets are isomorphic if and only if their coordinate rings are isomorphic as rings.

The next four exercises are the ring-theoretic versions of the exercises in the previous subsection.

Exercise 4.19.5. Consider Exercise 4.19.1. We have

$$\mathcal{O}(V_1) = \mathbb{C}[x, y]/\langle x \rangle,\ \mathcal{O}(V_2) = \mathbb{C}[x, y]/\langle x + y \rangle.$$

For the polynomial maps

$$P : V_1 \to V_2 \quad Q : V_2 \to V_1$$

with corresponding maps

$$P^* : \mathcal{O}(V_2) \to \mathcal{O}(V_1) \quad Q^* : \mathcal{O}(V_1) \to \mathcal{O}(V_2),$$

show the following:

(1) Let $f, g \in \mathbb{C}[x, y]$ agree on V_2, i.e., $f - g \in \langle x + y \rangle$. Show that $P^*(f) = P^*(g)$ on V_1.

(2) Show that P^* is a ring isomorphism by showing that Q^* is the inverse ring homomorphism.

Exercise 4.19.6. Using the notation from Exercise 4.19.2, show that

$$\mathbb{C}[t] \cong \mathbb{C}[x, y]/\langle x - y^2 \rangle$$

as rings by looking at the ring homomorphisms

$$P^* : \mathcal{O}(V_2) \to \mathcal{O}(V_1)$$

and

$$Q^* : \mathcal{O}(V_1) \to \mathcal{O}(V_2).$$

Exercise 4.19.7. Recalling Exercise 4.19.3, show $\mathbb{C}[x,y]/\langle x^2+y^2-1\rangle \cong \mathbb{C}[x,y]/\langle x^2-y^2-1\rangle$ as rings.

Exercise 4.19.8. Recalling Exercise 4.19.4, show

$$k[x,y,z]/\langle x+y, z-1\rangle \cong k[x,y,z]/\langle x-z^2, y+z\rangle$$

as rings.

Exercise 4.19.9. Let

$$V_1 = \mathrm{V}(I_1) \subset \mathbb{A}^n(k),\ V_2 = \mathrm{V}(I_2) \subset \mathbb{A}^m(k),\ V_3 = \mathrm{V}(I_3) \subset \mathbb{A}^i(k)$$

be three algebraic sets and suppose

$$P : V_1 \to V_2 \text{ and } \ Q : V_2 \to V_3$$

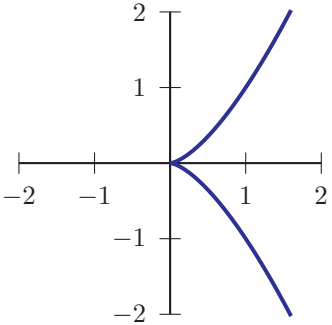
are polynomial maps.

(1) Show $(Q \circ P)^* = P^* \circ Q^*$.

(2) Show that if P is an isomorphism of varieties, then P^* is an isomorphism of rings.

Exercise 4.19.10. Let X and Y be algebraic sets and suppose that we have a ring isomorphism $\varphi : \mathcal{O}(Y) \to \mathcal{O}(X)$. Show that the algebraic sets X and Y are isomorphic.

The last series of exercises for this section deal with two varieties that are not isomorphic. Let $V_1 = \mathbb{A}^1(\mathbb{C})$ with coordinate t. Thus $\mathcal{O}(V_1) = \mathbb{C}[t]$. Our second variety is $V_2 = \mathrm{V}(x^3 - y^2) \subset \mathbb{A}^2(\mathbb{C})$, with $\mathcal{O}(V_2) = \mathbb{C}[x,y]/\langle x^3 - y^2\rangle$. The curve V_2 is the cuspidal cubic.



Thus the curve V_2 has a singular point while V_1 does not. If two varieties being isomorphic means intuitively that the two varieties

are the same up to a change of coordinates, we expect these varieties should not be isomorphic.

Exercise 4.19.11. Verify that $P(t) = (t^2, t^3)$ is a one-to-one polynomial map that maps V_1 onto V_2.

Despite the existence of this one-to-one polynomial map $P : V_1 \to V_2$, there is no inverse polynomial map $Q : V_2 \to V_1$. We show this by showing that the rings $\mathcal{O}(V_1)$ and $\mathcal{O}(V_2)$ are not isomorphic.

Exercise 4.19.12. Show that $\mathbb{C}[t] \ncong \mathbb{C}[x, y]/\langle x^3 - y^2 \rangle$ as rings. [Hint: Showing that P^* is not an isomorphism is not enough. You must show that there is *no* isomorphism between these rings. You may assume that the ring $\mathbb{C}[t]$ is a unique factorization domain and that both x and y are irreducible elements in the ring $\mathbb{C}[x, y]/\langle x^3 - y^2 \rangle$.]

4.20. Rational Maps

The goal of this section is to define another natural mapping of algebraic sets: rational maps.

There are two natural notions of equivalence in algebraic geometry: isomorphism and birationality. An isomorphism of varieties is given by polynomial maps while birational equivalence is given by rational maps. In this section we establish the correspondence between rational maps of varieties and homomorphisms of the associated function fields from Section 4.10.

4.20.1. Rational Maps. Let $V = \mathrm{V}(I)$ be an affine variety. Recall from Section 4.10 that the function field $\mathcal{K}_V$ of V is the field of fractions of the coordinate ring $\mathcal{O}_V$ of V, that is

$$\mathcal{K}_V = \left\{ \frac{f}{g} : f, g \in \mathcal{O}(V), g \notin I \right\} / \sim$$

where

$$\frac{f_1}{g_1} \sim \frac{f_2}{g_2} \quad \text{if and only if} \quad f_1 g_2 - f_2 g_1 \in I.$$

The elements of $\mathcal{K}_V$ are called *rational functions* on V.

Earlier we noted that a member f of the coordinate ring $\mathcal{O}_V$ is called a regular function and may be regarded as a function $f : V \to k$. Two members of $\mathcal{O}_V$ are equal if their difference is in $I(V)$, or equivalently, if they agree at every point in V. A difficulty arises with rational functions that does not arise with regular functions, in that a rational function is not, properly speaking, a function on V. By definition a rational function $F \in \mathcal{K}_V$ is of the form $\frac{f}{g}$ where $f, g \in \mathcal{O}_V$ and $g \notin I(V)$, so F is only partially defined on V. In particular, F is defined only at points p where $g(p) \neq 0$. This leads to the notion of a regular point.

Definition 4.20.1. A rational function $F \in \mathcal{K}_V$ is *regular* at $p \in V$ if there exist $f, g \in \mathcal{O}(V)$ such that F can be written in the form $F = \frac{f}{g}$ and $g(p) \neq 0$. The point p is called a *regular point* of F if F is regular at p. The collection of regular points of F is called the *domain of definition* of F and denoted $\mathrm{Dom}(F)$:

$$\mathrm{Dom}(F) = \{p \in V : F \text{ is regular at } p\}.$$

This means that we can regard a rational function $F \in \mathcal{K}_V$ as a function on its domain of definition. It is important to remember that elements of $\mathcal{K}_V$ are equivalence classes. F may have essentially different representatives as functions on V, though all representatives agree on (open) subsets of V. One consequence of this is that though F may be expressed as $\frac{f}{g}$ in $\mathcal{K}_V$, we cannot necessarily conclude that the domain of definition of F consists only of points for which g is nonzero. In particular, there may be points at which g vanishes that are in the domain of definition of F because there may be another representative in the same class which is defined at those points.

Exercise 4.20.1. Let $V = \mathrm{V}(xz - yw) \subset \mathbb{C}^4$, and let $F = \frac{x}{y}$. Show that there are regular points of F in $\mathrm{V}(y)$.

Exercise 4.20.2. Let V be a variety in $\mathbb{A}^n$ and $F \in \mathcal{K}_V$. Show that $\mathrm{Dom}(F)$ is a nonempty Zariski open subset of V.

This exercise shows that rational functions on V are defined on open subsets of V. Since any open subset of a variety V is dense, we see that the rational functions on V are defined on "most of" V. (To

say that an open subset U of V is *dense* means that any open subset of V intersects U.)

Definition 4.20.2. Suppose $V = \mathrm{V}(I) \subset \mathbb{A}^n$ is an affine variety. A *rational map* ϕ from V to $\mathbb{A}^m$ is an m-tuple of rational functions, that is, ϕ is given by $\phi(x_1, \ldots, x_n) = (F_1(x_1, \ldots, x_n), \ldots, F_m(x_1, \ldots, x_n))$, where $F_i \in \mathcal{K}_V$, $i = 1, \ldots, m$. Alternatively,

$$\phi(x_1, \ldots, x_n) = \left(\frac{f_1(x_1, \ldots, x_n)}{g_1(x_1, \ldots, x_n)}, \ldots, \frac{f_m(x_1, \ldots, x_n)}{g_m(x_1, \ldots, x_n)} \right),$$

where $f_i, g_i \in \mathcal{O}_V$, $i = 1, 2, \ldots, m$, and none of the g_i are in I.

Definition 4.20.3. We say ϕ is *regular* at a point $p \in V$ if each F_i is regular at p. Thus, the *domain of definition* of a rational map ϕ, denoted $\mathrm{Dom}(\phi)$, consists of all points in the domains of each F_i; that is,

$$\mathrm{Dom}(\phi) = \bigcap_{i=1}^{m} \mathrm{Dom}(F_i).$$

Since a rational map is defined only on a dense subset of V, we use the notation $\phi : V \dashrightarrow W$ to denote rational maps.

Exercise 4.20.3. Let $\phi : \mathbb{C}^2 \to \mathbb{C}^3$ be given by

$$\phi(x_1, x_2) = \left(\frac{x_1 + x_2}{x_1 - x_2}, \frac{x_1^2 + x_2}{x_1}, \frac{x_1 x_2^3}{x_1 + 3x_2} \right).$$

The rational map ϕ is not defined on three lines in $\mathbb{C}^2$. Draw these three lines as lines in $\mathbb{R}^2$.

Definition 4.20.4. The *image* of a rational map $\phi : V \dashrightarrow \mathbb{A}^m$ is the set

$$\phi(V) = \{\phi(p) \in \mathbb{A}^m : p \in V \text{ and } \phi \text{ is regular at } p\}.$$

Definition 4.20.5. Suppose $V = \mathrm{V}(I) \subset \mathbb{A}^n$ and $W = \mathrm{V}(J) \subset \mathbb{A}^m$ are affine varieties. A *rational map* $\phi : V \dashrightarrow W$ is a rational map $\phi : V \dashrightarrow \mathbb{A}^m$ such that $\phi(V) \subset W$.

Exercise 4.20.4. Show that the rational map

$$\phi(t) = \left(\frac{-2t}{1 + t^2}, \frac{1 - t^2}{1 + t^2} \right)$$

is a rational map from the line $\mathbb{C}$ to the circle $\mathrm{V}(x^2+y^2-1)$. Find the points on the line $\mathbb{C}$ where ϕ is not defined.

Exercise 4.20.5. The above rational map $\phi(t) = \left(\frac{-2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)$ was not made up out of thin air but reflects an underlying geometry. Let L be any line in the plane $\mathbb{C}^2$ through the point $(0,1)$ with slope t. Then the equation for this line is $y = tx+1$. First, draw a picture in $\mathbb{R}^2$ of the circle $\mathrm{V}(x^2+y^2-1)$ and the line L. Using the quadratic equation, show that the two points of intersection are $(0,1)$ and $\left(\frac{-2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)$, for a fixed slope t. Explain the underlying geometry of the map ϕ for when the slope t is zero.

4.20.2. Birational Equivalence. As we noted in the introduction there are two notions of equivalence of varieties. We have already studied isomorphisms, so we now turn our attention to birational equivalence. Recall that a morphism $\phi : V \to W$ is an isomorphism if there exists an inverse morphism ψ, i.e., a morphism $\psi : W \to V$ such that $\phi \circ \psi = \mathrm{Id}|_W$ and $\psi \circ \phi = \mathrm{Id}|_V$. To understand inverse morphisms we first had to understand compositions of morphisms. The definition of a birational map will follow the same template where we define composition of rational maps in the obvious way (which we make precise below), but a difficulty arises with rational maps that is not present in the case of isomorphisms, namely, compositions of rational maps may not be defined.

Exercise 4.20.6. Let $\psi : \mathbb{A}^1 \dashrightarrow \mathbb{A}^2$ be defined by $\psi(x) = (x,-x)$ and $\phi : \mathbb{A}^2 \dashrightarrow \mathbb{A}^1$ be defined by $\phi(x,y) = \frac{x}{x+y}$. Show that the domain of definition of $\phi \circ \psi$ is empty.

Suppose $\psi : V_1 \dashrightarrow V_2$ and $\phi : V_2 \dashrightarrow V_3$ are rational maps of affine varieties. Since ϕ is not defined on all of V_2, a problem arises when the image of ψ is not "big enough." Hence, we need to restrict our attention to only those rational maps whose images are sufficiently large. A natural question is "what is big enough?"

Definition 4.20.6. A rational map $\phi : V \dashrightarrow W$ is called *dominant* if $\phi(V)$ is dense in W.

We can now define the composition of two rational maps of affine varieties. Let $V_1 \subset \mathbb{A}^{n_1}$, $V_2 \subset \mathbb{A}^{n_2}$, and $V_3 \subset \mathbb{A}^{n_3}$ be affine varieties

with corresponding ideals I_1, I_2, and I_3. Suppose $\psi : V_1 \dashrightarrow V_2$ and $\phi : V_2 \dashrightarrow V_3$ are rational maps and that ψ is dominant. Let $(x_1, \ldots, x_{n_1})$ be affine coordinates on $\mathbb{A}^{n_1}$ and $(y_1, \ldots, y_{n_2})$ be affine coordinates on $\mathbb{A}^{n_2}$. Since ϕ and ψ are rational maps, we can write

$$\psi(x_1, \ldots, x_{n_1}) = (G_1(x_1, \ldots, x_{n_1}), \ldots, G_{n_2}(x_1, \ldots, x_{n_1})),$$

where $G_i \in \mathcal{K}_{V_1}$ and

$$\phi(y_1, \ldots, y_{n_2}) = (F_1(y_1, \ldots, y_{n_2}), \ldots, F_{n_3}(y_1, \ldots, y_{n_2})),$$

where $F_i \in \mathcal{K}_{V_2}$. Then

$$\phi \circ \psi(x_1, \ldots, x_{n_1}) = (F_1(G_1, \ldots, G_{n_2}), \ldots, F_{n_3}(G_1, \ldots, G_{n_2})).$$

We now check that this definition makes sense.

Exercise 4.20.7. Let $V_1 \subset \mathbb{A}^{n_1}$, $V_2 \subset \mathbb{A}^{n_2}$, and $V_3 \subset \mathbb{A}^{n_3}$ be affine varieties with corresponding ideals I_1, I_2, and I_3. Suppose $\psi : V_1 \dashrightarrow V_2$ and $\phi : V_2 \dashrightarrow V_3$ are rational maps and that ψ is dominant. Show that $\phi \circ \psi : V_1 \dashrightarrow V_3$ is a rational map.

Now we have that if $\psi : V_1 \dashrightarrow V_2$ is dominant, then $\phi \circ \psi : V_1 \dashrightarrow V_3$ is a rational map defined on a dense open subset of V_1. A consequence of this exercise is that composition of rational maps and homomorphisms of function fields are related in a very natural way. We will explore this further in the next subsection, but first we will define a birational map.

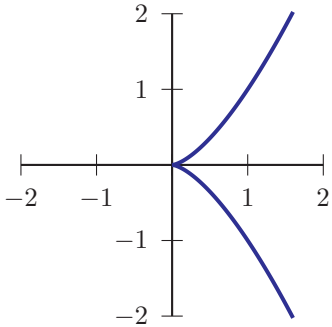
Definition 4.20.7. Suppose $V \subset \mathbb{A}^n(k)$ and $W \subset \mathbb{A}^m(k)$ are affine varieties. A dominant rational map $\phi : V \dashrightarrow W$ is called a *birational map* if there exists a dominant rational map $\psi : W \dashrightarrow V$ such that $\psi \circ \phi = \mathrm{Id}|_V$ and $\phi \circ \psi = \mathrm{Id}|_W$, where defined.

Definition 4.20.8. Two algebraic varieties $V \subset \mathbb{A}^n(k)$ and $W \subset \mathbb{A}^m(k)$ are *birationally equivalent* or *birational* if there exist birational maps between them. A variety that is birational to an affine space $\mathbb{A}^n$ is called *rational*.

Exercise 4.20.8. Show that the affine line $\mathbb{A}^1$ is birational to the circle $\mathrm{V}(x^2 + y^2 - 1)$ by finding an inverse to the rational map $\phi(t) = \left(\frac{-2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)$. [Hint: Recall that the map ϕ was obtained geometrically using the lines $y = tx + 1$ through the point $(0, 1)$ for various

slopes t, and then finding the line's second point of intersection with the circle.]

Exercise 4.20.9. Consider the cuspidal cubic curve $\mathrm{V}(y^2 - x^3)$ in the plane $\mathbb{A}^2$.



(1) Show that the map $\phi(t) = (t^2, t^3)$ maps the affine line $\mathbb{A}^1$ to the curve $\mathrm{V}(y^2 - x^3)$.

(2) Find a rational map $\psi : \mathrm{V}(y^2 - x^3) \dashrightarrow \mathbb{A}^1$ that is the inverse to the map ϕ.

Thus $\mathbb{A}^1$ and $\mathrm{V}(y^2 - x^3)$ are birational, even though they are not isomorphic as we have shown in Exercise 4.19.12.

4.20.3. Birational Equivalence and Field Isomorphisms. The goal of this subsection is to explore the relationships between rational maps of varieties and homomorphisms of function fields. In particular, we will establish the following theorem.

Theorem 4.20.10. Let $V = \mathrm{V}(I) \subset \mathbb{A}^n$ and $W = \mathrm{V}(J) \subset \mathbb{A}^m$ be two algebraic varieties. Then V and W are birational if and only if the function fields $\mathcal{K}_V$ and $\mathcal{K}_W$ are field isomorphic.

In Exercise 4.19.12 we showed that $\mathbb{A}^1$ and $\mathrm{V}(y^2 - x^3)$ are not isomorphic because their coordinate rings are not isomorphic. In Exercise 4.20.9 we established that $\mathbb{A}^1$ and $\mathrm{V}(y^2 - x^3)$ are birational. In this exercise we will show that their function fields are isomorphic.

Exercise 4.20.11. Let $V = \mathrm{V}(y^2 - x^3)$ be the cuspidal cubic in the plane $\mathbb{A}^2$.

(1) Let $\phi : \mathbb{A}^1 \dashrightarrow V$ be defined by $\phi(t) = (t^2, t^3)$. Show that the map $\phi^* : \mathcal{K}_V \to k(t)$ defined by $\phi^*(F) = F \circ \phi$ is a field homomorphism.

(2) Show that $y = \left(\frac{y}{x}\right)^3$ and $x = \left(\frac{y}{x}\right)^2$ in the field $\mathcal{K}_V$.

(3) Show that $k(t)$ and $\mathcal{K}_V$ are isomorphic as fields by showing that $(\phi^*)^{-1} = \psi^*$, where ψ^* is the field homomorphism associated to the rational map ψ found in Exercise 4.20.9.

The next series of exercises will provide a proof that affine varieties V and W are birational if and only if the function fields $\mathcal{K}_V$ and $\mathcal{K}_W$ are isomorphic.

Exercise 4.20.12. Let $V \subset \mathbb{A}^m$ and $W \subset \mathbb{A}^n$ be affine varieties and let $\phi : V \dashrightarrow W$ be a rational map. Show that there is a natural ring homomorphism

$$\phi^* : \mathcal{O}(W) \to \mathcal{K}_V.$$

Exercise 4.20.13. Let $V \subset \mathbb{A}^m$ and $W \subset \mathbb{A}^n$ be affine varieties and let $\phi : V \dashrightarrow W$ be a dominant rational map. Show that there is a natural field homomorphism

$$\phi^* : \mathcal{K}_W \to \mathcal{K}_V.$$

Exercise 4.20.14. Let $V \subset \mathbb{A}^m$ and $W \subset \mathbb{A}^n$ be affine varieties and let $\alpha : \mathcal{K}_W \to \mathcal{K}_V$ be a field homomorphism. Show that there exists a unique dominant rational map $\phi : V \dashrightarrow W$ such that $\phi^* = \alpha$.

Exercise 4.20.15. Let $\phi : \mathbb{A}^1 \dashrightarrow \mathbb{A}^2$ be defined by $\phi(t) = \left(t, \frac{1}{t}\right)$.

(1) Show that ϕ is not dominant.

(2) Show that $\phi^* : k[x, y] \to k(t)$ is a ring homomorphism.

(3) Show that $\phi^* : k(x, y) \to k(t)$ is not a field homomorphism.

Exercise 4.20.16. Suppose $\psi : V_1 \dashrightarrow V_2$ and $\phi : V_2 \dashrightarrow V_3$ are rational maps and that ψ is dominant. Show that $(\phi \circ \psi)^* : \mathcal{K}_{V_3} \to \mathcal{K}_{V_1}$ and $\psi^* \circ \phi^* : \mathcal{K}_{V_3} \to \mathcal{K}_{V_1}$ are the same.

Exercise 4.20.17. Show that $\phi : V \dashrightarrow W$ is a birational map if and only if $\phi^* : \mathcal{K}_W \to \mathcal{K}_V$ is a field isomorphism.

4.20.4. Blow-ups and Rational Maps. In Section 3.7 we saw that the blow-up of the origin $(0,0)$ in $\mathbb{C}^2$ is the space obtained by replacing the origin by the set of all complex lines in $\mathbb{C}^2$ through the origin. In coordinates, the blow-up consists of two copies of $\mathbb{C}^2$ that are patched together correctly. This section shows how these patchings can be viewed as appropriate birational maps.

Let $U = \mathbb{C}^2$, with coordinates u_1, u_2, and $V = \mathbb{C}^2$, with coordinates v_1, v_2, be the two complex planes making up the blow-up. Denote by $Z = \mathbb{C}^2$, with coordinates z_1, z_2, the original $\mathbb{C}^2$ whose origin is to be blown-up.

From Section 3.7, we have the polynomial maps

$$\pi_1 : U \to Z \text{ and } \pi_2 : V \to Z$$

given by

$$\pi_1(u_1, u_2) = (u_1, u_1 u_2) = (z_1, z_2)$$

and

$$\pi_2(v_1, v_2) = (v_1 v_2, v_2) = (z_1, z_2).$$

Exercise 4.20.18.

(1) Find the inverse maps

$$\pi_1^{-1} : Z \dashrightarrow U \text{ and } \pi_2^{-1} : Z \dashrightarrow V.$$

(2) Find the points in Z where the maps π_1^{-1} and π_2^{-1} are not defined.

(3) Conclude that U and Z are birational under π_1, as are V and Z under π_2.

Exercise 4.20.19. Find the maps

$$\pi_2^{-1} \circ \pi_1 : U \dashrightarrow V$$

and

$$\pi_1^{-1} \circ \pi_2 : V \dashrightarrow U.$$

Conclude that U and V are birational under these maps.

4.21. Products of Affine Varieties

The goal of this section is to show that the Cartesian product of affine varieties is again an affine variety. We also study the topology and function theory of the product of two affine varieties.

4.21.1. Product of Affine Spaces. In analytic geometry, the familiar xy-plane $\mathbb{R}^2$ is constructed as the Cartesian product of two real lines, $\mathbb{R} \times \mathbb{R}$, and thus is coordinatized by ordered pairs of real numbers. It is natural to ask whether the same construction can be used in algebraic geometry to construct higher-dimensional affine spaces as products of lower-dimensional ones.

Clearly we can identify $\mathbb{A}^2(k)$ with $\mathbb{A}^1(k) \times \mathbb{A}^1(k)$ as sets. However, this identification is insufficient to prove that $\mathbb{A}^2(k)$ is isomorphic to $\mathbb{A}^1(k) \times \mathbb{A}^1(k)$ as varieties.

Exercise 4.21.1. Let $\mathcal{O}(\mathbb{A}^n(k)) = k[x_1, \ldots, x_n]$ and $\mathcal{O}(\mathbb{A}^m(k)) = k[y_1, \ldots, y_m]$. Show that $\mathcal{O}(\mathbb{A}^{n+m}(k)) \cong k[x_1, \ldots, x_n, y_1, \ldots, y_m]$, where the latter is, by definition, the ring of regular functions on the product $\mathbb{A}^n(k) \times \mathbb{A}^m(k)$.

Frequently, when we form the product of topological spaces X and Y, the new space $X \times Y$ is endowed with the product topology. This topology has as its basis all sets of the form $U \times V$ where $U \subset X$ and $V \subset Y$ are open. In these exercises, the Zariski topology on the product $X \times Y$ will be compared to the product topology to determine that the Zariski topology is strictly finer.

Exercise 4.21.2. (This is very similar to Hartshorne [**Har77**], Exercise I.1.4.) In Exercise 4.21.1, you have shown that $\mathbb{A}^n(k) \times \mathbb{A}^m(k) \cong \mathbb{A}^{n+m}(k)$. In particular, $\mathbb{A}^1(k) \times \mathbb{A}^1(k) \cong \mathbb{A}^2(k)$.

(1) Describe an open set in the product topology on $\mathbb{A}^1(k) \times \mathbb{A}^1(k)$.

(2) Is an open set in the product topology on $\mathbb{A}^1(k) \times \mathbb{A}^1(k)$ also open in the Zariski topology of $\mathbb{A}^1(k) \times \mathbb{A}^1(k) \cong \mathbb{A}^2(k)$?

(3) Find an open set of the Zariski topology of $\mathbb{A}^1(k) \times \mathbb{A}^1(k) \cong \mathbb{A}^2(k)$ that is not open in the product topology.

This shows that the Zariski topology is *strictly finer* than the product topology on $\mathbb{A}^1(k) \times \mathbb{A}^1(k) \cong \mathbb{A}^2(k)$.

4.21.2. Product of Affine Varieties. Let $X \subset \mathbb{A}^n(k)$ and $Y \subset \mathbb{A}^m(k)$ be affine varieties. The Cartesian product of X and Y, $X \times Y$, can naturally be viewed as a subset of the Cartesian product $\mathbb{A}^n(k) \times \mathbb{A}^m(k)$.

Exercise 4.21.3. Let $X = \mathrm{V}(x_2 - x_1) \subset \mathbb{A}^2(k)$ and $Y = \mathrm{V}(y_1) \subset \mathbb{A}^2(k)$. Describe $X \times Y$ and show that it is a closed subset of $\mathbb{A}^4(k)$.

Exercise 4.21.4. If $X = \mathrm{V}(I) \subset \mathbb{A}^n(k)$ and $Y = \mathrm{V}(J) \subset \mathbb{A}^m(k)$ are algebraic sets, show that $X \times Y \subset \mathbb{A}^{n+m}(k)$ is also an algebraic set.

Let $X \subset \mathbb{A}^n(k)$ and $Y \subset \mathbb{A}^m(k)$ be affine subvarieties. Then $X \times Y$ is an algebraic subset of $\mathbb{A}^{n+m}(k)$. Endow $X \times Y$ with the subspace topology for the Zariski topology on $\mathbb{A}^{n+m}(k)$. (That is, the open sets of $X \times Y$ are sets of the form $(X \times Y) \cap U$ where U is Zariski open in $\mathbb{A}^{n+m}(k)$.) This is called the *product* of the affine varieties X and Y.

We now want to prove that the product of affine varieties is again an affine variety, which requires that we prove the product of irreducible sets is irreducible.

Exercise 4.21.5. Let $x_0 \in X$ be a (closed) point. Show that $\{x_0\} \times Y = \{(x_0, y) \in X \times Y : y \in Y\}$ is a subvariety of $X \times Y$ isomorphic to Y as a variety. Similarly, for any closed point $y_0 \in Y$, $X \times \{y_0\}$ is a subvariety of $X \times Y$ isomorphic to X.

In particular, if X is irreducible, so is $X \times \{y_0\}$ for each $y_0 \in Y$.

Our next goal is to show that if X and Y are irreducible, then $X \times Y$ is irreducible. While standard, we are basing these problems on Klaus Hulek's presentation in *Elementary Algebraic Geometry* [**Hul03**].

For the rest of this subsection, we assume that both $X \subset \mathbb{A}^n(k)$ and $Y \subset \mathbb{A}^m(k)$ are irreducible. Suppose that $X \times Y = Z_1 \cup Z_2$, where both Z_1 and Z_2 are algebraic sets in $\mathbb{A}^{n+m}(k)$ with $Z_1 \subsetneq X \times Y$ and $Z_2 \subsetneq X \times Y$.

Exercise 4.21.6. For any fixed $y_0 \in Y$ let $Y_0 = \mathbb{A}^n(k) \times \{y_0\}$. Show that

$$X \times \{y_0\} = Z_1 \cap Y_0$$

or

$$X \times \{y_0\} = Z_2 \cap Y_0.$$

From this, we deduce that $X \times \{y_0\}$ is entirely contained in one of Z_1 or Z_2 for each $y_0 \in Y$. Set

$$W_1 = \{y \in Y : X \times \{y\} \subset Z_1\}$$

and

$$W_2 = \{y \in Y : X \times \{y\} \subset Z_2\}.$$

Then

$$Y = W_1 \cup W_2.$$

Exercise 4.21.7. Show that if both W_1 and W_2 are Zariski closed, then $X \times Y$ is irreducible.

To finish the argument, we need to show that W_1 and W_2 are closed. Here are the necessary steps.

Exercise 4.21.8. For each $x \in X$, show that

$$W_i^x = \{y \in Y : (x, y) \in Z_i\}$$

is closed.

Exercise 4.21.9. Show that each W_i is closed. [Hint: Use that the intersection of closed sets, such as the ones in the previous problem, is closed.]

Thus, if X and Y are affine varieties, so is their product, $X \times Y$.

4.21.3. Products and Morphisms.

Exercise 4.21.10. Let $X \subset \mathbb{A}^n(k)$ and $Y \subset \mathbb{A}^m(k)$ be affine varieties.

(1) Show that $(x, y) \mapsto x$ is a morphism of affine varieties $\rho_X : X \times Y \to X$, called the projection onto the first factor.

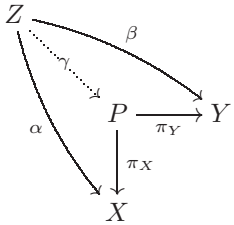
(2) Similarly, show that $(x, y) \mapsto y$ is a morphism, which we will denote by $\rho_Y : X \times Y \to Y$ and call the projection onto the second factor.

Exercise 4.21.11. Show that $\rho_X : X \times Y \to X$ and $\rho_Y : X \times Y \to Y$ are both *open morphisms*, i.e., if $U \subset X \times Y$ is an open subset, then $\rho_X(U)$ is an open subset of X and $\rho_Y(U)$ is an open subset of Y.

Exercise 4.21.12. Must ρ_X and ρ_Y also be *closed morphisms*, i.e., must the images of a closed set C in $X \times Y$ be closed in X and in Y?

Exercise 4.21.13. Suppose $\varphi : Z \to X$ and $\psi : Z \to Y$ are morphisms of affine varieties. Show that there is a well-defined morphism $\pi : Z \to X \times Y$ so that $\varphi = \rho_X \circ \pi$ and $\psi = \rho_Y \circ \pi$, where $\rho_X : X \times Y \to X$ and $\rho_Y : X \times Y \to Y$ are the projection morphisms.

This is the *universal property* for the product of varieties: Given X and Y, a variety P with morphisms $\pi_X : P \to X$ and $\pi_Y : P \to Y$ is the **product** of X and Y if, for any variety Z with morphisms $\alpha : Z \to X$ and $\beta : Z \to Y$, there is a unique morphism $\gamma : Z \to P$ so that



is a commutative diagram.

Therefore, if Q is another variety having this property, there are unique maps $\delta : P \to Q$, $\zeta : Q \to P$, $\pi : P \to P$ and $\varepsilon : Q \to Q$ by the universal property. Clearly, π, ε must both be the identity morphisms of P and Q, respectively. However, $\zeta \circ \delta : P \to P$ also satisfies the property of the arrow from P to itself, so that $\zeta \circ \delta = \pi$ is the identity on P. Similarly, $\delta \circ \zeta : Q \to Q$ is the identity morphism of Q, so ζ and δ are invertible morphisms which establish an isomorphism $P \cong Q$. Hence the product of two varieties is unique up to isomorphism.

Chapter 5

Projective Varieties

The key to this chapter is that projective space $\mathbb{P}^n$ is the natural ambient space for much of algebraic geometry. We will be extending last chapter's work on affine varieties to the study of algebraic varieties in projective space $\mathbb{P}^n$. We will see that in projective space we can translate various geometric objects into the language of graded rings. While affine varieties correspond to ideals in commutative rings, we will show that projective varieties in $\mathbb{P}^n$ correspond to homogeneous ideals.

You will observe that this chapter is much shorter than previous chapters. This is not because projective varieties are less important. Rather it is a reflection of the fact that, quite often, problems involving projective varieties can be reduced to the affine case and treated with methods developed in Chapter 4.

5.1. Definition of Projective Space

This section gives the basic definitions for projective n-space $\mathbb{P}^n(k)$.

In Chapter 1, we saw that all smooth conics in the complex projective plane $\mathbb{P}^2$ can be viewed as the "same." In Chapter 2, we saw that all smooth cubics in $\mathbb{P}^2$ can be viewed as tori. In Chapter 3,

we saw that curves of degree e and of degree f must intersect in exactly ef points, counted with multiplicity, provided we work in $\mathbb{P}^2$. All of this suggests that affine space $\mathbb{A}^n$ is not the natural place to study geometry; instead, we want to define some notion of projective n-space.

Let k be a field. (You can comfortably replace every k with the complex numbers $\mathbb{C}$, at least for most of this book.)

Definition 5.1.1. Let $a = (a_0, \ldots, a_n)$, $b = (b_0, \ldots, b_n) \in \mathbb{A}^{n+1}(k) - \{(0, \ldots, 0)\}$. We say that a *is equivalent to* b, denoted $a \sim b$, if there exists a $\lambda \neq 0$ in the field k such that

$$(a_0, \ldots, a_n) = \lambda(b_0, \ldots, b_n).$$

Exercise 5.1.1. In $\mathbb{A}^5 - \{(0, \ldots, 0)\}$, show

(1) $(1, 3, 2, 4, 5) \sim (3, 9, 6, 12, 15)$,

(2) $(1, 3, 2, 4, 5) \not\sim (3, 9, 6, 13, 15)$.

Exercise 5.1.2. Show that $\sim$ is an equivalence relation on $\mathbb{A}^n(k) - \{(0, \ldots, 0)\}$.

Definition 5.1.2. Projective n-space over the field k is

$$\mathbb{P}^n(k) = \left(\mathbb{A}^{n+1}(k) - \{(0, \ldots, 0)\}\right) / \sim .$$

We denote the equivalence class corresponding to a point $(a_0, \ldots, a_n)$ (with at least one $a_i \neq 0$) by

$$(a_0 : a_1 : \cdots : a_n).$$

We call the $(a_0 : a_1 : \cdots : a_n)$ *homogeneous coordinates* for $\mathbb{P}^n(k)$.

Exercise 5.1.3. Referring to Exercise 1.4.7, explain why $\mathbb{P}^n(k)$ can be thought of as the set of all lines through the origin in $\mathbb{A}^{n+1}(k)$.

We now want to examine the relationship between $\mathbb{A}^n(k)$ and $\mathbb{P}^n(k)$. There is a natural way to cover $\mathbb{P}^n(k)$ with $n + 1$ copies of $\mathbb{A}^n(k)$.

Exercise 5.1.4. Let $(a_0 : a_1 : a_2 : a_3 : a_4 : a_5) \in \mathbb{P}^5$. Suppose that $a_0 \neq 0$. Show that

$$(a_0, a_1, a_2, a_3, a_4, a_5) \sim \left(1, \frac{a_1}{a_0}, \frac{a_2}{a_0}, \frac{a_3}{a_0}, \frac{a_4}{a_0}, \frac{a_5}{a_0}\right).$$

Definition 5.1.3. Let $(x_0 : x_1 : \cdots : x_n)$ be homogeneous coordinates on $\mathbb{P}^n(k)$. Define the i^{th} *affine chart* to be

$$U_i = \{(x_0 : x_1 : \cdots : x_n) : x_i \neq 0\}.$$

Exercise 5.1.5. Prove that every element in $\mathbb{P}^n(k)$ is contained in at least one U_i. Thus the $n+1$ sets U_i, for $i = 0, \ldots, n$, cover $\mathbb{P}^n(k)$.

Exercise 5.1.6. Show that there is exactly one point in $\mathbb{P}^n(k)$ that is not in $U_1 \cup U_2 \cup \cdots \cup U_n$. Identify this point.

In the affine case, there is a natural way to link spaces with different dimensions: $\mathbb{A}^n(k)$ can be embedded in $\mathbb{A}^{n+1}(k)$ by mapping an n-tuple to an $(n+1)$-tuple with the last coordinate set equal to 0. Let's extend this so we can embed a projective space into a higher dimensional one.

Exercise 5.1.7. Show that we can map $\mathbb{P}^1(k)$ to the set of all points in $\mathbb{P}^n(k)$ that are not in $U_2 \cup U_3 \cup \cdots \cup U_n$.

Exercise 5.1.8. Show that we can map $\mathbb{P}^2(k)$ to the set of all points in $\mathbb{P}^n(k)$ that are not in $U_3 \cup U_4 \cup \cdots \cup U_n$.

Since there are $n+1$ copies of $\mathbb{A}^n(k)$ embedded in $\mathbb{P}^n(k)$, we need a way to move from one chart to another.

Definition 5.1.4. For $0 \leq i \leq n$, define maps $\phi_i : U_i \to \mathbb{A}^n(k)$ by

$$\phi_i(x_0 : x_1 : \cdots : x_n) = \left(\frac{x_0}{x_i}, \frac{x_1}{x_i}, \ldots, \widehat{x_i}, \ldots, \frac{x_n}{x_i}\right),$$

where $\widehat{x_i}$ means that x_i is omitted.

Exercise 5.1.9. For $\mathbb{P}^n(k)$, show for each i that $\phi_i : U_i \to \mathbb{A}^n(k)$ is

(1) one-to-one

(2) onto.

Since ϕ_i is one-to-one and onto, there is a well-defined inverse

$$\phi_i^{-1} : \mathbb{A}^n(k) \to \mathbb{P}^n(k).$$

Exercise 5.1.10. For $\phi_2^{-1} : \mathbb{A}^5 \to \mathbb{P}^5$, show that

$$\phi_2^{-1}(7, 3, 11, 5, 6) = (14 : 6 : 2 : 22 : 10 : 12).$$

Define maps $\psi_{ij} : \phi_j(U_i \cap U_j) \to \phi_i(U_i \cap U_j)$ by $\psi_{ij} = \phi_i \circ \phi_j^{-1}$ for $0 \leq i, j \leq n$.

Exercise 5.1.11. Explain how each ψ_{ij} is a rational map from $\mathbb{A}^n(k)$ to $\mathbb{A}^n(k)$.

Exercise 5.1.12. Show that the map $\psi_{02} : \mathbb{A}^2(k) \to \mathbb{A}^2(k)$ is

$$\psi_{02}(x_1, x_2) = \left(\frac{x_2}{x_1}, \frac{1}{x_1}\right).$$

Describe the set on which ψ_{02} is undefined.

Exercise 5.1.13. Explicitly describe $\psi_{12} : \mathbb{A}^2(k) \to \mathbb{A}^2(k)$. In other words, find $\psi_{12}(x_1, x_2)$. Describe the set on which ψ_{12} is undefined.

Exercise 5.1.14. Write explicitly the map $\psi_{02} : \phi_2(U_0 \cap U_2) \subset \mathbb{A}^n(k) \to \phi_0(U_0 \cap U_2) \subset \mathbb{A}^n(k)$ in coordinates $(x_1, x_2, \ldots, x_n)$. Describe the set on which ψ_{02} is undefined.

Exercise 5.1.15. Show that $\psi_{ij} \circ \psi_{jk} = \psi_{ik}$.

Exercise 5.1.16. Show that $\psi_{ij} \circ \psi_{jk} \circ \psi_{ki} = \mathrm{Id}$.

For those who have had topology, the above exercises show that $\mathbb{P}^n$ is a manifold. However, our primary interest in $\mathbb{P}^n(k)$ is as the natural ambient space for algebraic geometry. As in Chapter 4, we are seeking a dictionary between algebraic sets in $\mathbb{P}^n(k)$ and the sets of functions that vanish on them. Similar to our prior experience with $\mathbb{P}^2$ in Chapter 1, we will see that when working in projective n-space, we need to restrict our attention to homogeneous functions. The next section treats homogeneous polynomials and ideals in graded rings, which is the algebraic background needed to study algebraic geometry in projective space.

5.2. Graded Rings and Homogeneous Ideals

This section discusses why we need to consider graded rings and homogeneous ideals as the natural ring-theoretic objects to associate to projective varieties.

We want to study varieties in $\mathbb{P}^n(k)$, but first we must see why we cannot naively use the zero loci of arbitrary polynomials.

Exercise 5.2.1. Let

$$P(x_0, x_1, x_2, x_3, x_4, x_5) = x_0 - x_1x_2x_3x_4x_5.$$

(1) Show that
$$P(1, 1, 1, 1, 1, 1) = 0.$$
(2) Show that
$$P(2, 2, 2, 2, 2, 2) \neq 0.$$
(3) Show that
$$(1, 1, 1, 1, 1, 1) \sim (2, 2, 2, 2, 2, 2)$$
so that the two points in $\mathbb{C}^6$ will define the same point in $\mathbb{P}^5$.
(4) Conclude that $\{(x_0 : \ldots : x_5) \in \mathbb{P}^5 : P(x_0, \ldots, x_5) = 0\}$ is not a well-defined set.

As we have seen before, the key is to consider homogeneous polynomials.

Exercise 5.2.2. Let

$$P(x_0, x_1, x_2, x_3, x_4, x_5) = x_0^5 - x_1x_2x_3x_4x_5.$$

(1) Show that
$$P(1, 1, 1, 1, 1, 1) = 0.$$
(2) Show that
$$P(2, 2, 2, 2, 2, 2) = 0.$$
(3) Show that if $P(x_0, \ldots, x_5) = 0$, then for all $\lambda \in k$ we have
$$P(\lambda x_0, \ldots, \lambda x_5) = 0.$$
(4) Conclude that $\{(x_0 : \ldots : x_5) \in \mathbb{P}^5 : P(x_0, \ldots, x_5) = 0\}$ is a well-defined set.

The reason why the zero locus of $x_0^5 - x_1x_2x_3x_4x_5$ is a well-defined subset of $\mathbb{P}^5$ is that both terms x_0^5 and $x_1x_2x_3x_4x_5$ have degree five.

Definition 5.2.1. A polynomial for which each of its terms has the same degree is called *homogeneous*.

Exercises 5.2.1 and 5.2.2 suggest that we should consider only homogeneous polynomials to do algebraic geometry in projective space.

Exercise 5.2.3. If $f \in k[x_0, \ldots, x_n]$ is a homogeneous polynomial of degree d, then $f(\lambda x_0, \lambda x_1, \ldots, \lambda x_n) = \lambda^d f(x_0, x_1, \ldots, x_n)$ for every $\lambda \neq 0$ in the base field k.

Thus, even though the value of f at a point $P \in \mathbb{P}^n$ is not well-defined, the set of points at which f vanishes is well-defined. Hence, we restrict our attention to the zero loci of homogeneous polynomials when working in projective space $\mathbb{P}^n$.

First, we notice that when we work with homogeneous polynomials we gain additional structure on the ring $k[x_0, x_1, \ldots, x_n]$. Specifically, we can break up the polynomial ring $k[x_0, x_1, \ldots, x_n]$ in a natural way. Define R_d to be the set of all homogeneous polynomials of degree d in $k[x_0, x_1, \ldots, x_n]$. Note that the zero polynomial is in every R_d for $d \geq 0$.

Exercise 5.2.4. Let $R = k[x, y, z]$.

(1) Let $f = x + 2y$ and $g = x - z$. Show that $f + g$ and $f - g$ are in R_1 and $fg \in R_2$.

(2) Let $h = x^2 + yz$. Show that fh and gh are in R_3 and $h^2 \in R_4$.

Exercise 5.2.5. Let $R = k[x_0, x_1, \ldots, x_n]$.

(1) What is R_0?

(2) Show that if $f \in R_0$ and $g \in R_d$, then $fg \in R_d$.

(3) Show that for $f, g \in R_1$, $f + g \in R_1$ and $fg \in R_2$.

(4) Show that for $f, g \in R_d$, $f + g \in R_d$ and $fg \in R_{2d}$.

We can generalize Exercises 5.2.4 and 5.2.5 to show that $k[x_0, x_1, \ldots, x_n]$ is a graded ring.

Definition 5.2.2. A *graded ring* is a ring R together with a collection of subgroups R_d, $d \geq 0$, of the additive group R such that $R = \bigoplus_{d \geq 0} R_d$ and, for all $d, e \geq 0$, $R_d \cdot R_e \subseteq R_{d+e}$.

Exercise 5.2.6. As before, let $R = k[x_0, x_1, \ldots, x_n]$ with R_d the homogeneous polynomials of degree d.

(1) Prove that R_d is a group under addition.
(2) Prove for any $d, e \geq 0$, $R_d \cdot R_e \subseteq R_{d+e}$.
(3) Prove $k[x_0, x_1, \ldots, x_n] = \bigoplus_{d \geq 0} R_d$.

This notion of grading for a ring extends to ideals in the ring. Since we are interested in projective space and hence homogeneous polynomials, we define the related notion of a graded ideal.

Definition 5.2.3. An ideal I of a graded ring $R = \bigoplus_{d \geq 0} R_d$ is called *homogeneous* or *graded* if and only if $I = \bigoplus_{d \geq 0} (I \cap R_d)$.

Exercise 5.2.7. Determine whether each ideal of $k[x, y, z]$ is homogeneous.

(1) $\langle x - yz \rangle$
(2) $\langle x^2 - yz \rangle$
(3) $\langle x - yz, x^2 - yz \rangle$
(4) $\langle x^2 - yz, y^3 - xz^2 \rangle$

The next exercise gives us three equivalent descriptions for a homogeneous ideal.

Exercise 5.2.8. Prove that the following are equivalent.

(1) I is a homogeneous ideal of $k[x_0, \ldots, x_n]$.
(2) I is generated by homogeneous polynomials.
(3) If $f = \sum f_i \in I$, where each f_i is homogeneous, then $f_i \in I$ for each i.

The exercises in the rest of this section provide general results about graded rings and practice working with them.

Exercise 5.2.9. Let I be a homogeneous ideal in $R = k[x_0, \ldots, x_n]$. Prove the quotient ring R/I is a graded ring.

Exercise 5.2.10. Let $R = k[x, y, z]$ and $I = \langle x^2 - yz \rangle$. Show how to write R/I as a graded ring.

Exercise 5.2.11. Let $R = k[x, y, z, w]$ and $I = \langle xw - yz \rangle$. Show how to write R/I as a graded ring.

Exercise 5.2.12. Let $R = k[x, y, z]$ and let $I = \langle x^2 \rangle$, $J = \langle x, y \rangle$. Determine whether each ideal is homogeneous.

(1) $I \cap J$
(2) $I + J$
(3) IJ
(4) $\mathrm{Rad}(I) = \{f : f^m \in I \text{ for some } m > 0\}$

We can generalize these results to the intersections, sums, products, and radicals of any homogeneous ideals.

Exercise 5.2.13. Let A be an index set, and let $\{I_\alpha : \alpha \in A\}$ be a collection of homogeneous ideals in $k[x_0, \ldots, x_n]$. Also let J and J' be homogeneous ideals in $k[x_0, \ldots, x_n]$.

(1) Prove $\bigcap_{\alpha \in A} I_\alpha$ is homogeneous.
(2) Prove $\sum_{\alpha \in A} I_\alpha$ is homogeneous.
(3) Prove JJ' is homogeneous.
(4) Prove $\mathrm{Rad}(J)$ is homogeneous.

We will see that, as in the affine case, prime ideals correspond to irreducible varieties. The next exercise shows that to prove a homogeneous ideal is prime, it is sufficient to restrict to homogeneous elements.

Exercise 5.2.14. Let I be a homogeneous ideal in $R = k[x_0, \ldots, x_n]$. Prove that I is a prime ideal if and only if $fg \in I$ implies $f \in I$ or $g \in I$ for all homogeneous polynomials f, g.

5.3. Projective Varieties

In this section we will see that the V–I correspondence for affine varieties developed in Chapter 4 extends to projective varieties, but here the ideals must be homogeneous.

5.3.1. Algebraic Sets. To define varieties in $\mathbb{P}^n(k)$, we start with the zero sets of homogeneous polynomials.

Definition 5.3.1. Let S be a set of homogeneous polynomials in $k[x_0, \ldots, x_n]$. The *zero set of* S is

$$\mathrm{V}(S) = \{p \in \mathbb{P}^n(k) : f(p) = 0 \ \forall f \in S\}.$$

A subset of $\mathbb{P}^n(k)$ is called an *algebraic set* if it is the zero set of some set of homogeneous polynomials.

Exercise 5.3.1. Describe the zero sets $\mathrm{V}(S)$ in $\mathbb{P}^2$ for each set S.

(1) $S = \left\{x^2 + y^2 - z^2\right\}$
(2) $S = \left\{x^2, y\right\}$
(3) $S = \left\{x^2 + y^2 - z^2, x^2 - y^2 + z^2\right\}$

Exercise 5.3.2. Describe the algebraic sets in $\mathbb{P}^1$.

Exercise 5.3.3. Show that each set of points X is an algebraic set by finding a set of polynomials S so that $X = \mathrm{V}(S)$.

(1) $X = \{(0:1)\} \subset \mathbb{P}^1$
(2) $X = \{(0:0:1), (0:1:0), (1:0:0)\} \subset \mathbb{P}^2$
(3) $X = \{(1:1:1:1)\} \subset \mathbb{P}^3$

While in this book we are primarily concerned with varieties over $\mathbb{C}$, it is interesting to see how the algebraic sets vary with different base fields k.

Exercise 5.3.4. Let $I = \langle x^2 + y^2 \rangle \subset k[x, y]$.

(1) Find $\mathrm{V}(I)$ for $k = \mathbb{C}$.
(2) Find $\mathrm{V}(I)$ for $k = \mathbb{R}$.
(3) Find $\mathrm{V}(I)$ for $k = \mathbb{Z}_2$, the field with two elements.

Exercise 5.3.5. Let S be a set of homogeneous polynomials and let I be the ideal generated by the elements in S. Prove that $\mathrm{V}(I) = \mathrm{V}(S)$. This shows that every algebraic set is the zero set of a homogeneous ideal.

Exercise 5.3.6. Prove that every algebraic set is the zero set of a finite number of homogeneous polynomials. (The Hilbert Basis Theorem, Theorem 4.3.4, will be useful here.)

Exercise 5.3.7. We call the ideal $\langle x_0, x_1, \ldots, x_n\rangle \subset k[x_0, x_1, \ldots, x_n]$ the "irrelevant" maximal ideal of $k[x_0, x_1, \ldots, x_n]$. Prove that this is a maximal ideal and describe $\mathrm{V}(\langle x_0, x_1, \ldots, x_n\rangle)$. Why do we say that $\langle x_0, x_1, \ldots, x_n\rangle$ is irrelevant?

Exercise 5.3.8. Show if I and J are homogeneous ideals in $k[x_0, \ldots, x_n]$ with $I \subset J$, then $\mathrm{V}(I) \supset \mathrm{V}(J)$.

Exercise 5.3.9. Let I and J be homogeneous ideals in $k[x_0, x_1, \ldots, x_n]$.

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

(2) Prove $\mathrm{V}(I + J) = \mathrm{V}(I) \cap \mathrm{V}(J)$.

Exercise 5.3.10. Let I be a homogeneous ideal in $k[x_0, \ldots, x_n]$. Prove that $\mathrm{V}(\mathrm{Rad}(I)) = \mathrm{V}(I)$.

5.3.2. Ideals of Algebraic Sets. In the previous subsection we saw that algebraic sets in $\mathbb{P}^n(k)$ can always be expressed in the form $\mathrm{V}(I)$ for some homogeneous ideal I in $k[x_0, \ldots, x_n]$. In this subsection, we will complete the V–I correspondence by showing that homogeneous ideals arise from algebraic sets. We begin with a definition.

Definition 5.3.2. Let X be an algebraic set in $\mathbb{P}^n(k)$. The ideal of X is the homogeneous ideal $I(X)$ generated by the set

$$\{f \in k[x_0, \ldots, x_n] : f \text{ is homogeneous, } f(p) = 0 \text{ for all } p \in X\}.$$

Exercise 5.3.11. Let X be an algebraic set in $\mathbb{P}^n$. Prove that $I(X)$ is a homogeneous ideal.

Exercise 5.3.12. Find the ideal $I(X)$ for each algebraic set X.

(1) $X = \{(1 : 1)\}$ in $\mathbb{P}^1$

(2) $X = \mathrm{V}(\{x^2\})$ in $\mathbb{P}^2$

(3) $X = \mathrm{V}(\{x_0x_2 - x_1x_3, x_0 - x_3\})$ in $\mathbb{P}^3$

In Chapter 4 we proved Hilbert's Nullstellensatz (Theorem 4.4.1) for an affine algebraic variety $\mathrm{V}(I)$ over an algebraically closed field k, $I(\mathrm{V}(I)) = \mathrm{Rad}(I)$. To prove the projective version of this result, we will compare the projective and affine varieties corresponding to a

given homogeneous ideal. For a homogeneous ideal $J \subseteq k[x_0, \ldots, x_n]$, let

$$\mathrm{V}_a(J) = \left\{p \in \mathbb{A}^{n+1}(k) : f(p) = 0 \ \forall f \in J\right\}$$

be the *affine zero set* of the ideal J. Recall

$$I(\mathrm{V}_a(J)) = \{f \in k[x_0, \ldots, x_n] : f(p) = 0 \ \forall p \in \mathrm{V}_a(J)\}$$

is the ideal of polynomials vanishing on the affine variety $\mathrm{V}_a(J)$. Note that we do not require that the polynomials in $I(\mathrm{V}_a(J))$ be homogeneous.

Exercise 5.3.13. Let $J = \langle x - y\rangle \subseteq k[x, y]$.

(1) Find the affine zero set $\mathrm{V}_a(J) \subset \mathbb{A}^2(k)$.

(2) Find $I(\mathrm{V}_a(J))$ and show that this ideal is homogeneous.

(3) Show that $I(\mathrm{V}(J)) = \mathrm{Rad}(J)$.

Exercise 5.3.14. Let $J = \langle x - y, y + z\rangle \subseteq k[x, y, z]$.

(1) Find the affine zero set $\mathrm{V}_a(J) \subset \mathbb{A}^3(k)$.

(2) Find $I(\mathrm{V}_a(J))$ and show that this ideal is homogeneous.

(3) Show that $I(\mathrm{V}(J)) = \mathrm{Rad}(J)$.

Exercise 5.3.15. Let $J = \langle xy, yz, xz\rangle \subseteq k[x, y, z]$.

(1) Find the affine zero set $\mathrm{V}_a(J) \subset \mathbb{A}^3(k)$.

(2) Find $I(\mathrm{V}_a(J))$ and show that this ideal is homogeneous.

(3) Show that $I(\mathrm{V}(J)) = \mathrm{Rad}(J)$.

Exercise 5.3.16. Let J be a homogeneous ideal in $k[x_0, \ldots, x_n]$.

(1) Prove that $(a_0, \ldots, a_n) \in \mathrm{V}_a(J)$ if and only if $(\lambda a_0, \ldots, \lambda a_n) \in \mathrm{V}_a(J)$ for all $\lambda \neq 0$ in k.

(2) Let $I(\mathrm{V}_a(J)) = \{f \in k[x_0, \ldots, x_n] : f(p) = 0 \ \forall p \in \mathrm{V}_a(J)\}$ be the ideal of polynomials vanishing on the affine variety $\mathrm{V}_a(J)$. Prove that $I(\mathrm{V}_a(J))$ is in fact homogeneous and $I(\mathrm{V}_a(J)) = I(\mathrm{V}(J))$.

(3) Use Hilbert's Nullstellensatz to conclude that $I(\mathrm{V}(J)) = \mathrm{Rad}(J)$.

Exercise 5.3.17. Let J be a homogeneous ideal. Prove that $\mathrm{V}(J) = \emptyset$ if and only if $\langle x_0, x_1, \ldots, x_n\rangle \subseteq \mathrm{Rad}(J)$.

Exercise 5.3.18. Let X be an algebraic set in $\mathbb{P}^n(k)$. Show that $\mathrm{V}(I(X)) = X$.

Exercise 5.3.19. Let X and Y be algebraic sets in $\mathbb{P}^n(k)$. Show that $X \subset Y$ if and only if $I(Y) \subset I(X)$.

5.3.3. Irreducible Algebraic Sets and Projective Varieties. As in Chapter 4, we say that an algebraic set V is *reducible* if $V = V_1 \cup V_2$, where V_1 and V_2 are distinct algebraic sets with $V_1 \subsetneq V$ and $V_2 \subsetneq V$. An algebraic set that is not reducible is said to be *irreducible*.

Definition 5.3.3. A *projective variety* is defined to be an irreducible algebraic subset of $\mathbb{P}^n$ for some n.

Exercise 5.3.20. Determine whether each algebraic set in $\mathbb{P}^n$ is irreducible (and thus a projective variety).

(1) $\mathrm{V}(\langle x_0 \rangle)$

(2) $\mathrm{V}(\langle x_0 x_1 \rangle)$

(3) $\mathrm{V}(\langle x_1, x_2, \ldots, x_n \rangle)$

Exercise 5.3.21. Let $V \subset \mathbb{P}^n$ be an algebraic set.

(1) Suppose that V is reducible, say $V = V_1 \cup V_2$ where V_1 and V_2 are distinct algebraic sets with $V_1 \subsetneq V$ and $V_2 \subsetneq V$. Show that there are polynomials $f_1 \in I(V_1)$ and $f_2 \in I(V_2)$ such that $f_1 f_2 \in I(V)$ but $f_1, f_2 \notin I(V)$. Conclude that $I(V)$ is not a prime ideal.

(2) Prove that if $I(V)$ is a homogeneous ideal in $k[x_0, x_1, \ldots, x_n]$ which is not prime, then V is a reducible algebraic set in $\mathbb{P}^n$.

Therefore, an algebraic set V in $\mathbb{P}^n$ is a projective variety if and only if its ideal $I(V)$ is a homogeneous prime ideal in $k[x_0, x_1, \ldots, x_n]$. This establishes a correspondence between the projective varieties in $\mathbb{P}^n$ and the homogeneous prime ideals in the graded ring $k[x_0, x_1, \ldots, x_n]$ other than the ideal $J = \langle x_0, x_1, \ldots, x_n \rangle$. (Recall that J is called the irrelevant ideal, since $\mathrm{V}(J) = \emptyset$.)

Exercise 5.3.22. Show that $\mathbb{P}^n(k)$ is a projective variety when k is infinite.

Exercise 5.3.23. Determine whether each algebraic set V is a projective variety in $\mathbb{P}^2$ by determining whether $I(V)$ is prime.

(1) $\mathrm{V}(\langle xy \rangle)$
(2) $\mathrm{V}(\langle xy - z^2 \rangle)$
(3) $\mathrm{V}(\langle x^2 \rangle)$

Exercise 5.3.24. Suppose $V = V_1 \cup V_2$ is a reducible algebraic set in $\mathbb{P}^n$. Show that $I(V) = I(V_1) \cap I(V_2)$.

Exercise 5.3.25. Suppose V is an algebraic set in $\mathbb{P}^n$. Show that V is the union of a finite number of projective varieties.

5.3.4. The Zariski Topology. Analogous with affine varieties, the collection of algebraic sets will be the closed sets for the Zariski topology on $\mathbb{P}^n$, which we now show.

Exercise 5.3.26.

(1) Show that $\emptyset$ and $\mathbb{P}^n$ are algebraic sets in $\mathbb{P}^n$.
(2) Show that the union of a finite number of algebraic sets in $\mathbb{P}^n$ is again an algebraic set.
(3) Show that the intersection of an arbitrary collection of algebraic sets in $\mathbb{P}^n$ is again an algebraic set.
(4) Conclude that the algebraic sets in $\mathbb{P}^n$ form the collection of closed sets for a topology on $\mathbb{P}^n$: the *Zariski topology*.

Exercise 5.3.27. For the Zariski topology on $\mathbb{P}^1$:

(1) Show that $\{(0:1), (1:0)\}$ is a closed set.
(2) Find an open neighborhood of $\{(1:1)\}$.
(3) Describe the closed sets in $\mathbb{P}^1$.
(4) Find a basis of open sets for $\mathbb{P}^1$.

Exercise 5.3.28. For the Zariski topology on $\mathbb{P}^n$:

(1) Show that the sets $\mathbb{P}^n - \mathrm{V}(f)$, for homogeneous $f \in k[x_0, \ldots, x_n]$, form a basis for the Zariski topology on $\mathbb{P}^n$.
(2) Show that this topology is not Hausdorff. (Recall that a topological space is *Hausdorff* if for every pair of distinct points there exist disjoint open sets containing them.)

5.3.5. Parametrizing Conics via Projective Varieties. In Chapter 1 we studied conics in the plane and classified them up to equivalence in $\mathbb{R}^2$, $\mathbb{C}^2$ and $\mathbb{P}^2$. While there are infinitely many conics in $\mathbb{P}^2$, we showed that there were only three classes up to a projective change of coordinates: smooth, crossing lines, and double line conics. In this subsection, however, we will not exploit projective changes of coordinates but instead find a space whose points corresponds to conics in $\mathbb{P}^2$. That is, we seek a space such that there is a bijection between the set of its points and the set of conics in $\mathbb{P}^2$. Moreover, we hope the geometry of this space will provide insights regarding the family of conics in $\mathbb{P}^2$.

We begin by recalling that a conic in $\mathbb{P}^2$ is the zero set of a homogeneous polynomial of degree two, $\mathrm{V}(Ax^2+Bxy+Cy^2+Dxz+Eyz+Fz^2)$, where not all coefficients can be zero. Observe that $\mathrm{V}(Ax^2+Bxy+Cy^2+Dxz+Eyz+Fz^2) = \mathrm{V}(\lambda Ax^2+\lambda Bxy+\lambda Cy^2+\lambda Dxz+\lambda Eyz+\lambda Fz^2)$ for any $\lambda \neq 0$ in k. Thus, a conic in $\mathbb{P}^2$ is determined by its homogeneous polynomial up to scalar multiple, and this polynomial is determined by its coefficients. Therefore, we may identify the conic

$$\mathrm{V}(Ax^2+Bxy+Cy^2+Dxz+Eyz+Fz^2) \subset \mathbb{P}^2$$

with the point

$$(A : B : C : D : E : F) \in \mathbb{P}^5.$$

Conversely, for every point $(a_0 : a_1 : a_2 : a_3 : a_4 : a_5) \in \mathbb{P}^5$, we have the corresponding conic $\mathrm{V}(a_0x^2+a_1xy+a_2y^2+a_3xz+a_4yz+a_5z^2)$ in $\mathbb{P}^2$ since not all coordinates of the point in $\mathbb{P}^5$ can be zero. Therefore, the projective space $\mathbb{P}^5$ may be viewed as a *parameter space* for the family of all conics in $\mathbb{P}^2$.

Exercise 5.3.29. Show that the set of singular conics in $\mathbb{P}^2$ corresponds to an algebraic set in $\mathbb{P}^5$.

Thus the smooth conics form a Zariski open set in the space of all conics in $\mathbb{P}^2$. This implies that a "generic" conic in $\mathbb{P}^2$ will be smooth, and we begin to gain geometric insight from our parameter space.

Exercise 5.3.30. Show that the set of all conics in $\mathbb{P}^2$ that pass through a given point p corresponds to an algebraic set in $\mathbb{P}^5$.

Exercise 5.3.31. Fix a point p and a line ℓ through p in $\mathbb{P}^2$. Show that the set of all conics in $\mathbb{P}^2$ that are either tangent to ℓ or singular at p corresponds to an algebraic set in $\mathbb{P}^5$.

This is only a brief introduction to parameter spaces. If we change our focus from conics in $\mathbb{P}^2$ to lines in $\mathbb{P}^3$ or curves in $\mathbb{P}^n$, rather than $\mathbb{P}^5$ and some of its projective subvarieties, we would be led to Grassmann varieties and more general moduli spaces. Finding and studying such spaces is on the cusp of current research in algebraic geometry and its interaction with modern physics via string theory. In this book, we must be content with this limited exposure to such problems and now turn our attention to the study of functions on projective varieties.

5.4. Functions, Tangent Spaces, and Dimension

In this section we will study functions on projective varieties and use them to define the tangent space at a point and the dimension of a variety.

5.4.1. The Function Field of a Projective Variety. We now define a field of functions on a projective variety, as we did for curves in Section 3.4 and affine varieties in Section 4.10. Let $V \subseteq \mathbb{P}^n$ be a projective variety. We have seen that polynomial functions are not well-defined on projective space in Section 5.2, so instead we consider ratios of homogeneous polynomials of the same degree. These ratios will determine functions on $\mathbb{P}^n$.

Exercise 5.4.1. Let f and g be homogeneous polynomials of the same degree. Show that

$$\frac{f(\lambda x_0, \ldots, \lambda x_n)}{g(\lambda x_0, \ldots, \lambda x_n)} = \frac{f(x_0, \ldots, x_n)}{g(x_0, \ldots, x_n)}.$$

Thus $\dfrac{f}{g}$ is a well-defined function at all points $p \in \mathbb{P}^n$ with $g(p) \neq 0$.

Definition 5.4.1. Let $V \subset \mathbb{P}^n$ be a projective variety with ideal $I(V)$. The *function field* of V, $\mathcal{K}_V$, is the set of all ratios

$$\frac{f(x_0, \ldots, x_n)}{g(x_0, \ldots, x_n)}$$

modulo the relation $\sim$, where

(1) f and g are homogeneous polynomials of the same degree,

(2) $g \notin I(V)$ (which is a way of guaranteeing that g, the denominator, is not identically zero on the variety V), and

(3) $\dfrac{f_1}{g_1} \sim \dfrac{f_2}{g_2}$ if $f_1 g_2 - f_2 g_1 \in I(V)$.

Compare this with our definitions of function fields for plane curves (Definition 3.4.5) and affine varieties (Definition 4.10.1). The proofs of the next two exercises are similar to these earlier cases.

Exercise 5.4.2. Prove that $\sim$ is an equivalence relation.

Exercise 5.4.3. Prove that $\mathcal{K}_V$ is a field.

Exercise 5.4.4. Show that $\dfrac{f_1}{g_1} \sim \dfrac{f_2}{g_2}$ if and only if $\dfrac{f_1}{g_1}$ and $\dfrac{f_2}{g_2}$ are identical as functions on their common domain in V.

Thus elements of $\mathcal{K}_V$ may be viewed as functions on the projective variety V, but it is often the case that they are not defined on all of V but only on an open subset. We say that two rational functions are equal when they are identical on some open subset of the variety.

Exercise 5.4.5. Let $V = \mathrm{V}(\langle x^2 - yz \rangle)$ in $\mathbb{P}^2$.

(1) Show that $\dfrac{x}{z}$ is defined on an open subset U of V, and thus $\dfrac{x}{z}$ defines a function from U to the base field k.

(2) Show that $\dfrac{x}{z} = \dfrac{y}{x}$ on V and find an open subset of V where they agree.

Exercise 5.4.6. Let $V = \mathrm{V}(\langle x_0 x_2 - x_1^2, x_1 x_3 - x_2^2, x_0 x_3 - x_1 x_2 \rangle)$ in $\mathbb{P}^3$.

(1) Show that $\dfrac{x_0}{x_2}$ is defined on an open subset U of V, and thus defines a function from U to the base field k.

(2) Show that $\dfrac{x_0}{x_2} = \dfrac{x_1}{x_3}$ in $\mathcal{K}_V$.

Exercise 5.4.7. Let V be a projective variety in $\mathbb{P}^n$ and let $h = \dfrac{f}{g}$, where f and g are homogeneous polynomials of the same degree. Show that h is defined on an open subset U of V and thus defines a function from U to the base field k.

As in Chapter 4, we are interested in the functions on a variety because these functions give us an algebraic tool for studying the geometry of the space. In addition to its field of functions, an affine variety has a coordinate ring and, at each of its points, a local ring that provides information about the geometry of the variety near the point. In Section 4.19 we proved two affine varieties are isomorphic if and only if they have isomorphic coordinate rings. While we do not have the same property for projective varieties and their coordinate rings, the study of their local rings is still valuable (compare to the affine case in Section 4.13). This leads to the next definition.

Definition 5.4.2. Let V be a projective variety and let p be a point in V. The *local ring* of V at p is

$$\mathcal{O}_p(V) = \left\{ \frac{f}{g} \in \mathcal{K}_V : g(p) \neq 0 \right\}.$$

We call elements of $\mathcal{O}_p(V)$ *regular functions* at p.

Exercise 5.4.8. Let $V = \mathrm{V}(\langle x^2 - yz \rangle)$ and let $p = (0 : 1 : 0) \in V$. Show that $h = \dfrac{z}{x}$ is in $\mathcal{O}_p(V)$ by finding homogeneous polynomials f and g with $g(p) \neq 0$ and $h = \dfrac{f}{g}$ on an open set containing p.

Exercise 5.4.9. Verify that $\mathcal{O}_p(V)$ is a local ring with unique maximal ideal

$$\mathfrak{m}_p = \{h \in \mathcal{O}_p(V) : h(p) = 0\}.$$

Compare with Exercise 4.13.7.

The similarities to the affine case that we have seen thus far are not merely coincidental. The affine charts $\{U_i\}$ of $\mathbb{P}^n$ give a covering of a variety $V \subseteq \mathbb{P}^n$ by open sets as we proved in Section 5.1. That is, if $V \subset \mathbb{P}^n$ is an algebraic set, then each $V_i = U_i \cap V$ is open in V and

$V = \bigcup V_i$. Furthermore, since V is closed in $\mathbb{P}^n$, V_i is closed in U_i. Hence each V_i will be an algebraic set in the affine space $U_i \cong \mathbb{A}^n$, and we can generally reduce local problems on V to problems on one of the V_i, which means that we can use our tools from Chapter 4.

Exercise 5.4.10. Let V be an algebraic set in $\mathbb{P}^n$.

(1) Show that $\mathcal{K}_V \cong \mathcal{K}_{V_i}$ for each $i = 0, 1, \ldots, n$. [Hint: Consult Section 3.4.]

(2) Suppose that $p \in V$ belongs to V_j. Show that $\mathcal{O}_p(V) \cong \mathcal{O}_p(V_j)$.

This generalizes our result from Section 3.4, where we showed the equivalence of the function fields of a projective curve and an affine piece of that curve.

5.4.2. Tangent Spaces and Dimension. As in Chapter 4, we will define the tangent space to a projective variety V at a point $p \in V$ and use these tangent spaces to define the dimension of V. This again demonstrates the geometric insight provided by studying functions on a space.

Recall that a *derivation* is a map $L : R \to S$ from a k-algebra R to a k-algebra S with the following properties:

(i) L is k-linear, i.e., $L(af+bg) = aL(f)+bL(g)$ for all $a, b \in k$ and $f, g \in R$,

(ii) L obeys the Leibniz rule, $L(fg) = gL(f) + fL(g)$ for all $f, g \in R$.

Let V be a projective variety and let $p \in V$ be one of its points. Then the local ring $\mathcal{O}_p(V) = \left\{ \dfrac{f}{g} \in \mathcal{K}_V : g(p) \neq 0 \right\}$ is a k-algebra that captures the behavior of functions on V near p. As in Section 4.14, we define the tangent space of V at p in terms of derivations on the local ring $\mathcal{O}_p(V)$.

Definition 5.4.3. The *tangent space* of the projective variety V at a point p is the vector space

$$T_pV = \{L : \mathcal{O}_p(V) \to k \ : \ L \text{ is a derivation}\}.$$

By Exercise 5.4.10, if $V \subset \mathbb{P}^n$ is a projective variety then $\mathcal{O}_p(V) \cong \mathcal{O}_p(V_i)$ for any affine piece $V_i = U_i \cap V$ of V containing p, where the U_i are the affine charts of $\mathbb{P}^n$. Therefore, the tangent space T_pV is equal to $T_pV_i = \{\text{derivations } L : \mathcal{O}_p(V_i) \to k\}$, the tangent space to V_i at p. Moreover, since V is closed in $\mathbb{P}^n$, $V_i = V \cap U_i$ is closed in $U_i \cong \mathbb{A}^n$; thus it is an affine variety. Hence the computation of tangent spaces for projective varieties can always be done by dehomogenizing to an affine chart and using our various methods from Section 4.14. We do this explicitly in some cases.

Exercise 5.4.11. In $\mathbb{P}^2(\mathbb{C})$, let C be the curve given by the homogeneous polynomial $F(x_0, x_1, x_2) = x_1x_2 - x_0^2 = 0$.

(1) Verify that $p = (2 : 4 : 1)$ is on C.

(2) Let $C_0 = C \cap U_0$ and compute the tangent line T_pC_0.

(3) Let $C_2 = C \cap U_2$ and compute the tangent line T_pC_2.

You may think that T_pC_0 and T_pC_2 are not the same. This is because we dehomogenized to different affine charts. However, we will see that they are the same when we rehomogenize to projective lines in $\mathbb{P}^2$.

Exercise 5.4.12. Consider the curve $C = \mathrm{V}(x_1x_2 - x_0^2) \subset \mathbb{P}^2(\mathbb{C})$ and the point $p = (2 : 4 : 1) \in C$ from the previous exercise.

(1) Homogenize the equations for T_pC_0 and T_pC_2 and show that they define the same projective line in $\mathbb{P}^2(\mathbb{C})$.

(2) Show that the projective line from Part (1) is given by

$$x_0\frac{\partial F}{\partial x_0}(p) + x_1\frac{\partial F}{\partial x_1}(p) + x_2\frac{\partial F}{\partial x_2}(p) = 0.$$

Exercise 5.4.13. Consider the surface $V = \mathrm{V}(x_0x_1 - x_2x_3)$ in $\mathbb{P}^3$.

(1) Verify that $p = (1 : 1 : 1 : 1)$ is on V.

(2) Dehomogenize to the affine chart U_0 to compute the tangent plane T_pV_0.

(3) Homogenize T_pV_0 to a projective plane in $\mathbb{P}^3$ and compare this to the zero set of

$$x_0\frac{\partial F}{\partial x_0}(p) + x_1\frac{\partial F}{\partial x_1}(p) + x_2\frac{\partial F}{\partial x_2}(p) + x_3\frac{\partial F}{\partial x_3}(p) = 0.$$

Exercise 5.4.14. Consider the algebraic surface $V = \mathrm{V}(x_0^3 - x_1x_2x_3)$ in $\mathbb{P}^3$.

(1) Verify that $p = (1 : 1 : 1 : 1)$ is on V.

(2) Dehomogenize to the affine chart U_0 to compute the tangent plane T_pV_0.

(3) Dehomogenize to the affine chart U_3 to compute the tangent plane T_pV_3.

(4) Homogenize both T_pV_0 and T_pV_3 to show they give the same projective plane in $\mathbb{P}^3$. Compare this to the zero set of

$$x_0\frac{\partial F}{\partial x_0}(p) + x_1\frac{\partial F}{\partial x_1}(p) + x_2\frac{\partial F}{\partial x_2}(p) + x_3\frac{\partial F}{\partial x_3}(p) = 0.$$

Since T_pV is the same as the tangent space T_pV_i, where $V_i = V \cap U_i$ is an affine variety containing p, it follows that T_pV is a vector space over k from Exercise 4.14.6. Using this result, we can define the dimension of V as follows.

Definition 5.4.4. Let $V \subseteq \mathbb{P}^n$ be a projective variety. Then the *dimension* of V is the minimum dimension of T_pV over all points $p \in V$, where the dimension of T_pV refers to its dimension as a vector space.

Exercise 5.4.15. Show that the dimension of $\mathbb{P}^n$ is n.

5.5. Rational and Birational Maps

In this section, we explore rational maps between projective varieties.

5.5.1. Rational Maps. Elements of the function field $\mathcal{K}_V$ are defined on open subsets of the variety and define functions from these open subsets to the base field k. We now extend this idea to create maps from V to projective space.

Definition 5.5.1. Let V be a projective variety. A *rational map* from V to $\mathbb{P}^m$ is a function $h : V \dashrightarrow \mathbb{P}^m$ given by

$$h(p) = (h_0(p) : h_1(p) : \cdots : h_m(p)),$$

where $h_0, h_1, \ldots, h_m \in \mathcal{K}_V$ and at least one of the $h_i(p)$ is nonzero.

As in the case of rational functions in $\mathcal{K}_V$, a rational map may be defined only on an open subset of V. Specifically, the domain of $h = (h_0 : h_1 : \cdots : h_m)$ consists of only those points $p \in V$ where each h_i is defined and at least one of the $h_i(p)$ is nonzero.

Exercise 5.5.1. Prove that the above definition gives a well-defined function from an open subset of V to $\mathbb{P}^m$.

Exercise 5.5.2. Let $h : \mathbb{P}^1 \dashrightarrow \mathbb{P}^2$ be defined by

$$h((x : y)) = \left(\frac{x^2}{y^2} : \frac{x}{y} : 1\right).$$

(1) Determine the domain U of h.

(2) Show that the function $a((x : y)) = (x^2 : xy : y^2)$ agrees with h on U and is defined on all of $\mathbb{P}^1$.

Exercise 5.5.3. Let $V = \mathrm{V}(\langle x_0^2 + x_1^2 - x_2^2 \rangle)$ in $\mathbb{P}^2$, and let $h_0 = \dfrac{x_0}{x_2}$, $h_1 = \dfrac{x_1}{x_2}$.

(1) Determine the domain U of the rational map $h : V \dashrightarrow \mathbb{P}^1$ defined by $h(p) = (h_0(p) : h_1(p))$.

(2) Show that the function $(x_0 : x_1 : x_2) \mapsto (x_0 : x_1)$ agrees with h on U.

We can generalize the idea of the previous two exercises to see that rational maps may be defined using homogeneous polynomials rather than ratios.

Exercise 5.5.4. Let h be a rational map $h : V \dashrightarrow \mathbb{P}^m$, so h is defined as

$$h(p) = (h_0(p) : h_1(p) : \ldots : h_m(p)),$$

where $h_i = \dfrac{f_i}{g_i}$ with f_i, g_i both homogeneous polynomials of degree d_i, for $0 \leq i \leq m$.

(1) Show that

$$h(p) = (g(p)h_0(p) : g(p)h_1(p) : \cdots : g(p)h_m(p))$$

for any homogeneous polynomial g with $g(p) \neq 0$.

(2) Prove that any rational map $h : V \dashrightarrow \mathbb{P}^m$ can be written as
$$h(p) = (a_0(p) : a_1(p) : \cdots : a_m(p)),$$
where $a_0, a_1, \ldots, a_m$ are homogeneous polynomials of the same degree.

In some cases a rational map h will be defined on all of V rather than an open subset.

Definition 5.5.2. A rational map $h : V \dashrightarrow \mathbb{P}^m$ is called *regular at a point* $p \in V$ if there is an open neighborhood U of p on which h can be represented by rational functions $\frac{f_0}{g_0}, \frac{f_1}{g_1}, \ldots, \frac{f_m}{g_m}$ such that $g_i(q) \neq 0$ for each i and all $q \in U$ and at least one $f_i(p) \neq 0$. A rational map that is regular at all points $p \in V$ is called a *morphism*. When this is the case, we write $h : V \to \mathbb{P}^m$ with a solid arrow.

Exercise 5.5.5. Let $h : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ be defined by $(x_0x_1 : x_0x_2 : x_1x_2)$.

(1) Find all points p where h is regular.

(2) Describe the pre-images of the points $(0 : 0 : 1), (0 : 1 : 0)$, and $(1 : 0 : 0)$.

Exercise 5.5.6. Let $V = \mathrm{V}(\langle x_0x_3 - x_1x_2 \rangle) \subset \mathbb{P}^3$ and let $h : V \dashrightarrow \mathbb{P}^1$ be defined by $h((x_0 : x_1 : x_2 : x_3)) = (x_0 : x_2)$. Prove that h is a morphism and that its image is all of $\mathbb{P}^1$.

So far we have considered functions from a variety to projective space, but we are often interested in functions to another projective variety. We write
$$h : V \dashrightarrow W$$
when h is a rational map whose image lies in the projective variety W. Similarly, we write $h : V \to W$ when h is a morphism whose image lies in W.

Exercise 5.5.7. Prove that the rational map $h : \mathbb{P}^1 \dashrightarrow \mathbb{P}^2$ defined by
$$h\left((a_0 : a_1)\right) = (a_0 : a_1 : a_1)$$
is a morphism whose image lies in the line $\mathrm{V}(x_1 - x_2)$ in $\mathbb{P}^2$.

Thus, in the exercise above, $h : \mathbb{P}^1 \to \mathrm{V}(x_1 - x_2)$ is a morphism from the projective line to the projective variety $\mathrm{V}(x_1 - x_2) \subset \mathbb{P}^2$.

As in the affine case, the product of two projective varieties is a variety. In the next two exercises we will consider morphisms from products of projective spaces into a larger projective space.

Exercise 5.5.8. Define the *Segre embedding* of the product $\mathbb{P}^1 \times \mathbb{P}^1$ to be

$$\psi : \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^3$$

given by $\psi((a_0 : a_1), (b_0 : b_1)) = (a_0b_0 : a_0b_1 : a_1b_0 : a_1b_1)$.

(1) Show that ψ is well-defined.

(2) Let Y be the image of ψ in $\mathbb{P}^3$. Show that Y is an algebraic set.

Exercise 5.5.9. We now consider the product of the projective spaces $\mathbb{P}^k$ and $\mathbb{P}^\ell$. Define the *Segre embedding* $\psi : \mathbb{P}^k \times \mathbb{P}^\ell \to \mathbb{P}^{(k+1)(\ell+1)-1}$ by

$$\begin{aligned}\psi((a_0 : \cdots : a_k), (b_0 : \cdots : b_\ell)) \\ = (a_0b_0 : a_0b_1 : \cdots : a_0b_\ell : a_1b_0 : a_1b_1 : \cdots : a_kb_\ell).\end{aligned}$$

(1) Show that ψ is well-defined from $\mathbb{P}^k \times \mathbb{P}^\ell$ to $\mathbb{P}^{(k+1)(\ell+1)-1}$.

(2) Let Y be the image of ψ in $\mathbb{P}^{(k+1)(\ell+1)-1}$. Show that Y is an algebraic set.

5.5.2. Birational Maps. We next consider rational maps that have a rational inverse. These maps may be defined only on open subsets, but we shall see that this is sufficient to provide an important notion of equivalence between varieties.

Definition 5.5.3. Let $\varphi : V \dashrightarrow W$ be a rational map between projective varieties V and W such that there is a rational map $\psi : W \dashrightarrow V$ with the property $\psi \circ \varphi(p) = p$ for all points p in a non-empty open subset of V and $\varphi \circ \psi(q) = q$ for all points q in a non-empty open subset of W. We say that φ is a *birational map* with rational inverse ψ, and the varieties V and W are *birational*.

In Chapter 4, we worried about rational maps that were dominant. Because we are explicitly working with non-empty open subsets, these maps are dominant.

Exercise 5.5.10. Let $V = \mathrm{V}(\langle x_0 \rangle) \subset \mathbb{P}^2$ and let $\varphi : V \dashrightarrow \mathbb{P}^1$ be defined by

$$\varphi((x_0 : x_1 : x_2)) = (x_1 : x_2).$$

Prove that φ is a birational map.

Exercise 5.5.11. Let $V = \mathrm{V}(\langle x_0 + x_1 + x_2 + x_3 \rangle) \subset \mathbb{P}^3$. Show that V and $\mathbb{P}^2$ are birational.

Exercise 5.5.12. Define a rational map $\varphi : \mathbb{P}^1 \to \mathbb{P}^2$ by

$$\varphi((x_0 : x_1)) = (x_0^2 : x_0 x_1 : x_1^2).$$

(1) Show that the image of φ is a plane conic.

(2) Find the rational inverse of φ.

Exercise 5.5.13. Define a rational map $\varphi : \mathbb{P}^1 \to \mathbb{P}^3$ by

$$\varphi((x_0 : x_1)) = (x_0^3 : x_0^2 x_1 : x_0 x_1^2 : x_1^3).$$

(1) Find the image V of φ. (This image is called a twisted cubic curve.)

(2) Find the rational inverse from V to $\mathbb{P}^1$.

We now generalize the previous two exercises to construct morphisms from $\mathbb{P}^1$ to various projective spaces. The next two exercises follow Hartshorne [**Har77**], Exercise I.2.12.

Exercise 5.5.14. The d-uple embedding of $\mathbb{P}^1$:

(1) Fix a degree $d > 0$. How many monomials in the variables x_0 and x_1 of degree d exist? Call this number N and list the monomials in some order, $m_1, \ldots, m_N$.

(2) Show that $(x_0 : x_1) \mapsto (m_1 : \cdots : m_N)$ is a well-defined morphism from $\mathbb{P}^1$ to $\mathbb{P}^{N-1}$. This is called the *d-uple embedding* of $\mathbb{P}^1$.

(3) Let Y be the image of the 4-uple embedding of $\mathbb{P}^1$. Show that Y is an algebraic set.

Exercise 5.5.15. We generalize further to d-uple embeddings from $\mathbb{P}^n$.

(1) Fix a degree $d > 0$. How many monomials in the variables $x_0, x_1, \ldots, x_n$ of degree d exist? Call this number N and list the monomials in some order, $m_1, \ldots, m_N$.

(2) Show that $(x_0 : x_1 : \cdots : x_n) \mapsto (m_1 : \cdots : m_N)$ is a well-defined morphism from $\mathbb{P}^n$ to $\mathbb{P}^{N-1}$. This is called the *d-uple embedding* of $\mathbb{P}^n$.

(3) Let Y be the image of the 2-uple embedding of $\mathbb{P}^2$ in $\mathbb{P}^5$. This is called the *Veronese surface*. Show that Y is an algebraic set in $\mathbb{P}^5$.

5.5.3. Function Fields of Birational Varieties. We next state the correspondence between birational varieties and isomorphic function fields. As before, the projective case is similar to the affine case. The goal of this subsection is to prove:

Theorem 5.5.16. Projective varieties V and W are birational if and only if their function fields $\mathcal{K}_V$ and $\mathcal{K}_W$ are isomorphic.

Exercise 5.5.17. Let V and W be projective varieties and $\varphi : V \dashrightarrow W$ a rational map. Prove that φ induces a homomorphism between the function fields, $\varphi^* : \mathcal{K}_W \to \mathcal{K}_V$, defined by $\varphi^*(h) = h \circ \varphi$.

Exercise 5.5.18. Let $V \subset \mathbb{P}^n$ and $W \subset \mathbb{P}^m$ be projective varieties and let $\alpha : \mathcal{K}_W \to \mathcal{K}_V$ be a field homomorphism. Show that there exists a unique rational map $\varphi : V \dashrightarrow W$ whose image is dense in W such that $\varphi^* = \alpha$.

Exercise 5.5.19. Prove Theorem 5.5.16.

We now look at some examples of this correspondence.

Exercise 5.5.20. Let $\varphi : \mathbb{P}^1 \to C$ be the rational map $\varphi((x_0 : x_1)) = (x_0^2 : x_0x_1 : x_1^2)$ from Exercise 5.5.12, where C is the plane conic $C = \mathrm{V}(xz - y^2)$.

(1) Explicitly write out the map $\varphi^* : \mathcal{K}_C \to \mathcal{K}_{\mathbb{P}^1}$.

(2) Let ψ be the rational inverse to φ. Explicitly write out the map $\psi^* : \mathcal{K}_{\mathbb{P}^1} \to \mathcal{K}_C$.

(3) Verify that $\varphi^* \circ \psi^*$ is the identity on $\mathcal{K}_{\mathbb{P}^1}$ and $\psi^* \circ \varphi^*$ is the identity on $\mathcal{K}_C$.

Exercise 5.5.21. Let $V = \mathrm{V}(y_0y_3 - y_1y_2, y_0y_2 - y_1^2, y_2^2 - y_1y_3) \subset \mathbb{P}^3$ be the image of the birational map $\varphi : \mathbb{P}^1 \to V$, $\varphi((x_0 : x_1)) = (x_0^3 : x_0^2x_1 : x_0x_1^2 : x_1^3)$, the twisted cubic curve from Exercise 5.5.13. Find the map on the function fields $\mathcal{K}_V \to \mathcal{K}_{\mathbb{P}^1}$ induced by φ.

5.6. Proj(R)

Given a commutative graded ring R, we define $\mathrm{Proj}(R)$, the projective analogue of $\mathrm{Spec}(R)$ from Chapter 4.

We next define the projective counterpart of the prime spectrum $\mathrm{Spec}(R)$. The *Proj* construction is an important initial step in the study of projective schemes associated to graded rings. We will state only the definition and look at several examples of how this construction relates to projective varieties.

Let R be a graded ring, which for our purposes will be mainly $k[x_0, \ldots, x_n]$ or a quotient of this polynomial ring. As before with projective varieties, we are interested in *homogeneous* ideals apart from the irrelevant ideal. (Recall that the irrelevant ideal of $k[x_0, \ldots, x_n]$ is $\langle x_0, x_1, \ldots, x_n \rangle$; for a general graded ring R we call the ideal generated by all elements of positive degree *irrelevant*.)

Define $\mathrm{Proj}(R)$ to be the set of all homogeneous prime ideals in R that do not contain the irrelevant ideal. This plays the role for projective varieties that Spec plays for affine varieties, providing a dictionary between graded rings with their homogeneous ideals and projective varieties with their algebraic sets.

The set $\mathrm{Proj}(R)$ is given the Zariski topology as follows. For any homogeneous ideal H in R, define

$$Z(H) = \{I \in \mathrm{Proj}(R) : H \subseteq I\},$$

the set of homogeneous prime ideals containing H (again excluding the irrelevant ideal). As in the construction of the Zariski topology on $\mathrm{Spec}(R)$, we say that the sets $Z(H)$ are *closed* in $\mathrm{Proj}(R)$. Recall that open sets are defined to be complements of closed sets, thus of the form $\mathrm{Proj}(R) - Z(H)$ for some homogeneous ideal H. In the next exercise we show that this defines a topology on $\mathrm{Proj}(R)$.

Exercise 5.6.1.

(1) Show that the empty set and Proj(R) are open.

(2) Prove that the arbitrary union of open sets of Proj(R) is also open.

(3) Prove that the intersection of a finite number of open sets is also open.

Exercise 5.6.2. Let $R = \mathbb{C}[x]$. Show that Proj(R) is a point.

Exercise 5.6.3. In this exercise we show how to obtain the projective line $\mathbb{P}^1$ as Proj(R) for the ring $R = \mathbb{C}[x_0, x_1]$.

(1) Let I be a homogeneous prime ideal in R such that I does not contain the irrelevant ideal $\langle x_0, x_1 \rangle$. Prove that either $I = \{0\}$ or I is generated by one linear polynomial.

(2) Show how the ideal $\langle x_0 \rangle$ corresponds to the point $(0 : 1) \in \mathbb{P}^1$. Prove that this ideal is maximal among those in Proj(R).

(3) Find the prime ideal I that corresponds to the point $(1 : 2)$ and prove that the set $\{I\}$ is closed in Proj(R).

(4) Find the prime ideal I that corresponds to the point $(a : b)$ and prove that the set $\{I\}$ is closed in Proj(R).

(5) Prove that every closed point of Proj(R) is a prime ideal in R that is maximal among those in Proj(R).

(6) Show that closed points of Proj(R) correspond to points in $\mathbb{P}^1$.

Exercise 5.6.4. In this exercise we show how to obtain the projective plane $\mathbb{P}^2$ as Proj(R) for the ring $R = \mathbb{C}[x_0, x_1, x_2]$.

(1) Show that the ideal $I = \langle x_0, x_1 \rangle$ corresponds to the point $(0 : 0 : 1) \in \mathbb{P}^2$. Prove that this ideal is maximal among those in Proj(R), so that $Z(I) = \{I\}$.

(2) Show that $Z(I) \neq \{I\}$ for the ideal $I = \langle x_0^2 + x_1^2 + x_2^2 \rangle$, by finding a point $P \in Z(I)$ with $P \neq I$.

(3) Find the prime ideal I that corresponds to the point $(1:2:3)$ and prove that the set $\{I\}$ is closed in $\mathrm{Proj}(R)$.

(4) Find the prime ideal I that corresponds to the point $(a:b:c)$ and prove that the set $\{I\}$ is closed in $\mathrm{Proj}(R)$.

(5) Prove that closed points of $\mathrm{Proj}(R)$ correspond to points in $\mathbb{P}^2$.

Exercise 5.6.5. In this exercise we show how to obtain $\mathbb{P}^n$ as $\mathrm{Proj}(R)$ for $R = \mathbb{C}[x_0, x_1, \ldots, x_n]$.

(1) Show that the ideal $I = \langle x_0, x_1, \ldots, x_{n-1}\rangle$ corresponds to the point $(0:0:\cdots:0:1) \in \mathbb{P}^n$. Prove that this ideal is maximal among those in $\mathrm{Proj}(R)$, so that $Z(I) = \{I\}$.

(2) Show that $Z(I) \neq \{I\}$ for the ideal $I = \langle x_0^2 + x_1^2 + \cdots + x_n^2\rangle$, by finding a point $P \in Z(I)$ with $P \neq I$.

(3) Find the prime ideal I that corresponds to the point $(a_0 : a_1 : \cdots : a_n)$, and prove that the set $\{I\}$ is closed in $\mathrm{Proj}(R)$.

(4) Prove that every closed point of $\mathrm{Proj}(R)$ corresponds to a point in $\mathbb{P}^n$.

As an extension of the previous exercises we next use the Proj construction to obtain a description of the parabola $x_0x_1 - x_2^2$ in $\mathbb{P}^2$. While this exercise provides some practice in using the definitions, it is not a recommended method for studying a parabola!

Exercise 5.6.6. Let $S = \mathbb{C}[x_0, x_1, x_2]/I$, where $I = \langle x_0x_1 - x_2^2\rangle$.

(1) As a chance to review some commutative algebra, prove that the homogeneous ideals of S correspond to homogeneous ideals of $\mathbb{C}[x_0, x_1, x_2]$ containing I.

(2) Show that the ideal $J = \langle x_0, x_2\rangle \subset S$ corresponds to the point $(0:1:0)$ on the parabola. Prove that the class of this ideal in $\mathrm{Proj}(S)$ is maximal among those not containing the irrelevant ideal, so that $Z(J) = \{J\}$.

(3) Find the prime ideal J that corresponds to the point $(-1:-1:1)$ on the parabola and prove that the set $\{J\}$ is closed in $\mathrm{Proj}(S)$.

(4) For an arbitrary point $(a : b : c)$ on the parabola, find the corresponding prime ideal J in S and prove that the set $\{J\}$ is closed in $\mathrm{Proj}(S)$.

(5) Show that the points of the parabola correspond to the closed points of $\mathrm{Proj}(S)$.

Chapter 6

The Next Steps: Sheaves and Cohomology

Sheaves and cohomology are two of the key mathematical ideas developed in the 20th-century. Their scope and power have fundamentally shaped current algebraic geometry and much more of modern mathematics. The goal of this chapter is to sketch the beginnings of sheaf theory. We will recast our study of divisors into the language of invertible sheaves. Finally, we will recast the statement of Riemann-Roch into the language of Čech cohomology of invertible sheaves. The underlying motivation for this chapter is to develop the necessary tools to pass from local to global information.

6.1. Intuition and Motivation for Sheaves

The goal of this section is to motivate our eventual definition of sheaves in terms of local versus global properties. As examples, we will review curve intersections and Riemann-Roch, and introduce the Mittag-Leffler problem of finding rational functions with prescribed poles on a curve.

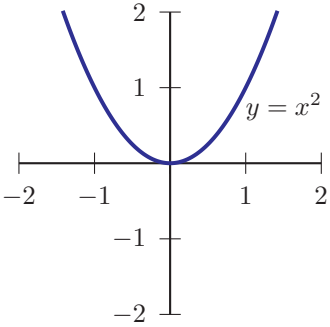
6.1.1. Local versus Global. We started this text with problems about conics in the plane $\mathbb{R}^2$ but saw that we needed to pass to the

complex projective plane $\mathbb{P}^2$. The rhetoric is that the conic in $\mathbb{R}^2$ (or in $\mathbb{C}^2$) is local, while the homogenized conic in $\mathbb{P}^2$ is global. This language is used since we form the complex projective plane $\mathbb{P}^2$ by patching (or gluing) together three copies of $\mathbb{C}^2$.

This patching or gluing is a powerful idea. With sheaves, we will again perform gluing operations, but this time we will be gluing functions rather than spaces together. The idea is almost the same. We need to describe how the functions overlap and be sure that they agree where they should. One of the roles sheaves will have to play for us is to record how functions can be pieced together from local parts to form larger wholes.

6.1.2. Local versus Global Curve Intersections. Bézout's Theorem is the quintessential global result. Here is why:

Exercise 6.1.1. Find a curve in $\mathbb{C}^2$ that intersects the curve $C = \mathrm{V}(y - x^2)$



in exactly one point, counting multiplicity.

This is an example of a local intersection, as it is happening in $\mathbb{C}^2$.

Exercise 6.1.2. Homogenize the two curves from the previous problem. Show that the two curves now must intersect in exactly two points.

The homogenized curve in $\mathbb{P}^2$ is the global version. The fact that the total intersection number must be two is thought of as a global result.

This is common. In $\mathbb{C}^2$, curves of degree d and e can intersect in any number of points, from zero to de, while the corresponding curves in $\mathbb{P}^2$ must intersect in exactly de points.

6.1.3. Local versus Global for Riemann-Roch. Let C be a smooth curve of genus g in $\mathbb{P}^2$. For any divisor D on C, Riemann-Roch (Theorem 3.6.47), states that

$$l(D) - l(K - D) = \deg(D) - g + 1.$$

Here $l(D)$ is the dimension of the vector space of all $f \in \mathcal{K}_C$ such that

$$\operatorname{div}(f) + D \geq 0.$$

Thus $l(D)$ is a measure of how many rational functions there are on the curve C with certain prescribed poles and zeros.

There is nothing to prevent us from trying to find affine analogues, namely for any affine curve C to ask for the dimension of the vector space of all $f \in \mathcal{K}_C$ such that

$$\operatorname{div}(f) + D \geq 0,$$

for a divisor D on C. But these vector spaces are quite different from the projective case, and no clean analogue to Riemann-Roch exists.

Exercise 6.1.3. Let C be the curve in $\mathbb{C}^2$ given by

$$y = x^2.$$

Let D be the divisor $-(0,0)$. Show that there is an $f \in \mathcal{K}_C$ such that

$$\operatorname{div}(f) + D \geq 0.$$

Exercise 6.1.4. Let C be the curve in $\mathbb{P}^2$ given by

$$yz = x^2$$

(the homogenization of the affine curve from the previous problem). Let D be the divisor $-(0:0:1)$. Show that there is no $f \in \mathcal{K}_C$ such that

$$\operatorname{div}(f) + D \geq 0.$$

6.1.4. Local versus Global for the Mittag-Leffler Problem. (This subsection requires a bit of complex analysis. If you want, whenever you see the term "meromorphic," just think ratios of polynomials.)

We will begin with a motivating example. Suppose f is a function whose Laurent series centered at a is given by $f(z) = \sum_{k=-\infty}^{\infty} c_k(z-a)^k$. The *principal part* of f at a is $\sum_{k=-\infty}^{-1} c_k(z-a)^k$. The function f has a *pole of order m at a* if the principal part of f at a is $\sum_{k=-m}^{-1} c_k(z-a)^k$, with $c_{-m} \neq 0$, that is, if the principal part of f at a is a finite sum.

Let Ω be an open subset of $\mathbb{C}$ and let $\{a_j\}$ be a sequence of distinct points in Ω such that $\{a_j\}$ has no limit point in Ω. For each integer $j \geq 1$ consider the rational function

$$P_j(z) = \sum_{k=1}^{m_j} \frac{c_{j,k}}{(z-a_j)^k}.$$

The Mittag-Leffler Theorem states that there exists a meromorphic function f on Ω, holomorphic outside of $\{a_j\}$, whose principal part at each a_j is $P_j(z)$ and which has no other poles in Ω. This theorem allows meromorphic functions on $\mathbb{C}$ to be constructed with an arbitrarily preassigned discrete set of poles.

Exercise 6.1.5. Find a meromorphic function f that has a pole of order 2 at the origin such that the residue of the origin is 0.

Exercise 6.1.6. Let $\omega_1, \omega_2 \in \mathbb{C}$ such that $\frac{\omega_1}{\omega_2} \notin \mathbb{R}$. Find a meromorphic function that has a pole at every point in the lattice $\Lambda = \{m\omega_1 + n\omega_2 \mid m, n \in \mathbb{Z}\}$.

Since we can construct functions with arbitrarily preassigned discrete sets of poles on $\mathbb{C}$, it is natural to ask the same question on a complex curve (which we know can be viewed as a real surface). Suppose X is a smooth complex projective curve (also called a Riemann surface.) Given a discrete set of points $\{a_j\}$ and a principal part $P_j(z)$ at each a_j, where z is a local affine coordinate, does there

exist a rational function f on X, defined outside $\{a_j\}$, whose principal part at each a_j is $P_j(z)$? Locally, there is such a function provided by the Mittag-Leffler Theorem, but whether there exists such a function defined globally is more subtle. This requires passing from local information to global information. The primary virtue of sheaves is that they provide a mechanism to deal with problems passing from local information to global information.

6.1.5. Local versus Global: the Sheaf of Regular Functions. Prior to giving the definition of sheaves, we will look at a concrete example of a sheaf that has the virtue of its ubiquitousness. In the next section, the reader will prove that the object we encounter here is indeed a sheaf.

Let X be an algebraic variety, either affine or projective. There is always the sheaf $\mathcal{O}_X$ of regular functions on X, defined by assigning to each Zariski open set U in X its ring of regular functions

$$\mathcal{O}_X(U) = \{\text{regular function on } U\}$$

and letting $r_{V,U}$, for $U \subset V \subset X$, be the restriction map. In fact, we have already been using the notation $\mathcal{O}_X$ throughout this book.

Exercise 6.1.7. Consider the projective line $\mathbb{P}^1$ with homogeneous coordinates $(x_0 : x_1)$. Let $U_0 = \{(x_0 : x_1) : x_0 \neq 0\}$. Show that the ring $\mathcal{O}_X(U_0)$ is isomorphic to the ring $\mathbb{C}[t]$. Show that $\mathcal{O}_X(\mathbb{P}^1)$ is isomorphic to $\mathbb{C}$, the constant functions.

The functions making up $\mathcal{O}_X(U_0)$ are viewed as local, while those making up $\mathcal{O}_X(\mathbb{P}^1)$ are global. This of course extends to any projective variety, as seen in the following example for curves.

Exercise 6.1.8. In $\mathbb{P}^2$, let

$$X = \{(x_0 : x_1 : x_2) : x_0^2 + 3x_1^2 - x_2^2 = 0\},$$

and let $U_0 = \{(x_0 : x_1 : x_2) \in X : x_0 \neq 0\}$. Show that $\mathcal{O}_X(U_0)$ is isomorphic to the ring $\mathbb{C}[s,t]/\langle 3s^2 - t^2 + 1\rangle$ but that $\mathcal{O}_X(X)$ is isomorphic to $\mathbb{C}$, the constant functions.

6.2. The Definition of a Sheaf

We first define presheaves and then define sheaves.

Suppose X is a topological space. Since we are interested in both the local and global structure of X, we wish to assign to each open set U of X a collection of data that is somehow characteristic of U. Since different kinds of algebraic structures can encode geometric information about a topological space, it is useful to introduce a concept that encompasses different ways of assigning algebraic structures to the space.

Definition 6.2.1. A *presheaf* $\mathcal{F}$ of rings of functions (or modules over rings) on X consists of a ring of functions (resp. module, etc.) $\mathcal{F}(U)$ for every open set $U \subset X$ and the ring homomorphism given by the restriction map (resp. module homomorphism, etc.) $r_{V,U} : \mathcal{F}(V) \to \mathcal{F}(U)$ for any two nested open subsets $U \subset V$ satisfying the following two conditions:

(i) $r_{U,U} = \mathrm{id}_{\mathcal{F}(U)}$

(ii) For open subsets $U \subset V \subset W$ one has $r_{W,U} = r_{V,U} \circ r_{W,V}$.

The elements of $\mathcal{F}(U)$ are called the *sections* of $\mathcal{F}$ over U and the map $r_{V,U}$ is called the *restriction map*, and $r_{V,U}(s)$ is often written $s\big|_U$.

For almost all of our examples, each $\mathcal{F}(U)$ will consist of some specified type of function defined on the open set U. In this type of case, when $U \subset V$, if f is a function with domain V, then $r_{V,U}(f)$ is simply the same function f, but now with domain restricted to the smaller open set U. Then the first axiom can be interpreted as requiring that the restriction of a function from a space to itself always returns the same function. That is, a trivial restriction should not change functions. The second axiom, in turn, says that the result of a sequence of restrictions should be identical to the single restriction from the initial to the final subspace. Again, in the context of restrictions of functions, this axiom is very natural. This also means that for the following exercises, where you are asked to show that various objects are presheaves, you just have to show that if $f \in \mathcal{F}(V)$, then

f with domain restricted to a smaller open set U is in $\mathcal{F}(U)$, or in other words, that the restriction map $r_{V,U}$ really does map elements of $\mathcal{F}(V)$ to elements of $\mathcal{F}(U)$. (This also means that the answers will not be that long.)

The building block for almost all sheaves in algebraic geometry is the sheaf of regular functions $\mathcal{O}_X$ on an algebraic variety X. We first show that $\mathcal{O}_X$ is at the least a presheaf.

Exercise 6.2.1. Suppose X is a variety, affine or projective. Show that its sheaf of regular functions $\mathcal{O}_X$, as in Section 6.1.5, is a presheaf as just defined.

Exercise 6.2.2. Suppose X is a topological space. For open U define

$$\mathcal{F}(U) = \{f : U \to \mathbb{Z} : f \text{ constant on connected components of } U\}$$

and let $r_{V,U}(f)$ be the restriction of f from V to U. Show that $\mathcal{F}$ is a presheaf of rings.

Exercise 6.2.3. Suppose X is a topological space. Define

$$C(U) = \{f : U \to \mathbb{C} : f \text{ is continuous}\}$$

and let $r_{V,U}(f)$ be the restriction of f from V to U. Show that C is a presheaf of rings.

Exercise 6.2.4. Suppose $X = \mathbb{C}$. Define

$$\mathcal{B}(U) = \{f : U \to \mathbb{C} : f \text{ is a bounded holomorphic function}\}$$

and let $r_{V,U}(f)$ be the restriction of f from V to U. Show that $\mathcal{B}$ is a presheaf of rings.

Presheaves enable us to assign to each open set of a topological space X an algebraic structure that describes the open set and how it fits inside of X. However, presheaves are top-down constructions; we can restrict information from larger to smaller sets. The problem of globalizing local data is not within the scope of the definition of a presheaf. That is, presheaves do not provide the means to deduce global properties from the properties we find locally in the open sets of X. The definition of a sheaf below is meant to resolve this, enabling us to pass data from global to local settings but also to patch local information together to establish global results when possible.

Definition 6.2.2. A presheaf $\mathcal{F}$ of rings of functions (or modules over rings) on X is called a *sheaf* of rings of functions (or modules over rings) if, for every collection U_i of open subsets of X with $U = \bigcup_i U_i$, the following two additional conditions are satisfied.

(iii) If $s, t \in \mathcal{F}(U)$ and $r_{U,U_i}(s) = r_{U,U_i}(t)$ for all i, then $s = t$.

(iv) If $s_i \in \mathcal{F}(U_i)$ and if for $U_i \cap U_j \neq \emptyset$ we have

$$r_{U_i,U_i\cap U_j}(s_i) = r_{U_j,U_i\cap U_j}(s_j)$$

for all i, j, then there exists $s \in \mathcal{F}(U)$ such that $r_{U,U_i}(s) = s_i$.

In light of the interpretation of functions and their restrictions, the new axioms for a sheaf are essential ingredients for inferring global information from local data. Axiom (iii) requires that two functions must be the same if they agree everywhere locally, i.e., if for every subset W of U, $s\big|_W = t\big|_W$, then $s = t$. Were this not true, then it would be impossible to construct a single global function on U from the parts of it we have on each of the U_i. Hence, axiom (iii) has to do with the uniqueness of global functions that we might construct from local data. Axiom (iv), in turn, has to do with the existence of such functions. Whenever we are given a collection of functions defined on various parts of X, we can patch them together to form a unique (due to axiom (iii)) function on X as long as this is feasible, i.e., two constituent functions s_i and s_j must agree wherever both are defined in X.

For our above presheaves, the patching is clear. The only reason that all of the above presheaves are not automatically sheaves is if the patched together function on the open set U is not an element of the corresponding presheaf.

Exercise 6.2.5. Let our presheaf $\mathcal{F}$ be a presheaf of functions, with $r_{V,U}(f)$ being the restriction map $f\big|U$. Show that axiom (iii) is equivalent to the following. If $s \in \mathcal{F}(U)$ such that $r_{U,U_i}(s) = 0$ for all i, then $s = 0$.

Exercise 6.2.6. Show that the presheaf $\mathcal{F}$ from Exercise 6.2.2 is a sheaf.

Exercise 6.2.7. Show that the presheaf C from Exercise 6.2.3 is a sheaf.

Exercise 6.2.8. Suppose X is a variety, affine or projective. Show that its sheaf of regular functions, $\mathcal{O}_X$, is a sheaf. (This is the key example for this section.)

Exercise 6.2.9. Show that the presheaf $\mathcal{B}$ from Exercise 6.2.4 is not a sheaf.

As we found in the last exercise, not all presheaves are sheaves. There is a construction, which we will describe now, that associates a sheaf to any presheaf in a universal way. The key distinction between a sheaf and a presheaf is the ability with a sheaf to assemble local data together to construct global results. Thus we first need to focus on the local data in a presheaf and force the construction of global information from it to construct the associated sheaf. To be as local as possible, we want to study the essense of a presheaf at a point.

As in the examples above, let us suppose that the elements of a presheaf $\mathcal{F}$ on X are functions. That is, an element $s \in \mathcal{F}(U)$ is a function on the open set U. Then the value $s(x)$ alone will not capture the essence of this function at x, for it is very likely that several distinct functions may have the same value at x. Hence we want to keep track of not only the value of s at x but also the values of s near x. This can be done by keeping track of the pair (U, s), where U is an open set containing x and $s \in \mathcal{F}(U)$. However, if V is any other open set containing x, then $U \cap V$ is one also and $(U \cap V, s|_{U\cap V})$ is really the same function near x that (U, s) is. So these two "local functions" at x should be identified with one another. In general, the pairs (U, s) and (V, t) are *equivalent* whenever there is a third open set W with $W \subset U \cap V$, $x \in W$, and $s|_W = t|_W$ in $\mathcal{F}(W)$.

Exercise 6.2.10. Let $\mathcal{F}$ be a presheaf of functions on a space X. Let $x \in X$ and let U and V be open sets containing the point x. Suppose $s \in \mathcal{F}(U)$ and $t \in \mathcal{F}(V)$. Set

$$(U, s) \sim (V, t)$$

whenever there is a third open set W with $W \subset U \cap V$, $x \in W$, and $s|_W = t|_W$ in $\mathcal{F}(W)$. Show that this is an equivalence relation.

Definition 6.2.3. If $\mathcal{F}$ is any presheaf on a topological space X and x is any point in X, the equivalence class of (U, s), where U is an open set of X containing x and $s \in \mathcal{F}(U)$, is denoted by s_x and is called the *germ* of the section s at x.

Definition 6.2.4. Let X be a topological space and let $\mathcal{F}$ be a presheaf on X. For a point $x \in X$, the *stalk* of $\mathcal{F}$ at x, denoted $\mathcal{F}_x$, consists of the germs s_x of sections at x for all open sets U containing x and all $s \in \mathcal{F}(U)$.

We can now explain how to extend any presheaf to an actual sheaf.

Definition 6.2.5. Using the stalks of a presheaf $\mathcal{F}$ on X, we construct the *sheaf associated to* $\mathcal{F}$, denoted $\mathcal{F}^+$, as follows. For any open set U, $\mathcal{F}^+(U)$ consists of all functions s from U to the union $\bigcup_{x \in U} \mathcal{F}_x$ of the stalks of $\mathcal{F}$ over points of U such that

(1) for each $x \in U$, $s(x) \in \mathcal{F}_x$

(2) for each $x \in U$, there is a neighborhood V of x, contained in U, and an element $\hat{s} \in \mathcal{F}(V)$, such that for all $y \in V$, the germ $\hat{s}_y$ of $\hat{s}$ at y is equal to $s(y)$.

This is an admittedly complicated definition. What we want is for our candidate sheaf $\mathcal{F}^+(U)$ to contain $\mathcal{F}(U)$ plus whatever extra that is needed to make it a sheaf. The next problem is showing how to interpret elements of $\mathcal{F}(U)$ as also being in the new $\mathcal{F}^+(U)$.

Exercise 6.2.11. Let $\mathcal{F}$ be a presheaf on a topological space X. Let $s \in \mathcal{F}(U)$. Interpret s as an element of $\mathcal{F}^+(U)$.

Exercise 6.2.12. Let $\mathcal{F}$ be a presheaf on a topological space X. Prove that $\mathcal{F}^+$ is a sheaf on X.

Exercise 6.2.13. Let $\mathcal{F}$ be a sheaf on a topological space X. Show that this sheaf is the same as our newly constructed sheaf $\mathcal{F}^+$.

Exercise 6.2.14. For the presheaf $\mathcal{B}$ of Exercise 6.2.4, show that its associated sheaf, $\mathcal{B}^+$, on $X = \mathbb{C}$ is the sheaf of holomorphic functions on $\mathbb{C}$. (The sheaf of holomorphic functions $\mathcal{H}$ on $\mathbb{C}$ is defined by setting for all open U

$$\mathcal{H}(U) = \{f : U \to \mathbb{C} : f \text{ is a holomorphic function}\}.$$

To work this problem you will need to know that two holomorphic functions that agree on any open set in $\mathbb{C}$ agree everywhere the functions are defined.)

6.3. The Sheaf of Rational Functions

The second most important sheaf in algebraic geometry is the sheaf of rational functions $\mathcal{K}_X$, whose definition is the goal of this section.

Let X be an algebraic variety, either affine or projective. Then X is equipped with its sheaf of regular functions, $\mathcal{O}_X$. There is another basic sheaf for every algebraic variety X, namely the function field sheaf $\mathcal{K}_X$, which plays the "sheaf-theoretic" role of the function field. Morally we want to think of $\mathcal{K}_X$ as the ratio of the functions in $\mathcal{O}_X$. The actual definition, though, is mildly subtle, as we will see. It is here, in fact, that we will need to use the difference between a presheaf and a sheaf.

We start by defining a presheaf $\mathcal{K}'_X$. For each open U in X, let $\mathcal{K}'_X(U)$ be the function field of the ring $\mathcal{O}_X(U)$, with the standard restriction map for functions. (Here we are using the Zariski topology; thus the various open U are complements of the zero loci for various polynomials.) Thus $\mathcal{K}'_X(U)$ consists of all ratios

$$\frac{f}{g},$$

with $f, g \in \mathcal{O}(U)$ and g not the zero function. The goal of the next series of exercises is to see why $\mathcal{K}'_X$ is only a presheaf and to motivate why we actually want to look at its associated sheaf.

Exercise 6.3.1. Let X be an algebraic variety, either affine or projective. Verify that $\mathcal{K}'_X$ is a presheaf of fields of functions on X.

We now concentrate on the space $\mathbb{P}^1$, which is covered by the two open sets $U_0 = \{(x_0 : x_1) : x_0 \neq 0\}$ and $U_1 = \{(x_0 : x_1) : x_1 \neq 0\}$. Then on U_0 we let $s = (x_1/x_0)$ be our affine coordinate, and on U_1 we let $t = (x_0/x_1)$ be our affine coordinate. On the overlap, $U_0 \cap U_1$, we have $s = (1/t)$.

Exercise 6.3.2. Show that $\mathcal{K}'_{\mathbb{P}^1}(U_0)$ is isomorphic to the field $\mathbb{C}(s)$ and that $\mathcal{K}'_{\mathbb{P}^1}(U_1)$ is isomorphic to the field $\mathbb{C}(t)$.

Exercise 6.3.3. Show that $\mathcal{K}'_{\mathbb{P}^1}(\mathbb{P}^1)$ is isomorphic to the field $\mathbb{C}$.

Exercise 6.3.4. Using that $(1/t) \in \mathcal{K}'_{\mathbb{P}^1}(U_1)$ and condition (iii) in the definition of a sheaf, show that $\mathcal{K}'$ cannot be a sheaf.

Definition 6.3.1. The *function field sheaf* $\mathcal{K}_X$ for an algebraic variety X is the sheaf associated to the presheaf $\mathcal{K}'_X$

Exercise 6.3.5. Show that $\mathcal{K}_{\mathbb{P}^1}(\mathbb{P}^1)$ is isomorphic to the field $\mathbb{C}(s)$.

6.4. Divisors

The goal of this problem set is to generalize the notion of divisor from being the finite formal sum of points on a complex curve to being the finite formal sum of codimension one subvarieties of an algebraic variety.

In this section, we revisit a familiar tool, divisors, from Chapter 3. We will see how divisors are intimately related to the special class of invertible sheaves in the next section and how this can be used to give a new presentation of the Riemann-Roch Theorem at the end of the chapter.

Recall from Chapter 3 that a divisor D on a curve C is a formal finite linear combination of points on C with integer coefficients, $D = n_1p_1 + n_2p_2 + \cdots + n_kp_k$ with $n_1, \ldots, n_k \in \mathbb{Z}$ and $p_1, \ldots, p_k \in C$. One might think a divisor on a variety X would be a formal finite sum of points as before. However, this turns out not to be the correct generalization. Recall the purpose of a divisor on a curve was to keep track of the zeros and poles of a single function. On a variety X, a function's zeros constitute an algebraic subvariety usually of dimension one less than the dimension of X. Thus, rather than adding points, we should add subsets that look like the zero sets of single functions on X. To be precise:

Definition 6.4.1. A *codimension one subvariety* of a variety X is a proper irreducible algebraic subset $Y \subset X$ such that there are no other proper irreducible algebraic subsets Z satisfying $Y \subsetneq Z \subsetneq X$.

Definition 6.4.2. Let X be an algebraic variety. A *divisor* D on X is a finite formal sum over the integers $\mathbb{Z}$ of codimension one subvarieties of X. The set of all divisors is denoted $\mathrm{Div}(X)$.

Let X be a curve in $\mathbb{P}^2$ and let p, q, r be points on X. Then an example of a divisor is

$$D = 3p - 5q + r.$$

The coefficients $3, -5, 1$ are just integers, while the points p, q, r are the codimension one subvarieties of X. We need to use the term "formal sum" since adding points makes no real sense.

An example of a divisor on $\mathbb{P}^2$, using the homogeneous coordinates x_0, x_1, x_2, would be

$$3(x_0^2+x_1x_2=0)-7(x_0^5+x_1^3x_2^2=0) = 3\mathrm{V}(x_0^2+x_1x_2)-7\mathrm{V}(x_0^5+x_1^3x_2^2).$$

The coefficients 3 and -7 are just integers, while the codimension one subvarieties are the curves $\mathrm{V}(x_0^2 + x_1x_2)$ and $\mathrm{V}(x_0^5 + x_1^3x_2^2)$.

As divisors are formal sums, we should be able to add them. Thus if $D_1 = 3p - 5q + r$ and $D_2 = 8q + 4s - 4t$ are two divisors on the curve X, define

$$\begin{aligned} D_1 + D_2 = &\ 3p - 5q + r + 8q + 4s - 4t \\ = &\ 3p + (-5 + 8)q + r + 4s - 4t \\ = &\ 3p + 3q + r + 4s - 4t. \end{aligned}$$

Exercise 6.4.1. Let X be an algebraic curve. Let $D_1 = \sum_{p\in X} n_p p$ and $D_2 = \sum_{p\in X} m_p p$, where the $n_p, m_p \in \mathbb{Z}$, be two divisors on X. If we define

$$D_1 + D_2 = \sum_{p\in X} (n_p + m_p)p,$$

show that $\mathrm{Div}(X)$ is an abelian group. (Note in the above sums for the divisors D_1 and D_2, that even though the sums are over all points $p \in X$, we are assuming that $n_p = m_p = 0$ for all but a finite number of points on X; this is what is meant in the definition of a divisor by the phrase "finite formal sum.")

Exercise 6.4.2. Let X be an algebraic variety. Let $D_1 = \sum n_V V$ and $D_2 = \sum m_V V$, where the $n_V, m_V \in \mathbb{Z}$, be two divisors on X.

Here both sums are over all codimension one subvarieties of X. If we define

$$D_1 + D_2 = \sum (n_V + m_V)V,$$

show that $\mathrm{Div}(X)$ is an abelian group.

Definition 6.4.3. A divisor $D = \sum n_V V$ is *effective* if, for all codimension one subvarieties V of X, we have $n_V \geq 0$. In this case we write $D \geq 0$.

We now want to link divisors with both the geometry of the variety X and functions defined on X. In particular, we want to associate to every element $f \in \mathcal{O}_X$ a divisor, which we will denote by $\mathrm{div}(f)$. This in turn will allow us to define, for every rational function $f/g \in \mathcal{K}_X$ (where $f, g \in \mathcal{O}_X$), the divisor

$$\mathrm{div}\left(\frac{f}{g}\right) = \mathrm{div}(f) - \mathrm{div}(g).$$

Let X be a curve in $\mathbb{P}^2$. Let $C = \mathrm{V}(P(x_0, x_1, x_2))$ be another curve in $\mathbb{P}^2$ that shares no components with X. Then define

$$D = X \cap C = \sum_{p \in X \cap C} m_p p,$$

where m_p is the intersection multiplicity of the intersection point. Since C shares no components with X, their intersection is a finite set of points, so D is a divisor on X. Since C is defined as the zero locus of the homogeneous polynomial P, then we can think of P as an element of $\mathcal{O}_X$. We use the notation

$$\mathrm{div}(P) = \sum_{p \in X \cap C} m_p p.$$

Exercise 6.4.3. Let $X = \mathrm{V}(x^2 + y^2 - z^2)$ be a conic in $\mathbb{P}^2$. If $C_1 = \mathrm{V}(x - y)$ and $C_2 = \mathrm{V}(y - z)$, show that the two corresponding divisors are

$$\begin{aligned} D_1 &= X \cap C_1 = \left(\frac{1}{\sqrt{2}} : \frac{1}{\sqrt{2}} : 1\right) + \left(-\frac{1}{\sqrt{2}} : -\frac{1}{\sqrt{2}} : 1\right) \\ D_2 &= X \cap C_2 = 2(0 : 1 : 1). \end{aligned}$$

Give a geometric interpretation for the coefficients in D_1 and D_2.

We now want to define on $\mathbb{P}^n$ divisors associated to homogeneous polynomials.

Definition 6.4.4. Let $x_0, \ldots, x_n$ be homogeneous coordinates for $\mathbb{P}^n$. Given a homogeneous polynomial $f(x_0, \ldots, x_n)$, factor f into its irreducible factors:

$$f(x_0, \ldots, x_n) = \prod f_i(x_0, \ldots, x_n)^{n_i}.$$

Then the *divisor* on $\mathbb{P}^n$ associated to f is

$$\operatorname{div}(f) = \sum n_i \mathrm{V}(f_i).$$

Exercise 6.4.4. Let

$$\begin{aligned} f(x_0, x_1, x_2) = {} & x_0^3 - x_0^2 x_1 + 5x_0^2 x_2 - x_0 x_1^2 - 2x_0 x_1 x_2 \\ & + 8x_0 x_2^2 + x_1^3 - 3x_1^2 x_2 + 4x_2^3. \end{aligned}$$

Show that

$$\operatorname{div}(f) = \mathrm{V}(x_0 + x_1 + x_2) + 2\mathrm{V}(x_0 - x_1 + 2x_2).$$

Finally, we now want to define the divisor $\operatorname{div}(f)$ on a variety X for any $f \in \mathcal{O}_X$. Our algebraic variety X is in either $\mathbb{C}^n$ or $\mathbb{P}^n$. Then we can think of f as a polynomial in n-variables if $X \subset \mathbb{C}^n$, or as a homogeneous polynomial in $(n+1)$-variables if $X \subset \mathbb{P}^n$. In either case, we will look at the irreducible components of the intersection:

$$X \cap \mathrm{V}(f) = V_1 \cup \cdots \cup V_k.$$

Though we have not yet defined the numbers m_{V_i}, we will want our eventual divisor to be

$$\operatorname{div}(f) = \sum m_{V_i} V_i.$$

Morally we want m_{V_i} to capture the order of vanishing of f along the component V_i. We will sketch the argument in the next paragraphs and problems.

Let V be an irreducible codimension one subvariety of an algebraic variety X. This means for any $p \in V$, there is an open affine set U containing p so that there is an irreducible function $g \in \mathcal{O}_X(U)$ such that

$$V \cap U = \mathrm{V}(g) \cap U.$$

(For a few more of the technical details for why we have to bring in this seemingly extraneous open set, see Chapter 1, Section 1 of [**GH94**].) Now let $f \in \mathcal{O}_X$. Define

$$m_V(f) = \max\{k \in \mathbb{Z} : g^k \text{ divides } f\}.$$

Definition 6.4.5. Let X be an algebraic variety and $f \in \mathcal{O}_X$. Then the *divisor associated to* f is

$$\operatorname{div}(f) = \sum m_V(f) \cdot V,$$

where the sum is over all codimension one irreducible subvarieties V of X.

We want to see that on $\mathbb{P}^1$, this new definition agrees with our earlier one.

Exercise 6.4.5. Let

$$f(x_0, x_1) = (x_0 - x_1)^2(x_0 - 2x_1).$$

Using the above definition, show that

$$\operatorname{div}(f) = 2(1:1) + (2:1).$$

In the next few sections, we will see that the following definition for linear equivalence for divisors will be important:

Definition 6.4.6. Let X be a projective variety. Divisors D_1 and D_2 are said to be *linearly equivalent* if there are two homogeneous polynomials f and g of the same degree such that

$$D_1 + \operatorname{div}\left(\frac{f}{g}\right) = D_2.$$

We denote this by

$$D_1 \sim D_2.$$

Exercise 6.4.6. On $\mathbb{P}^1$, show that $D_1 = (1:1)$ is linearly equivalent to $D_2 = (1:0)$.

Exercise 6.4.7. Let $\mathbb{P}^2$ have homogeneous coordinates x_0, x_1, x_2. Show that the divisors $D_1 = \mathrm{V}(x_0^2 + 3x_2^2)$ and $D_2 = \mathrm{V}(x_0^2)$ are linearly equivalent.

Exercise 6.4.8. Let $f(x_0, \ldots, x_n)$ be any homogeneous polynomial of degree d. Show that the divisors $D_1 = \mathrm{V}(f)$ and $D_2 = \mathrm{V}(x_0^d)$ are linearly equivalent.

Exercise 6.4.9. Let $f(x_0, \ldots, x_n)$ and $g(x_0, \ldots, x_n)$ be any two homogeneous polynomials of degree d. Show that the divisors $D_1 = \mathrm{V}(f)$ and $D_2 = \mathrm{V}(g)$ are linearly equivalent.

Exercise 6.4.10. Show that linear equivalence is indeed an equivalence relation on the group $\mathrm{Div}(X)$.

Definition 6.4.7. The group $\mathrm{Div}(X)$ divided out by the equivalence relation of linear equivalence is called the *Picard group*, or the *divisor class group*, of X.

Exercise 6.4.11. Let D_1 and D_2 be two divisors on $\mathbb{P}^1$. Show that $D_1 \sim D_2$ if and only if they have the same degree.

Exercise 6.4.12. Let D_1 and D_2 be two divisors on $\mathbb{P}^n$. Show that $D_1 \sim D_2$ if and only if they have the same degree.

Exercise 6.4.13. Show that the map

$$\deg : \mathrm{Div}(\mathbb{P}^n) \to \mathbb{Z}$$

given by

$$\deg\left(\sum n_V V\right) = \sum n_V$$

is a group homomorphism, treating $\mathbb{Z}$ as a group under addition.

Exercise 6.4.14. Show that the Picard group for $\mathbb{P}^n$ is isomorphic to the group $\mathbb{Z}$ under addition.

6.5. Invertible Sheaves and Divisors

In this section we link divisors with invertible sheaves, a special type of sheaf.

Definition 6.5.1. On an algebraic variety X, an *invertible sheaf* $\mathcal{L}$ is any sheaf so that there is an open cover $\{U_i\}$ of X such that $\mathcal{L}(U_i)$ is a rank-one $\mathcal{O}_X(U_i)$-module.[1]

[1] Modules are similar to vector spaces, which are always defined over a field of scalars such as $\mathbb{C}$. The scalars for modules, however, may be taken from an arbitrary

Thus for each open set U_i, we have $\mathcal{L}(U_i)$ is isomorphic to $\mathcal{O}_X(U_i)$ as an $\mathcal{O}_X(U_i)$-module.

We will first see how to intuitively associate to a divisor D an invertible sheaf, which we will denote by $\mathcal{L}_D$. Let $D = \sum n_V V$ be a divisor, where the V are codimension one subvarieties of X. We know that $n_V = 0$ for all but a finite number of V. We can cover X by open affine sets U_i so that for each i there is a rational function $f_i \in \mathcal{K}(U_i)$ such that

$$\operatorname{div}(f_i) = D \cap U_i.$$

In other words, the zeros and poles of f_i agree with the coefficients n_V of D.

Exercise 6.5.1. For the conic $X = \mathrm{V}(x^2 + y^2 - z^2)$ in $\mathbb{P}^2$, consider the divisor

$$D = \left(\frac{1}{\sqrt{2}} : \frac{1}{\sqrt{2}} : 1\right) + \left(-\frac{1}{\sqrt{2}} : -\frac{1}{\sqrt{2}} : 1\right) - (1 : i : 0)\,.$$

On the open set $U = \{(x : y : z) \mid z \neq 0\}$, show that if

$$f(x, y, z) = \frac{x}{z} - \frac{y}{z},$$

then

$$\operatorname{div}(f) \cap U = D \cap U.$$

Thus each divisor D can be thought of as not only a finite formal sum of codimension one subvarieties but also as some collection (U_i, f_i), where the $\{U_i\}$ are an open affine cover of X and each $f_i \in \mathcal{K}_X(U_i)$. It works out that these two methods are exactly equivalent when X is a smooth variety but are not necessarily the same when X is singular (though the proof of this fact is non-trivial). From this point forward, we will restrict our attention to smooth varieties.

Further, this definition of D depends on the choice of open cover, which is hardly unique. The key is that if we write D as some (U_i, f_i) or as some (V_j, g_j), for some other open cover $\{V_j\}$ with $g_j \in \mathcal{K}_X(V_j)$, then on the overlaps $U_i \cap V_j$ the $\frac{f_i}{g_j}$ have no zeros or poles.

---

ring, which is the key difference in the definition. The notion of dimension translates into that of rank for modules. A more detailed account of modules and rank can be found in [**DF03**] or [**Her75**].

Thus we can write a divisor D as

$$D = (U_i, f_i).$$

Definition 6.5.2. Given $D = (U_i, f_i)$, define the invertible sheaf $\mathcal{L}_D$ by setting

$$\mathcal{L}_D(U_i) = \left\{ \frac{g}{f_i} : g \in \mathcal{O}_X(U_i) \right\}.$$

Exercise 6.5.2. Suppose that

$$\frac{g}{f_i}, \frac{h}{f_i} \in \mathcal{L}_D(U_i).$$

Show that

$$\frac{g}{f_i} + \frac{h}{f_i} \in \mathcal{L}_D(U_i).$$

For any $\alpha \in \mathcal{O}_X(U_i)$, show that

$$\frac{\alpha g}{f_i} \in \mathcal{L}_D(U_i).$$

(This problem is explicitly showing that each $\mathcal{L}_D(U_i)$ is an $\mathcal{O}_X(U_i)$-module; it is not hard.)

For a divisor $D = (U_i, f_i)$, let

$$g_{ij} = \frac{f_i}{f_j}.$$

We know that on the intersection $U_i \cap U_j$, the functions g_{ij} have no zeros or poles.

Exercise 6.5.3. Show that on $U_i \cap U_j \cap U_k$, we have

$$g_{ij} g_{jk} g_{ki} = 1.$$

(For those who know about vector bundles, this means that the invertible sheaf $\mathcal{L}_D$—or for that matter the divisor D—can be thought of as a complex line bundle.)

There is another equivalent way of associating an invertible sheaf to a divisor D. Again let $D = \sum n_V V$, where each V is a codimension one subvariety of X. Let U be an open subset of X. Then we define

$$D\big|_U = \sum n_V (V \cap U).$$

For any $f \in \mathcal{K}_X(U)$, define $\operatorname{div}(f)\big|_U$ to be the divisor of zeros and poles of f on the open set U.

Definition 6.5.3. Define a sheaf $\mathcal{L}_D$ by setting, for each open set U of X,

$$\mathcal{L}_D(U) = \{f \in \mathcal{K}_X(U) \mid (\mathrm{div}(f) + D) \cap U \geq 0\}.$$

More colloquially, $\mathcal{L}_D(U)$ consists of those rational functions on U whose poles are no worse than $-D$.

Exercise 6.5.4. Let $D = (U_i, f_i)$ be a divisor on X. Let $\mathcal{L}_D$ be the invertible sheaf associated to D as constructed in Definition 6.5.2 and let $\mathcal{L}'_D$ be the invertible sheaf associated to D as described in Definition 6.5.3. Show that for each open set U in X, $\mathcal{L}_D(U) = \mathcal{L}'_D(U)$. Thus the definitions give two ways to associate the same invertible sheaf to D.

Exercise 6.5.5. For $\mathbb{P}^1$ with homogeneous coordinates $(x : y)$, let $D = (1 : 0)$. Let $U_1 = \{(x : y) : x \neq 0\}$ and $U_2 = \{(x : y) : y \neq 0\}$. Show that $\mathcal{L}_D(U_1)$ is isomorphic to the set of all rational functions of the form $\frac{f(t)}{t}$, where $f(t) \in \mathbb{C}[t]$. (Here let $t = y/x$.) By letting $s = x/y$, show that $\mathcal{L}_D(U_2)$ is isomorphic to $\mathbb{C}[s]$. Finally show that $\mathcal{L}_D(\mathbb{P}^1)$ is not empty.

Exercise 6.5.6. For $\mathbb{P}^1$ with homogeneous coordinates $(x : y)$, let $D = -(1 : 0)$. Let $U_1 = \{(x : y) : x \neq 0\}$ and $U_2 = \{(x : y) : y \neq 0\}$. Show that $\mathcal{L}_D(U_1)$ is isomorphic to the ideal $\{f(t) \in \mathbb{C}[t] : f(0) = 0\}$. (Here let $t = y/x$.) By letting $s = x/y$, show that $\mathcal{L}_D(U_2)$ is isomorphic to $\mathbb{C}[s]$. Finally show that $\mathcal{L}_D(\mathbb{P}^1)$ is empty.

6.6. Basic Homology Theory

Homology theory is presented as a means for measuring the non-exactness of sequences of rings or modules.

Homology and cohomology theories permeate a large part of modern mathematics. There is a serious start-up cost to understanding this machinery, but it is well worth the effort.

Suppose we have a collection of objects $\{M_i\}$, such as a bunch of rings of functions, modules, abelian groups, or vector spaces, for

$i = 0, 1, 2, \ldots$. Suppose that we have maps

$$d_i : M_i \to M_{i-1}$$

where each d_i is an appropriate map, meaning that if the M_i are rings, then the d_i are ring homomorphisms and if the M_i are vector spaces, then the d_i are linear transformations. We write these out as a sequence

$$\cdots \to M_{i+1} \to M_i \to M_{i-1} \to \cdots,$$

with the map from $M_i \to M_{i-1}$ given by d_i. We require for all i that

$$\text{Image}(d_i) \subset \text{Kernel}(d_{i-1}).$$

In other words,

$$d_{i-1} \circ d_i = 0 \text{ for all } i.$$

We call this a *complex*. Frequently the index i is left off, which leads $d_{i-1} \circ d_i = 0$ to be written as the requirement

$$d \circ d = 0.$$

Definition 6.6.1. A sequence

$$\cdots \to M_{i+1} \to M_i \to M_{i-1} \to \cdots$$

is *exact* if for all i we have

$$\text{Image}(d_i) = \text{Kernel}(d_{i-1}).$$

Exercise 6.6.1. Let

$$0 \to A_3 \to A_2 \to A_1 \to 0$$

be an exact sequence of either rings or vector spaces, with 0 denoting either the zero ring or the vector space of one point. Show that the map $A_3 \to A_2$ must be one-to-one and the map $A_2 \to A_1$ must be onto.

Exercise 6.6.2. Find group homomorphisms so that the corresponding sequence

$$0 \to \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to 0$$

is exact.

In the above, $\mathbb{Z}/2\mathbb{Z}$ denotes the "quotienting" of the integers by the even integers, and hence is the group of two elements $\{0, 1\}$.

Definition 6.6.2. Let

$$\cdots \to M_{i+1} \to M_i \to M_{i-1} \to \cdots$$

be a sequence of abelian groups or vector spaces. Then the i-th *homology* is

$$H_i = \text{Kernel}(d_{i-1})/\text{Image}(d_i).$$

Exercise 6.6.3. Show that a sequence of abelian groups or vector spaces is exact if and only if for all i we have $H_i = 0$. (This is just an exercise in applying definitions; there really is not much to show.)

Thus homology is a way of measuring the exactness of a complex.

6.7. Čech Cohomology

The bare bones of Čech cohomology is given. This allows us to study the Čech cohomology for divisors on algebraic varieties, which in turn allows us to state Riemann-Roch for curves in the language of Čech cohomology. This approach is what can be generalized to other types of algebraic varieties.

In the above section we discussed homology theory. To some extent, there is a dual theory called cohomology. It too is a measure of the non-exactness of a complex. We will not be concerned with the relation between homologies and cohomologies, but will instead just explicitly define the Čech cohomology of an invertible sheaf $\mathcal{L}$ on an algebraic variety X.[2]

Start with a finite open affine cover $\mathcal{U} = \{U_i\}$ of X, for $i = 1, \ldots, N$. For any collection $0 \le i_0 < i_1 < \cdots < i_p \le N$, let

$$U_{i_0 i_1 \cdots i_p} = U_{i_0} \cap U_{i_1} \cap \cdots \cap U_{i_p}.$$

We know that $\mathcal{L}(U_{i_0 i_1 \cdots i_p})$ is isomorphic to a rank-one $\mathcal{O}_X(U_{i_0 i_1 \cdots i_p})$-module. Then for each p, define

$$\mathcal{C}^p(\mathcal{U}, \mathcal{L}) = \prod_{(0 \le i_0 < i_1 < \cdots < i_p \le N)} \mathcal{L}(U_{i_0 i_1 \cdots i_p}).$$

[2] This whole section is heavily influenced by Chapter III.4 in Hartshorne [**Har77**].

We want to define a map

$$d : \mathcal{C}^p(\mathcal{U}, \mathcal{L}) \to \mathcal{C}^{p+1}(\mathcal{U}, \mathcal{L})$$

such that

$$d \circ d : \mathcal{C}^p(\mathcal{U}, \mathcal{L}) \to \mathcal{C}^{p+2}(\mathcal{U}, \mathcal{L})$$

is the zero map, which allows us to form a complex whose exactness we can measure. Following notation in Hartshorne [**Har77**], let $\alpha \in \mathcal{C}^p(\mathcal{U}, \mathcal{L})$. This means that $\alpha = (\alpha_{i_0 i_1 \cdots i_p})$. To define $d(\alpha)$ we need to specify, for each $(p+2)$-tuple $(i_0, i_1, \cdots, i_{p+1})$ with $0 \leq i_0 < i_1 < \cdots < i_{p+1} \leq N$, what the element $d(\alpha)_{i_0 i_1 \cdots i_{p+1}}$ should be. We set

$$d(\alpha)_{i_0 i_1 \cdots i_{p+1}} = \sum_{k=0}^{p+1} (-1)^k \alpha_{i_0 i_1 \cdots \widehat{i_k} \cdots i_{p+1}},$$

where the $\widehat{i_k}$ means that we delete the i_k term. Here $\alpha_{i_0 i_1 \cdots \widehat{i_k} \cdots i_{p+1}}$ stands for the restriction map

$$r_{U_{i_0 i_1 \cdots \widehat{i_k} \cdots i_{p+1}}, U_{i_0 i_1 \cdots i_k \cdots i_{p+1}}},$$

which exists since $\mathcal{L}$ is a sheaf.

In order to make this a bit more concrete, suppose that $\mathcal{U}$ consists of just three open sets U_0, U_1, U_2.

Exercise 6.7.1. Using

$$\begin{aligned} \mathcal{C}^0(\mathcal{U}, \mathcal{L}) &= \mathcal{L}(U_0) \times \mathcal{L}(U_1) \times \mathcal{L}(U_2) \\ \mathcal{C}^1(\mathcal{U}, \mathcal{L}) &= \mathcal{L}(U_{01}) \times \mathcal{L}(U_{02}) \times \mathcal{L}(U_{12}) \\ \mathcal{C}^2(\mathcal{U}, \mathcal{L}) &= \mathcal{L}(U_{012}), \end{aligned}$$

show that

$$d \circ d : \mathcal{C}^0(\mathcal{U}, \mathcal{L}) \to \mathcal{C}^2(\mathcal{U}, \mathcal{L})$$

is the zero map.

Exercise 6.7.2. Let $\alpha = (\alpha_0, \alpha_1, \alpha_2) \in \mathcal{C}^0(\mathcal{U}, \mathcal{L})$ be an element such that $d(\alpha) = 0$. Show that there must be a single element of $\mathcal{L}(X)$ that restricts to α_0 on the open set U_0, to α_1 on the open set U_1 and to α_2 on the open set U_2. This is why we say that something in the kernel of d acting on $\mathcal{C}^0(\mathcal{U}, \mathcal{L})$ defines a global section of the sheaf.

We return to the more general situation. Now that we have a definition for the map d, we have a complex

$$0 \to \mathcal{C}^0(\mathcal{U}, \mathcal{L}) \to \cdots \to \mathcal{C}^N(\mathcal{U}, \mathcal{L}) \to 0,$$

where the first map $0 \to \mathcal{C}^0(\mathcal{U}, \mathcal{L})$ just sends 0 to the zero element of $\mathcal{C}^0(\mathcal{U}, \mathcal{L})$ and the last map $\mathcal{C}^N(\mathcal{U}, \mathcal{L}) \to 0$ sends everything in $\mathcal{C}^N(\mathcal{U}, \mathcal{L})$ to zero.

Definition 6.7.1. The p-th *Čech cohomology group* for the sheaf $\mathcal{L}$ with respect to the open cover $\mathcal{U}$ is

$$H^p(\mathcal{U}, \mathcal{L})$$
$$= \left(\ker(d : \mathcal{C}^p(\mathcal{U}, \mathcal{L}) \to \mathcal{C}^{p+1}(\mathcal{U}, \mathcal{L}))/\mathrm{Im}(d : \mathcal{C}^{p-1}(\mathcal{U}, \mathcal{L}) \to \mathcal{C}^p(\mathcal{U}, \mathcal{L}))\right).$$

Thus Čech cohomology is a measure of the failure of exactness for the complex $0 \to \mathcal{C}^0(\mathcal{U}, \mathcal{L}) \to \cdots \to \mathcal{C}^N(\mathcal{U}, \mathcal{L}) \to 0$. This is highly dependent on the choice of open cover $\mathcal{U}$. If this choice really mattered, then Čech cohomology would not be that useful. Luckily, if each of the open sets $U_i \in \mathcal{U}$ is affine, we will always find that the Čech cohomology groups are isomorphic. (See Hartshorne [**Har77**], III.4.5, though if you go to this source directly from this section, it will be rough going, or see Griffiths and Harris [**GH94**], Chapter 0, Section 3, which is still not a walk in the park.)

One final theoretical point: It is the case that if D_1 and D_2 are linearly equivalent divisors on X, then the corresponding Čech cohomology groups must be isomorphic. This is usually written as

Theorem 6.7.3. If $D_1 \sim D_2$ for divisors on X, then for all d, we have

$$H^d(X, \mathcal{L}_{D_1}) = H^d(X, \mathcal{L}_{D_2}).$$

We do not prove this but will have some exercises showing this property. Recall in an earlier exercise that divisors up to linear equivalence on projective space $\mathbb{P}^r$ are classified by degree. It is common to replace $\mathcal{L}_D$, for a divisor D of degree n on $\mathbb{P}^r$, by the notation

$$\mathcal{O}(n).$$

Thus people frequently consider the Čech cohomology groups

$$H^d(\mathbb{P}^r, \mathcal{O}(n)),$$

which equal $H^d(\mathbb{P}^r, \mathcal{L}_D)$ for any divisor D of degree n.

We spend some time on $\mathbb{P}^1$. Let $(x_0 : x_1)$ be homogeneous coordinates on $\mathbb{P}^1$. There is a natural open cover $\mathcal{U} = \{U_0, U_1\}$ by setting

$$\begin{aligned} U_0 &= \{(x_0 : x_1) : x_0 \neq 0\} \\ U_1 &= \{(x_0 : x_1) : x_1 \neq 0\}. \end{aligned}$$

On U_0, let $s = \frac{x_1}{x_0}$ and on U_1, let $t = \frac{x_0}{x_1}$. On the overlap $U_0 \cap U_1$ we have

$$s = \frac{1}{t}.$$

Now consider the divisor $D = 2(1 : 0)$.

Exercise 6.7.4. Show that $D \cap U_0$ is described by $\mathrm{V}(s^2)$ and that $D \cap U_1$ is described by $\mathrm{V}(1)$ (which is a fancy way of writing the empty set). Show that $2(1 : 0)$ has an equivalent description as $\left\{(U_0, s^2), (U_1, 1)\right\}$.

Exercise 6.7.5. Use the notation from the above problem. Using that

$$\mathcal{L}_D(U) = \{f(s) \in \mathbb{C}(s) : (\mathrm{div}(f) + D) \cap U \geq 0\},$$

show that

$$\begin{aligned} \mathcal{L}_{2(1:0)}(U_0) &= \left\{ \frac{a_0 + a_1 s + \cdots + a_n s^n}{s^2} : a_0, \ldots, a_n \in \mathbb{C}, n \geq 0 \right\} \\ \mathcal{L}_{2(1:0)}(U_1) &= \left\{ b_0 + b_1 t + \cdots + b_m t^m : b_0, \ldots, b_m \in \mathbb{C}, m \geq 0 \right\}. \end{aligned}$$

On the overlap $U_{01} = U_0 \cap U_1$, we will write the restriction maps as

$$r_{U_0, U_{01}}(f(s)) = f(s)$$

and

$$r_{U_1, U_{01}}(g(t)) = g\left(\frac{1}{s}\right).$$

Exercise 6.7.6. Show that

$$d : \mathcal{C}^0(\mathcal{U}, \mathcal{L}_{2(1:0)}) \to \mathcal{C}^1(\mathcal{U}, \mathcal{L}_{2(1:0)})$$

is given by

$$d\left(\frac{a_0 + a_1 s + \cdots + a_n s^n}{s^2}, b_0 + b_1 t + \cdots + b_m t^m \right)$$

$$= b_0 + \frac{b_1}{s} + \cdots + \frac{b_m}{s^m} - \frac{a_0}{s^2} - \frac{a_1}{s} - a_2 - a_3 s + \cdots - a_n s^{n-2}.$$

Exercise 6.7.7. Show that

$$\left(\frac{a_0 + a_1 s + \cdots + a_n s^n}{s^2}, b_0 + b_1 s + \cdots + b_m t^m \right)$$

is in the kernel of the map d if and only if $a_k = 0$ and $b_k = 0$ for $k > 2$ and $a_0 = b_2, a_1 = b_1, a_2 = b_0$.

Exercise 6.7.8. Based on the previous exercise, explain why we can consider $H^0(\mathbb{P}^1, \mathcal{L}_{2(1:0)})$ as the set of all 2 degree homogeneous polynomials in x_0 and x_1, or in other words

$$H^0(\mathbb{P}^1, \mathcal{L}_{2(1:0)}) = \{ax_0^2 + bx_0x_1 + cx_1^2 : a, b, c \in \mathbb{C}\}.$$

Exercise 6.7.9. By similar reasoning, show that for all $d > 0$, we have

$$H^0(\mathbb{P}^1, \mathcal{L}_{d(1:0)}) = \{b_d x_0^d + b_{d-1} x_0^{d-1} x_1 + \cdots + b_0 x_1^d : a_k \in \mathbb{C}\}.$$

(This problem requires you to generalize the last five exercises. Thus it will take a bit to write up.)

Exercise 6.7.10. Show that

$$H^0(\mathbb{P}^1, \mathcal{L}_{-2(1:0)}) = 0.$$

(This involves showing that $\mathcal{L}_{-2(1:0)}(\mathbb{P}^1)$ is empty.)

Exercise 6.7.11. By similar reasoning, show that for all $d > 0$, we have

$$H^0(\mathbb{P}^1, \mathcal{L}_{-d(1:0)}) = 0.$$

The next step in the development of Čech cohomology for divisors would be to put the Riemann-Roch Theorem into this language. We will simply state the theorem:

Theorem 6.7.12 (Riemann-Roch Theorem)**.** Let X be a smooth curve and let D be a divisor on X. Then

$$\dim\, H^0(X, \mathcal{L}_D) - \dim\, H^1(X, \mathcal{L}_D) = \deg(D) + 1 - g.$$

The right-hand side is exactly what we had in Chapter 3. The key is showing that the left-hand side is equivalent to what we had earlier. Thus we would need to show that

$$l(D) = \dim H^0(X, \mathcal{L}_D),$$

which is not that hard, and

$$l(K - D) = \dim H^1(X, \mathcal{L}_D),$$

which does take work.

As the above is true only for curves, this is only the beginning. For example, there is a Riemann-Roch for surfaces:

Theorem 6.7.13 (Riemann-Roch for Surfaces)**.** Let X be a smooth projective surface and let D be a divisor on X. Then

$$\begin{aligned}&\dim H^0(X, \mathcal{L}_D) - \dim H^1(X, \mathcal{L}_D) + \dim H^2(X, \mathcal{L}_D)\\ &= \left(\frac{D \cdot D - D \cdot K}{2}\right) + 1 + p_a.\end{aligned}$$

The right-hand side means the following. Since in general divisors are linear combinations of codimension one subvarieties, divisors on surfaces are curves. The $D \cdot D$ denotes the intersection number of D with itself (such numbers have to be carefully defined). The divisor K is the surface analogue of the canonical divisor; thus $D \cdot K$ is the intersection number of the curves D and K. The p_a is something called the arithmetic genus.

The left-hand side, namely the alternating sum of the various dimensions of the Čech cohomology groups, is called the Euler characteristic of the divisor. In general, we have:

Definition 6.7.2. For a smooth projective variety X of dimension n, the *Euler characteristic* of a divisor D is

$$\chi(D) = \sum_{i=0}^{n} (-1)^k \dim H^k(X, \mathcal{L}_D).$$

All generalizations of Riemman-Roch have the form

$$\chi(D) = \text{some formula capturing geometry and topology.}$$

In this section, we saw how computations of Čech cohomology on $\mathbb{P}^1$ came down to the manipulation of polynomials, which is precisely how we started this book. The power of this section's machinery lies in how many different areas of mathematics (even those far from the joys of polynomial manipulation) can be recast and informed by the language of cohomology. For example, much of the work in algebraic geometry in the last part of the 20th century was developing the correct generalizations of Riemann-Roch. We predict mathematicians in the 21st century will continue this path, but now with an emphasis on the correct generalizations of cohomology theories. (For the expert, we are thinking "motives.") To the student, you are now on the cusp of the beginnings of current algebraic geometry.

Bibliography

[AG06] Avner Ash and Robert Gross. *Fearless Symmetry. Exposing the Hidden Patterns of Numbers*. Princeton University Press, Princeton, New Jersey, 2006. With a foreword by Barry Mazur.

[Arr06] Enrique Arrondo. Another elementary proof of the nullstellensatz. *American Mathematical Monthly*, 113:169–171, February 2006.

[AZ07] A. V. Akopyan and A. A. Zaslavsky. *Geometry of Conics*, volume 26 of *Mathematical World*. American Mathematical Society, Providence, RI, 2007. Translated from the 2007 Russian original by Alex Martsinkovsky.

[Bix98] Robert Bix. *Conics and Cubics. A Concrete Introduction to Algebraic Curves*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1998.

[Bum98] Daniel Bump. *Algebraic Geometry*. World Scientific Publishing Co. Inc., River Edge, NJ, 1998.

[CLO07] David Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, 3rd edition, 2007.

[DF03] David Dummit and Richard Foote. *Abstract Algebra*. Wiley, Hoboken, New Jersey, 3rd edition, 2003.

[Fis01] Gerd Fischer. *Plane Algebraic Curves*, volume 15 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2001. Translated from the 1994 German original by Leslie Kay.

[Fowler04] Kristine K. Fowler, editor. *Using the Mathematics Literature*, volume 66 of *Books in Library and Information Science*. Marcel Dekker Inc., New York, 2004. Includes recommended resources in algebraic and differential geometry by Thomas Garrity.

[Ful69] William Fulton. *Algebraic Curves. An Introduction to Algebraic Geometry*. Mathematics Lecture Notes Series. W. A. Benjamin, Inc., New York-Amsterdam, 1969. Notes written with the collaboration of Richard Weiss.

[GH94] Phillip Griffiths and Joseph Harris. *Principles of Algebraic Geometry*. Wiley Classics Library. John Wiley & Sons, New York, 1994. Reprint of the 1978 original.

[Gib98] C. G. Gibson. *Elementary Geometry of Algebraic Curves. An Undergraduate Introduction*. Cambridge University Press, Cambridge, 1998.

[Har77] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1977.

[Har95] Joe Harris. *Algebraic Geometry. A First Course*, volume 133 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Heidelberg. Corrected reprint of 1992 original edition, 1995.

[Has07] Brendan Hassett. *Introduction to Algebraic Geometry*. Cambridge University Press, Cambridge, 2007.

[Her75] I. N. Herstein. *Topics in Algebra*. Wiley, Hoboken, New Jersey, 2nd edition, 1975.

[Hol12] Audun Holme. *A Royal Road to Algebraic Geometry*. Springer, Heidelberg, 2012.

[Hul03] Klaus Hulek. *Elementary Algebraic Geometry*, volume 20 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2003. Translated from the 2000 German original by Helena Verrill.

[Hus87] Dale Husemöller. *Elliptic Curves*, volume 111 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1987. With an appendix by Ruth Lawrence.

[Ken] Keith Kendig. *Conics*, volume 29 of *The Dolciani Mathematical Expositions*, Mathematical Association of America, Washington, DC, 2005.

[Kir92] Frances Kirwan. *Complex Algebraic Curves*, volume 23 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1992.

[Kun05] Ernst Kunz. *Introduction to Plane Algebraic Curves*. Birkhäuser Boston Inc., Boston, MA, 2005. Translated from the 1991 German edition by Richard G. Belshoff.

[Mir95] Rick Miranda. *Algebraic Curves and Riemann Surfaces*, volume 5 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1995.

[Mum95] David Mumford. *Algebraic Geometry I. Complex Projective Varieties*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1976 edition.

[Mum99] David Mumford. *The Red Book of Varieties and Schemes*, volume 1358 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, expanded edition, 1999. Includes the Michigan lectures (1974) on curves and their Jacobians. With contributions by Enrico Arbarello.

[Per08] Daniel Perrin. *Algebraic Geometry. An Introduction*. Universitext. Springer-Verlag London Ltd., London, 2008. Translated from the 1995 French original by Catriona Maclean.

[Rei88] Miles Reid. *Undergraduate Algebraic Geometry*, volume 12 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1988.

[Sha94a] Igor R. Shafarevich. *Basic Algebraic Geometry. 1. Varieties in Projective Space*. Springer-Verlag, Berlin, 2nd edition, 1994. Translated from the 1988 Russian edition and with notes by Miles Reid.

[Sha94b] Igor R. Shafarevich. *Basic Algebraic Geometry. 2. Schemes and Complex Manifolds*. Springer-Verlag, Berlin, 2nd edition, 1994. Translated from the 1988 Russian edition by Miles Reid.

[Sil86] Joseph H. Silverman. *The Arithmetic of Elliptic Curves*, volume 106 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1st edition, 1986.

[SKKT00] Karen E. Smith, Lauri Kahanpää, Pekka Kekäläinen, and William Traves. *An Invitation to Algebraic Geometry*. Universitext. Springer-Verlag, New York, 2000.

[Uen99] Kenji Ueno. *Algebraic Geometry. 1. From Algebraic Varieties to Schemes*, volume 185 of *Translations of Mathematical Monographs*. Iwanami Series in Modern Mathematics. American Mathematical Society, Providence, RI, 1999. Translated from the 1997 Japanese original by Goro Kato.

[Uen01] Kenji Ueno. *Algebraic Geometry. 2. Sheaves and Cohomology*, volume 197 of *Translations of Mathematical Monographs*. Iwanami Series in Modern Mathematics. American Mathematical Society, Providence, RI, 2001. Translated from the 1997 Japanese original by Goro Kato.

Index

Published Titles in This Subseries

66 **Thomas Garrity, Richard Belshoff, Lynette Boos, Ryan Brown, Carl Lienert, David Murphy, Junalyn Navarra-Madsen, Pedro Poitevin, Shawn Robinson, Brian Snyder, and Caryn Werner,** Algebraic Geometry, 2013

63 **María Cristina Pereyra and Lesley A. Ward,** Harmonic Analysis, 2012

58 **Álvaro Lozano-Robledo,** Elliptic Curves, Modular Forms, and Their L-functions, 2011

51 **Richard S. Palais and Robert A. Palais,** Differential Equations, Mechanics, and Computation, 2009

49 **Francis Bonahon,** Low-Dimensional Geometry, 2009

33 **Rekha R. Thomas,** Lectures in Geometric Combinatorics, 2006

32 **Sheldon Katz,** Enumerative Geometry and String Theory, 2006

7 **Judy L. Walker,** Codes and Curves, 2000

3 **Roger Knobel,** An Introduction to the Mathematical Theory of Waves, 2000

2 **Gregory F. Lawler and Lester N. Coyle,** Lectures on Contemporary Probability, 1999

Algebraic Geometry has been at the center of much of mathematics for hundreds of years. It is not an easy field to break into, despite its humble beginnings in the study of circles, ellipses, hyperbolas, and parabolas.

This text consists of a series of exercises, plus some background information and explanations, starting with conics and ending with sheaves and cohomology. The first chapter on conics is appropriate for first-year college students (and many high school students). Chapter 2 leads the reader to an understanding of the basics of cubic curves, while Chapter 3 introduces higher degree curves. Both chapters are appropriate for people who have taken multivariable calculus and linear algebra. Chapters 4 and 5 introduce geometric objects of higher dimension than curves. Abstract algebra now plays a critical role, making a first course in abstract algebra necessary from this point on. The last chapter is on sheaves and cohomology, providing a hint of current work in algebraic geometry.

ISBN 978-0-8218-9396-8
9 780821 893968
STML/66

For additional information and updates on this book, visit
www.ams.org/bookpages/stml-66