San Jose State University
SJSU ScholarWorks

Master's Theses | Master's Theses and Graduate Research

Spring 2010

Elliptic Curves and Cryptography

Senorina Ramos Vazquez
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Vazquez, Senorina Ramos, "Elliptic Curves and Cryptography" (2010). *Master's Theses*. 3794.
DOI: https://doi.org/10.31979/etd.6fat-tnvm
https://scholarworks.sjsu.edu/etd_theses/3794

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

ELLIPTIC CURVES AND CRYPTOGRAPHY

A Thesis

Presented to

The Faculty of the Department of Mathematics

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Senorina R. Vázquez

May 2010

© 2010

Senorina R. Vázquez

ALL RIGHTS RESERVED

The Designated Thesis Committee Approves the Thesis Titled

ELLIPTIC CURVES AND CRYPTOGRAPHY

by

Senorina R. Vázquez

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

SAN JOSÉ STATE UNIVERSITY

May 2010

Dr. Timothy Hsu	Department of Mathematics
Dr. Marilyn Blockus	Department of Mathematics
Dr. Brian Peterson	Department of Mathematics

ABSTRACT

ELLIPTIC CURVES AND CRYPTOGRAPHY

by Senorina R. Vázquez

In this expository thesis we study elliptic curves and their role in cryptography. In doing so we examine an intersection of linear algebra, abstract algebra, number theory, and algebraic geometry, all of which combined provide the necessary background. First we present background information on rings, fields, groups, group actions, and linear algebra. Then we delve into the structure and classification of finite fields as well as construction of finite fields and computation in finite fields. We next explore logarithms in finite fields and introduce the Diffie-Hellman key exchange system. Subsequently, we take a look at the projective and affine planes and we examine the action of the general linear group of degree 3 (over K) on the points of the projective plane $P^2(K)$. We then explore the geometry of the projective plane with Desargues Theorem. Next, we study conics, quadratic forms, and methods of counting intersection of curves. Finally, we study forms of degree 3 and we are able to explore cubics and the group law on an elliptic curve which leads us to our ultimate goal of examining the role of elliptic curves in cryptography.

DEDICATION

I dedicate this work to my husband, Arturo, and three children: Clarissa, Isabelle, and Aarón. They endured many days and nights of neglect from me all in the name of this work, so I dedicate it to them with much love, respect, and admiration in the hope that it instills in my children the drive to pursue higher education. I also dedicate this work to my many nieces and nephews so that they may follow in my footsteps and dare to dream bigger. In particular, I dedicate this work to David whom graduates from high school this same year and begins his journey through college. I hope my accomplishment teaches him that there are no limits to the possibilities, and that it reminds him, my children, nieces, and nephews of those famous words of encouragement, "Querér es Podér."

ACKNOWLEDGEMENTS

I want to thank my husband, Arturo, whom I admire for his never-ending patience, support, and encouragement; for not telling me it was too late to go after this dream; and for reading my thesis and listening to me talk about it constantly. To my sister Sandra, and friends Blanca and Belinda, thanks for the "therapy" sessions that helped me remain sane and focused as a student, wife and mother. Leti, I thank you for planting the seed and then for your help and encouragement throughout this journey. Dr. Hsu, thanks for all the "red"! That is, thanks for pushing me to produce something to be proud of and as close to perfect as possible. Your guidance, patience, encouragement, and support were indispensable in nudging me forward one section at a time.

Gracias, familia, por apoyarme y por no decirme que era muy tarde para estudiar otra Maestria. Mil gracias a mi mami quien me ayuda mucho en mi hogar para que yo pueda seguir adelante con mis estudios. Mil gracias a mi papi que suele ser el companero de mis hijos cuando me ocupo con mi trabajo. Sin ustedes ninguna parte de esta maestria, ni de esta tesis, hubiera sido posible.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

Figure

CHAPTER 1

INTRODUCTION

1.1 Motivation

According to Koblitz [Kob06], cryptography is the study of methods that allow us to send messages in disguise so that only the intended receiver is able to read them. Thus in cryptography one examines cryptosystems, which consist of the enciphering functions and deciphering functions, as well as the messages they encipher and decipher. Historically, much of this study focused on private key cryptosystems where the sender and receiver agreed on private keys for sending messages and receiving messages, and was primarily used for military and diplomatic reasons. However, with these cryptosystems, anyone who knew enough to decipher messages could not only "break a code" but also determine the enciphering key. Enciphering and deciphering were considered equivalent sciences in a cryptosystem until the 1970's, when Whitfield Diffie and Martin Hellman invented public key cryptography.

In a public key cryptosystem someone who has the ability to encipher cannot feasibly use the enciphering key to derive the deciphering key. That is, in this case the function used to encipher (called a "trapdoor function") is not invertible. A famous cryptosystem that incorporated one of these functions is the RSA cryptosystem and is based on the tremendous difficulty of factoring large composite

integers whose prime factors are not known. However, the notion of "trapdoor function" is relative to current advances in computer technology. For example, an enciphering function regarded as a trapdoor function in 1994 may lose its trapdoor status in 2014 as new algorithm discoveries surface. Furthermore, public key cryptosystems tend to be slower to implement than private key systems so it is often preferred to use a combination of the two: a public key system to exchange their keys for use with a private key system.

Public key cryptosystems are also advantageous because they allow us to publish a directory of enciphering keys for each user of a cryptosystem with a predetermined algorithm. When someone then wants to send a private message to a specific user they must simply look up their enciphering key and use it with the agreed upon algorithm. Then only the intended recipient will have the deciphering key and be able to read the message. In recent years the applications of cryptography have expanded to include any activity where communication systems (such as the internet) play a role. For example, the need to maintain security and confidentiality with data, records, and electronic financial transactions have led to recent applications of cryptography. In fact, the need for newer and more secure trapdoor functions has merged number theory and algebraic geometry to create a new area of application for elliptic curves in cryptography. More precisely, the theory of elliptic curves over finite fields provides a plethora of flexible opportunities in cryptography.

As wireless connectivity expands and cellular phones evolve to miniature hand-held powerful devices with access to a wide variety of applications, developers need to secure the data that is transmitted to and from the devices without compromising the performance of the device. Those designing applications need to consider the challenges of transmitting an encryption key and its effect on

performance [CER]. For this reason Elliptic Curve Cryptography (ECC) has become extremely appealing because it has the ability to use much smaller keys than RSA to provide the same level of security [Edg06].

As we shall see in chapter 7, an elliptic curve over a finite field is defined as follows.

Definition 1.1.1. Assume K is an algebraically closed field not of characteristic 2 and let $g(x)$ be a cubic with no repeated roots where $g(x) = ax^3 + bx^2 + cx + d$. Let $f(x, y) = y^2 - g(x)$ and $F(X, Y, Z) = ZY^2 - aX^3 - bZX^2 - cZ^2X - dZ^3$ be irreducible cubics so that neither contains a line or a conic. Suppose $E \in K[X, Y, Z]$ is a cubic form defining a nonempty plane curve $C : (E = 0) \subset P^2(K)$. Then the set

$$\begin{aligned} E &= \{(X : Y : Z) \in P^2(K) | F(X, Y, Z) = 0\} \\ &= \{(x, y) \in A^2(K) | f(x, y) = 0\} \cup \{(0 : 1 : 0)\} \end{aligned} \tag{1.1}$$

is called an **elliptic curve**.

An elliptic curve group consists of the points on an elliptic curve coupled with a group law, also defined in chapter 7.

At the heart of every cryptosystem is a computationally difficult to solve mathematical problem. In particular, it is the discrete logarithm problem that is the basis for the security of many cryptosystems and Elliptic Curve Cryptography (ECC) is no exception. ECC relies on the difficulty of solving the discrete logarithm problem for the group of an elliptic curve over some finite field [CER]. Elliptic curves have cryptographic appeal because one can "multiply" a point by a number (as in repeated addition) to produce another point on the curve, but another cannot easily figure out what number was used to multiply. This paper will seek to explain how this is done.

1.2 Outline

We now give a summary of the organization of this paper. In order to fully understand Elliptic Curve Cryptography (ECC) we need to study elliptic curves over finite fields so we begin our journey with a review of finite fields. Then we briefly study the cryptography aspect of this paper and how it relates to finite fields. Our study next turns to the projective and affine planes, since elliptic curves are sometimes better understood in a projective setting. We then begin our study of conics and finally move on to study cubics.

Specifically, in Chapter 2 we present background information on rings, fields, groups, group actions and linear algebra. In Chapter 3 we delve into the structure and classification of finite fields as well as construction of finite fields and computation in finite fields. In Chapter 4 we explore logarithms in finite fields and introduce the Diffie-Hellman key exchange system. Chapter 5 begins with a look at the projective and affine planes then we examine the action of the general linear group of degree three (over K) on the points of $P^2(K)$, a projective plane. Finally, in Chapter 5 we explore the connections between the projective and affine planes with Desargues Theorem.

Chapters 6 and 7 specifically deal with conics and cubics. In Chapter 6 we begin with affine and projective conics, then move into quadratic forms and methods of counting intersection of curves. Finally, in Chapter 7 we study forms of degree 3 and we are able to explore cubics and the group law of an elliptic curve which leads us to our final goal of examining the role of elliptic curves in cryptography.

CHAPTER 2

BACKGROUND

2.1 Rings and Fields

In this section, we follow Gallian [Gal06] and Fraleigh [Fra03].

Definition 2.1.1. The **characteristic** of a ring R is defined as the smallest integer n such that $nr = 0$ for any $r \in R$. If no such integer exists, then R is said to be of characteristic 0 and is denoted by $\text{char}\, R = 0$.

We need to define a derivative without using the concept of limits.

Definition 2.1.2. Let F be a field, and let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \tag{2.1}$$

belong to $F[x]$. The **derivative** of $f(x)$, denoted by $f'(x)$, is the polynomial

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + ... + 2a_2 x + a_1 \tag{2.2}$$

in $F[x]$.

Using Definition 2.1.2 we can establish the following properties.

Lemma 2.1.3. *Let $f(x), g(x) \in F[x]$ and let $a \in F$. Then*

(1) $(f(x) + g(x))' = f'(x) + g'(x)$,

(2) $(af(x))' = af'(x)$,

(3) $(f(x)g(x))' = f(x)g'(x) + g(x)f'(x)$, *and*

(4) $f'(g(t), h(t)) = f_x(g(t), h(t))g'(t) + f_y(g(t), h(t))h'(t)$. □

Proof. We refer to Chapter 20 in Gallian for a proof [Gal06]. □

We also require a definition of a partial derivative:

Definition 2.1.4. Let F be a field, and let

$$f(x, y) = a_n x^n + a_{n-1} x^{n-1} y + ... + a_1 x y^{n-1} + a_0 y^n \tag{2.3}$$

belong to $F[x, y]$. The **partial derivative** of $f(x, y)$ with respect to x, denoted by $f_x(x, y)$, is the polynomial

$$f_x(x, y) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} y + ... + 2 a_2 x y^{n-2} + a_1 y^{n-1} \tag{2.4}$$

in $F[x, y]$.

Definition 2.1.5. A nonempty subset A of a ring R is an **ideal** of R if

(1) $a - b \in A$ whenever $a, b \in A$, and

(2) ra and ar are in A whenever $a \in A$ and $r \in R$.

Definition 2.1.6. A **prime ideal** A of a commutative ring R is a proper ideal of R such that $a, b \in R$ and $ab \in A$ imply $a \in A$ or $b \in A$. A proper ideal A of a commutative ring R is a **maximal ideal** of R if, whenever B is an ideal of R and $A \subseteq B \subseteq R$, then $B = A$ or $B = R$.

Theorem 2.1.7. *Let R be a ring with unity* 1. *If* 1 *has order n under addition, then* $\operatorname{char} R = n$. *If* 1 *has infinite order under addition, then* $\operatorname{char} R = 0$.

Proof. See Gallian, Chapter 13, Theorem 13.3 for a proof [Gal06]. □

Theorem 2.1.8. *If R is an integral domain, then* $\operatorname{char} R = 0$ *or p, a prime.*

Proof. See Gallian, Chapter 13, Theorem 13.4 for a proof [Gal06]. □

Definition 2.1.9. A **principal ideal domain** is an integral domain R in which every ideal has the form $\langle a \rangle = \{ra | r \in R\}$ for some a in R.

Definition 2.1.10. A field E is an **extension field** of a field F if F is a subfield of E.

Definition 2.1.11. Let E be an extension field of F and let $f(x) \in F[x]$ then E is a **splitting field** of $f(x)$ over F if $f(x)$ can be written as a product of linear factors in $E[x]$ and there is no proper subfield of E in which $f(x)$ can be written as a product of linear factors. A polynomial that can be factored into linear factors in a field F is said to **split** over F.

Next we establish a criterion for multiple zeros of a polynomial over a field F.

Theorem 2.1.12. *A polynomial $f(x)$ over a field F has a multiple zero in some extension E if and only if $f(x)$ and $f'(x)$ have a common factor of positive degree in $E[x]$.*

Proof. From Gallian [Gal06], Chapter 20, Theorem 20.8, if a is a multiple zero of $f(x)$ in some extension E, then there is a $g(x)$ in $E[x]$ such that $f(x) = (x-a)^2 g(x)$. Thus by Lemma 2.1.3 we have

$$f'(x) = (x-a)^2 g'(x) + 2(x-a)g(x) = (x-a) \cdot [(x-a)g'(x) + 2g(x)]. \tag{2.5}$$

Hence $x - a$ is a factor of both $f(x)$ and $f'(x)$ in the extension E of F. Conversely, suppose that $f(x)$ and $f'(x)$ have a common factor of positive degree. Let a be a

zero of the common factor. Then a is a zero of $f(x)$ and $f'(x)$. Since a is a zero of $f(x)$, there is a polynomial $q(x)$ such that $f(x) = (x-a)q(x)$. Then $f'(x) = (x-a)q'(x) + q(x)$ and $0 = f'(a) = q(a)$. Thus, $x-a$ is a factor of $q(x)$ and a is a multiple zero for $f(x)$. □

Lemma 2.1.13. *If F is a field of characteristic p, then $(a+b)^p = a^p + b^p$ for any a and b in F.*

Proof. For proof we refer to Hungerford [Hun97]. □

Theorem 2.1.14. *Let R be a commutative ring with unity and let A be an ideal of R. Then R/A is a field if and only if A is maximal.*

Proof. We refer to Fraleigh, section 27, for a proof [Fra03]. □

Theorem 2.1.15. *Let F be a field. Then $F[x]$ is a principal ideal domain.*

Proof. Refer to Chapter 16, pg 297 in Gallian for a proof [Gal06]. □

Theorem 2.1.16. *Let F be a field. If $f(x) \in F[x]$, then $f(x)$ is reducible if and only if $g(x)$ divides $f(x)$ for some irreducible polynomial $g(x)$ (such that $deg(g) \leq (degf)/2$).*

Proof. By definition, a nonconstant,nonzero polynomial $f(x) \in F[x]$ is reducible over F if and only if $f(x)$ can be expressed as a product $g(x)h(x)$ of two polynomials $g(x)$ and $h(x)$ in $F[x]$, both of lower degree than the degree of $f(x)$. Also, since $deg(f) = deg(g) + deg(h)$, we can't have both $deg(g), deg(h) > (degf)/2$. So, by interchanging $g(x)$ and $h(x)$ if necessary, $g(x)$ divides $f(x)$. □

Theorem 2.1.17. *Let F be a field. If $f(x) \in F[x]$ and deg $f(x) = 2$ or 3, then $f(x)$ is reducible over F if and only if $f(x)$ has a zero in F.*

Proof. Refer to Chapter 17, pg 304 in Gallian for a proof [Gal06]. □

Theorem 2.1.18. *Let F be a field and let $f(x)$ be a nonconstant polynomial in $F[x]$. Then there exists a splitting field for $f(x)$ over F.*

Proof. We refer to Chapter 20, pg 355 in Gallian for a proof [Gal06]. □

Theorem 2.1.19. *Let F be a field and let $f(x) \in F[x]$. Then any two splitting fields of $f(x)$ over F are isomorphic.*

Proof. We refer to Chapter 20, pg 360 in Gallian for a proof [Gal06]. □

2.2 Groups and Group Actions

In this section, we follow Gallian [Gal06] as well as Fraleigh [Fra03].

Definition 2.2.1. The **order** of an element a in an additive group $(G, +)$ is the smallest positive integer n such that $na = 0$. The **order** of an element a in a multiplicative group $(G, \cdot)$ is the smallest positive integer n such that $a^n = 1$.

Definition 2.2.2. Let F be a field. Then $F^* = F \setminus \{0\}$ is the multiplicative group of units in F.

Theorem 2.2.3. (Theorem of Lagrange) *Let G be a finite group and H be a subgroup of G. Then the order of H divides the order of G.*

Proof. We refer to Fraleigh, Section 10, for a proof [Fra03]. □

Corollary 2.2.4. *The order of an element of a finite group divides the order of the group.*

Proof. We refer to Fraleigh, Section 10, for a proof [Fra03]. □

Theorem 2.2.5. *Let G be a finitely generated abelian group. Then G is isomorphic to a direct product of cyclic groups in the form*

$$G \cong \mathbb{Z}_{{p_1}^{a_1}} \oplus \mathbb{Z}_{{p_2}^{a_2}} \oplus \ldots \oplus \mathbb{Z}_{{p_n}^{a_n}} \oplus \mathbb{Z} \oplus \mathbb{Z} \ldots \oplus \mathbb{Z} \tag{2.6}$$

where the p_i are not necessarily distinct primes and the a_i are positive integers. The direct product is unique up to ordering of the factors.

Proof. We refer to Fraleigh, section 11, for a proof [Fra03]. □

Definition 2.2.6. Let X be a set. A **permutation** of X is a bijection from X to itself. The **permutation group** of X, or S_X, is the collection of all permutations of X under the operation of composition.

Definition 2.2.7. An **action** of a group G on a set X is a homomorphism from G to S_X. An equivalent formulation is that an action is

$$\begin{aligned} G \times X &\to X \\ (g, x) &\mapsto gx \end{aligned} \tag{2.7}$$

such that $g(hx) = gh(x)$ for any two elements $g, h \in G$ and any point $x \in X$. We also have to require that $ex = x$ for any $x \in X$, where $e \in G$ is the identity element.

Definition 2.2.8. We say G acts transitively on S if for any $s, t \in S$ there exists a $g \in G$ such that $gs = t$.

Theorem 2.2.9. *G acts transitively on S if and only if there is an $s_0 \in S$ such that for any $s \in S$ there exists a $g \in G$ where $gs = s_0$.*

Proof. The $\Rightarrow$ direction is clear by definition. In the opposite direction, let $s, t \in S$. We know that there exist $g, h \in G$ such that $gs = s_0$ and $ht = s_0$, so $(h^{-1}g)s = h^{-1}(gs) = h^{-1}s_0 = t$. Thus there exists $f = h^{-1}g \in G$ such that $fs = t$ for any $s, t \in S$, so G acts transitively on S. □

2.3 Linear Algebra

The present section will follow Friedberg, Insel, and Spence [SF03].

Theorem 2.3.1. *Let K be a field. Let V and W be vector spaces with subspaces V_1 and W_1 respectively. If $T : V \to W$ is linear, then $T(V_1)$ is a subspace of W and $\{x \in V | T(x) \in W_1\}$ is a subspace of V.*

Proof. By definition $T(V_1) = \{T(x) | x \in V_1\}$. We know $0_V \in V_1$ since V_1 is a subspace, and $T(0_V) = 0_W$ since T is a linear transformation. Thus $0_W \in T(V_1)$. Let $x, y \in V_1$, so $T(x), T(y) \in T(V_1)$. Then $T(x + y) = T(x) + T(y)$ since T is linear, but $x + y \in V_1$ because V_1 is a subspace. Hence $T(x + y) = T(x) + T(y) \in T(V_1)$. Now let $c \in K$ and $x \in V_1$. Then $cx \in V_1$ since V_1 is a subspace. So $T(cx) \in T(V_1)$, but $T(cx) = cT(x)$ because T is a linear transformation. Thus $cT(x) \in T(V_1)$ and $T(V_1)$ is a subspace of W.

Let $Z = \{x \in V | T(x) \in W_1\}$. Since V is a vector space and T is a linear transformation, we know that $T(0_V) = 0_W$. Furthermore, $0_W \in W_1$ since W_1 is a subspace, so $0_V \in Z$. Next let $x_1, x_2 \in Z \subseteq V$. Then $T(x_1), T(x_2) \in W_1$ and since W_1 is a subspace we have $T(x_1) + T(x_2) \in W_1$ which implies that $T(x_1 + x_2) \in W_1$ since T is linear. Hence $x_1 + x_2 \in Z$. Finally, let $c \in K$ and $x \in Z$. Then $cx \in V$ since V is a vector space and $T(x) \in W_1$ by definition of Z. So $cT(x) \in W_1$ since W_1 is a subspace which implies that $T(cx) \in W_1$ since T is linear. Therefore $cx \in Z$ and Z is a subspace of V. □

Lemma 2.3.2. *Let T be an invertible linear transformation from V to W. Then V is finite-dimensional if and only if W is finite-dimensional and $dim(V) = dim(W)$.*

Proof. Refer to Section 2.4, page 101, of Friedberg, Insel, and Spence [SF03]. □

Theorem 2.3.3. *Given two subspaces V and W of a finite dimensional vector space, the subspace $V + W$ is finite-dimensional and*
$dim(V + W) = dim(V) + dim(W) - dim(V \cap W)$.

Proof. Refer to Section 1.6 of Friedberg, Insel, and Spence [SF03]. □

Theorem 2.3.4. *Let V and W be vector spaces over K and suppose that $\{u_1, ..., u_n\}$ is a basis for V. Then for $\{w_1, ..., w_n\}$ in W, there exists exactly one linear transformation $T : V \to W$ such that $T(u_i) = w_i$ for $1 \leq i \leq n$. Furthermore, if $w_1, w_2, ..., w_n$ is also a basis for W, then T is invertible (i.e., T is an isomorphism).*

Proof. Refer to Section 2.1 of Friedberg, Insel, and Spence [SF03]. □

Corollary 2.3.5. *If $\{u_1, ..., u_k\}$ and $\{w_1, ..., w_k\}$ are linearly independent subsets of V, there exists an invertible linear transformation, $T : V \to V$ such that $T(u_i) = w_i$ for $1 \leq i \leq k$.*

Proof. Complete each linearly independent set to a basis of V by Corollary 2(c) to the Replacement Theorem on pg 45 of Friedberg, Insel, and Spence [SF03]. Then by Theorem 2.3.4, there exists a linear transformation, T, that maps each u_i to the corresponding w_i. For the same reason, there exists a linear transformation L, that maps each w_i to the corresponding u_i. Then for $1 \leq i \leq n$, $TL(w_i) = T(u_i) = w_i$ and $LT(u_i) = L(w_i) = u_i$, so $LT = TL = \mathrm{id}$, by uniqueness in Theorem 2.3.4. □

Definition 2.3.6. Suppose that V and W are finite-dimensional vector spaces with ordered bases $\beta = v_1, v_2, ..., v_n$ and $\gamma = w_1, w_2, ..., w_m$, respectively. Let $T : V \to W$ be linear. Then for each $j, 1 \leq j \leq n$, there exist unique scalars $a_{ij} \in K, 1 \leq i \leq m$, such that

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \tag{2.8}$$

for $1 \leq j \leq n$. We call the $m \times n$ matrix A defined by $A_{ij} = a_{ij}$ the **matrix representation of T in the ordered bases β and γ**, and write $A = [T]_\beta^\gamma$. If $V = W$ and $\beta = \gamma$, then we write $A = [T]_\beta$.

Theorem 2.3.7. *Let V, W, and Z be finite-dimensional vector spaces with ordered bases α, β, γ, respectively. Let $T : V \to W$ and $U : W \to Z$ be linear transformations. Then $[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$.*

Proof. See Friedberg, Insel, and Spence, Section 2.3, [SF03]. □

Theorem 2.3.8. *Let V and W be finite-dimensional vector spaces with ordered bases β and γ, respectively. Let $T : V \to W$ be linear. Then T is invertible if and only if $[T]_\beta^\gamma$ is invertible, and furthermore, in that case $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$.*

Proof. See Friedberg, Insel, and Spence, Section 2.4, [SF03]. □

Theorem 2.3.9. *Let A be an $m \times n$ matrix with entries from F. Then the left-multiplication transformation $L_A : F^n \to F^m$, defined by left multiplication by the matrix A on elements of F^n, is linear. Furthermore, if β and γ are the standard ordered bases for F^n and F^m, respectively, then we have $[L_A]_\beta^\gamma = A$.*

Proof. See Friedberg, Insel, and Spence, Section 2.3, [SF03]. □

Definition 2.3.10. The group of all invertible 3 by 3 matrices over a field K is denoted by $GL_3(K)$.

Theorem 2.3.11. *The group of all invertible linear transformations from $K^3 \to K^3$ is isomorphic to $GL_3(K)$.* □

Proof. Let β be a basis for K^3 and let ϕ be the map from the group of all invertible linear transformations to $GL_3(K)$, $T \mapsto [T]_\beta$. Then by Theorem 2.3.8, ϕ is well-defined, and by Theorem 2.3.7, ϕ is a homomorphism. Finally, by Theorem 2.3.9, ϕ is an isomorphism, since $A \mapsto L_A$ is ϕ^{-1}. □

CHAPTER 3

FINITE FIELDS

3.1 The Structure of Finite Fields

In the present section we will follow the presentations of Fraleigh [Fra03], Gallian [Gal06] and Koblitz [Kob06]. Before we move to the classification of finite fields it is necessary to establish that any finite field has order p^n where p is a prime.

Theorem 3.1.1. *Let F be a finite field. Then $|F| = p^n$ for some prime p.*

Proof. The characteristic of F must be either 0 or a prime p, by Theorem 2.1.8. It cannot be 0, because this would imply that F contains a copy of the integers and would not be finite; thus F must have characteristic p. Let $a \in F^*$. Then $ap = 0$; In fact, because p is prime, it is the smallest positive integer that makes this equation true, so the order of a is p by definition. Also, by Theorem 2.2.5 we know that

$$(F, +) \cong \mathbb{Z}_{{p_1}^{a_1}} \oplus \mathbb{Z}_{{p_2}^{a_2}} \oplus \ldots \oplus \mathbb{Z}_{{p_n}^{a_n}} \tag{3.1}$$

for some primes p_i and positive integers a_i. Furthermore, if ${p_i}^{a_i} \neq p$ then there is an $a \in F$ such that $|a| = {p_i}^{a_i}$, which contradicts the earlier stated fact that all elements in F^* have order p. Thus

$$(F, +) \cong \mathbb{Z}_p \oplus \mathbb{Z}_p \oplus \ldots \oplus \mathbb{Z}_p, \tag{3.2}$$

and $\mid F \mid = p^n$. □

Now that we have looked at the structure of the additive group of a finite field we will examine the structure of its multiplicative group.

Theorem 3.1.2. *The multiplicative group of nonzero elements in a finite field is cyclic.*

Proof. Let F be a finite field. By the Fundamental Theorem of Finitely Generated Abelian Groups F^* is isomorphic to a direct product $C_1 \times C_2 \times C_3 \times ... \times C_r$ where each C_i is a multiplicative cyclic group of prime power d_i. Let m be the least common multiple of all the d_i. In particular, $m \leq d_1 d_2 ... d_r$. If $a_i \in C_i$, then $a_i^{d_i} = 1$, so $a_i^m = 1$, since each d_i divides m. Thus for all elements $b \in F^*$, we have $b^m = 1$, so every element of F^* is a zero of the polynomial $x^m - 1$. However, F^* has $d_1 d_2 ... d_r$ elements, while $x^m - 1$ can have at most m zeros in the field F, so $m \geq d_1 d_2 ... d_r$. Hence $m = d_1 d_2 ... d_r$, so the primes involved in the prime powers are distinct and thus relatively prime. Consequently, F^* is cyclic of order m. □

3.2 The Classification of Finite Fields

In this section we refer to Fraleigh [Fra03], Gallian [Gal06], Hungerford [Hun97], and Koblitz [Kob06]. We will show that every element of a finite field F of order p^n is a zero of the polynomial $f(x) = x^{p^n} - x$, and that the set of zeros of $f(x)$ is closed under addition, subtraction, multiplication, and division. These two results are the last ones necessary to finally reach our objective of classifying finite fields.

Theorem 3.2.1. *If F is a field of order p^n, then every element of F is a zero of $f(x) = x^{p^n} - x$.*

Proof. We know that F^* is a cyclic group of order $p^n - 1$. Also, by Corollary 2.2.4, we know that for any element $a \in F^*$, $a^{|F^*|} = a^{p^n - 1} = 1$. But then $a^{p^n} = a$ and

$a^{p^n} - a = 0$. So $f(a) = 0$ for $a \in F^*$, and $f(0) = 0$. Thus every element of F is a zero of $f(x) = x^{p^n} - x$. □

Consequently, a field F of order p^n is the splitting field of $f(x)$ over $\mathbb{Z}_p$, as seen next.

Theorem 3.2.2. *Let E be the splitting field of $f(x) = x^{p^n} - x$ over $\mathbb{Z}_p$. Then the set F of zeros of $f(x)$ in E contains 0 and 1, is closed under addition, subtraction, multiplication, and division (by nonzero elements). Thus F is a subfield of E in which $f(x)$ splits, and therefore $F = E$.*

Proof. Let a and $b \in E$ be zeros of $f(x)$. Note that E has characteristic p since it is an extension of $\mathbb{Z}_p$, which means that $(a+b)^{p^n} = a^{p^n} + b^{p^n}$, by Lemma 2.1.13. Since a and b are zeros of $f(x)$ we know that $a^{p^n} = a$ and $b^{p^n} = b$. Thus we may conclude that $(a+b)^{p^n} = (a+b)$. Then $f(a+b) = (a+b)^{p^n} - (a+b) = (a+b) - (a+b) = 0$ implies that we have closure under addition (and similarly, subtraction). To show closure under multiplication we note that
$f(ab) = (ab)^{p^n} - (ab) = a^{p^n}b^{p^n} - (ab) = ab - (ab) = 0$. Finally, for division, we use again the fact that if a and b are zeros of $f(x)$ then $a^{p^n} = a$ and $b^{p^n} = b$. We can thus conclude that if $b \neq 0$ we have $\frac{a^{p^n}}{b^{p^n}} = \frac{a}{b}$, or in other words, $\frac{a^{p^n}}{b^{p^n}} - \frac{a}{b} = 0$. Hence $f(\frac{a}{b}) = 0$. Also, since $f(1) = 0$ and $f(0) = 0$ we see that $0, 1 \in F$, therefore F is a subfield containing all the zeros of $f(x)$. Also, every element of F is a zero, so F is the smallest possible field containing the zeros of $f(x)$, so by definition of a splitting field it is E itself, the splitting field of $f(x)$ over $\mathbb{Z}_p$. □

We can now prove the main theorem of this section which serves to classify finite fields.

Corollary 3.2.3. *For each prime p and each positive integer n, there is, up to isomorphism, a unique finite field of order p^n.*

Proof. If F is any finite field such that $|F| = p^n$, then F is the splitting field of $f(x) = x^{p^n} - x$ over $\mathbb{Z}_p$, by Theorems 3.2.1 and 3.2.2. So if it exists, F must be unique.

Conversely, if E is the splitting field of $f(x) = x^{p^n} - x$ over $\mathbb{Z}_p$, then by Theorem 3.2.2, E contains a subfield F of order p^n (which is actually E itself). □

In summary, we know that every finite field is of order p^n, and that for each p^n there exists a finite field of that order. Then from Theorem 3.1.1we know that

$$(F, +) \cong \mathbb{Z}_p \oplus \mathbb{Z}_p \oplus \ldots \oplus \mathbb{Z}_p, \tag{3.3}$$

and by Theorem 3.1.2 we know that

$$(F^*, \cdot) \cong \mathbb{Z}_{p^n - 1} \tag{3.4}$$

.

3.3 Construction of Finite Fields

The present section follows Gallian [Gal06] and Fraleigh [Fra03] in its presentation.

Theorem 3.3.1. *If $f(x) \in \mathbb{Z}_p[x]$ is irreducible and of degree k, then $\mathbb{Z}_p[x]/\langle f(x)\rangle$ is a finite field of order p^k.*

Proof. Let $f(x) \in \mathbb{Z}_p[x]$ be an irreducible polynomial. Then the ideal $\langle f(x)\rangle \subseteq \mathbb{Z}_p[x]$. Let I be any ideal of $\mathbb{Z}_p[x]$ such that $\langle f(x)\rangle \subseteq I \subseteq \mathbb{Z}_p[x]$. Since $\mathbb{Z}_p$ is a field, by Theorem 2.1.15 it follows that $\mathbb{Z}_p[x]$ is a principal ideal domain, so there

exists $g(x) \in \mathbb{Z}_p[x]$ such that $I = \langle g(x) \rangle$. Since I contains $\langle f(x) \rangle$ it follows that $f(x) \in I$. That is, $f(x) = g(x)h(x)$, for some $h(x) \in \mathbb{Z}_p[x]$. However, $f(x)$ is irreducible in $\mathbb{Z}_p[x]$ so it follows that either $g(x)$ or $h(x)$ is a unit (a constant). If $g(x)$ is a unit, then $I = \langle g(x) \rangle = \mathbb{Z}_p[x]$. If $h(x)$ is a unit, then it follows that $f(x)$ and $g(x)$ are associates, so $\langle f(x) \rangle = \langle g(x) \rangle = I$. Hence $\langle f(x) \rangle$ is maximal in $\mathbb{Z}_p[x]$. Now by Theorem 2.1.14, since $\langle f(x) \rangle$ is maximal in $\mathbb{Z}_p[x]$ we can conclude that $\mathbb{Z}_p[x]/\langle f(x) \rangle$ is a field.

Now, since $f(x)$ is of degree k, it follows that the cosets in $\mathbb{Z}_p[x]/\langle f(x) \rangle$ can be represented uniquely by polynomials of at most degree $k-1$. Let $a_0 + a_1 x + \ldots + a_{k-1}x^{k-1} \in \mathbb{Z}_p[x]/\langle f(x) \rangle$. Then there are p choices in Z_p for each a_i and there are k places where a_i appears, thus there are a total of p^k possibilities for the polynomial. Hence $\mathbb{Z}_p[x]/\langle f(x) \rangle$ is a finite field of order p^k. □

Example 3.3.2. Constructing a field of 4 elements. Take $\mathbb{Z}_2$ as the base field and the polynomial $f(x) = x^2 + x + 1$. Since $f(x)$ has degree 2, by Theorem 2.1.17 it suffices to show that neither 1 or 0 are zeros of $f(x)$ to prove it is irreducible. However $f(0) = 1$ and $f(1) = 1$, so $f(x)$ is irreducible over $\mathbb{Z}_2$. Now by Theorem 3.3.1, we know that $\mathbb{Z}_2[x]/\langle x^2 + x + 1 \rangle$ is a finite field of 4 elements. The 4 elements are polynomials of degree 1 or less, given by $0 + 0x$, $0 + 1x$, $1 + 0x$, and $1 + 1x$. That is, the elements are $0, 1, x$, and $1 + x$.

Example 3.3.3. Constructing a field of 9 elements. Take $\mathbb{Z}_3$ as the base field and the polynomial $f(x) = x^2 + 1$. Since $f(x)$ is of degree 2, by Theorem 2.1.17 it suffices to show that neither $2, 1$ or 0 are zeros of $f(x)$ to prove it is irreducible. But $f(0) = 1$, $f(1) = 2$, and $f(2) = 2$ so $f(x)$ is irreducible over $\mathbb{Z}_3$. Now by Theorem 3.3.1 we know that $\mathbb{Z}_3[x]/\langle x^2 + 1 \rangle$ is a finite field of 9 elements. The 9 elements are polynomials of degree 1 or less, given by $0 + 0x$, $0 + 1x$, $0 + 2x$,$1 + 0x$,

$1+1x$, $1+2x$,$2+0x$, $2+1x$, and $2+2x$. That is, the elements are $0, 1, 2, x, 2x, 1+1x, 1+2x, 2+1x$, and $2+2x$.

Example 3.3.4. Constructing a field of 8 elements. Take $\mathbb{Z}_2$ as the base field and the polynomial $f(x) = x^3 + x + 1$. Since $f(x)$ is of degree 3, by Theorem 2.1.17 it suffices to show that neither 1 or 0 are zeros of $f(x)$ to prove it is irreducible. However, $f(0) = 1$ and $f(1) = 1$ so $f(x)$ is irreducible over $\mathbb{Z}_2$. Now by Theorem 3.3.1 we know that $\mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle$ is a finite field of 8 elements. The 8 elements are polynomials of degree 2 or less, given by $0, 1, x, x+1, x^2, x^2+1, x^2+x$, and $x^2 + x + 1$.

3.4 Computing in Finite Fields

In this section we will follow Gallian [Gal06] and Fraleigh [Fra03]. We begin by looking at the multiplication tables for $\mathbb{Z}_2[x]/\langle x^2 + x + 1\rangle$ and $\mathbb{Z}_3[x]/\langle x^2 + 1\rangle$.

Table 3.1: A Multiplication Table for the Finite Field $\mathbb{Z}_2[x]/\langle x^2 + x + 1\rangle$

$\times$	0	1	x	$1+x$
0	0	0	0	0
1	0	1	x	$1+x$
x	0	x	$1+x$	1
$1+x$	0	$1+x$	1	x

In Table 3.2 we replace 2 with -1 for convenience and we see that the order of $1+x, 1-x, -1+x$, and $-1-x$ is 8. Thus each of these four elements are generators of the multiplicative group. For example, $(1+x)^2 = -x$, $(1+x)^3 = 1-x$, $(1-x)^4 = -1$, $(1+x)^5 = -1-x$, $(1+x)^6 = x$, $(1+x)^7 = -1+x$ and $(1+x)^8 = 1$. So the order of $1+x$ is 8 and it generates the whole group.

Table 3.2: A Multiplication Table for the Finite Field $\mathbb{Z}_3[x]/\langle x^2+1\rangle$

$\times$	1	-1	x	$-x$	$1+x$	$1-x$	$-1+x$	$-1-x$
1	1	-1	x	$-x$	$1+x$	$1-x$	$-1+x$	$-1-x$
-1	-1	1	$-x$	x	$-1-x$	$-1+x$	$1-x$	$1+x$
x	x	$-x$	-1	1	$-1+x$	$1+x$	$-1-x$	$1-x$
$-x$	$-x$	x	1	-1	$1-x$	$-1-x$	$1+x$	$-1+x$
$1+x$	$1+x$	$-1-x$	$-1+x$	$1-x$	$-x$	-1	1	x
$1-x$	$1-x$	$-1+x$	$1+x$	$-1-x$	-1	x	$-x$	1
$-1+x$	$-1+x$	$1-x$	$-1-x$	$1+x$	1	$-x$	x	-1
$-1-x$	$-1-x$	$1+x$	$1-x$	$-1+x$	x	1	-1	$-x$

CHAPTER 4

DISCRETE LOGS AND THE DIFFIE-HELLMAN KEY EXCHANGE

4.1 Logarithms in Finite Fields

In this section we refer to Gallian [Gal06] and Koblitz [Kob06] and we make use of the following definition:

Definition 4.1.1. Let F be a finite field and $F^* = \langle a \rangle$ with $|F^*| = p^n - 1$. Then $c = a^b$ if and only if the **discrete logarithm** of c to the base a is b, $b = \log_a c$, where $c \in F^*$ and $b \in \mathbb{Z}_{p^n-1}$.

Example 4.1.2. Constructing a field of 32 elements. Take $\mathbb{Z}_2$ as the base field and the polynomial $f(x) = x^5 + x^3 + 1$. Since neither 1 or 0 are zeros of $f(x)$ we know that $f(x)$ has no linear factors over $\mathbb{Z}_2$. By Theorem 2.1.16, if $f(x)$ is reducible then it must be divisible by an irreducible polynomial of degree 2. The only possible polynomials of degree 2 over Z_2 are $x^2, x^2 + 1, x^2 + x$ and $x^2 + x + 1$, and the only irreducible polynomial is the last one since $x^2 + 1 = (x + 1)^2$. Dividing we get that $x^5 + x^3 + 1 = (x^2 + x + 1)(x^3 + x^2 + x) + (x + 1)$, where the non-zero remainder confirms that $f(x)$ is not divisible by a polynomial of degree 2. Therefore $f(x)$ cannot be written as a product of a quadratic and a cubic, nor does it have linear factors, and thus is irreducible over $\mathbb{Z}_2$. Now by Theorem 3.3.1 we know that $\mathbb{Z}_2[x]/\langle x^5 + x^3 + 1\rangle$ is a finite field of 32 elements. The 32 elements are polynomials

of degree 4 or less, of the form $ax^4 + bx^3 + cx^2 + dx + e$, where $a, b, c, d, e \in \mathbb{Z}_2$.

Using Definition 4.1.1, we can create Table 4.1, the table of logarithms to the base x for the group $\mathbb{Z}_2[x]/\langle x^5 + x^3 + 1\rangle$.

Let $F = \mathbb{Z}_2[x]/\langle x^5 + x^3 + 1\rangle$. Then from Table 4.1 we notice that x generates F^*. Moreover, the order of each element must divide 31 by Corollary 2.2.4 since the order of F^* is 31. So each non-identity element is of order 31 and will in turn generate the entire group F^*.

We can also use Table 4.1 to calculate products by adding the logarithms modulo 31, as in the next example.

Example 4.1.3. To compute $(x^2 + x + 1)(x^4 + x^2 + x)$ we find $log_x(x^2 + x + 1) = 22$ and $log_x(x^4 + x^2 + x) = 27$ and add them modulo 31 to get $22 + 27 = 49 \equiv 18(mod 31)$. Then $log_x(g(x)) = 18$ corresponds to the polynomial $g(x) = x^4 + x^3 + 1$ which is thus the product of $(x^2 + x + 1)(x^4 + x^2 + x)$.

4.2 The Diffie-Hellman Key Exchange System

In the present section we refer to the presentation by Koblitz [Kob06]. The Diffie-Hellman method is a means of exchanging secret keys over insecure channels. For example, if Aarón and Clarissa wish to communicate secretly over an insecure channel they may use the Diffie-Hellman method to exchange a secret key over this channel and then use the key to encrypt their messages to each other and thus be able to communicate secretly over the insecure channel. The Diffie-Hellman method enables us to exchange a secret key over an insecure channel where everyone can listen, even our eavesdropper Isabelle, yet it is nearly impossible to construct the secret key if you are an eavesdropper. The security of the Diffie-Hellman Key Exchange depends on the difficulty of solving the discrete logarithm problem for

Table 4.1: A Table of logarithms for the field $\mathbb{Z}_2[x]/\langle x^5 + x^3 + 1\rangle$

x^n, multiplicative	$g(x)$, additive	$log_x(g(x))$
x	x	1
x^2	x^2	2
x^3	x^3	3
x^4	x^4	4
x^5	$x^3 + 1$	5
x^6	$x^4 + x$	6
x^7	$x^3 + x^2 + 1$	7
x^8	$x^4 + x^3 + x$	8
x^9	$x^4 + x^3 + x^2 + 1$	9
x^{10}	$x^4 + x + 1$	10
x^{11}	$x^3 + x^2 + x + 1$	11
x^{12}	$x^4 + x^3 + x^2 + x$	12
x^{13}	$x^4 + x^2 + 1$	13
x^{14}	$x + 1$	14
x^{15}	$x^2 + x$	15
x^{16}	$x^3 + x^2$	16
x^{17}	$x^4 + x^3$	17
x^{18}	$x^4 + x^3 + 1$	18
x^{19}	$x^4 + x^3 + x + 1$	19
x^{20}	$x^4 + x^3 + x^2 + x + 1$	20
x^{21}	$x^4 + x^2 + x + 1$	21
x^{22}	$x^2 + x + 1$	22
x^{23}	$x^3 + x^2 + x$	23
x^{24}	$x^4 + x^3 + x^2$	24
x^{25}	$x^4 + 1$	25
x^{26}	$x^3 + x + 1$	26
x^{27}	$x^4 + x^2 + x$	27
x^{28}	$x^2 + 1$	28
x^{29}	$x^3 + x$	29
x^{30}	$x^4 + x^2$	30
x^{31}	1	31

large numbers and is summarized in the Diffie-Hellman assumption.

Definition 4.2.1. The **Diffie-Hellman Assumption**: Let $G = \langle g \rangle$ be a cyclic group. It is computationally infeasible to compute g^{ab} knowing only g, g^a and g^b.

According to Koblitz [Kob06] it is impractical to use the Diffie-Hellman method solely for transmitting a secret message because it is very slow when compared to a classical cryptosystem such as a shift cypher. So a compromise is made and the Diffie-Hellman method can be used to randomly determine an encoding key for a classical system. We first review the role of discrete logarithms in finite fields and then apply them to the Diffie-Hellman method for determining a key.

Example 4.2.2. Using Definition 4.1.1 and the cyclic group Z_{19}^* with generator 2 we can compute the discrete logarithm of any element in the group. For example, the discrete logarithm of 7 to the base 2 is 6. That is, $2^6 = 7$. Similarly, $\log_2(13) = 5$ since $2^5 = 32 \equiv 13 \pmod{19}$.

We also saw how discrete logarithms can be calculated in Example 4.1.3 in finite fields such as $F = \mathbb{Z}_2[x]/\langle x^5 + x^3 + 1 \rangle$, or more precisely, in F^*.

Now let $G = \langle g \rangle$ be a large cyclic group and let $N = |\langle g \rangle|$. In the Diffie-Hellman method it is assumed that N is public knowledge. Suppose that Aarón and Clarissa want to agree on a key with which to encode their future correspondence so that Isabelle, the evesdropper, cannot decipher their messages.

(1) First, Aarón selects a random integer a between 1 and N and makes public g^a while keeping a secret.

(2) Also, Clarissa selects a random integer c between 1 and N and makes public g^c and keeps c a secret.

(3) Aarón can now use the public g^c and raises it to his secret a power to acquire g^{ac}.

(4) Similarly, Clarissa uses the public g^a raised to her secret c to also acquire g^{ac}.

(5) Isabelle, the third party eavesdropper, is only aware of N, g, g^a and g^c and according to the Diffie-Hellman assumption, it is computationally infeasible to compute g^{ac} given only this information.

(6) Thus, Aarón and Clarissa can agree to encrypt their correspondence to each other with the key g^{ac}.

In this way these two people can agree to use the key g^{ac} as an encryption key for private correspondence.

CHAPTER 5

PROJECTIVE AND AFFINE PLANES

In this chapter we will follow Samuel's [Sam88] and Fulton's [Ful08] presentation of Algebraic Geometry. We will describe the projective and affine plane according to both.

5.1 The Projective Plane

We begin with a well known definition.

Definition 5.1.1. An **incidence structure** is a triple $(\mathcal{P}, \mathcal{L}, \mathcal{I})$, where $\mathcal{P}$ is a set (the **points**), $\mathcal{L}$ is a set (the **lines**), and $\mathcal{I}$ is a relation between $\mathcal{P}$ and $\mathcal{L}$. If $P \in \mathcal{P}$ and $L \in \mathcal{L}$ and $P \sim L$, then we say P is incident with L (P lies on L).

Definition 5.1.2. Let $\mathcal{X}$ be a set and $\mathcal{P} \subseteq \mathcal{X}, \mathcal{L} \subseteq \mathcal{X}$. We say $P \in \mathcal{P}$ and $L \in \mathcal{L}$ are incident if and only if $P \in L$. In this case we use $(\mathcal{P}, \mathcal{L})$, not $(\mathcal{P}, \mathcal{L}, \mathcal{I})$.

We also have [Sam88]:

Definition 5.1.3. A **projective plane** is an incidence structure $(\mathcal{P}, \mathcal{L})$ such that the following axioms are satisfied:

Axiom 5.1.4. *Two distinct points in $\mathcal{P}$ belong to exactly one line $\mathcal{L}$.*

Axiom 5.1.5. *Two distinct lines in $\mathcal{L}$ have exactly one common point.*

Next we define an equivalence relation on K^3.

Definition 5.1.6. Let K be a field and let $\sim$ be the relation on K^3 defined by saying that $(x_0, x_1, x_2) \sim (y_0, y_1, y_2)$ if and only if there is a $\lambda \in K^*$ such that $x_i = \lambda y_i$ for any i.

Theorem 5.1.7. *The relation $\sim$ defined above is an equivalence relation on K^3.*

Proof. (1) Since K is a field, $1 \in K^*$. Thus $x_i = 1y_i$ for any i and $(x_0, x_1, x_2) \sim (x_0, x_1, x_2)$.

(2) Suppose $(x_0, x_1, x_2) \sim (y_0, y_1, y_2)$. Then there exists a $\lambda \in K^*$ such that $x_i = \lambda y_i$ for any i. Then $\lambda^{-1} \in K$ since K is a field and $\lambda \neq 0$ by definition. Then $\lambda^{-1}x_i = \lambda^{-1}\lambda y_i$ for any i. So we have $\lambda^{-1}x_i = y_i$ for any i and therefore $(y_0, y_1, y_2) \sim (x_0, x_1, x_2)$.

(3) Suppose $(x_0, x_1, x_2) \sim (y_0, y_1, y_2)$ and $(y_0, y_1, y_2) \sim (z_0, z_1, z_2)$. Then there exist $\lambda, \kappa \in K^*$ such that $x_i = \lambda y_i$ and $y_i = \kappa z_i$ for any i. So $x_i = \lambda(\kappa z_i) = (\lambda\kappa)z_i$, for any i, where $\lambda\kappa \in K^*$. Thus $(x_0, x_1, x_2) \sim (z_0, z_1, z_2)$. □

Following Fulton [Ful08], we can define $P^2(K)$ as follows:

Definition 5.1.8. $P^2(K)$ is an incidence structure given by:

- Points $[x_0 : x_1 : x_2]$ in $P^2(K)$ are the equivalence classes of $(x_0, x_1, x_2) \in K^3 \setminus \{\mathbf{0}\}$ (where $\mathbf{0} = (0, 0, 0)$) under the equivalence relation $\sim$ defined above. These equivalence classes (or points) in $P^2(K)$ are thus (deleted) 1-dimensional subspaces of K^3 given by

$$\{\lambda(x, y, z) | \lambda \neq 0\} \tag{5.1}$$

for fixed $(x, y, z) \in K^3$, $(x, y, z) \neq (0, 0, 0)$.

- Lines in $P^2(K)$ are defined as (deleted) 2-dimensional subspaces of K^3 of the form

$$\{\lambda_1(x_0, x_1, x_2) + \lambda_2(y_0, y_1, y_2) | \lambda_1, \lambda_2 \in K\} \setminus \{\mathbf{0}\} \tag{5.2}$$

  where (x_0, x_1, x_2) and (y_0, y_1, y_2) are linearly independent vectors.

In this description, $P^2(K)$ consists of all (deleted) lines passing through the origin, so we adopt the convention that a point of $P^2(K)$ "is" a 1-dimensional subspace of K^3 and that a line of $P^2(K)$ "is" a 2-dimensional subspace of K^3.

Let us now verify that $P^2(K)$ satisfies Definition 5.1.3, our definition of a projective plane.

Theorem 5.1.9. *Two distinct points in $P^2(K)$ are contained on a unique line in $P^2(K)$.*

Proof. Let W and V be distinct points in $P^2(K)$. Then V and W represent distinct 1-dimensional subspaces in K^3. So $\dim(V) = \dim(W) = 1$ and $\dim(V \cap W) = 0$ since V and W only intersect at the origin in K^3 by definition. Thus by Theorem 2.3.3 $\dim(V + W) = 2$, so $V + W$ is a 2-dimensional subspace of K^3. Consequently, $V + W$ is a line in $P^2(K)$ containing both V and W.

Now suppose U is any 2-dimensional subspace of K^3 containing V and W. Then by above, $V + W$ is a subspace of U of the same dimension. Thus $U = V + W$. □

Theorem 5.1.10. *Two distinct lines in $P^2(K)$ have exactly one common point in $P^2(K)$.*

Proof. Let V and W be two distinct lines in $P^2(K)$. Then V and W represent two distinct 2-dimensional subspaces of K^3. That is, $\dim(V) = \dim(W) = 2$. Then $\dim(V + W) = 3$ since V and W span a 3-dimensional subspace; otherwise, if we

suppose $\dim(V+W)=2$ it would imply that our subspaces V and W are not distinct, a contradiction. So by Theorem 2.3.3 $\dim(V\cap W)=2+2-3=1$. That is, $V\cap W$ is a 1-dimensional subspace in K^3 and so $V\cap W$ is a point in $P^2(K)$ by definition.

Now suppose U is any 1-dimensional subspace in K^3 contained in both V and W. Then $U\subseteq V\cap W$. So by above, $\dim(V\cap W)=1$, and U is a subspace of $V\cap W$ of the same dimension and thus they must be the same subspace. Hence our 1-dimensional intersection in K^3 is unique which means that our common point in $P^2(K)$ is unique. □

Corollary 5.1.11. *$P^2(K)$ is a projective plane.* □

5.2 The Affine Plane

The Affine Plane can be defined as follows [Ful08]:

Definition 5.2.1. The Affine Plane $A^2(K)$ is the incidence structure given by:

- The points in $A^2(K)$ are 2-tuples of elements of K.
- The lines of $A^2(K)$ are of the form $\{(a,b)+t(c,d)|t\in K\}=L$ for fixed $a,b,c,d\in K$, $(c,d)\neq(0,0)$.

Lemma 5.2.2. *Let $L=\{(a,b)+t(c,d)|t\in K\}$. If we replace (a,b) with any point on L, and replace (c,d) with any non-zero scalar multiple of itself, we still have the same line.*

Proof. Let $L=\{(a,b)+t(c,d)|t\in K\}$, take $(a',b')=(a,b)+t_0(c,d)$, let $(c',d')=\lambda(c,d),\lambda\neq 0$, and let $L'=\{(a',b')+t(c',d')|t\in K\}$. Then

$$\begin{aligned} L' &= \{(a,b)+t_0(c,d)+t\lambda(c,d)|t\in K\} \\ &= \{(a,b)+(t_0+t\lambda)(c,d)|t\in K\}, \end{aligned} \tag{5.3}$$

and since $t_0 + t\lambda$ will run through all values of K, $L' = L$. □

Definition 5.2.3. Two lines in $A^2(K)$ are said to be **parallel** if they do not intersect. We write $L_1 \| L_2$ to indicate that L_1 is parallel to L_2.

Definition 5.2.4. The **line at infinity**, L_∞, in $P^2(K)$ is defined as the set of all equivalence classes of the form $[x : y : 0]$, that is, the set of equivalence classes of all 3-tuples in which the third coordinate is 0.

Definition 5.2.5. Two lines in $P^2(K)$ are said to be **parallel** if they intersect at a point on L_∞.

Let $f : A^2(K) \to P^2(K)$ be defined by:

$$f(x, y) = [x : y : 1]. \tag{5.4}$$

As we shall see (section (3) of Theorem 5.2.6), for every line $L \in A^2(K)$ there exists a line $L' \in P^2(K)$ such that $f(L) = L' \backslash (L' \cap L_\infty)$. Thus we can define the **completion** of $f(L)$ by $\overline{f(L)} = L'$.

Theorem 5.2.6. *Let K be a field. Then for the map f defined in* (5.4)*, we have*

(1) $P^2(K) = f(A^2(K)) \cup L_\infty$.

(2) f *is one to one.*

(3) If L is a line in $A^2(K)$, then for some line L' in $P^2(K)$,
$f(L) = L' \setminus (L' \cap L_\infty)$.

(4) If L is a line in $P^2(K)$, then the preimage $f^{-1}(L)$ is a line in $A^2(K)$.

(5) Let L_1 and $L_2 \in A^2(K)$. Then $L_1 \| L_2$ if and only if $\overline{f(L_2)} \| \overline{f(L_2)}$.

Proof. (1) Let $[x_0 : y_0 : z_0] \in P^2(K)$. If $z_0 \neq 0$ then

$$[x_0 : y_0 : z_0] = \left[\frac{x_0}{z_0} : \frac{y_0}{z_0} : 1\right] = f\left(\frac{x_0}{z_0}, \frac{y_0}{z_0}\right). \tag{5.5}$$

If $z_0 = 0$ then $[x_0 : y_0 : 0] \in L_\infty$. Therefore $P^2(K) \subseteq f(A^2(K)) \cup L_\infty$. Also, by definition of f, $f(A^2(K)) \subseteq P^2(K)$, and we know $L_\infty \subseteq P^2(K)$, so $f(A^2(K)) \cup L_\infty \subseteq P^2(K)$. Thus $P^2 = f(A^2(K)) \cup L_\infty$.

(2) Suppose $f(x_1, y_1) = f(x_2, y_2)$. Then $[x_1 : y_1 : 1] = [x_2 : y_2 : 1]$ and so we can say that for some $\lambda \in K^*$ it is true that $(x_1, y_1, 1) = \lambda(x_2, y_2, 1)$. So $(x_1, y_2, 1) = (\lambda x_2, \lambda y_2, \lambda)$, which implies that $\lambda = 1$. Thus $x_1 = x_2$ and $y_1 = y_2$ and f is one-to-one.

(3) Suppose L is a line in $A^2(K)$; that is,

$$L = \{(a, b) + t(c, d)\} = \{(a + tc, b + td) | t \in K\} \tag{5.6}$$

where $(c, d) \neq (0, 0)$. Then

$$\begin{aligned} f(L) &= \{[a + tc : b + td : 1] | t \in K\} \\ &= \{\lambda(a + tc, b + td, 1) | \lambda \in K^*, t \in K\} \end{aligned} \tag{5.7}$$

since by definition of our equivalence relation it is true that $\lambda \neq 0$. Consequently,

$$f(L) = \{\lambda(a, b, 1) + \lambda t(c, d, 0) | \lambda \in K^*, t \in K\} \tag{5.8}$$

which is a two-dimensional subspace spanned by the linearly independent vectors $(a, b, 1)$ and $(c, d, 0)$, minus all elements of $f(L)$ of the form $k(c, d, 0)$, where $k \in K$.

(4) Suppose L is a line in $P^2(K)$. That is,

$$L = \{\lambda_1(x_0, x_1, x_2) + \lambda_2(y_0, y_1, y_2) | \lambda_1, \lambda_2 \in K\} \backslash \{\mathbf{0}\}, \tag{5.9}$$

where (x_0, x_1, x_2) and (y_0, y_1, y_2) are linearly independent. Since $L \neq L_\infty$, we can assume that at least one of x_2 or y_2 is non-zero. Without loss of generality, assume $x_2 \neq 0$. Then

$$L = \{\lambda_1 \left(\frac{x_0}{x_2}, \frac{x_1}{x_2}, 1\right) + \lambda_2(y_0, y_1, y_2) | \lambda_1, \lambda_2 \in K\} \backslash \{\mathbf{0}\}. \tag{5.10}$$

Let $a = \dfrac{x_0}{x_2}$ and $b = \dfrac{x_1}{x_2}$, $c = y_0 - y_2 a$, $d = y_1 - y_2 b$, and note that $(y_0, y_1, y_2) - y_2(a, b, 1) = (y_0 - y_2 a, y_1 - y_2 b, 0) = (c, d, 0)$. Then $(a, b, 1)$ and $(c, d, 0)$ are linearly independent and

$$L = \{\lambda_1(a, b, 1) + \lambda_2(c, d, 0) | \lambda_1, \lambda_2 \in K\} \backslash \{\mathbf{0}\}. \tag{5.11}$$

Let $L_1 = \{(a, b) + t(c, d)\}$. This implies that

$$\begin{aligned} L \cap f(A^2(K)) &= \{\lambda_1(a, b, 1) + \lambda_2(c, d, 0) | \lambda_1, \lambda_2 \in K, \lambda_1 \neq 0\} \\ &= f(L_1), \end{aligned} \tag{5.12}$$

according to the previous proof. So $L \cap f(A^2(K))$ is the image of L_1 in $A^2(K)$ and this means that the pre-image of $L \cap f(A^2(K))$ is L_1, since f is one to one.

(5) Now suppose that $L_1, L_2 \in A^2(K)$ such that $L_1 \| L_2$. Then since f is one-to one, $f(L_1) \cap f(L_2) = \emptyset$ where $\overline{f(L_1)}$ and $\overline{f(L_2)} \in P^2(K)$. So $\overline{f(L_1)} \cap \overline{f(L_2)} \in L_\infty$ since two lines in $P^2(K)$ must intersect at a point. So $\overline{f(L_1)} \| \overline{f(L_2)}$ by definition of parallel in $P^2(K)$.

Finally, suppose that $\overline{f(L_1)} \| \overline{f(L_2)}$. Then $\overline{f(L_1)} \cap \overline{f(L_2)} = a \in L_\infty$ which implies that $L_1 \cap L_2 = \emptyset$ in $A^2(K)$ and thus $L_1 \| L_2$ in $A^2(K)$. □

Thus we can embed $A^2(K)$ in $P^2(K)$ by defining $f(x, y) = [x : y : 1]$. Then

$$P^2(K) = \{[x : y : 1]\} \cup \{[x : y : 0]\} = f(A^2(K)) \cup L_\infty, \tag{5.13}$$

and we can conclude that the geometry of $P^2(K) \setminus L_\infty$ is the geometry of $A^2(K)$, which allows for some necessary manipulations in our proof of Desargue's Theorem in Section 5.4.

5.3 The Action of $GL_3(K)$ on the points of $P^2(K)$

In this section we apply the definition of group action to the group $GL_3(K)$ and the set $P^2(K)$. (See Definition 2.2.7 for background.)

We shall assume the following convention:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (x, y, z). \tag{5.14}$$

Let A be a 3×3 matrix in $GL_3(K)$ and let $[x : y : z]$ be a point in $P^2(K)$. Then the action of $GL_3(K)$ on the points of $P^2(K)$ is the map

$$GL_3(K) \times P^2(K) \to P^2(K) \tag{5.15}$$

given by

$$A[x : y : z] = A \begin{bmatrix} x \\ y \\ z \end{bmatrix} / \sim . \tag{5.16}$$

We must first verify that this action is well defined. Assume $(x, y, z) \sim (x', y', z')$. Then $(x', y', z') = \lambda(x, y, z)$ for some $\lambda \in K^*$. Then

$$A[\lambda x : \lambda y : \lambda z] = A \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda z \end{bmatrix} = \lambda A \begin{bmatrix} x \\ y \\ z \end{bmatrix} \sim A \begin{bmatrix} x \\ y \\ z \end{bmatrix} . \tag{5.17}$$

So $A(x, y, z) \sim A(x', y', z')$ and the action is well defined.

Theorem 5.3.1. *The action of $GL_3(K)$ on the points of $P^2(K)$ preserves lines; more precisely, for any $A \in GL_3(K)$, if W is a subspace of K^3, then AW is a subspace and $\dim(AW) = \dim W$. So the action of $GL_3(K)$ preserves $P^2(K)$.*

Proof. By Theorem 2.3.1 and Lemma 2.3.2, an invertible transformation preserves subspaces and dimension. So by Theorem 2.3.11, so do the elements of $GL_3(K)$. □

Theorem 5.3.2. *The action of $GL_3(K)$ on the points of $P^2(K)$ is transitive.*

Proof. Suppose $P_1 = \langle u_1 \rangle$ and $P_2 = \langle v_1 \rangle$ are points of $P^2(K)$. Then by Corollary 2.3.5 there exists an invertible linear transformation $T : K^3 \to K^3$ such that $T(u_1) = v_1$. Then since this transformation can be represented by matrix multiplication by Theorem 2.3.11, we have shown that $GL_3(K)$ acts transitively on the points of $P^2(K)$. □

Theorem 5.3.3. *The action of $GL_3(K)$ on $P^2(K)$ is transitive on lines.*

Proof. Let S be a basis for L_1, a two-dimensional subspace of K^3, and let W be a basis for L_2, another two-dimensional subspace of K^3. Then by Corollary 2.3.5 there exists a linear transformation T such that $T(S) = W$, so $T(L_1) = L_2$. Thus $GL_3(K)$ acts transitively on the lines of $P^2(K)$. □

Theorem 5.3.4. *Given (P_i, L_i), $i = 1, 2$, where L_i is a line and P_i is a point not on L_i, there exists $g \in GL_3(K)$ such that $gP_1 = P_2$ and $gL_1 = L_2$.*

Proof. Let $\{u_1, u_2\}$ be a basis for L_1 and let $P_1 = u_3$. Since P_1 is not on L_1 it follows that $S = \{u_1, u_2, u_3\}$ is a linearly independent subset of K^3 and thus is a basis for K^3. Similarly, bases for L_2 and P_2, $\{w_1, w_2\}$ and $\{w_3\}$, together form another basis W, for K^3. Then by Theorem 2.3.4 there exists a linear transformation T such that $T(u_i) = w_i$, so there exists the required $g \in GL_3(K)$ such that $gP_1 = P_2$ and $gL_1 = L_2$. □

Corollary 5.3.5. *Given $(P, L), P$ not on L, there exists $g \in GL_3(K)$ such that $gL = L_\infty$ and $gP = O$, where O is the origin.* □

Theorem 5.3.6. *Given (P_i, L_i), $i = 1, 2$, where L_i is a line and P_i is a point on L_i, there exists $g \in GL_3(K)$ such that $gP_1 = P_2$ and $gL_1 = L_2$.*

Proof. Let $P_1 = \langle u_1 \rangle$ be a nonzero point on L_1, then extend $\{u_1\}$ to a basis of L_1. Similarly, let $P_2 = \langle w_1 \rangle$ be a point on L_2, then extend $\{w_1\}$ to a basis of L_2. Then by Theorem 2.3.5 there exists a linear transformation T such that $T(u_i) = w_i$. □

5.4 Desargues Theorem

This section follows the presentation by Samuel [Sam88] and we use the notation $L_{AA'}$ for the line through A and A' in either $A^2(K)$ or $P^2(K)$.

Definition 5.4.1. For $A, O \in A^2(K)$, $A \neq \mathbf{0}$, the vector $\overrightarrow{OA} = A - O \in K^2$.

Note that $L_{AA'} = \{A + t\overrightarrow{AA'} | t \in K\}$. (See Definition 5.2.1).

Definition 5.4.2. In $A^2(K)$, a **parallelogram** is 4 points A, A', B, B' such that no three points are collinear, $L_{AA'} \| L_{BB'}$, and $L_{AB} \| L_{A'B'}$.

Lemma 5.4.3. *The points O, A, B are collinear with $O \neq A, O \neq B$ if and only if there exists λ such that $\overrightarrow{OA} = \lambda\overrightarrow{OB}$. (Therefore if $\overrightarrow{OA}$ and $\overrightarrow{OB}$ are independent, then O, A, and B are not collinear.)*

Proof. $\Rightarrow$

Let $L = \{O + t(c, d)\}$ be the line through points O, A, and B. Define $A = O + t_1(c, d)$ and $B = O + t_2(c, d)$ such that $t_1, t_2 \neq 0$. Then

$$\overrightarrow{OA} = A - O = O + t_1(c, d) - O = t_1(c, d) \tag{5.18}$$

and

$$\overrightarrow{OB} = B - O = O + t_2(c,d) - O = t_2(c,d). \tag{5.19}$$

Thus

$$\overrightarrow{OA} = \frac{t_1}{t_2}\overrightarrow{OB}, \tag{5.20}$$

so we may take $\lambda = \dfrac{t_1}{t_2}$.

$\Leftarrow$

If $\overrightarrow{OA}$ and $\overrightarrow{OB}$ are dependent then there exists $\lambda \in K, \lambda \neq 0$, such that $\overrightarrow{OA} = \lambda\overrightarrow{OB} = \lambda(B - O) = \lambda B - \lambda O$. So we get $\lambda O + \overrightarrow{OA} = \lambda B$. That is, $O + \dfrac{1}{\lambda}\overrightarrow{OA} = B$, so $B \in L_{OA}$. □

Lemma 5.4.4. *For $A \neq B, A' \neq B'$, the following are equivalent:*

(1) $L_{AB} \| L_{A'B'}$.

(2) $\overrightarrow{AB} = a\overrightarrow{A'B'}$, for some $a \in K, a \neq 0$, and no three of the points A, B, A', and B' are collinear.

Proof. Let $L_{AB} = \{A + t\overrightarrow{AB} | t \in K\}$ and $L_{A'B'} = \{A' + s\overrightarrow{A'B'} | s \in K\}$. There are three cases to consider: The case where $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are independent, then the case where $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are dependent with $L_{AB} \cap L_{A'B'} \neq \emptyset$, and finally the case where $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are dependent with $L_{AB} \cap L_{A'B'} = \emptyset$. In each of the three cases we will show that the truth value of (1) and (2) are the same.

CASE 1: Suppose $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are independent. Then $P \in L_{AB} \cap L_{A'B'}$ if and only if $A + t\overrightarrow{AB} = A' + s\overrightarrow{A'B'}$ for some $s, t \in K$. That is, $t\overrightarrow{AB} - s\overrightarrow{A'B'} = A' - A = \overrightarrow{AA'}$ and since $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are independent they span K^2 so there exists $s, t \in K$ such that $\overrightarrow{AA'} = t\overrightarrow{AB} - s\overrightarrow{A'B'}$. Thus there exists a $P \in L_{AB} \cap L_{A'B'}$ and (1) is false. Furthermore, if $\overrightarrow{AB} = a\overrightarrow{A'B'}$, then $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are dependent and our assumption is contradicted, so (2) must be false also.

CASE 2: Suppose $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are dependent and $L_{AB} \cap L_{A'B'} \neq \emptyset$. Since $L_{AB} \cap L_{A'B'} \neq \emptyset$ we have that (1) is false. Since $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are dependent there exists $a \in K$ such that $\overrightarrow{AB} = a\overrightarrow{A'B'}$. Again we have that $P \in L_{AB} \cap L_{A'B'}$ if and only if $A + t\overrightarrow{AB} = A' + s\overrightarrow{A'B'}$ for some $s, t \in K$. That is, $t\overrightarrow{AB} - s\overrightarrow{A'B'} = A' - A$, or $ta\overrightarrow{A'B'} - s\overrightarrow{A'B'} = A' - A$. This implies that $A = A' + (s - ta)\overrightarrow{A'B'}$ and so $A \in L_{A'B'}$. Thus the three points A, A', B' are collinear and (2) is false.

CASE 3: Suppose $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are dependent and $L_{AB} \cap L_{A'B'} = \emptyset$. Then $L_{AB} \cap L_{A'B'} = \emptyset$ implies (1) is true by definition of parallel. Also, $\overrightarrow{AB} = a\overrightarrow{A'B'}$ is true by definition of dependence. Furthermore suppose (for example) $A' \in L_{AB}$; then $L_{AB} \cap L_{A'B'} \neq \emptyset$ is a contradiction so $L_{AB} \cap L_{A'B'} = \emptyset$ implies that no 3 points are collinear and (2) is true. Hence both conditions have the same truth value in each of the 3 cases thereby proving equivalence. □

Lemma 5.4.5. *The points A, A', B and B' form a parallelogram if and only if $\overrightarrow{AB} = \overrightarrow{A'B'}$ and no three points are collinear.*

Proof. $\Rightarrow$

By Lemma 5.4.4 we know that $\overrightarrow{AB} = a\overrightarrow{A'B'}$ and $\overrightarrow{AA'} = b\overrightarrow{BB'}$, since $L_{AB} \| L_{A'B'}$ and $L_{AA'} \| L_{BB'}$. Also, $\overrightarrow{AB'} = \overrightarrow{AA'} + \overrightarrow{A'B'}$ by Definition 5.4.1, and similarly $\overrightarrow{AB'} = \overrightarrow{AB} + \overrightarrow{BB'}$. Thus $\overrightarrow{AA'} + \overrightarrow{A'B'} = \overrightarrow{AB} + \overrightarrow{BB'}$. By substituting we get $b\overrightarrow{BB'} + \overrightarrow{A'B'} = a\overrightarrow{A'B'} + \overrightarrow{BB'}$. After a little algebra we have $(b-1)\overrightarrow{BB'} + (1-a)\overrightarrow{A'B'} = 0$ and since points A', B, and B' are not collinear we know that $\overrightarrow{BB'}$ and $\overrightarrow{A'B'}$ are linearly independent by Lemma 5.4.3 so it must be that $b - 1 = 0$ and $1 - a = 0$. Thus $a = 1$, $b = 1$ and $\overrightarrow{AB} = \overrightarrow{A'B'}$.

$\Leftarrow$

Suppose $\overrightarrow{AB} = \overrightarrow{A'B'}$. Then $\overrightarrow{AB'} = \overrightarrow{AA'} + \overrightarrow{A'B'} = \overrightarrow{AB} + \overrightarrow{BB'}$ by Definition 5.4.1. Substituting we arrive at $\overrightarrow{AA'} + \overrightarrow{AB} = \overrightarrow{AB} + \overrightarrow{BB'}$. Thus

$\overrightarrow{AA'} = \overrightarrow{BB'}$ and no three points are collinear, which implies that $L_{AA'} \| L_{BB'}$ by Lemma 5.4.4. Similarly, $L_{AB} \| L_{A'B'}$. □

Theorem 5.4.6. *(Desargues Theorem) In $P^2(K)$, let L, L', and L'' be distinct lines having a common point O. If $A, B \in L, A', B' \in L'$, and $A'', B'' \in L''$ are points distinct from one another and from O, then the three intersection points $I = L_{AA'} \cap L_{BB'}, J = L_{AA''} \cap L_{BB''}$, and $K = L_{A'A''} \cap L_{B'B''}$ are collinear.*

Proof. By Corollary 5.3.5 we can assume that L_{IJ} is the line at infinity since there exists a $g \in GL_3(K)$ that would move L_{IJ} to the line at infinity. Then there are two cases to consider: The case where O is on L_∞, and the case when O is not on L_∞.

Case 1: Let L, L', and L'' lie in the projective plane such that O is not on L_{IJ}. By Corollary 5.3.5 we can assume O is the origin in the Affine Plane since there exists a $g \in GL_3(K)$ that would move said point to the origin. In the Affine Plane, we have that O, A and B are collinear, so by Lemma 5.4.3, there exists an $a \in K$ such that

$$\overrightarrow{OB} = a\overrightarrow{OA}. \tag{5.21}$$

Similarly,

$$\overrightarrow{OB'} = a'\overrightarrow{OA'} \tag{5.22}$$

and

$$\overrightarrow{OB''} = a''\overrightarrow{OA''} \tag{5.23}$$

for some $a', a'' \in K$. Then since I is at infinity, $L_{AA'} \| L_{BB'}$ (by Definition 5.2.3), $\overrightarrow{AA'} = \overrightarrow{OA'} - \overrightarrow{OA}$, and $\overrightarrow{BB'} = \overrightarrow{OB'} - \overrightarrow{OB}$, there exists a $c \in K$ such that

$$\overrightarrow{OB} - \overrightarrow{OB'} = c(\overrightarrow{OA} - \overrightarrow{OA'}). \tag{5.24}$$

That is,

$$a\overrightarrow{OA} - a'\overrightarrow{OA'} = c\overrightarrow{OA} - c\overrightarrow{OA'}, \tag{5.25}$$

which implies that $a = a' = c$ since $\overrightarrow{OA}$ and $\overrightarrow{OA'}$ are linearly independent because L, and L' are distinct lines. Similarly, $a = a''$, since J is at infinity and $L_{AA''} \| L_{BB''}$. Then from (5.21), (5.22), and (5.23), we get that

$$\overrightarrow{OB''} - \overrightarrow{OB'} = a(\overrightarrow{OA''} - \overrightarrow{OA'}) \tag{5.26}$$

That is,

$$\overrightarrow{B''B'} = a\overrightarrow{A''A'}, \tag{5.27}$$

which implies that $L_{B''B'} \| L_{A''A'}$ (Lemma 5.4.4) and thus $K \in L_{IJ}$.

Case 2: Let L, L', and L'' lie in $P^2(K)$ such that O is on L_{IJ}. Then by Definition 5.2.5, the lines L, L', and L'' are parallel, since the three lines have the point O in common. Furthermore, since L_{IJ} is the line at infinity we have $I = L_{AA'} \cap L_{BB'} \in L_\infty$, thus $L_{AA'} \| L_{BB'}$. Similarly $J = L_{AA''} \cap L_{BB''} \in L_\infty$ and $L_{AA''} \| L_{BB''}$. Then by using Lemma 5.4.5 on parallelogram $ABB'A'$, we know that $\overrightarrow{AB} = \overrightarrow{A'B'}$, since $L_{AB} \| L_{A'B'}$ and $L \| L'$. Similarly, by using Lemma 5.4.5 on parallelogram $ABB''A''$, we know that $\overrightarrow{AB} = \overrightarrow{A''B''}$, since $L_{AB} \| L_{A''B''}$ and $L \| L''$. Thus $\overrightarrow{AB} = \overrightarrow{A'B'} = \overrightarrow{A''B''}$. Again by using Lemma 5.4.5, $A'B'B''A''$ is a parallelogram so we know that $\overrightarrow{A'B'} = \overrightarrow{A''B''}$ implies $L_{A'A''} \| L_{B'B''}$ so that $K = L_{A'A''} \cap L_{B'B''} \in L_\infty$ as well. Therefore I, J, and K all lie on L_∞.

□

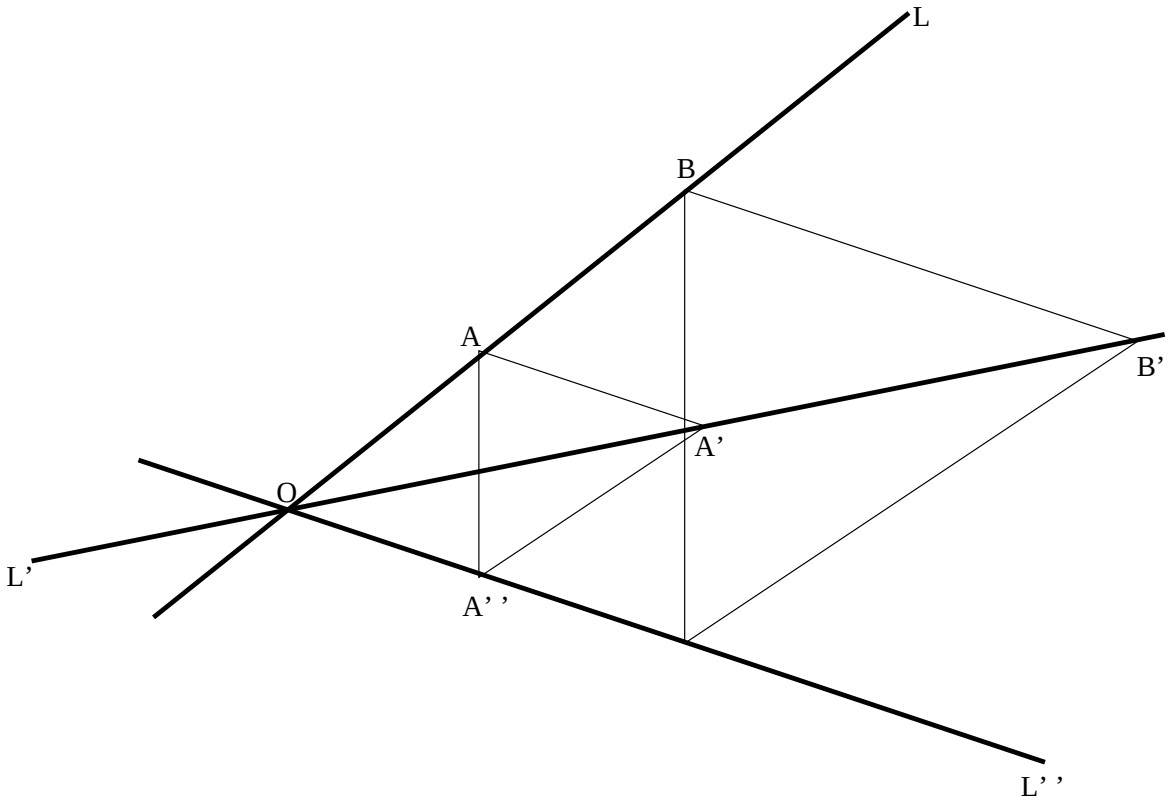

Figure 5.1: Desargues Case 1

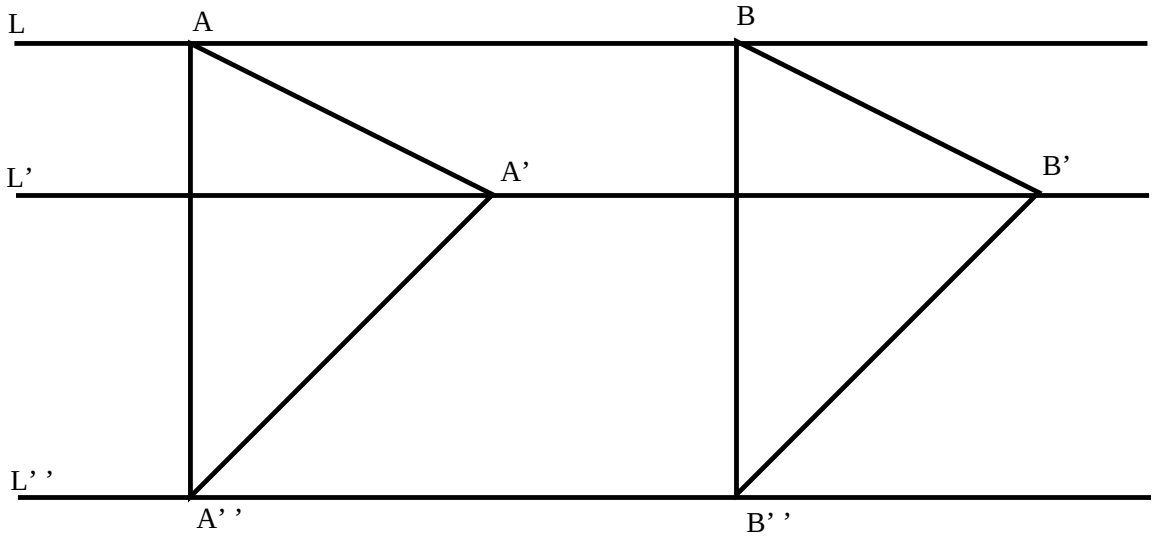

Figure 5.2: Desargues Case 2

CHAPTER 6

CONICS

The present chapter follows Reid's [Rei90] presentation of Conics in Chapter 1 of his book.

6.1 Affine and Projective Conics

Definition 6.1.1. A **homogeneous polynomial** is a polynomial whose monomials with nonzero coefficients all have the same (total) degree.

Definition 6.1.2. Given a quadratic polynomial

$$q(x, y) = ax^2 + bxy + cy^2 + dx + ey + f, \tag{6.1}$$

with coefficients in K, the *corresponding homogeneous polynomial* is

$$Q(X, Y, Z) = aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2, \tag{6.2}$$

where $Q(X, Y, Z) = Z^2 q(\frac{X}{Z}, \frac{Y}{Z})$ and $q(x, y) = Q(x, y, 1)$.

Definition 6.1.3. An affine **conic** C_A in $A^2(K)$ is a curve given by

$$C_A := \{(x, y) | q(x, y) = 0\}, \tag{6.3}$$

where $q(x, y)$ is of the form in (6.1).

Definition 6.1.4. A projective **conic** C_P in $P^2(K)$ is a curve given by

$$C_P := \{(X : Y : Z) | Q(X, Y, Z) = 0\}, \tag{6.4}$$

where $Q(X, Y, Z)$ is of the form in (6.2). The condition $Q(X, Y, Z) = 0$ is well defined on the equivalence class $(X : Y : Z)$ because $Q(\lambda X, \lambda Y, \lambda Z) = \lambda^2 Q(X, Y, Z)$ for any $\lambda \in K$.

The above correspondence gives the "same" curve because of the following theorem.

Theorem 6.1.5. *The image of the affine conic C_A under the map f defined in* (5.4) *is equal to the intersection of the image of the affine plane with the projective conic C_P. That is,*

$$f(C_A) = C_P \cap f(A^2(K)). \tag{6.5}$$

Proof. Suppose $[X : Y : Z] \in f(C_A)$. Then there exists $(x, y) \in A^2(K)$ such that $f(x, y) = [x : y : 1] = [X : Y : Z]$ and $0 = q(x, y) = Q(x, y, 1)$ since (x, y) is in C_A. Therefore $f(C_A) \subseteq C_P \cap A^2(K)$.

Now suppose $[X : Y : Z] \in C_P \cap A^2(K)$. By definition this means that $Z \neq 0$, otherwise the point would be on the line at infinity, and $Q(X, Y, Z) = 0$. Let $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$. Then $q(x, y) = q(\frac{X}{Z}, \frac{Y}{Z}) = \frac{1}{Z^2} Q(X, Y, Z) = 0$ since $Z \neq 0$. So $(x, y) \in C_A$ and $f(x, y) = [x : y : 1] = [\frac{X}{Z} : \frac{Y}{Z} : 1]$. However, $Z(\frac{X}{Z}, \frac{Y}{Z}, 1) = (X, Y, Z)$, thus $[\frac{X}{Z} : \frac{Y}{Z} : 1] = [X : Y : Z]$. Hence $[X : Y : Z] \in f(C_A)$ and $C_P \cap A^2(K) \subseteq f(C_A)$. □

6.2 Quadratic Forms

The following definitions on quadratic forms are based on Friedberg, Insel, and Spence [SF03]. Quadratic forms are of interest to us because these are precisely

the homogeneous polynomials of degree 2.

Definition 6.2.1. Let V be a vector space over a field K. A function H from the set $V \times V$ of ordered pairs of vectors to K is called a **bilinear form** on V if H is linear in each of the variables when the other variable is held fixed.

Definition 6.2.2. A bilinear form H on a vector space V is **symmetric** if $H(x, y) = H(y, x)$ for all $x, y \in V$.

Definition 6.2.3. Let V be a vector space over K. A function $Q : V \to K$ is called a **quadratic form** if there exists a symmetric bilinear form H such that

$$Q(\mathbf{x}) = H(\mathbf{x}, \mathbf{x}) \tag{6.6}$$

for all $\mathbf{x} \in V$. If the characteristic of K is not 2, there is a one-to one correspondence between symmetric bilinear forms and quadratic forms given by Equation (6.6) and

$$H(\mathbf{x}, \mathbf{y}) = \frac{1}{2}[Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})]. \tag{6.7}$$

For the remainder of this section we will assume the characteristic of K is not 2, that V is a finite dimensional vector space over K, and that H is a symmetric bilinear form on V.

Definition 6.2.4. The symmetric bilinear form H is **non-degenerate** if for any $v \in V, v \neq 0$, there exists a $w \in V$ such that $H(v, w) \neq 0$. Otherwise H is **degenerate**.

Definition 6.2.5. Let $x \in V$, then x is **isotropic** if $H(x, x) = 0$. Otherwise x is **non-isotropic**.

Lemma 6.2.6. *For any $x \in V$, $H(0, x) = 0$.*

Proof. Since $H(0,x) = H(0+0,x) = H(0,x) + H(0,x)$, $0 = H(0,x)$. □

Lemma 6.2.7. *If all $x \in V$ are isotropic, then H is degenerate. (So if H is non-degenerate there exists an $x \in V$ such that x is non-isotropic.)*

Proof. Since all $x \in V$ are isotropic we know $H(x,x) = 0$ for any $x \in V$. Then $H(x+y, x+y) = 0$, so

$$H(x,x) + H(x,y) + H(y,x) + H(y,y) = 0 \tag{6.8}$$

since H is bilinear. However, $H(x,x) = 0 = H(y,y)$ by assumption so $H(x,y) + H(y,x) = 0$. That is, $2H(x,y) = 0$ since H is symmetric. Since K is not of characteristic 2, $H(x,y) = 0$ for any $x, y \in V$ and H is degenerate. □

Definition 6.2.8. Let W be a subspace of V. Then

$$W^{\perp} = \{v \in V | H(v,w) = 0 \text{ for all } w \in W\}. \tag{6.9}$$

Lemma 6.2.9. *Assuming H is non-degenerate, if $w \in V$ is non-isotropic, then*

(1) dim $\langle w \rangle^{\perp}$ = dim $V - 1$,

(2) $V = \langle w \rangle \oplus \langle w \rangle^{\perp}$, and

(3) $H|_{\langle w \rangle^{\perp}}$ is non-degenerate.

Proof. (1) Let $T_w(x) = H(x,w)$. We see that T_w is a linear transformation from V to K since H is a bilinear form, and $\langle w \rangle^{\perp} = \ker(T_w)$. We know that $\dim(K) = 1$, so $\operatorname{rank}(T_w) \leq 1$. However, $T_w(w) = H(w,w) \neq 0$ since w is non-isotropic, so $\operatorname{rank}(T_w) = 1$. By the Dimension Theorem, $\dim V = \operatorname{rank} T_w + \operatorname{nullity} T_w$, so $\dim V = 1 + \operatorname{nullity} T_w$. That is $\dim V - 1 = \operatorname{nullity} T_w = \dim \langle w \rangle^{\perp}$.

(2) For $x \in V$, let $b = \dfrac{H(x,w)}{H(w,w)}$ so $bH(w,w) = H(x,w)$. Then $H(x,w) - bH(w,w) = 0$. That is, $H(x - bw, w) = 0$ since H is bilinear, so $x - bw \in \langle w \rangle^{\perp}$. That is, $x \in \langle w \rangle^{\perp} + \langle w \rangle$ since $bw \in \langle w \rangle$. Thus $V = \langle w \rangle + \langle w \rangle^{\perp}$.

Now let $x \in \langle w \rangle \cap \langle w \rangle^{\perp}$. Then $x = bw$ because x is in $\langle w \rangle$ and $H(bw, w) = 0$ since $bw = x \in \langle w \rangle^{\perp}$. Thus $bH(w,w) = 0$ since H is bilinear, but $H(w,w) \neq 0$. Hence $b = 0, x = 0$, and $\langle w \rangle \cap \langle w \rangle^{\perp} = \{0\}$.

(3) Let $v \in \langle w \rangle^{\perp}$. We know there exists $y \in V$ such that $H(v,y) \neq 0$ since H is non-degenerate, where $y = bw + x$ for some $x \in \langle w \rangle^{\perp}$. Then $H(v, bw + x) = bH(v,w) + H(v,x) \neq 0$, since H is bilinear. But $H(v,w) = 0$ by definition of $\langle w \rangle^{\perp}$. Hence $H(w,x) \neq 0$ and H is non-degenerate. □

Definition 6.2.10. An **orthogonal basis** is a basis $\{v_i\}$ of V such that $H(v_i, v_i) \neq 0$ and $H(v_i, v_j) = 0$ for $i \neq j$.

Theorem 6.2.11. *If H is non-degenerate on a finite dimensional vector space V of dimension n, then V has an orthogonal basis.*

Proof. We proceed by induction on the dimension of V.

Let $\dim V = 1$. By Lemma 6.2.7, there exists $x \neq 0 \in V$ such that x is non-isotropic and $H(x,x) \neq 0$. Let $\{v_1 = x\}$ be a basis of V, then $H(v_1, v_1) = H(x,x) \neq 0$ and $H(v_i, v_j) = 0$ for $i \neq j$, holds vacuously.

Let $\dim V = n > 1$. Again there exists $x \neq 0 \in V$ such that x is non-isotropic and $H(x,x) \neq 0$. By Lemma 6.2.9, $V = \langle x \rangle \oplus \langle x \rangle^{\perp}, \dim \langle x \rangle^{\perp} = n - 1$, and $H|_{\langle x \rangle^{\perp}}$ is non-degenerate. By induction there exists a basis $\{v_1, ..., v_{n-1}\}$ of $\langle v \rangle^{\perp}$ such that $H(v_i, v_i) \neq 0$ for $1 \leq i \leq n-1$, and $H(v_i, v_j) = 0$ for $i \neq j$ and $1 \leq i, j \leq n-1$.

Then $\{v_1, ..., v_{n-1}, x\}$ is a basis of $V = \langle x \rangle \oplus \langle x \rangle^\perp$ and $H(x, x) \neq 0$ since x is non-isotropic and $H(x, v_i) = 0$ for any i since $v_i \in \langle x \rangle^\perp$. □

Lemma 6.2.12. *Assuming H is non-degenerate, let W be a subspace of V with $\dim W = m$. Suppose $H|_W$ is non-degenerate. Then*

(1) $\dim W^\perp = \dim V - \dim W$,

(2) $V = W \oplus W^\perp$, *and*

(3) $H|_{W^\perp}$ *is non-degenerate.*

Proof. (1) Assume $\{w_1, w_2, ..., w_m\}$ is an orthogonal basis for W, by Theorem 6.2.11. Let $T_W(x) = \begin{bmatrix} H(x, w_1) \\ H(x, w_2) \\ \vdots \\ H(x, w_m) \end{bmatrix}$. We see that T_W is linear since

$$\begin{aligned} T(ax + y) &= \begin{bmatrix} H(ax + y, w_1) \\ H(ax + y, w_2) \\ \vdots \\ H(ax + y, w_m) \end{bmatrix} = \begin{bmatrix} H(ax, w_1) \\ H(ax, w_2) \\ \vdots \\ H(ax, w_m) \end{bmatrix} + \begin{bmatrix} H(y, w_1) \\ H(y, w_2) \\ \vdots \\ H(y, w_m) \end{bmatrix} \\ &= a \begin{bmatrix} H(x, w_1) \\ H(x, w_2) \\ \vdots \\ H(x, w_m) \end{bmatrix} + \begin{bmatrix} H(x, w_1) \\ H(y, w_2) \\ \vdots \\ H(y, w_m) \end{bmatrix}. \end{aligned} \tag{6.10}$$

We also see that $v \in W^{\perp}$ if and only if

$$T(v) = \begin{bmatrix} H(v, w_1) \\ H(v, w_2) \\ \vdots \\ H(v, w_m) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{6.11}$$

Thus $\ker T_W = W^{\perp}$. Now to find the rank of T_W consider

$$T_W(\frac{1}{a}w_i) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ H(\frac{1}{a}w_i, w_i) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \frac{1}{a}H(w_i, w_i) \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \frac{1}{a}a \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{6.12}$$

So the standard basis vectors are in the image of T_W which means their span is also in the image and so the image of T_W has dimension m. Thus the rank of T_W is m. So by the Dimension Theorem, $\dim W^{\perp} = \dim V - \dim W$.

(2) Let $x \in V$, and let $v = x - b_1 w_1 - ... - b_m w_m$, where $b_i = \dfrac{H(x, w_i)}{H(w_i, w_i)}$. Then

$$\begin{aligned} H(v, w_i) &= H(x - b_1 w_1 - ... - b_m w_m, w_i) \\ &= H(x, w_i) - b_1 H(w_1, w_i) - ... - b_i H(w_i, w_i) - ... - b_m H(w_m, w_i) \\ &= H(x, w_i) - b_i H(w_i, w_i) = 0 \end{aligned} \tag{6.13}$$

so $v = x - b_1 w_1 - ... - b_m w_m \in W^{\perp}$, and $x \in W^{\perp} + W$ since $b_1 w_1 + ... + b_m w_m \in W$. Hence $V = W + W^{\perp}$.

Now let $x \in W \cap W^{\perp}$. Since $x \in W$, we know that $x = b_1 w_1 + ... + b_m w_m$. Also, since $x \in W^{\perp}$ we know that $H(x, w_i) = 0$. That is,

$$\begin{aligned} 0 = H(x, w_i) &= H(b_1 w_1 + ... + b_m w_m, w_i) \\ &= b_1 H(w_1, w_i) + b_2 H(w_2, w_i) + ... + b_i H(w_i, w_i) + ... + b_m H(w_m, w_i) \\ &= b_i H(w_i, w_i), \end{aligned} \tag{6.14}$$

by Theorem 6.2.11. So $b_i H(w_i, w_i) = 0$ and we know $H(w_i, w_i) \neq 0$. Thus $b_i = 0$ and $W \cap W^{\perp} = \{0\}$.

(3) For $v \in \langle w \rangle^{\perp}$ there exists $y \in V$ such that $H(v, y) \neq 0$ since H is non-degenerate on V. But $y = w + x$ where $w \in W$ and $x \in \langle w \rangle^{\perp}$. Then $H(v, w + x) = H(v, w) + H(v, x) \neq 0$, since H is linear. But $H(v, w) = 0$ by definition of $\langle w \rangle^{\perp}$. Hence $H(v, x) \neq 0$ and $H|_{W^{\perp}}$ is non-degenerate.

□

Theorem 6.2.13. *Let K be a field of characteristic $\neq 2$, and V a 3-dimensional K-vector space. Let Q be a non-degenerate quadratic form on V and let H be the symmetric non-degenerate bilinear form associated to Q. Let*

$$C_p = \{(X : Y : Z) | Q(X, Y, Z) = 0\} \neq \emptyset. \tag{6.15}$$

Then C_p is equivalent to $\{(X : Y : Z) | XZ - Y^2 = 0\}$.

Proof. Since $C_p \neq \emptyset$ there exists an $e_1 \in V$ such that $Q(e_1) = H(e_1, e_1) = 0$. Furthermore there exists a vector $w \in V$ such that $H(e_1, w) \neq 0$ because H is non-degenerate. Let $H(e_1, w) = a$. Then if $z = \frac{1}{a} w$, we have

$$H(e_1, z) = H(e_1, \frac{1}{a} w) = \frac{1}{a} H(e_1, w) = \frac{1}{a} a = 1, \tag{6.16}$$

since H is bilinear. Now let $e_3 = z - \frac{b}{2}e_1$, where $b = H(z,z)$. Then

$$\begin{aligned} H(e_3,e_3) &= H(z - \frac{b}{2}e_1, z - \frac{b}{2}e_1) \\ &= H(z,z) - \frac{b}{2}H(z,e_1) - \frac{b}{2}H(e_1,z) + \left(\frac{b}{2}\right)^2 H(e_1,e_1) \\ &= H(z,z) - \frac{b}{2}H(e_1,z) - \frac{b}{2}H(e_1,z) + 0, \end{aligned} \tag{6.17}$$

since H is symmetric and $H(e_1,e_1) = 0$. Then $H(e_3,e_3) = H(z,z) - bH(e_1,z)$ but $H(e_1,z) = 1$, so $H(e_3,e_3) = H(z,z) - b = 0$. Next we note that

$$H(e_1,e_3) = H(e_1, z - \frac{b}{2}e_1) = H(e_1,z) - \frac{b}{2}H(e_1,e_1) = 1 \tag{6.18}$$

since $H(e_1,z) = 1$ and $H(e_1,e_1) = 0$.

So e_1 and e_3 are both isotropic and $H(e_1,e_3) = 1$. Furthermore, suppose e_1 and e_3 are linearly dependent. Then there exists λ such that $e_1 = \lambda e_3$, so we have

$$1 = H(e_1,e_3) = H(\lambda e_3, e_3) = \lambda H(e_3,e_3) = 0, \tag{6.19}$$

which is a contradiction. Thus e_1 and e_3 are independent vectors. Now we can consider the 2-dimensional subspace $\langle e_1, e_3\rangle$ and we claim that $H|_{\langle e_1,e_3\rangle}$ is non-degenerate. Let $u \in \langle e_1,e_3\rangle$ such that $u \neq 0$. Then $u = ae_1 + be_3$. If $a \neq 0$ then

$$H(u,e_3) = H(ae_1 + be_3, e_3) = aH(e_1,e_3) + bH(e_3,e_3) = a. \tag{6.20}$$

If $a = 0, b \neq 0$ then

$$H(u,e_1) = H(ae_1 + be_3, e_1) = aH(e_1,e_1) + bH(e_1,e_3) = b. \tag{6.21}$$

Therefore $H|_{\langle e_1,e_3\rangle}$ is non-degenerate.

Now let us take $e_2 \in \langle e_1,e_3\rangle^\perp, e_2 \neq 0$. By Theorem 6.2.12 we know that $H|_{\langle e_1,e_3\rangle^\perp}$ is non-degenerate so $a = H(e_2,e_2) \neq 0$. Let $e_1' = \frac{-a}{2}e_1$; then in the basis

$\{e_1', -e_2, e_3\}$ we have

$$\begin{aligned}
Q(X,Y,Z) &= Q(Xe_1' - Ye_2 + Ze_3) \\
&= Q(\frac{-a}{2}Xe_1 - Ye_2 + Ze_3) \\
&= H(\frac{-a}{2}Xe_1 - Ye_2 + Ze_3, \frac{-a}{2}Xe_1 - Ye_2 + Ze_3) \\
&= H(\frac{-a}{2}Xe_1, \frac{-a}{2}Xe_1) - H(\frac{-a}{2}Xe_1, Ye_2) + H(\frac{-a}{2}Xe_1, Ze_3) \\
&\quad - H(Ye_2, \frac{-a}{2}Xe_1) + H(Ye_2, Ye_2) - H(Ye_2, Ze_3) \\
&\quad + H(Ze_3, \frac{-a}{2}Xe_1) - H(Ze_3, Ye_2) + H(Ze_3, Ze_3) \\
&= H(\frac{-a}{2}Xe_1, Ze_3) + H(Ze_3, \frac{-a}{2}Xe_1) + H(Ye_2, Ye_2) \\
&= H(\frac{-a}{2}Xe_1, Ze_3) + H(\frac{-a}{2}Xe_1, Ze_3) + H(Ye_2, Ye_2) \\
&= 2H(\frac{-a}{2}Xe_1, Ze_3) + Y^2H(e_2, e_2) \\
&= 2\frac{-a}{2}XZH(e_1, e_3) + Y^2a \\
&= -aXZ + aY^2, \\
&= aXZ - aY^2,
\end{aligned} \tag{6.22}$$

with $a \neq 0$. Therefore $Q(X,Y,Z) = 0$ if and only if $XZ - Y^2 = 0$. □

Since we are interested in studying the zeroes of Q we can simplify our desired points to the set

$$C = \{(X:Y:Z) | XZ - Y^2 = 0\} = \{(X:Y:Z) | XZ = Y^2\} \subseteq P^2(K). \tag{6.23}$$

Let $\phi : P^1(K) \to C$ be given by $\phi(U:V) = (U^2 : UV : V^2)$. Define

$$\psi : C \to P^1(K) \tag{6.24}$$

by

$$\psi(X:Y:Z) = \begin{cases} (X:Y) & \text{if } Z \neq 0 \\ (Y:Z) & \text{if } X \neq 0 \end{cases} \tag{6.25}$$

If $X \neq 0$ and $Z \neq 0$, then $(X : Y) = (ZX : ZY) = (Y^2 : ZY) = Y(Y : Z) = (Y : Z)$, so they are in the same equivalence class and so ψ is consistent.

Theorem 6.2.14. *The maps ϕ and ψ are well-defined and inverses of each other.*

Proof. For $\lambda \neq 0$,

$$\begin{aligned} \phi(\lambda U : \lambda V) &= (\lambda^2 U^2 : \lambda^2 UV : \lambda^2 V) \\ &= \lambda^2 (U^2 : UV : V^2) = (U^2 : UV : V^2) = \phi(U : V). \end{aligned} \tag{6.26}$$

Similarly,

$$\begin{aligned} \psi(\lambda X : \lambda Y : \lambda Z) &= \begin{cases} (\lambda X : \lambda Y) & \text{if } X \neq 0 \\ (\lambda Y : \lambda Z) & \text{if } Z \neq 0 \end{cases} \\ &= \begin{cases} \lambda(X : Y) & \text{if } X \neq 0 \\ \lambda(Y : Z) & \text{if } Z \neq 0 \end{cases} \\ &= \lambda \psi(X : Y : Z). \end{aligned} \tag{6.27}$$

If $X \neq 0$ and $Z \neq 0$, then

$$(X : Y) = (ZX : ZY) = (Y^2 : ZY) = Y(Y : Z) = (Y : Z), \tag{6.28}$$

so ψ is consistent.

We show that the image of ϕ is contained in C. Let
$(U^2 : UV : V^2) \in \phi(P^1(K))$. Then
$XZ - Y^2 = U^2V^2 - (UV)^2 = (UV)^2 - (UV)^2 = 0$, so $(U^2 : UV : V^2) \in C$ and thus
$\operatorname{Im} \phi \subseteq C$.

Last we show that ϕ and ψ are inverse maps. On one hand,

$$\begin{aligned}
\psi\phi(U:V) &= \psi(U^2 : UV : V^2) \\
&= \begin{cases} (U^2 : UV) & \text{if } U \neq 0 \\ (UV : V^2) & \text{if } V \neq 0 \end{cases} \\
&= \begin{cases} U(U:V) & \text{if } U \neq 0 \\ V(U:V) & \text{if } V \neq 0 \end{cases} \\
&= (U:V).
\end{aligned} \tag{6.29}$$

Similarly, suppose $(X : Y : Z) \in C$. Then

$$\begin{aligned}
\phi\psi(X:Y:Z) &= \begin{cases} \phi(X:Y) & \text{if } X \neq 0 \\ \phi(Y:Z) & \text{if } Z \neq 0 \end{cases} \\
&= \begin{cases} (X^2 : XY : Y^2) & \text{if } X \neq 0 \\ (Y^2 : YZ : Z^2) & \text{if } Z \neq 0. \end{cases} \\
&= \begin{cases} (\frac{X^2}{X} : \frac{XY}{X} : \frac{XZ}{X}) & \text{if } X \neq 0 \\ (\frac{Y^2}{Z} : \frac{YZ}{Z} : \frac{Z^2}{Z}) & \text{if } Z \neq 0. \end{cases} \\
&= \begin{cases} (X : Y : \frac{Y^2}{X}) & \text{if } X \neq 0 \\ (\frac{Y^2}{Z} : Y : Z) & \text{if } Z \neq 0. \end{cases} \\
&= \phi(X:Y:Z)
\end{aligned} \tag{6.30}$$

because $XZ = Y^2$ on C. □

6.3 Counting Intersections of Curves

In this section we follow Reid's [Rei90] presentation of plane conics.

Definition 6.3.1. A **homogeneous polynomial of degree d** in two variables is given by $F(U,V) = a_dU^d + a_{d-1}U^{d-1}V + ... + a_iU^iV^{d-i} + ... + a_0V^d$. The associated polynomial in 1 variable to $F(U,V)$ is $f(u) = a_du^d + a_{d-1}u^{d-1} + ... + a_iu^i + ... + a_0$.

Definition 6.3.2. Let F be of degree d. Suppose $F(U,V) = 0, (U,V) \neq (0,0)$.

(1) If $V \neq 0$, then the **multiplicity** of $(U:V)$ as a zero of F is defined to be the multiplicity of $\alpha = \dfrac{U}{V} \in K$ as a zero of f.

(2) If $V = 0$, then the **multiplicity** $(U:V)$ as a zero of F is defined to be $d - \deg f$.

So the multiplicity of a zero of F of the form $(\alpha, 1)$ is the greatest power of $(U - \alpha V)$ dividing F, and at $(1,0)$ it is the greatest power of V dividing F.

Lemma 6.3.3. *Let $F(U,V)$ be a nonzero homogeneous polynomial of degree d in U,V. Then F has at most d zeros on $P^1(K)$. Furthermore, if K is algebraically closed, then F has exactly d zeros on $P^1(K)$ provided these are counted with multiplicities as defined above.*

Proof. Let m_∞ be the multiplicity of the zero of F at $(1:0)$. Then by definition, $d - m_\infty$ is the degree of the inhomogeneous polynomial f. Then the lemma reduces to the fact that a polynomial in one variable has at most $\deg f$ roots, by a corollary to the Division Algorithm for $K[x]$ (see Fraleigh [Fra03]). □

Definition 6.3.4. Let the number of elements in a set A be denoted by $\#A$.

Theorem 6.3.5. *Let $L \subset P^2(K)$ be a line defined by*

$$L = \{U(a_1, b_1, c_1) + V(a_2, b_2, c_2) \mid U, V \in K\} \setminus \{0\}, \tag{6.31}$$

let $C \subset P^2(K)$ be a non-degenerate conic defined by

$$C = \{(X:Y:Z) \mid XZ - Y^2 = 0\} \setminus \{0\}, \tag{6.32}$$

and let $D \in P^2(K)$ be a curve defined by

$$D = \{(X, Y, Z) \mid G_d(X, Y, Z) = 0\}, \tag{6.33}$$

where G is a homogeneous polynomial of degree d in X, Y, and Z. Assume $L \nsubseteq D$ and $C \nsubseteq D$. Then $\#\{L \cap D\} \leq d$ and $\#\{C \cap D\} \leq 2d$.

Proof. From the definition of L, we know that any point on L is of the form $(Ua_1 + Va_2, Ub_1 + Vb_2, Uc_1 + Vc_2)$. To find where these points intersect D we substitute and get $G_d(Ua_1 + Va_2, Ub_1 + Vb_2, Uc_1 + Vc_2)$ which is a homogeneous polynomial in U, V of degree d. So by Lemma 6.3.3, this implies there are at most d zeros; that is, $\#\{L \cap D\} \leq d$.

By Theorem 6.2.14, $C = (U^2 : UV : V^2)|U, V \in K\} \setminus \{0\}$, so to find the intersection points of C with D we substitute and get $G_d(U^2, UV, V^2)$. This is a homogeneous polynomial of degree $2d$, so by Lemma 6.3.3, $\#\{C \cap D\} \leq 2d$. □

Lemma 6.3.6. *Suppose C_1 is a degenerate conic. Then C_1 is either a point, a line, or two lines.*

Proof. Suppose C_1 is degenerate. Let H be the bilinear form associated with the quadratic form that defines C_1. Then since H is degenerate there exists a nonzero $e_3 \in V$ such that $H(v, e_3) = 0$ for any $v \in V = K^3$. Extend $\{e_3\}$ to a basis $\{e_1, e_2, e_3\}$ of V. Then H is either non-degenerate on $W = \langle e_1, e_2 \rangle$ or degenerate on W.

Case 1: Suppose H is non-degenerate on W. Then either there exists a nonzero isotropic vector $e_1' \in W$, or not.

Case 1a: Suppose there exists a nonzero isotropic vector, e_1', i.e., $H(e_1', e_1') = 0$. Then there exists a vector $w \in W$ such that $H(e_1', w) \neq 0$ because H

is non-degenerate on W. Let $H(e_1', w) = a \neq 0$; then if $z = \frac{1}{a}w$, we have

$$H(e_1', z) = H(e_1', \frac{1}{a}w) = \frac{1}{a}H(e_1', w) = \frac{1}{a}a = 1, \tag{6.34}$$

since H is bilinear. Now let $e_2' = z - \frac{b}{2}e_1'$, where $b = H(z, z)$. Then

$$\begin{aligned} H(e_2', e_2') &= H(z - \frac{b}{2}e_1', z - \frac{b}{2}e_1') \\ &= H(z, z) - \frac{b}{2}H(z, e_1') - \frac{b}{2}H(e_1', z) + (\frac{b}{2})^2 H(e_1', e_1') \\ &= H(z, z) - \frac{b}{2}H(e_1', z) - \frac{b}{2}H(e_1', z) + 0, \end{aligned} \tag{6.35}$$

since H is symmetric and $H(e_1', e_1') = 0$. Then $H(e_2', e_2') = H(z, z) - bH(e_1', z)$ but $H(e_1', z) = 1$, so $H(e_2', e_2') = H(z, z) - b = 0$. Next we note that

$$H(e_1', e_2') = H(e_1', z - \frac{b}{2}e_1') = H(e_1', z) - \frac{b}{2}H(e_1', e_1') = 1 \tag{6.36}$$

since $H(e_1', z) = 1$ and $H(e_1', e_1') = 0$. So e_1' and e_2' are both isotropic and $H(e_1', e_2') = 1$. Then

$$\begin{aligned} Q(Xe_1' + Ye_2' + Ze_3) &= H(Xe_1' + Ye_2' + Ze_3, Xe_1' + Ye_2' + Ze_3) \\ &= H(Xe_1', Xe_1') + H(Xe_1', Ye_2') + H(Xe_1', Ze_3) \\ &\quad + H(Ye_2', Xe_1') + H(Ye_2', Ye_2') + H(Ye_2', Ze_3) \\ &\quad + H(Ze_3, Xe_1') + H(Ze_3, Ye_2') + H(Ze_3, Ze_3) \\ &= 2H(Xe_1', Ye_2') \\ &= 2XY, \end{aligned} \tag{6.37}$$

since $H(e_3, v) = H(e_1', e_1') = H(e_2', e_2') = 0$ and $H(e_1', e_1') = 1$. So in the basis $\{e_1', e_2', e_3\}$,

$$C = \{(X, Y, Z) | XY = 0\} = \{(X, Y, Z) | X = 0\} \cup \{(X, Y, Z) | Y = 0\}, \tag{6.38}$$

the union of two projective lines.

Case 1b: Suppose there is no nonzero isotropic vector in W. Then for any $v \in W, H(v,v) \neq 0$. Let $w \in W = \langle e_1, e_2 \rangle$. Then for $w = Xe_1 + Ye_2 \in W$ we have

$$\begin{aligned} Q(Xe_1 + Ye_2 + Ze_3) &= H(w + Ze_3, w + Ze_3) \\ &= H(w,w) + H(w, Ze_3) + H(Ze_3, w) + H(Ze_3, Ze_3) \\ &= H(w,w), \end{aligned} \tag{6.39}$$

since $H(v, e_3) = 0$ for any $v \in V$. So $Q(Xe_1 + Ye_2 + Z_3) = H(w,w) = 0$ if and only if $w = 0$, and the only solution is the projective point $(0:0:1)$.

Case 2: Suppose H is degenerate on W. Then there exists $e_2' \in W$ such that for any $v \in W$ we have $H(e_2', v) = 0$. Extend $\{e_2'\}$ to a basis $\{e_1', e_2'\}$ of W so that $\{e_1', e_2', e_3\}$ is a basis of V. Then

$$\begin{aligned} Q(Xe_1' + Ye_2' + Ze_3) &= H(Xe_1' + Ye_2' + Ze_3, Xe_1' + Ye_2' + Ze_3) \\ &= H(Xe_1', Xe_1') + H(Xe_1', Ye_2') + H(Xe_1', Ze_3) \\ &\quad + H(Ye_2', Xe_1') + H(Ye_2', Ye_2') + H(Ye_2', Ze_3) \\ &\quad + H(Ze_3, Xe_1') + H(Ze_3, Ye_2') + H(Ze_3, Ze_3) \\ &= H(Xe_1', Xe_1') \\ &= X^2 H(e_1', e_1'), \end{aligned} \tag{6.40}$$

where $H(e_1', e_1') \neq 0$ since Q is not the zero polynomial. Thus

$$C = \{(X,Y,Z) | X^2 = 0\} = \{(X,Y,Z) | X = 0\}, \tag{6.41}$$

a projective line. □

Theorem 6.3.7. *If $|K| \geq 4$, $\operatorname{char} K \neq 2$, and $P_1, ..., P_5 \in P^2(K)$ are distinct points such that no 4 are collinear, then there exists at most one conic through $P_1, ..., P_5$.*

Proof. Suppose that C_1 and C_2 are non-empty conics with $C_1 \neq C_2$ such that $C_1 \cap C_2 \supset \{P_1, ..., P_5\}$. Then we must consider two cases: At least one C_i is non-degenerate or both C_i are degenerate.

Case 1: Suppose at least one of the conics is non-degenerate. Without loss of generality, let C_1 be non-degenerate. Then by Theorem 6.2.13 C_1 is equivalent to the parametrized curve $\{(U^2, UV, V^2) | (U, V) \in P^1(K)\}$. By Theorem 6.3.5 we know that if $C_1 \nsubseteq C_2$, then $\#\{C_1 \cap C_2\} \leq 2d = 4$, but we have 5 points in $C_1 \cap C_2$ so it must be that $C_1 \subset C_2$. Let Q_2 be the equation of C_2. Then $Q_2(U^2, UV, V^2) = 0$ for all $(U, V) \in P^1(K)$, where

$$Q_2(X, Y, Z) = a_1X^2 + a_2XY + a_3Y^2 + a_4Z^2 + a_5XZ + a_6YZ, \tag{6.42}$$

since Q_2 is a homogeneous polynomial of degree 2. To solve for a_i we assume our field has at least 4 elements in it, $\{0, 1, -1, \alpha\}$, and we evaluate Q_2 at the points $(0, 0, 1)$, $(1, 0, 0)$, $(1, 1, 1)$, $(1, -1, 1)$, $(\alpha^2, \alpha, 1)$. Then

$$Q_2(0, 0, 1) = a_4 = 0, \tag{6.43}$$

$$Q_2(1, 0, 0) = a_1 = 0, \tag{6.44}$$

$$Q_2(1, 1, 1) = a_2 + a_3 + a_5 + a_6 = 0, \tag{6.45}$$

$$Q_2(1, -1, 1) = -a_2 + a_3 + a_5 - a_6 = 0, \tag{6.46}$$

$$Q_2(\alpha^2, \alpha, 1) = \alpha^3 a_2 + \alpha^2 a_3 + \alpha^2 a_5 + \alpha a_6 = 0. \tag{6.47}$$

By adding (6.45) and (6.46) we get $2a_3 + 2a_5 = 0$. Since our field is not of characteristic 2, $a_3 = -a_5$. Substituting this relationship into (6.45) implies that $a_2 = -a_6$, and substituting it into (6.47) yields

$$\alpha^3 a_2 + \alpha a_6 = \alpha^2 a_2 + a_6 = 0, \tag{6.48}$$

since $\alpha \neq 0$. However, since $a_2 = -a_6$, we can write this last equation as

$$-\alpha^2 a_6 + a_6 = a_6(1 - \alpha^2) = a_6(1 + \alpha)(1 - \alpha) = 0, \tag{6.49}$$

where $\alpha \neq \pm 1$. Thus it must be true that $a_6 = 0 = a_2$, and

$$Q_2 = a_3Y^2 + a_5XZ = -a_5Y^2 + a_5XZ = a_5(XZ - Y^2), \tag{6.50}$$

since $a_3 = -a_5$. So we have that Q_2 is a multiple of $(XZ - Y^2)$, which contradicts our assumption that $C_1 \neq C_2$.

Case 2: Suppose both C_1 and C_2 are degenerate. Then by Lemma 6.3.6 C_1 and C_2 are either a point, a line, or two lines. If C_i is a point, then all five points, $P_1, ..., P_5$ are the same point, contradicting the assumption that the points are distinct. If C_i is a line, then $\{P_1, ..., P_5\} \subseteq C_i$ since C_i passes through the points, thus contradicting the fact that no four points are collinear. If C_i is 2 lines, then let $C_1 = L_a \cup L_b, C_2 = L_c \cup L_d$. Since all 5 points pass through both conics we can deduce without loss of generality that if
$P_1 \in L_a \cap L_c, P_2 \in L_b \cap L_d, P_3 \in L_a \cap L_d, P_4 \in L_b \cap L_c$, then $P_5 \in L_a \cap L_c$. So L_a and L_c pass through both points P_1 and P_5 and therefore must be the same line. Let $L_0 = L_a = L_c$. Then

$$C_1 \cap C_2 = (L_0 \cup L_b) \cap (L_0 \cup L_d) = L_0 \cup (L_b \cap L_d), \tag{6.51}$$

where L_b and L_d intersect at exactly one point and consequently, the remaining four points must lie on L_0 which contradicts the fact that no 4 points are collinear.

Thus our assumption that $C_1 \neq C_2$ must be false for non-empty C_1, C_2 and so there exists at most one conic through $P_1, ..., P_5$. □

Definition 6.3.8. Let S_2 be the vector space of quadratic forms of degree 2 in three variables over K. Then,

$$S_2 = \{a_1X^2 + a_2Y^2 + a_3Z^2 + a_4XY + a_5XZ + a_6YZ | a_i \in K\}, \tag{6.52}$$

where $\{X^2, Y^2, Z^2, XY, XZ, YZ\}$ is a basis of S_2 and so S_2 is of dimension 6. Furthermore,

$$S_2(P_1, ..., P_n) = \{Q \in S_2 \mid Q(P_i) = 0 \text{ for } 1 \leq i \leq n\}. \tag{6.53}$$

is the subspace of quadratic forms of degree 2 passing through the points $P_1, ..., P_n$.

Corollary 6.3.9. $\dim S_2(P_1, ..., P_n) \geq 6 - n$.

Proof. Let $P_i = (X_i : Y_i : Z_i)$. Then

$$\begin{aligned} &S_2(P_1, ..., P_n) \\ &= \left\{ Q \;\middle|\; \begin{array}{c} a_1X_1^2 + a_2Y_1^2 + a_3Z_1^2 + a_4X_1Y_1 + a_5X_1Z_1 + a_6Y_1Z_1 = 0 \\ a_1X_2^2 + a_2Y_2^2 + a_3Z_2^2 + a_4X_2Y_2 + a_5X_2Z_2 + a_6Y_2Z_2 = 0 \\ \vdots \\ a_1X_n^2 + a_2Y_n^2 + a_3Z_n^2 + a_4X_nY_n + a_5X_nZ_n + a_6Y_nZ_n = 0 \end{array} \right\} \end{aligned} \tag{6.54}$$

for $a_i \in K$. Then by Rank-Nullity we have

$$\operatorname{rank} \begin{pmatrix} X_1^2 & Y_1^2 & Z_1^2 & X_1Y_1 & X_1Z_1 & Y_1Z_1 \\ X_2^2 & Y_2^2 & Z_2^2 & X_2Y_2 & X_2Z_2 & Y_2Z_2 \\ \vdots & & & & & \\ X_n^2 & Y_n^2 & Z_n^2 & X_nY_n & X_nZ_n & Y_nZ_n \end{pmatrix} + \dim(S_2(P_1, ..., P_n)) = 6, \tag{6.55}$$

where the rank is at most n, and thus $\dim(S_2(P_1, ..., P_n)) \geq 6 - n$. □

Corollary 6.3.10. *If $n \leq 5$ and no 4 of $P_1, ...P_n$ are collinear, then* $\dim S_2(P_1, ...P_n) = 6 - n$.

Proof. By Theorem 6.3.7, if $n = 5$, then $\dim S_2(P_1, ..., P_n) \leq 1$, and by Corollary 6.3.9, $\dim S_2(P_1, ..., P_n) \geq 1$, so the Corollary is true in this case. If $n \leq 4$

then we can add points $P_{n+1}, ..., P_5$ to $P_1, ..., P_n$ where we still have no 4 points collinear, and since each point imposes at most one new linear condition, this gives

$$1 = \dim S_2(P_1, ..., P_5) \geq \dim S_2(P_1, ..., P_n) - (5 - n). \tag{6.56}$$

So $6 - n \geq \dim S_2(P_1, ..., P_n)$, and by Corollary 6.3.9, equality holds. □

CHAPTER 7

CUBICS

In this chapter we follow the presentation of Cubics and the group law in chapter 2 of Reid's book [Rei90].

7.1 Forms of Degree 3

First, we expand on the previous chapter's definition for forms of degree 2. Suppose S_3 is the set of all forms Q of degree 3 in three variables over K. Then

$$\begin{aligned} S_3 = \{&a_1X^3 + a_2Y^3 + a_3Z^3 + a_4X^2Y + a_5X^2Z \\ &+ a_6Y^2X + a_7Y^2Z + a_8Z^2X + a_9Z^2Y + a_{10}ZXY \mid a_i \in K\}. \end{aligned} \tag{7.1}$$

So a basis for S_3 is $\{X^3, Y^3, Z^3, X^2Y, X^2Z, Y^2X, Y^2Z, Z^2X, Z^2Y, ZXY\}$ and $\dim S_3 = 10$.

Definition 7.1.1. For $P_1, ..., P_n \in P^2(K)$, let
$S_3(P_1, ..., P_n) = \{Q \in S_3 \mid Q(P_i) = 0 \text{ for } i = 1, ..., n\}$ be the subspace of forms of degree 3 passing through points $P_1, ..., P_n$.

Theorem 7.1.2. $\dim S_3(P_1, ..., P_n) \geq 10 - n$.

Proof. Let $P_i = (X_i : Y_i : Z_i)$. Each of the conditions $Q(P_i) = Q(X_i, Y_i, Z_i) = 0$ is a

linear condition on the coefficients of Q, and so by Rank-Nullity we have

$$\operatorname{rank}\begin{pmatrix} X_1^3 & Y_1^3 & Z_1^3 & X_1Y_1Z_1 & X_1Z_1^2 & Y_1Z_1^2 & Y_1^2Z_1 & Y_1^2X_1 & X_1^2Z_1 & X_1^2Y_1 \\ X_1^3 & Y_2^3 & Z_2^3 & X_2Y_2Z_2 & X_2Z_2^2 & Y_2Z_2^2 & Y_2^2Z_2 & Y_2^2X_2 & X_2^2Z_2 & X_2^2Y_2 \\ \vdots & & & & & & & & & \\ X_n^3 & Y_n^3 & Z_n^3 & X_nY_nZ_n & X_nZ_n^2 & Y_nZ_n^2 & Y_n^2Z_n & Y_n^2X_n & X_n^2Z_n & X_n^2Y_n \end{pmatrix} + \dim(S_3(P_1, ..., P_n)) = 10, \tag{7.2}$$

where the rank is at most n, and thus $\dim(S_3(P_1, ..., P_n)) \geq 10 - n$. □

Lemma 7.1.3. *Suppose that $|K| = q \geq 7, \operatorname{char} K \neq 2$, and let $F \in S_3$.*

(1) Let $L \subset P^2(K)$ be a line. If $F \equiv 0$ on L, then F is divisible in $K[X, Y, Z]$ by the equation of L. That is, $F = HF_2$ where H is the equation of L and $F_2 \in S_2$.

(2) Let $C \subset P^2(K)$ be a nonempty non-degenerate conic. If $F \equiv 0$ on C, then F is divisible in $K[X, Y, Z]$ by the equation of C. That is, $F = QF_1$ where Q is the equation of C and $F_1 \in S_1$.

Proof. (1) By Theorem 5.3.3, we can assume $H = X$, i.e., $L = \{(X : Y : Z) \mid X = 0\}$. Then for $F \in S_3$, we can group all the monomials involving X and rewrite F as $F = XF_2 + G(Y, Z)$, where $G \in S_3$ and $F_2 \in S_2$. So $F \equiv 0$ on L if and only if $G(Y, Z) = 0$ for any $(Y : Z) \in P^1(K)$. Suppose $G(Y, Z) \neq 0$, by contradiction. Then by Lemma 6.3.3, $G(Y, Z)$ has 3 or fewer zeros on $P^1(K)$ since it has degree 3. However, $|L| = |P^1(K)| = q + 1 \geq 8$, since $q \geq 7$ by hypothesis, and $\deg G = 3 < 8$, so we have a contradiction. Thus $F = XF_2$.

(2) By Theorem 6.2.13, we can assume $Q = XZ - Y^2$. Next, by substituting $XZ - Q$ for Y^2 in F, we can rewrite $F = QF_1 + A(X, Z) + YB(X, Z)$,

where $A(X,Z) = a_0X^3 + a_1X^2Z + a_2XZ^2 + a_3Z^3$ is a homogeneous polynomial of degree 3 and $B(X,Z) = b_0X^2 + b_1XZ + b_2Z^2$ is a homogeneous polynomial of degree 2. Note that C is a parametrised conic given by $X = U^2, Y = UV, Z = V^2$. So $F \equiv 0$ on C if and only if $G(U,V) = A(U^2,V^2) + UVB(U^2,V^2) \equiv 0$ on C. That occurs if and only if

$$a_0U^6 + a_1U^4V^2 + a_2U^2V^4 + a_3V^6 + UV(b_0U^4 + b_1U^2V^2 + b_2V^4) = 0, \quad (7.3)$$

if and only if

$$a_0U^6 + a_1U^4V^2 + a_2U^2V^4 + a_3V^6 + b_0U^5V + b_1U^3V^3 + b_2UV^5 = 0, \quad (7.4)$$

if and only if

$$a_0U^6 + b_0U^5V + a_1U^4V^2 + b_1U^3V^3 + a_2U^2Z^4 + b_2UV^5 + a_3V^6 = 0. \quad (7.5)$$

But $G(U,V)$ is a polynomial of degree 6 so it has at most 6 zeros, if $G \neq 0$. However, we have $q \geq 7$ zeros by assumption, yet $6 < 7$ so we have a contradiction. Thus $A(X,Z) = B(X,Z) = 0$ and $F = QF_1$. □

Definition 7.1.4. A set of points are **conconic** if they all lie on a nondegenerate conic.

Corollary 7.1.5. *Let K be a finite field with $|K| \geq 7$. Let $L \subset P^2(K)$ be a line with equation $H = 0$, $C \subset P^2(K)$ be a nondegenerate conic with equation $Q = 0$, and suppose that points $P_1, ..., P_n \in P^2(K)$ are given.*

(1) If $P_1, ..., P_a \in L$, and $P_{a+1}, ...P_n \notin L$, with $a > 3$, then

$$S_3(P_1, ..., P_n) = HS_2(P_{a+1}, ..., P_n). \quad (7.6)$$

(2) If $P_1, ..., P_a \in C$, and $P_{a+1}, ..., P_n \notin C$, with $a > 6$, then

$$S_3(P_1, ..., P_n) = QS_1(P_{a+1}, ..., P_n). \quad (7.7)$$

Proof. (1) If F is homogeneous of degree 3, and the curve $D : (F = 0)$ meets L in points $P_1, ..., P_a$ with $a > 3$, then by Theorem 6.3.5 we must have $L \subset D$, and by Lemma 7.1.3 $F = HF_2$, for some F_2. Since $P_{a+1}, ..., P_n$ are not on L then $H(P_i) \neq 0$ for $a + 1 \leq i \leq n$, and we must have $F_2 \in S_2(P_{a+1}, ..., P_n)$.

(2) If F is homogeneous of degree 3, and the curve $D : (F = 0)$ meets C in points $P_1, ..., P_a$ with $a > 6$ then by the contrapositive of Theorem 6.3.5, $C \subset D$, and by Lemma 7.1.3 $F = QF_1$, for some F_1. Since $P_{a+1}, ..., P_n$ are not on C, then $Q(P_i) \neq 0$ for $a + 1 \leq i \leq n$, and we must have $F_1 \in S_1(P_{a+1}, ..., P_n)$.

□

Theorem 7.1.6. *Let K be a field with $|K| \geq 6$ with* $\operatorname{char} K \neq 2$ *and $P_1, ..., P_8 \in P^2(K)$ distinct points. Suppose that no 4 of $P_1, ..., P_8$ are collinear, and no 7 of $P_1, ..., P_8$ lie on a non-degenerate conic. Then*

$$\dim S_3(P_1, ..., P_8) = 2. \tag{7.8}$$

Proof. By Theorem 7.1.2, $\dim S_3(P_1, ..., P_8) \geq 2$.

It remains to show that $\dim S_3(P_1, ..., P_8) \leq 2$.

Case 1: No 3 points are collinear and no 6 points are conconic. Suppose by contradiction that $\dim S_3(P_1, ..., P_8) \geq 3$ and let $P_9, P_{10} \in L$ where $L = L_{P_1P_2}$, the line determined by P_1 and P_2. Then

$$\dim S_3(P_1, ..., P_{10}) \geq \dim S_3(P_1, ..., P_8) - 2, \tag{7.9}$$

since we added two points, so $\dim S_3(P_1, ..., P_{10}) \geq 1$ since $S_3(P_1, ..., P_8) \geq 3$ by assumption. Since S_3 is a non-zero vector space there exists a nonzero cubic, $F \in S_3(P_1, ..., P_n)$, passing through all 10 points. Then since $P_1, P_2, P_9, P_{10} \in L$ and

since no 3 of $P_1, ..., P_8$ are collinear we have $P_3, ..., P_8$ are not on L so by Corollary 7.1.5, $F = HQ$ for some $Q \in S_2(P_3, ..., P_8)$. If Q is non-degenerate this implies that $P_3, ..., P_8$ are conconic, a contradiction. Similarly if Q is a point, line pair or a double line, then at least 3 of the points are collinear, thus a contradiction. So we must have $\dim S_3(P_2, ..., P_8) = 2$.

Case 2a: Suppose P_1, P_2, P_3 all lie on the line $L : (H = 0)$. Let $P_9 \in L$, where $P_4, ..., P_8$ are not on L because no 4 points are collinear by assumption.

Then by Corollary 7.1.5, $S_3(P_1, ...P_9) = HS_2(P_4, ..., P_8)$. Also, since no 4 of $P_4, ...P_8$ are collinear, we have $\dim S_2(P_4, ..., P_8) = 1$ by Corollary 6.3.10. Then by Corollary 7.1.5, $HS_2(P_4, ..., P_8) = S_3(P_1, ..., P_9)$ is a bijective linear transformation that applies H to $S_2(P_4, ..., P_8)$ so that the dimension of the image must equal the dimension of the range. So, $\dim S_3(P_1, ..., P_9) = 1$, which implies that $\dim S_3(P_1, ..., P_8) \leq 2$ since we add at most one dimension by taking away one point. Consequently, $\dim S_3(P_1, ..., P_8) = 2$.

Case 2b: Suppose $P_1, ..., P_6$ all lie on a non-degenerate conic, $C : (Q = 0)$. Since $|K| \geq 6$, choose $P_9 \in C$ distinct from $P_1, ..., P_6$. By Corollary 7.1.5 we have $S_3(P_1, ..., P_9) = QS_1(P_7, P_8)$ where $\dim S_1(P_7, P_8) = 1$ since we know that only one line passes through 2 points. Furthermore, since $QS_1(P_7, P_8) = S_3(P_1, ..., P_9)$ is a bijective linear transformation from $S_1(P_7, P_8)$ to $S_3(P_1, ..., P_9)$, the dimension of the image and the domain are the same. Thus $\dim S_3(P_1, ..., P_9) = 1$ which implies that $\dim S_3(P_1, ..., P_8) \leq 2$. Consequently, $S_3(P_1, ..., P_8) = 2$. □

Corollary 7.1.7. *Assume $|K| \geq 9$. Let C_1, C_2 be two cubic curves with $C_1 \cap C_2 = \{P_1, ..., P_9\}$, all distinct points. Then a cubic D passing through $P_1, ..., P_8$ also passes through P_9.*

Proof. If 4 of the points $P_1, ..., P_8$ were collinear on line L, then each of C_1 and C_2

would meet L in ≥ 4 points, and thus contain L by the contrapositive of Theorem 6.3.5, which contradicts the assumption that $C_1 \cap C_2 = \{P_1, ..., P_9\}$. Suppose 7 of the points $P_1, ..., P_7$ are conconic on D; then each of C_1, C_2 would meet D in ≥ 7 points, and thus contain D again by Theorem 6.3.5 which contradicts the assumption on $C_1 \cap C_2$ again. Thus there are no 4 points of $P_1, ..., P_8$ collinear and no 7 points of $P_1, ..., P_8$ conconic, so by Theorem 7.1.6 we can conclude that $\dim S_3(P_1, ..., P_8) = 2$. Let F_1, F_2 be the equations of C_1, C_2, then $\{F_1, F_2\}$ is a basis of $S_3(P_1, ..., P_8)$ since $\dim S_3(P_1, ..., P_8) = 2$. Let G be the equation of D, where $G = \lambda F_1 + \gamma F_2$. Now $F_1(P_9) = F_2(P_9) = 0$, so $G(P_9) = 0$ and therefore D also passes through P_9. □

7.2 Cubics and the Group Law

We follow the construction of the group law in section 2.8 of Reid's book [Rei90]. Our goal is to define an abelian group operation on an elliptic curve E but we will need to establish some facts and assumptions first.

Definition 7.2.1. Let $C : \{(x, y) \in A^2(K) | f(x, y) = 0\}$ and $P = (a, b)$ be a point on C. Then P is a **singular point** if $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Theorem 7.2.2. *Let $C : \{(x, y) | f(x, y) = 0\}$ in $A^2(K)$. Let $P = (a, b)$ be a point on C, so that $f(P) = 0$. Let*

$$L = \{(x, y) | f_x(P)(x - a) + f_y(P)(y - b) = 0\}. \tag{7.10}$$

If P is a nonsingular point of C, then P is a repeated root of $f|_L = 0$.

Proof. Let $c = f_x(a, b)$ and $d = f_y(a, b)$. So at least one of c or d is nonzero. Without loss of generality we can assume $d \neq 0$. Then $L : c(x - a) + d(y - b) = 0$ and $y = -\frac{c}{d}(x - a) + b$ for any $(x, y) \in L$. Let $g(x) = f|_L = f\left(x, -\frac{c}{d}(x - a) + b\right)$.

Then $g(a) = f(a,b) = 0$, since $P = (a,b)$ is on C and $f(P) = 0$. Next we compute by Lemma 2.1.3

$$g'(x) = f_x\left(x, -\frac{c}{d}(x-a)+b\right) + f_y\left(x, -\frac{c}{d}(x-a)+b\right)\left(-\frac{c}{d}\right). \tag{7.11}$$

Then

$$\begin{aligned} g'(a) &= f_x\left(a, -\frac{c}{d}(a-a)+b\right) + f_y\left(a, -\frac{c}{d}(a-a)+b\right)\left(-\frac{c}{d}\right) \\ &= f_x(a,b) + f_y(a,b)\left(-\frac{c}{d}\right) \end{aligned} \tag{7.12}$$

where $c = f_x(a,b)$ and $d = f_y(a,b)$. Thus

$$g'(a) = f_x(a,b) + f_y(a,b)\left(-\frac{f_x(a,b)}{f_y(a,b)}\right) = 0. \tag{7.13}$$

So $(x-a)$ is a common factor for $g(x)$ and $g'(x)$ and thus a is a repeated root of $g(x) = f|_L$ by Theorem 2.1.12. □

Definition 7.2.3. Let $C : \{(x,y) \in A^2(K) | f(x,y) = 0\}$ and let $P = (a,b) \in C$. If P is nonsingular then the **tangent line to the curve** C **at** P is given by

$$T_P(C) = f_x(P)(x-a) + f_y(P)(y-b) = 0. \tag{7.14}$$

Theorem 7.2.4. *Let K be an algebraically closed field not of characteristic 2. If $C : f(x,y) = y^2 - g(x) = 0$, with $g(x)$ a cubic in x and with no repeated roots, then every point on $C \in A^2(K)$ is nonsingular.*

Proof. The proof is by contradiction. Suppose (a,b) is singular. Then $f_y(x,y) = 2y$ and $f_y(a,b) = 2b = 0$, so $b = 0$. So $f(a,b) = f(a,0) = -g(a) = 0$. Furthermore, $f_x(x,y) = -g'(x)$ and $f_x(a,b) = -g'(a) = 0$, since (a,b) is singular. Now $x = a$ is a zero of $g(x)$ and of $g'(x)$, thus a is a repeated root of $g(x)$ by Theorem 2.1.12. Therefore, if $g(x)$ has no repeated roots, then every point is nonsingular. □

To pursue defining the group of an elliptic curve E we continue with the following definition.

Definition 7.2.5. Assume K is an algebraically closed field not of characteristic 2 and let $g(x)$ be a cubic with no repeated roots where $g(x) = ax^3 + bx^2 + cx + d$. Let $f(x, y) = y^2 - g(x)$ and $F(X, Y, Z) = ZY^2 - aX^3 - bZX^2 - cZ^2X - dZ^3$ be irreducible cubics so that neither contains a line or a conic. Suppose $E \in K[X, Y, Z]$ is a cubic form defining a nonempty plane curve $C : (E = 0) \subset P^2(K)$. Then the set

$$\begin{aligned} E &= \{(X : Y : Z) \in P^2(K) | F(X, Y, Z) = 0\} \\ &= \{(x, y) \in A^2(K) | f(x, y) = 0\} \cup \{(0 : 1 : 0)\} \end{aligned} \tag{7.15}$$

is called an **elliptic curve**.

Now we define our zero element, $\mathcal{O}$.

Definition 7.2.6. Let $\mathcal{O} = (0 : 1 : 0)$ be the point at infinity on E. Then we define $T_{\mathcal{O}}(E) = L_\infty$ (see Chapter 5, Section 1). We note that $F|_{L_\infty}$ has a triple root at $\mathcal{O}$.

Next we define the addition of points on E by first defining the third point of intersection and its negative.

Definition 7.2.7. Let $L_{PQ}(E)$ denote the third point on $L_{PQ} \cap E$. That is:

(1) If $P \neq Q, L_{PQ} \neq T_P E, L_{PQ} \neq T_Q E$, then there exists a genuine third point on E that is neither P nor Q and which we define as $L_{PQ}(E)$.

(2) If $P \neq Q, L_{PQ} = T_P E$, then we define $L_{PQ}(E) = P$.

(3) If $P = Q, L_{PQ} = T_P E = T_Q E$, then either there exists a genuine third point that we define as $L_{PQ}(E)$ and is neither P nor Q, or $P = Q = \mathcal{O}$ in which case we define $L_{PQ}(E) = \mathcal{O}$.

Definition 7.2.8. For any point P we define $-P = L_{P\mathcal{O}}(E)$.

Theorem 7.2.9. *If $P = (a, b) = (a : b : 1) \in E$ then $-P = (a, -b)$. Furthermore, $-\mathcal{O} = \mathcal{O}$.*

Proof. We observe that

$$\begin{aligned} L_{P\mathcal{O}} &= \{(X : Y : Z) | X - aZ = 0\} \\ &= \{(x, y) | x = a\} \cup \{(0 : 1 : 0)\}. \end{aligned} \tag{7.16}$$

Then $L_{P\mathcal{O}} \cap E = \{(a, y) | f(a, y) = y^2 - g(a) = 0\}$. This implies that $b^2 = g(a)$ and so $(-b)^2 = g(a)$ as well. Thus the points on $L_{P\mathcal{O}} \cap E$ are

$$\{P = (a, b), \mathcal{O}, -P = (a, -b)\}. \tag{7.17}$$

□

Corollary 7.2.10. *For any point P, we have $-(-P) = P$.*

Proof. We know $L_{P\mathcal{O}}(E) = -P$. Furthermore, the point $-(-P) = L_{(-P)\mathcal{O}}(E)$, but $L_{(-P)\mathcal{O}} = L_{P\mathcal{O}}$ by definition of $-P$. Thus, $L_{(-P)\mathcal{O}}(E) = P$. □

Finally we arrive at our definition of a sum of two points on a conic.

Definition 7.2.11. We define an operation "+" on E by $P + Q = -L_{PQ}(E)$.

The construction of $(A + B) + C$ above can be seen in the graphic that follows.

Theorem 7.2.12. *The set $(E, +)$ is an abelian group where the identity is $\mathcal{O}$ and the inverse of P is $-P$.*

Proof. We need to check for associativity, identity, inverses, and commutativity under the operation in E [Gal06]. Let P and Q be points on E.

(1) $P + Q = -L_{PQ}(E) = -L_{QP}(E) = Q + P$, so $(E, +)$ is commutative.

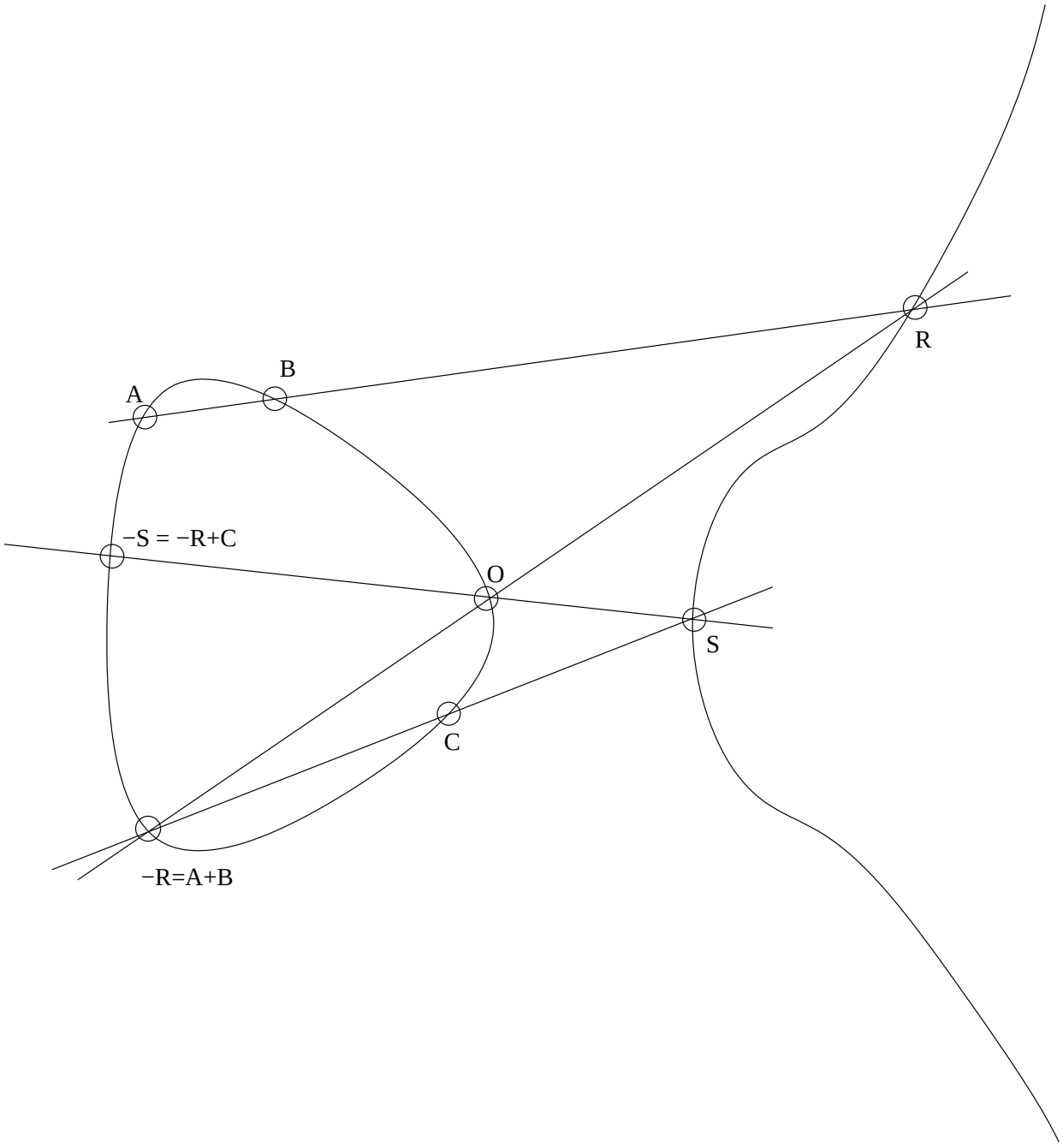

Figure 7.1: Cubic curve and its group law

(2) $P + \mathcal{O} = -L_{P\mathcal{O}}(E) = -(-P)$, by definition of $-P$ and thus $-(-P) = P$, so $\mathcal{O}$ is the identity element.

(3) $-P + P = -L_{(-P)P}(E) = -\mathcal{O}$ by definition of $-P$ as the third point on the line through $\mathcal{O}$ and P, and thus $-\mathcal{O} = \mathcal{O}$. So each point has an additive inverse.

(4) Proof of Associativity for a special case follows (for a complete proof see Silverman [Sil86]). Let A, B and C be points on E. We begin by constructing $(A + B) + C$.

- Let $L_{AB}(E) = R$.
- Then $L_{R\mathcal{O}}(E) = -R = A + B$, by definition.
- Now let $L_{(-R)C}(E) = S$.
- Then $L_{S\mathcal{O}}(E) = -S = (A + B) + C$.

Next we construct $A + (B + C)$.

- Let $L_{BC}(E) = Q$.
- Then $L_{Q\mathcal{O}}(E) = -Q = B + C$, by definition.
- Now let $L_{(-Q)A}(E) = T$.
- Then $L_{T\mathcal{O}}(E) = -T = A + (B + C)$.

So we need to show that $-S = -T$, but it is sufficient to show that $S = T$.

Let $D_1 = L_{AB} \cup L_{Q\mathcal{O}} \cup L_{(-R)C}$ and $D_2 = L_{BC} \cup L_{R\mathcal{O}} \cup L_{(-Q)A}$. The equations of 3 lines multiplied together yield a cubic, so D_1, D_2 are cubics. Then

$$E \cap D_1 = \{A, B, R, Q, \mathcal{O}, -Q, -R, C, S\} \tag{7.18}$$

and

$$E \cap D_2 = \{B, C, Q, R, \mathcal{O}, -R, -Q, A, T\} \tag{7.19}$$

where these are the only possible points in $E \cap D_1$ and $E \cap D_2$ because F is irreducible so $E \cap D_2 = (E \cap L_{BC}) \cup (E \cap L_{R\mathcal{O}}) \cup (E \cap L_{(-Q)A})$, and similarly $E \cap D_1 = (E \cap L_{AB}) \cup (E \cap L_{Q(O)}) \cup (E \cap L_{(-R)C})$. Note that the first 8 points of each intersection are distinct and in common (here we make our assumption that the 9 points of the first intersection are distinct), so by Corollary 7.1.7, D_2 passes through the 9th point as well. That is, $S = T$ and associativity is true. □

Corollary 7.2.13. *Let k be a subfield of K. Let E be an elliptic curve defined by $y^2 = ax^3 + bx^2 + cx + d$ with $a, b, c, d \in k$. Then*

$$E(k) = \{(x, y) \in k^2 | y^2 = ax^3 + bx^2 + cx + d\} \cup \{\mathcal{O}\} \tag{7.20}$$

is a subgroup of $E(K)$.

Recall that a subset H of a group G is a subgroup of G if and only if H is closed under the operation of G, the identity element of G is in H, and for any element in H it is true that its inverse is also in H [Fra03].

Proof. We first show that if $P \in E(k)$, then $-P \in E(k)$. Let $y \in k$; then $-y \in k$ since k is a field. So if $P = (x, y) \in E(k)$, then $-P = (x, -y) \in E(k)$ since $(-y)^2 = ax^3 + bx^2 + cx + d$. Also, $(O) \in E(k)$ by definition.

We now want to show that if $P, Q \in E(k)$ then their third point of intersection on E, $L_{PQ}(E)$, is also in $E(k)$, since it then follows that $P + Q = -L_{PQ}(E) \in E(k)$.

CASE 1: Suppose $P \neq Q$ such that $P, Q \in E(k)$. Let $P = (\ell, m)$ and $Q = (n, p)$. Note that if $\ell = n$ then $L_{PQ}(E) = \mathcal{O} \in E(k)$. Otherwise,

$L_{PQ} : y = p + \dfrac{m-p}{\ell - n}(x - n)$ and

$$L_{PQ} \cap E : \left(p + \frac{m-p}{\ell - n}(x - n)\right)^2 = ax^3 + bx^2 + cx + d. \tag{7.21}$$

So $L_{PQ} \cap E$ is the solution set to

$$0 = ax^3 + bx^2 + cx + d - \left(p + \frac{m-p}{\ell - n}(x - n)\right)^2, \tag{7.22}$$

a cubic equation in x. Since we know that $(x = \ell)$ and $(x = n)$ lie in the intersection they must be roots of the equation so we may factor them out and this leaves a third root, $(x = r)$, with r in k. Furthermore, we substitute $x = r$ into $y = p + \dfrac{m-p}{\ell - n}(x - n)$ to get our y-coordinate and so $y \in k$ as well.

CASE 2: Suppose $P = Q$, then L_{PQ} is the tangent line to the curve E at P which is given by the equation $0 = f_x(\ell, m)(x - \ell) + f_y(\ell, m)(y - m)$, where $f_x(\ell, m) = 3a\ell^2 + 2b\ell + c \in k$ and $f_y(\ell, m) = 2m \in k$. This equation can be simplified to one of the form $y = \lambda x + v$ for some $\lambda, v \in k$. Substituting into the equation for E we find that $L_{PQ} \cap E$ has a double root at point $P = Q$ so we may factor out $(x = \ell)$ twice and we are left with a third root, $(x = r)$, with $r \in k$. Again, we substitute $x = r$ into $y = p + \dfrac{m-p}{\ell - n}(x - n)$ to get our y-coordinate and so $y \in k$ as well.

So the third point $L_{PQ}(E) = (r, y)$ in $L_{PQ} \cap E$, is also in $E(k)$.

Thus $E(k)$ is a subgroup of $E(K)$. □

Example 7.2.14. Let E be the elliptic curve defined by $y^2 = x^3 - x$ over the field F_{11} (i.e., $E = E(F_{11})$). Note that the square numbers (mod 11) are $0^2 = 0, 1^2 = 1, 2^2 = 4, 3^2 = 9, 4^2 = 5$ and $5^2 = 3$. Computation shows that the values of x that give us a perfect square on the right side of the equation, (mod 11), are

$x = 0, 1, -1, -2, -3, 4, -5$, which yield the 12 points on E:

$$\begin{gathered}(0,0),(1,0),(-1,0),(-2,4),(-2,-4),(-3,3),\\(-3,-3),(4,4),(4,-4),(-5,1),(-5,-1),\mathcal{O}.\end{gathered} \tag{7.23}$$

Example 7.2.15. Let E be the elliptic curve defined by $y^2 = x^3 - x$ over the field F_{11}. Let $A = (0,0), B = (1,0)$ and $C = (-5,1)$. We will illustrate the associativity of points with points A, B, and C. Note that $L_{AB} : y = 0$, so

$$L_{AB} \cap E : 0^2 = x^3 - x = x(x^2 - 1), \tag{7.24}$$

where $x = 0, 1, -1$ are the three roots to this equation. However, the first two roots correspond to points A and B, and thus the third root yields the third point of $L_{AB} \cap E$. That is, $L_{AB}(E) = (-1, 0)$. Thus $A + B = -L_{AB}(E) = (-1, 0)$. Now

$$L_{(A+B)C} : y = \frac{-1}{4}(x+5)+1 = -3(x+5)+1 = -3x-14 = -3x-3 \pmod{11}, \tag{7.25}$$

so

$$L_{(A+B)C} \cap E : (-3(x+1))^2 = x^3 - x = x(x^2 - 1). \tag{7.26}$$

Simplifying both sides of this equation we see that $(x+1)(9x+9) = (x^2 - x)(x+1)$ and thus $0 = (x+1)(x^2 - 10x - 9)$ where after division by $(x+5)$ we note that $x = -1, -5, 4$ are the three roots to the cubic equation $L_{(A+B)C} \cap E$. However, the first two roots correspond to points $A + B$ and C, respectively, so the third root yields the third point of $L_{(A+B)C} \cap E$. That is, since $y = -3(4) - 3$, $L_{(A+B)C}(E) = (4, -4)$. Thus,

$$(A + B) + C = -L_{(A+B)C}(E) = (4, 4). \tag{7.27}$$

Next we will calculate $A + (B + C)$. Note that

$$L_{BC} : y = \frac{1}{-6}(x - 1) = -2(x - 1) \pmod{11}, \tag{7.28}$$

so

$$L_{BC} \cap E : (-2(x-1))^2 = x^3 - x = x(x^2-1), \tag{7.29}$$

Simplifying both sides we arrive at $(x-1)(4x-4) = (x^2+x)(x-1)$ and finally $0 = (x^2-3x+4)(x-1)$ where upon dividing by $(x+5)$ we note that $x = 1, -5, -3$ are the three roots to the cubic $L_{BC} \cap E$. However, the first two roots correspond to points B and C, and thus the third root yields the third point of $L_{BC} \cap E$. That is, since $y = -2(-3-1)$, $L_{BC}(E) = (-3,-3)$. Thus $B+C = -L_{BC}(E) = (-3,3)$. Now

$$L_{(B+C)A} : y = -x, \tag{7.30}$$

so

$$L_{(B+C)A} \cap E : (-x)^2 = x(x^2-1). \tag{7.31}$$

Simplifying both sides of this equation we see that $0 = (x^2-x-1)x$ where after division by $(x+3)$ we note that $x = 0, -3, 4$ are the three roots to the cubic $L_{(B+C)A} \cap E$. However, the first two roots correspond to points $B+C$ and A, respectively, so the third root yields the third point of $L_{(B+C)A} \cap E$. That is, since $y = -4$, $L_{(B+C)A}(E) = (4,-4)$. Thus,

$$A + (B+C) = -L_{(B+C)A}(E) = (4,4). \tag{7.32}$$

So we have illustrated that $(A+B)+C = A+(B+C)$.

Example 7.2.16. Let E be the elliptic curve defined by $y^2 = x^3 - x$ over the field F_{11}. Let $P = (4,4)$. Then $T_P(E) : -3(x-4) + 8(y-4) = 0$ or $y = \dfrac{3x-2}{8}$. Furthermore, $L_{PP} \cap E : \left(\dfrac{3x-2}{8}\right)^2 = x^3 - x$. Simplifying this equation we arrive at $0 = 2(x^3 - x^2 + 4x + 2)$ and after division by $(x-4)$ twice we note that the third root is also $(x = 4)$. Thus since $y = \left(\dfrac{3(4)-2}{8}\right)$, $L_{PP}(E) = (4,4)$ and so $P+P = (4,-4) = -P$. So we have that $(P+P)+P = (-P)+P = \mathcal{O}$, that is, the order of P is 3.

7.3 Elliptic Curve Analog of the Diffie-Hellman Key Exchange

In the present section we refer to the presentation by Koblitz [Kob06] on the application of the Diffie-Hellman Key Exchange to elliptic curves. Recall that the Diffie-Hellman method enables us to exchange a secret key over an insecure channel where everyone can listen, even an eavesdropper, and that the security of the Diffie-Hellman Key Exchange depends on the difficulty of solving the discrete logarithm problem for large numbers.

Let E be an elliptic curve defined over a finite field F_q.

Definition 7.3.1. The **discrete log Diffie-Hellman assumption on** E: If E is an elliptic curve over F_q, B is a point on E, and P is another point on E that is known to be equal to xB for some integer x, then for large q and suitable E, it is computationally infeasible to find x such that $xB = P$, where $xB = B + B + ... + B$ (x times).

However, we must recall that the Diffie-Hellman method is very slow when compared to a classical cryptosystem such as a shift cypher, so it is most effective when used to share an encoding key for a classical cryptosystem.

In the Diffie-Hellman analogue for elliptic curves it is assumed that F_q and E is public knowledge. Suppose that Aarón and Clarissa want to agree on a key with which to encode their future correspondence so that Isabelle, the evesdropper, cannot decipher their messages.

(1) First, Aarón and Clarissa publicly choose a point $B \in E$, that is preferably a generator of E but not necessarily, so long as the subgroup generated by B is large.

(2) Then, Aarón selects a random integer a and makes public aB while keeping

a secret.

(3) Also, Clarissa selects a random integer c and makes public cB while keeping c a secret.

(4) Aarón can now use the public cB to compute the key $a(cB) = acB$.

(5) Similarly, Clarissa uses the public aB to compute the same key $c(aB) = acB$.

(6) Finally, Isabelle, the third party eavesdropper, is only aware of q, E, B, aB and cB and according to the discrete log problem, it is computationally infeasible to compute acB given only this information.

(7) Thus, Aarón and Clarissa can agree to encrypt their correspondence to each other with the key acB.

Given the nature of elliptic curves and their group law, the Diffie-Hellman Key Exchange system is very secure and utilized for maintaining privacy over the internet.

BIBLIOGRAPHY

[CER] CERTICOM, *ECC tutorial: Elliptic curve groups and the discrete logarithm problem*, http://certicom.com/.

[Edg06] J. Edge, *An introduction to elliptic curve cryptography*, http://lwn.net/Articles/174127/, 2006.

[Fra03] J. Fraleigh, *A first course in abstract algebra*, 7th ed., Addison Wesley, 2003.

[Ful08] W. Fulton, *Algebraic curves: An introduction to algebraic geometry*, William Fulton, 2008.

[Gal06] J. Gallian, *Contemporary abstract algebra*, 6th ed., Houghton Mifflin, 2006.

[Hun97] T. Hungerford, *Abstract algebra, an introduction*, 2nd ed., Brookes/Cole, 1997.

[Kob06] N. Koblitz, *A course in number theory and cryptography*, 2nd ed., Springer, 2006.

[Rei90] M. Reid, *Undergraduate algebraic geometry*, Cmbridge University Press, 1990.

[Sam88] P. Samuel, *Projective geometry*, Springer-Verlag, 1988.

[SF03] L E. Spence S. Friedberg, A. J. Insel, *Linear algebra*, 4th ed., Prentice Hall, 2003.

[Sil86] J. H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, 1986.