Elliptic curves over finite fields and their pairings

Matan Prasma

December 22, 2022

Contents

1 Introduction

These notes grew as part of a math seminar I gave in Aragon Association during 2022. Since the construction of Miller's algorithm [Mil], the Cryptography community started to use Elliptic curves and their pairing extensively. By now, many publicly available code libraries allow one to efficiently compute Elliptic curves over finite fields and evaluate their pairings. However, compared to Machine Learning, where the mathematical pre-requisites consist of Linear Algebra, Calculus and basic Statistics, Elliptic curves require more background and are usually taught at a master level in pure Mathematics. This state of affairs poses a challenge to engineers and others who wish to understand the mathematical building blocks.

To assist overcoming the challenge mentioned above, these notes aim to give a self-contained, rigorous and elementary account of most of the material required for pairing-based Cryptography. I collected material from several standard sources, and sometimes formulated elementary arguments to replace non-elementary explanations I found in the literature. In particular, I completely avoid relying on Galois Theory or Algebraic Geometry unlike most textbooks on the subject.

I'd like to thank Amir Taaki, Alex Kampa, Artem Grigor, Roger Baig and Arnaucube for many useful comments during the writing of this manuscript.

2 Naive Set Theory

2.1 Sets and functions

Slogan. *Sets are the machine code of modern Mathematics.*

On a fundamental level, modern Math is built on Set Theory. From that point of view, a **Set** S is a collection of elements such that for every object x in

our 'universe' we can determine whether x is an element of S, denoted $x \in S$ or that x is not an element of S, denoted $x \notin S$.

When we want to specify the elements of a set S, we do so with curly brackets and commas separating between elements e.g. $S = \{a, b, c\}$. If S has finite number of elements (or just 'finite') we denote by $|S|$ the number of elements of S. Of course, S need not be finite, and in this case, we need a rule in order to specify the elements of S, e,g. $S = \{n | n \geq 2\}$ or if the rule is obvious, we can write $S = \{2, 3, 4, ...\}$. For sets A, B we write $A \subseteq B$ if $\forall a \in A, a \in B$ and say that A is included in B. The basic operations on sets include **union**

$$A \cup B = \{x | x \in A \vee x \in B\},$$

intersection

$$A \cap B = \{x | x \in A \wedge x \in B\}$$

and **complement** (or subtraction)

$$A \setminus B = \{x | x \in A \wedge x \notin B\}.$$

Remark 2.1. More generally, let I be a set that we refer to as an 'index set'. Suppose that for every $i \in I$ we are given a set U_i. Then we can form the union

$$\bigcup_{i \in I} U_i = \{x | \exists i \in I : x \in U_i\}$$

and the intersection

$$\bigcap_{i \in I} U_i = \{x | \forall i \in I : x \in U_i\}.$$

Our fundamental assumption is that there exist a special set, called the **empty set** and denoted $\varnothing$ that has no elements. More formally, we can write

$$\varnothing = \{x | x \neq x\}$$

and observe that for every set A we have $\varnothing \subseteq A$. Using the empty set, we can in fact define all natural numbers as follows:

$$0 := \varnothing,$$

$$1 := \{\varnothing\},$$

$$2 := \{\varnothing, \{\varnothing\}\} = \{0, 1\},$$

...,

$$n := \{0, 1, ..., n-1\}.$$

Example 2.2. Some special sets and their notation are given below.

1. $\mathbb{N} = \{1, 2, 3, ...\}$ the natural numbers.
2. $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$ the integers.
3. $\mathbb{Q} = \{\frac{a}{b} | a, b \in \mathbb{Z} \wedge b \neq 0\}$ the rational numbers.

4. $\mathbb{R}$ the real numbers.

The elements of a set have no particular order and we remove duplicated elements so that for example $\{a,b,c\} = \{a,c,a,b\}$. In case we wish to talk about **ordered elements** we can do the following: Given to objects, a,b we can consider the set of two elements

$$O_{ab} = \{\{a\},\{a,b\}\}.$$

If a',b' are any two other elements, we have that

Observation 2.3. $O_{ab} = O_{a'b'}$ if and only if $a = a'$ and $b = b'$.

We refer to the set O_{ab} as the **ordered pair** of a,b and denote it $(a,b) := O_{ab}$. The definition of O_{ab} could have been different and we gave the one above as a convention. In fact all we need from it is the property stated in 2.3.

Using the notion of ordered pairs we can make the following

Definition 2.4. Given two sets A,B, the **Cartesian product** of A and B is the set

$A \times B := \{(a,b) | a \in A, b \in B\}$.

For example, if $A = \{0,1\}$ and $B = \{1,2\}$ then

$$A \times B = \{(0,1),(0,2),(1,1),(1,2)\}.$$

As another example, if $A = B = \mathbb{R}$, then

$$A \times B = \{(x,y) | x,y \in \mathbb{R}\} =: \mathbb{R}^2$$

i.e. the two dimensional plane aka the X-Y plane.

Example 2.5. Let

$$A = [2,4] = \{x \in \mathbb{R} | 2 \leq x \leq 4\} \subseteq \mathbb{R}$$

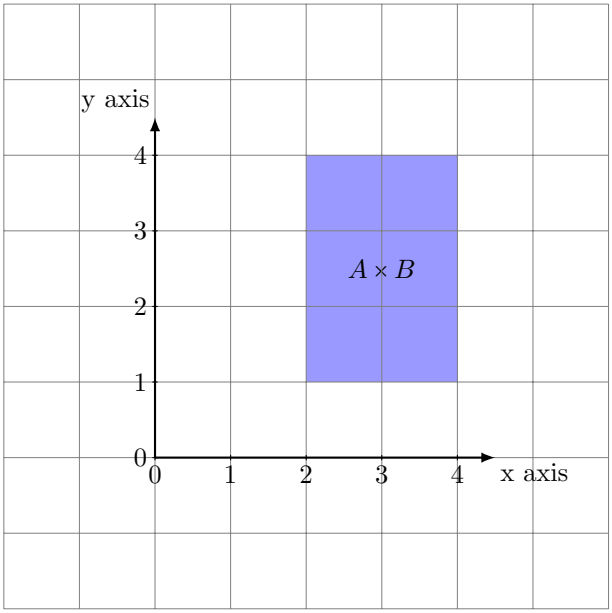
and

$$B = [1,4] = \{x \in \mathbb{R} | 1 \leq x \leq 3\} \subseteq \mathbb{R}$$

be a pair of intervals.
Then

$$A \times B \subseteq \mathbb{R}^2$$

can be depicted as follows:



In order to 'move' between sets, we need functions. Given sets A, B a function f from A to B is a rule that assigns for each $a \in A$ a unique element in B, denoted $f(a)$. We denote such a function by $f : A \longrightarrow B$. Informally speaking, we can think of the function f as a machine with set of inputs A and a set of possible outputs B; the machine f assigns to each input $a \in \mathcal{A}$ a unique output $f(a) \in B$. But what is a rule? to give a more formal definition we go as follows:

Definition 2.6. A function $f : A \longrightarrow B$ is a subset $f \subseteq A \times B$ that satisfy two properties:

1. for all $a \in A$, there is $b \in B$, also denoted as $b := f(a)$ such that $(a, b) \in f$ (so that f is defined on all elements in A).
2. if $(a, b) \in f$ and $(a, b') \in f$ then $b = b'$ (so that f gives a unique element in B for every element in A).

The set A is called the **domain** of f and the set B is the **range** of f.

Example 2.7. Let $A = \varnothing$ and B any set. Then $A \times B = \varnothing$ so that there can be at most one function $f : \varnothing \longrightarrow B$, namely the one corresponding to $\varnothing \subseteq A \times B$. One can see that the conditions of Definition 2.6 are vacantly satisfied. This function is called the empty function. On the other hand, if $A \neq \varnothing$ and $B = \varnothing$ we have again $A \times B = \varnothing$ but now $\varnothing \subseteq A \times B$ is not a function $A \longrightarrow B$ since condition 1 of Definition 2.6 is not satisfied (since there is at least one element a in A).

We say two functions $A \underset{f'}{\overset{f}{\rightrightarrows}} B$ (with the same domain and range) are

equal, denoted $f = f'$, if for every $a \in A$, $f(a) = f(a')$ or in other words if $f = f' \subseteq A \times B$ as sets.

Definition 2.8. Suppose $f : A \longrightarrow B$ and $g : B \longrightarrow C$ are functions. We define their **composition** $g \circ f$ to be the function $g \circ f : A \longrightarrow C$ specified by the rule

$$(g \circ f)(a) \coloneqq g(f(a))$$

for all $a \in A$.

Exercise 2.9. In light of the definition above...

1. Prove that $g \circ f$ is indeed a function.
2. Suppose in addition that $h : C \longrightarrow D$ is another function. Prove that there is an equality of functions $h \circ (g \circ f) = (h \circ g) \circ f$.

2.2 Isomorphisms

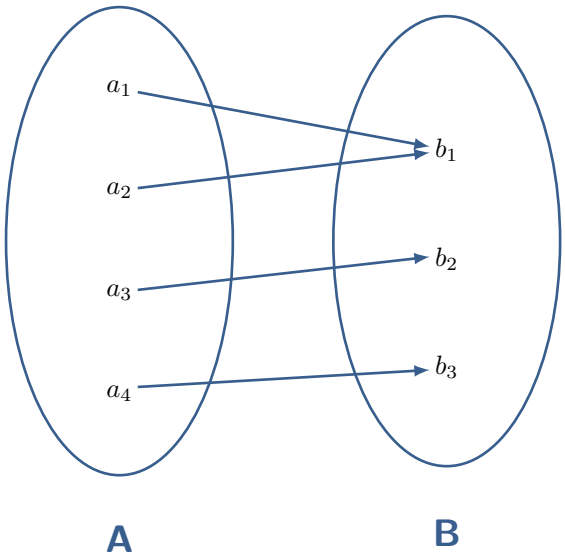
Definition 2.10. Let A, B be (possibly infinite) sets. A function $f : A \longrightarrow B$ is called:

1. **monomorphism** if for all $a \neq a' \in A$, we have $f(a) \neq f(a') \in B$.
2. **epimorphism** if for every $b \in B$, there exists some $a \in A$ such that $f(a) = b$.
3. **isomorphism** if it is monomorphism and epimorphism.

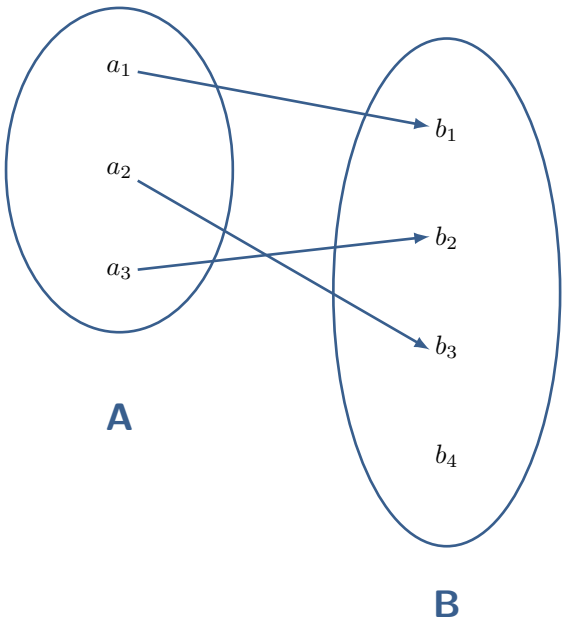
Remark 2.11. It also customary to call a monomorphism an 'injection' or 'one-to-one' and an epimorphism a 'surjection' or 'onto'. A synonym for isomorphism is 'bijection'.

The intuition for Definition 2.10 often comes from the case of finite sets as the following examples demonstrate.

Example 2.12. The function below is epimorphism but not monomorphism:



Example 2.13. The function below is monomorphism but not epimorphism:



In light of the examples above, we can formulate the following

Proposition 2.14. *Let A, B be finite sets.*

1. *There exists a monomorphism function $f : A \longrightarrow B$ iff $|A| \leq |B|$.*
2. *There exists an epimorphism function $f : A \longrightarrow B$ iff $|A| \geq |B|$.*

3. *There exists an isomorphism* $f : A \longrightarrow B$ *iff* $|A| = |B|$

Proof. Left for the reader. □

Proposition 2.15. *Let* A, B, C *be sets and let* $f : A \longrightarrow B$ *and* $g : B \longrightarrow C$ *be two functions.*

1. *if* f *and* g *are monomorphisms, then so is* $g \circ f$.
2. *if* f *and* g *are epimorphisms, then so is* $g \circ f$.
3. *if* f *and* g *are isomorphisms, then so is* $g \circ f$.

Proof.

1. to show that $g \circ f : A \longrightarrow C$ is monomorphism, let $a \neq a' \in A$. Since f is monomorphism, $f(a) \neq f(a')$. Thus, since g is monomorphism, $g(f(a)) \neq g(f(a'))$ and we are done.
2. to show that $g \circ f : A \longrightarrow C$ is epimorphism, let $c \in C$ be an arbitrary element. Since $g : B \longrightarrow C$ is epimorphism, there exists $b \in B$ such that $g(b) = c$. Since $f : A \longrightarrow B$ is epimorphism, there exists $a \in A$ such that $f(a) = b$. It follows that

   $$(g \circ f)(a) = g(f(a)) = g(b) = c$$

   and we are done.
3. this follows from the other two claims.

□

It is useful to notice that an epimorphism (a surjective functions) are "right cancallable", in the following sense:

Lemma 2.16. *Let* A, B *be sets and* $f : A \longrightarrow B$ *an epimorphism. Then for any set* C *and functions* $g, h : B \rightrightarrows C$ *such that* $g \circ f = h \circ f$, *we have* $g = h$.

Proof. Let $b \in B$ be any element. Since f is an epimorphism, there exists $a \in A$ such that $f(a) = b$. Therefore,

$$g(b) = g(f(a)) = h(f(a)) = h(b).$$

Since b was arbitrary, $g = h$. □

Exercise 2.17. Dually to Lemma 2.16, a monomorphism (an injective function) between two sets is "left cancallable". Formulate the precise statement that expresses this idea and prove it.

Let A be a set. Then the function $A \longrightarrow A$ given by $a \mapsto a$ is called the **identity function** of A and denoted as id_A. We have:

Theorem 2.18. *Let $f : A \longrightarrow B$ be an isomorphism. Then there is a function $f^{-1} : B \longrightarrow A$ such that $f \circ f^{-1} = \mathrm{id}_B$ and $f^{-1} \circ f = \mathrm{id}_A$.*

Remark 2.19. The function f^{-1} in Theorem 2.18 is called the inverse of f.

Proof. Let $b \in B$ be an arbitrary element. Since f is an epimorphism, there exists $a \in A$ such that $f(a) = b$. We claim that there is a unique $a \in A$ with that property: otherwise, there would be $a' \neq a \in A$ such that $f(a) = b = f(a')$ which would contradict the assumption that f is a monomorphism. Since that a is unique, we define $f^{-1}(b) = a$ and we obtain a function $f^{-1} : B \longrightarrow A$. By the definition of f^{-1}, we see that for any $a \in A$,

$$(f^{-1} \circ f)(a) = f^{-1}(f(a)) = a$$

so that $f^{-1} \circ f = \mathrm{id}_A$. By that same definition, for any $b \in B$, $f^{-1}(b) = a$ where $a \in A$ is the unique element such that $f(a) = b$. Thus, for any $b \in B$,

$$(f \circ f^{-1})(b) = f(f^{-1}(b)) = f(a) = b$$

so that $f \circ f^{-1} = \mathrm{id}_B$. □

Example 2.20.

1. Let $A = \{1, 2, 3\}$, $B = \{a, b, c\}$ and suppose $f : A \longrightarrow B$ is given by $f(1) = c$, $f(2) = b$, $f(3) = a$. Then f is an isomorphism and its inverse $f^{-1} : B \longrightarrow A$ is given by $f^{-1}(a) = 3$, $f^{-1}(b) = 2$, $f^{-1}(c) = 1$. It is easy to check that $f \circ f^{-1} = \mathrm{id}_B$ and $f^{-1} \circ f = \mathrm{id}_A$.

2. Let $A = [0, 1]$ and $B = [3, 5]$ and define $f : [0, 1] \longrightarrow [3, 5]$ by $f(x) = 2x + 3$. Then f is an isomorphism whose inverse $f^{-1} : [3, 5] \longrightarrow [0, 1]$ is given by $f^{-1}(y) = \frac{y-3}{2}$. Observe that for any $x \in [0, 1]$, $(f^{-1} \circ f)(x) = f^{-1}(f(x)) = f^{-1}(2x+3) = \frac{(2x+3)-3}{2} = x$ so that $f^{-1} \circ f = \mathrm{id}_{[0,1]}$. Similarly $f \circ f^{-1} = \mathrm{id}_{[3,5]}$.

2.3 Equivalence Relations

We saw how one can construct the natural numbers $\mathbb{N}$ as sets. How does one go about defining, for example, the rational numbers as sets? One immediate problem is that a rational number doesn't have a unique representation. For example $\frac{1}{2} = \frac{2}{4} = \frac{3}{6} = \ldots$. Informally speaking, we would like to say that $\frac{1}{2}$ is 'equivalent' to $\frac{2}{4}$ and so on.

Definition 2.21. Let X be a set. A **relation** R on X is a subset $R \subseteq X \times X$. When $(x, y) \in R$ we denote $x \sim_R y$ or $x \sim y$ if R is understood from the context.

Example 2.22. A function $f : X \longrightarrow X$ is a relation on X: $x \sim y$ iff $y = f(x)$.

We are interested in a particular type of relations:

Definition 2.23. A relation R on X is called an **equivalence relation** if it satisfies the following properties:

1. Reflexive: for any $x \in X$, $x \sim x$.

2. Symmetric: for any $x, y \in X$, $x \sim y$ iff $y \sim x$.

3. Transitive: for any $x, y, z \in X$, if $x \sim y$ and $y \sim z$ then $x \sim z$.

Example 2.24. Let $X = \mathbb{Z}$ be the integers and let $n \in \mathbb{N}$ be a natural number. The relation $= (\ mod\ n)$ is an equivalence relation on $\mathbb{Z}$:

1. for any $x \in \mathbb{Z}$, $x = x\ (\ mod\ n)$.

2. for any $x, y \in \mathbb{Z}$, $x = y\ (\ mod\ n)$ iff $y = x\ (\ mod\ n)$.

3. for any $x, y, z \in \mathbb{Z}$, if $x = y\ (\ mod\ n)$ and $y = z\ (\ mod\ n)$ then $x = z\ (\ mod\ n)$

Construction 2.25. Let R be an equivalence relation on X. For $x \in X$ we denote $[x] := \{y \in X | x \sim y\}$ and call it the **equivalence class** of x. Not that if $xa, b \in [x]$, then $a \sim x$ and $b \sim x$. By symmetry, $x \sim b$ and by transitivity $a \sim b$. All elements in $[x]$ are equivalent to each other. It follows that $[x] = [y]$ iff $x \sim y$. If $[x] \bigcap [y] \neq \varnothing$ and $z \in [x] \bigcap [y]$ then $z \sim x$ and $z \sim y$. By symmetry, $x \sim z$ and by transitivity $x \sim y$ so that $[x] = [y]$. In other words, two equivalence classes $[x], [y]$ are either equal or disjoint.

We can consider the collection of equivalence classes of elements of X, $Q = \{[x] | x \in X\}$. Note that this description of Q includes many repetitions since $[x] = [y]$ whenever $x \sim y$. Clearly, every x belongs to some set in Q, namely $[x]$. Thus, the collection Q forms a **partition** of X in that $\bigcup_{x \in X} [x] = X$ and each pair $[x] \neq [y]$ satisfies $[x] \bigcap [y] = \varnothing$.

We will denote

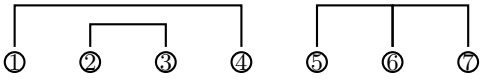
$$X / \backsim := \{[x] | x \in X\}$$

and refer to it as the **quotient set** of X by R. Note that we always have a function $q : X \longrightarrow X / \backsim$ given by $q(x) = [x]$. We call q the **quotient map**. Furthermore, the set of equivalence classes $\{[x] | x \in X\}$ forms a **partition** of X in that $X = \bigcup_{x \in X} [x]$ and any two equivalence classes $[x], [y]$ are either equal or disjoint.

Construction 2.26. Conversely, suppose we have a partition of X, $\{U_i\}_{i \in I}$ (where I is an 'index' set), i.e. $X = \bigcup_{i \in I} U_i$ and for any $i, j \in I$ U_i, U_j are either equal or disjoint. Then we can define an equivalence relation on X by declaring $x \sim y$ if and only if $x, y \in U_i$ for some i. One readily verifies that this is indeed an equivalence relation on X.

Example 2.27. The picture below describes a partition of the set

$$X = \{1, 2, 3, 4, 5, 6, 7\} :$$



This partition corresponds to the equivalence relation given by

$$1 \sim 4,$$

$$2 \sim 3$$

and

$$5 \sim 6 \sim 7.$$

The quotient set is given by definition as

$$X/\backsim = \{[1],[2],[3],[4],[5],[6],[7]\}$$

and after omitting repetitions we get

$$X/\backsim = \{[1],[2],[5]\}.$$

note that we could also write, for example,

$$X/\backsim = \{[4],[3],[6]\}$$

since $[1] = [4],\ [2] = [3],\ [5] = [6]$

Example 2.28. Consider $X = \mathbb{Z}$ with the equivalence relation $= (\ mod\, n)$. The equivalence classes of the quotient set can be represented as

$$\mathbb{Z}/\backsim = \{[0],[1],...[n-1]\}$$

which is a set of size n. One typically denotes

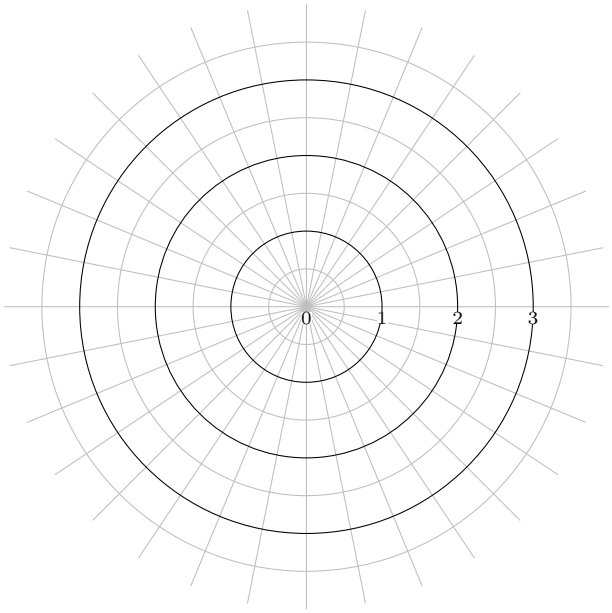
$$\mathbb{Z}/\backsim := \mathbb{Z}_n = \{0,1,...,n-1\}.$$

Example 2.29. Let $I = [0,1]$ be the unit interval and define a relation $\sim$ by: $x \sim x$ for any $x \in I$ and $0 \sim 1$. One easily verifies that this is an equivalence relation with equivalence classes

$$I/\backsim = \{[x] | 0 < x < 1\} \bigcup \{[0]\}.$$

We can identify $I/\sim$ with the circle $S^1 \subseteq \mathbb{C}$ via the map $w : I/\backsim \longrightarrow S^1$ given by $w(x) = e^{2\pi i x}$ (note that $w(0) = w(1)$ so this map is well-defined).

Example 2.30. Define a relation on the plane $\mathbb{R}^2$ by setting $(x,y) \sim (z,w)$ iff $x^2 + y^2 = z^2 + w^2$. It is easy to see that this is an equivalence relation. For example, for any $(x,y) \in \mathbb{R}^2$ we have $(x,y) \sim (x,y)$ since $x^2 + y^2 = x^2 + y^2$ so the relation is reflexive. What are the equivalence classes of $\sim$? Let $P = (x_0, y_0)$ be an arbitrary fixed point and denote $r_0 = x_0^2 + y_0^2$. The equivalence class of P is the set of all points (x,y) such that $x^2 + y^2 = r_0$ namely the circle with centre at the origin $O = (0,0)$ and radius r_0. When $P = O = (0,0)$ we have $r_0 = 0$ and the equivalence class of P is the singeleton $\{(0,0)\}$ (that can be viewed as a circle with radius 0). Note that the collection of all these circles is a partition of the plane $\mathbb{R}^2$ as can be seen below:



Exercise 2.31. Describe the quotient set of the equivalence relation in Example 2.30. Is it isomorphic to some known set?

3 Groups

3.1 Groups and subgroups

One of the most fundamental objects in Mathematics is a Group. The notion of a Group goes back to Galois, who studied solutions to polynomial equations and used groups in order to develop **Galois Theory**.

Definition 3.1. A **Group** consists of the data of a set G with a chosen element $e \in G$ (called unit) together with a binary operation (i.e. a function) $\mu : G \times G \longrightarrow G$ satisfying the following conditions:

1. (unitary) For every $x \in G$, $\mu(x, e) = \mu(e, x) = x$.
2. (associativity) for every $x, y, z \in G$, $\mu(\mu(x, y), z) = \mu(x, \mu(y, z))$.
3. (invertability) for every $x \in G$, there exists an element $x^{-1} \in G$ (called the inverse of x) such that $\mu(x, x^{-1}) = \mu(x^{-1}, x) = e$

We typically denote the binary operation by $\cdot$ although it need not come from a multiplication of numbers. Under this notation, the conditions of Definition 3.1 read:

1. For every $x \in G$, $x \cdot e = e \cdot x = x$.

2. For every $x, y, z \in G$, $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

3. For every $x \in G$, there exists an element $x^{-1} \in G$ such that $x \cdot x^{-1} = e = x^{-1} \cdot x$

Exercise 3.2. Let G be a group. We claim that the unit element e is unique: if there exists an element $e' \in G$ such that for every $x \in G$, $x \cdot e' = e' \cdot x = x$ then since e is a unit element $e \cdot e' = e'$ and since e' is a unit element $e \cdot e' = e$ and it follows that $e = e'$.

In a similar fashion, prove that

1. for any $x \in G$, the inverse x^{-1} is unique: if there exists $\tilde{x}^{-1} \in G$ such that $x \cdot \tilde{x}^{-1} = e = \tilde{x}^{-1} \cdot x$ then $x^{-1} = \tilde{x}^{-1}$.

2.

Examples 3.3. Let us sketch a few common groups.

1. Choose an object e and let $G = \{e\}$. Then G has an obvious group structure and is called the **trivial group**.

2. The integers with addition, $(\mathbb{Z}, +, 0)$ is a group. The inverse of $x \in \mathbb{Z}$ is $-x$.

3. The natural numbers with addition $(\mathbb{N}, +, 0)$ is not a group since the inverse axiom is not satisfied.

4. The integers modulu n $(\mathbb{Z}_n, +\ (\ mod\ n), 0)$ is a group.

5. Let $\mathbb{C} = \{x + iy | x, y \in \mathbb{R}\}$ be the complex numbers. For $z = x + iy$ and $z' = x' + iy'$ in $\mathbb{C}$ the multiplication zz' is defined as

$$zz' := (x + iy)(x' + iy') = (xx' - yy') + i(xy' + x'y) \in \mathbb{C}.$$

Recall the Polar representation of $z \in \mathbb{C}$ is given by $z = r(\cos\theta + i\sin\theta)$ where r is the radius and $\theta \in [0, 2\pi]$. If $z' = r'(\cos\theta' + i\sin\theta')$ then

$$zz' = rr'(\cos\theta + i\sin\theta)(\cos\theta' + i\sin\theta') =$$

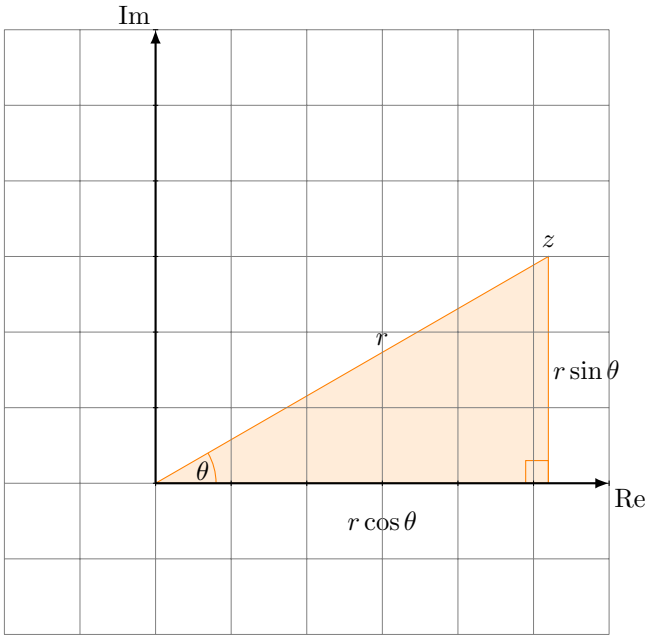
$$rr'(\cos\theta\cos\theta' - \sin\theta\sin\theta' + i[\sin\theta\cos\theta' + \sin\theta'\cos\theta])$$

$$= rr'(\cos(\theta + \theta') + i\sin(\theta + \theta'))$$

where the last equality is obtained from trigonometric identities:

$$\cos(\theta + \theta') = \cos\theta\cos\theta' - \sin\theta\sin\theta',$$

$$\sin(\theta + \theta') = \sin\theta\cos\theta' + \sin\theta'\cos\theta.$$



The exponential form is

$$z = re^{i\theta} := r(\cos\theta + i\sin\theta).$$

If $z' = r'e^{i\theta'}$ then from the discussion above, multiplication of complex numbers takes the form $z \cdot z' = r \cdot r'e^{i(\theta+\theta')}$. In light of this, we define the (complex) **unit circle** $S^1 \subseteq \mathbb{C}$ as

$$S^1 = \{z = e^{i\theta} | \theta \in \mathbb{R}\}.$$

Note that the radius of an element of S^1 is $r = 1$ as expected. Furthermore, multiplication of complex numbers that belong to S^1 yields a complex number in S^1:

$$e^{i\theta} \cdot e^{i\theta'} = e^{i(\theta+\theta')} \in S^1.$$

The element $1 = 1 + 0i = e^{i\cdot 0}$ satisfies $e^{i\theta} \cdot 1 = e^{i\theta}$ and for $z = e^{i\theta}$ we can take $z^{=1} = e^{-i\theta}$ to have $z \cdot z^{-1} = 1$. This means that $(S^1, \cdot, 1)$ is a group. Unlike the groups previously described, S^1 carries a **geometric structure** as well.

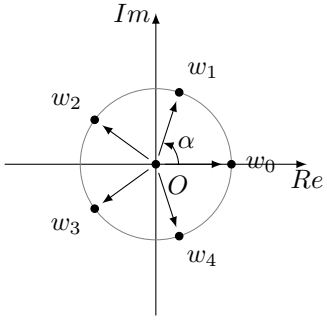
6. Let n be a natural number and consider the polynomial $p(x) = x^n - 1$. When we look for zeros of $p(x)$ over the complex numbers, i.e. $z \in \mathbb{C}$ such that $p(z) = 0$ we find that there are exactly n distinct such numbers, given by

$$w_k = e^{\frac{2k\pi i}{n}}$$

for $k = 0, 1, ..., n-1$. The set

$$\mu_n(\mathbb{C}) = \{w_k\}_{k=0}^{n-1}$$

is called the n**th roots of unit**. When $n = 5$, for example, these roots can be depicted as follows:



If $w, w' \in \mu_n(\mathbb{C})$ then $(ww')^n = w^n w'^n = 1 \cdot 1 = 1$ so that ww' is a zero of $p(x) = x^n - 1$ hence belongs to $\mu_n(\mathbb{C})$. Furthermore, for $k \neq 0$ and $w_k = e^{\frac{2k\pi i}{n}} \in \mu_n(\mathbb{C})$, we have

$$w_{n-k} = e^{\frac{2(n-k)\pi i}{n}}$$

and thus

$$w_k w_{n-k} = e^{\frac{i(2k\pi + 2(n-k)\pi)}{n}} = e^{i2\pi} = 1.$$

It follows that $w_{n-k} = w_k^{-1}$ and so $\mu_n(\mathbb{C})$ is a group under complex multiplication.

The groups in the examples above satisfy the property that the operation $\cdot$ is commutative, i.e. that for any x, y, $x \cdot y = y \cdot x$. Such groups are called **abelian**.

Let us see an example of a non-abelian group:

Example 3.4. Let $[n] = \{1, 2, ..., n\}$ be a set of size n. Consider the collection of all functions $\sigma : [n] \longrightarrow [n]$ that are isomorphisms of $[n]$ (i.e. monomorphism and epimorphism). We denote

$$\mathbb{S}_n = \{\sigma : [n] \longrightarrow [n] | \sigma \; is \text{ isomorphism}\}.$$

We define a binary operation on $\mathbb{S}_n$ by assigning for $\sigma, \tau \in \mathbb{S}_n$ their composition $\mu(\sigma, \tau) := \tau \circ \sigma$. Note that we swapped the order of σ and τ to account for the property of composition – $(\tau \circ \sigma)(i) = \tau(\sigma(i))$ – so that σ 'acts' first. The unit element is the identity function $\mathrm{id}_{[n]} : [n] \longrightarrow [n]$. Associativity axiom follows from associativity of composition, and the inverse axiom holds by Theorem 2.18 that says that an isomorphism has an inverse function. The group $\mathbb{S}_n$ is called the **symmetric group** (or the **permutation group**) on n letters. Note that since composition of function is not commutative, the binary operation on $\mathbb{S}_n$ is not commutative as well. We say that $\mathbb{S}_n$ is a **non-abelian group**.

Let $n = 3$ and consider the symmetric group $\mathbb{S}_3$. We can represent an element $\sigma \in \mathbb{S}_n$ in the form

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ \sigma(1) & \sigma(2) & \sigma(3) \end{pmatrix}$$

which means that 1 is mapped to $\sigma(1)$, 2 is mapped to $\sigma(2)$ and 3 is mapped to $\sigma(3)$.

Suppose $\sigma \in \mathbb{S}_3$ is given by $1 \mapsto 2,\ 2 \mapsto 3,\ 3 \mapsto 1$ and $\tau \in \mathbb{S}_3$ is given by $1 \mapsto 3,\ 2 \mapsto 1,\ 3 \mapsto 2$. Then we can depict $\sigma \cdot \tau$ as

$$\sigma \cdot \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \mathrm{id}_{[3]} = e$$

Exercise 3.5. Find a pair of elements $\alpha, \beta \in \mathbb{S}_3$ such that $\alpha \cdot \beta \neq \beta \cdot \alpha$ and prove your claim in a similar way to the calculation in Example 3.4.

For the next example, we need to invoke a

Theorem 3.6 (Fermat's little theorem)**.** *If p is prime and $1 \leq a \leq p-1$ then $a^{p-1} = 1 \ (\ mod\, p)$.*

whose proof relies on a

Lemma 3.7. *For any $x, y \in \Bbbk_p$ and any n,*

$$(x+y)^{p^n} = x^{p^n} + y^{p^n} \ (\ mod\, p).$$

Proof. To see the claim for $n = 1$ write

$$(x+y)^p = \sum_{i=0}^{p} \binom{p}{i} x^i y^{p-i}$$

and observe that p divides $\binom{p}{i} = \frac{p!}{i!(p-i)!}$ for every $0 < i < p$ so we are left with $(x+y)^p = x^p + y^p \ (\ mod\, p)$. If $n = 2$ then $(x+y)^{p^2} = ((x+y)^p)^p = (x^p + y^p)^p = x^{p^2} + y^{p^2}$ by a repeated application of the $n = 1$ case. The proof continues by straightforward induction.

□

Proof of Theorem 3.6. We will prove an equivalent statement, that for any $a \in \Bbbk_p^\times$, $a^p = a \ (\ mod\, p)$ by induction on a. If $a = 1$ the statement is clear. Suppose the statement is true for a and consider the case of $a+1$. Then $(a+1)^p = a^p + 1^p = a + 1 \ (\ mod\, p)$ by Lemma 3.7 and the induction hypothesis so we are done. □

Example 3.8. Let p be a prime and consider

$$\Bbbk_p^\times = \{1, 2, ..., p-1\} = \Bbbk_p \setminus \{0\}$$

with the operation $\times \ (\ mod\ p)$ and unit element 1. Then $\Bbbk_p^\times$ is a group. Unitary and associativity axioms clearly hold. To prove the inverse axiom, we revoke Theorem 3.6 to get

$$a(a^{p-2}) = 1 \ (\ mod\ p)$$

so that

$$a^{-1} = a^{p-2}.$$

3.2 Homomorphisms

As will be the case with many of the future Mathematical objects we encounter, a group is a set together with an additional structure, namely a binary operation which in turns is subject to some conditions (associativity, unitary, existence of inverses). In order to 'move' between groups we need a function between their underlying sets that respects that structure.

Definition 3.9. Let G, H be groups. A function $f : G \longrightarrow H$ is called a (group) **homomorphism** if for any $x, y \in G$, $f(xy) = f(x)f(y)$. Note that the multiplication of the left-hand side is that of G and the one on the right-hand side is that of H. A homomorphism of groups $f : G \longrightarrow H$ is called an **isomorphism** if f is an isomorphism as a function between the underlying sets of G and H. In the latter case we write $f : G \xrightarrow{\cong} H$ or simply $G \cong H$.

Remark 3.10. It follows from the definition that a homomorphism $f : G \longrightarrow H$ must satisfy $f(e) = e$ as we have for any $x \in G$: $f(x) = f(ex) = f(e)f(x)$ and multiplying this equation by $f(x)^{-1}$ we get the desired. Similarly, a homomorphism f automatically satisfies $f(x^{-1}) = f(x)^{-1}$ for any $x \in G$ since

$$e_H = f(e_G) = f(xx^{-1}) = f(x)f(x^{-1})$$

and multiplying the equation by $f(x)^{-1}$ gives the desired result.

Exercise 3.11. Let n, m be positive integers. Show that the map $\mathbb{Z}_{nm} \longrightarrow \mathbb{Z}_n$ given by $x \mapsto x \ mod \ n$ is a homomorphism. If we consider the map $\mathbb{Z}_{nm+1} \longrightarrow \mathbb{Z}_n$ given by the same formula, would it still be a homomrphism?

Exercise 3.12. If $f : G \longrightarrow H$ is an isomorphism of groups, prove that the inverse $f^{-1} : H \longrightarrow G$ is a group homomorphism as well, hence an isomorphism of groups in itself.

Informally speaking, if $f : G \cong H$ is an isomorphism of groups, then G and H have 'are the same up to a change of symbols' (and one such change of symbols is given by f). Formally, every group theoretic property holds for G if and only if it holds for H. For example, if G and H are finite, then the isomorphism f transforms the multiplication table of G exactly to that of H. For this reason, we will generally consider two isomorphic groups to be the 'same'.

Examples 3.13.

1. Let $n \in \mathbb{N}$ and $f : \mathbb{Z} \longrightarrow \mathbb{Z}_n$ be defined by $f(x) = x \ (\ mod \ n)$. Then f is a homomorphism since for any $x, y \in \mathbb{Z}$ we have $f(x+y) = (x+y) \ (\ mod \ n) = f(x) + f(y) \ (\ mod \ \ n)$.

2. For any pair of groups G, H we have the **trivial homomorphism** $G \longrightarrow H$ given by $x \in G \mapsto e$.

3. Let μ_n be the group of nth roots of unity and define a function $f : \mathbb{Z}_n \longrightarrow \mu_n$ by
$$f(k) = e^{\frac{2k\pi i}{n}}.$$

To see that this is a homomorphism let $k, k' \in \mathbb{Z}_n$ and write $k+k' = qn+r$. Then in $\mathbb{Z}_n$ we have $k+k' = r$ so that

$$f(k+k') = f(r) = e^{\frac{2r\pi i}{n}}$$

and on the other hand,

$$f(k)f(k') = e^{\frac{2(k+k')\pi i}{n}} = e^{2q\pi i}e^{\frac{2r\pi i}{n}} = e^{\frac{2r\pi i}{n}} = f(r) = f(k+k').$$

Note that f is clearly a monomorphism and an epimorphism so that it is an isomorphism.

4. there is no non-trivial homomorphism $f : \mathbb{Z}_n \longrightarrow \mathbb{Z}$: if there was, let $x \in \mathbb{Z}_n$ be such that $f(x) \neq 0 \in \mathbb{Z}$. In $\mathbb{Z}_n$ he n–fold sum

$$\underbrace{x + x + ... + x}_{n-times} =: nx$$

equals zero but since f is a homomorphism we have

$$0 = f(0) = f(x + x + ... + x) = f(x) + f(x) + ... + f(x)$$

which is an n–fold sum of non-zero elements – contradiction.

It will be useful for us to have a way of building new groups from old ones.

Construction 3.14. Let $G = (G, \cdot_G, e_G)$ and $H = (H, \cdot_H, e_H)$ be two groups and consider the Cartesian product

$$G \times H = \{(x, y) | x \in G \wedge y \in H\}.$$

Define a binary operation on $G \times H$ as follows: if $(x, y), (x', y') \in G \times H$,

$$(x, y) \cdot (x', y') := (x \cdot_G x', y \cdot_H y').$$

Proposition 3.15. *Under the binary operation defined above, $G \times H$ is a group with unit element (e_G, e_H) and the inverse for $(x, y) \in G \times H$ is given by (x^{-1}, y^{-1}). We will often denote this group by*

$$G \oplus H.$$

Proof. Left as an exercise. □

The last two examples in Examples 3.3 of the groups S^1 and μ_n suggest an interesting phenomena. We have $\mu_n \subseteq S^1$ and the group operation in both is given by multiplication of complex numbers.

Definition 3.16. Let $G = (G, \cdot, e)$ be a group and $H \subseteq G$ a subset. We say that H is a **subgroup** of G if $(H, \cdot, e)$ is a group under the restricted binary operation and the unit element of G. In that case, we denote $H \leq G$.

Equivalently, a subset $H \subseteq G$ of a group G is a subgroup if the following conditions hold:

1. $e \in H$.
2. for any $x, y \in H$, $xy \in H$.
3. for any $x \in H$, $x^{-1} \in H$.

Examples 3.17.

1. Let $G = \mathbb{Z}_{10}$. The set $H = \{0, 5\}$ is a subgroup of G. Note first that $|H| \| |G|$. Second, note that the set $H' = \{0, 4\}$ is not a subgroup of G since $4 + 4 = 8 \ (\ mod\ 10)$ and $8 \notin H'$.
2. As mentioned, for any natural number n, the group of nth roots of unity μ_n is a subgroup of the unit circle S^1.
3. Let $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ be the multiplicative group of the field of complex numbers. Then S^1 is a subgroup of $\mathbb{C}^\times$. Since μ_n is a subgroup of S^1, it follows that $\mu_n \leq \mathbb{C}^\times$.
4. The group $\mathbb{Z}_n = \{0, 1, ..., n-1\}$ can be viewed as a subset of the group $\mathbb{Z}$ but $\mathbb{Z}_n \nleq \mathbb{Z}$. The reason is that the binary operation in $\mathbb{Z}$ is $+$ whereas in $\mathbb{Z}_n$ it is $+ \ (\ mod\ n)$. Thus, for example we have $n - 1 \in \mathbb{Z}$ but $(n-1) + (n-1) = 2n - 2 \notin \mathbb{Z}_n$.
5. Let $(G, e_G), (H, e_H)$ be groups. The Cartesian product $G \times H$ is a group by Proposition 3.15 where the multiplication is done coordinate-wise. The subsets $G \times \{e_H\} = \{(g, e_H) | g \in G\}$ and $\{e_G\} \times H = \{(e_G, h) | h \in H\}$ are both subgroups of $G \times H$. Note that $G \times \{e_H\} \cong G$ and $\{e_g\} \times H \cong H$.

Remark 3.18. Note that if H is a subgroup of G, then H is itself a group and the inclusion $H \subseteq G$ can be viewed as a group homomorphism $H \hookrightarrow G$.

We are now ready to perform a central

Construction 3.19. Let G be a group and $H \leq G$ a subgroup. For an element $x \in G$, we define the (left) **coset**

$$xH := \{xh | h \in H\}.$$

Note first that if $x \in H$ then $xH = H$: if $xh \in xH$ then since $h \in H$ and H is closed under multiplication, $xh \in H$ so that $xH \subseteq H$. Conversely, if $h \in H$ then since $x \in H$ and H is a group, $x^{-1}h \in H$. Thus, $h = x(x^{-1}h) \in xH$ so that $H \subseteq xH$ and we have $H = xH$.

Proposition 3.20. *For a group G and a subgroup $H \leq G$, the set of cosets $\{xH\}_{x \in G}$ form a partition of G, i.e.,*

$$G = \bigcup_{x \in G} xH$$

and for any $x, y \in G$, the cosets xH, yH are either equal or disjoint.

Proof. First, for any $x \in G$ we have $x = x \cdot e \in xH$ since $e \in H$. Thus, $G = \bigcup_{x \in G} xH$. Second, let $x, y \in G$ and suppose that xH, yH are not disjoint. Then there is

$$\alpha \in xH \bigcap yH$$

so that there are $h, h' \in H$ such that $\alpha = xh = yh'$. Then

$$x = yh'h^{-1} \in yH$$

and

$$y = xh(h')^{-1} \in xH.$$

If $a = xh'' \in xH$ then

$$a = yh'h^{-1}h'' \in yH$$

and if $b = yh'''$ then

$$b = xh(h')^{-1}h'''$$

so $b \in xH$. It follows that $xH = yH$. □

Lemma 3.21. *For $H \leq G$ as above, and any $x, y \in G$, we have $xH = yH$ iff $x^{-1}y \in H$.*

Proof. If $xH = yH$ then there is $h \in H$ such that $y = y \cdot e = xh$, so that $x^{-1}y = h \in H$. Conversely, if $x^{-1}y \in H$ then $y = x(x^{-1}y) \in xH$ and since xH, yH are either equal or disjoint by Proposition 3.20 we deduce that $xH = yH$. □

Corollary 3.22. *Under the conditions above, we have an equivalence relation on G defined by $a \sim b$ iff $a, b \in xH$ for some $x \in G$. We denote the quotient set by G/H.*

Suppose $H \leq G$ is a subgroup and let $x, y \in G$ be such that $xH \neq yH$. Define a function $f : xH \longrightarrow yH$ by $f(xh) = yh$. We claim that f is an isomorphism of sets: clearly, f is an epimorphism since every element $yh \in yH$ satisfies $f(xh) = yh$. The function f is also a monomorphism: if $xh, xh' \in xH$ and $f(xh) = f(xh')$ then by definition $yh = yh'$ and multiplying the last equation by y^{-1} we get $h = h'$ so that $xh = xh'$.

Now suppose G is finite. Then for each $x \in G$, xH is a finite set and the discussion above means that for any $x, y \in G$, $|xH| = |yH|$. Thus, the set G can be partitioned to sets of equal size $G = \bigcup_{x \in G} xH$. We get the following

Theorem 3.23 (Lagrange)**.** *Let G be a finite group and $H \leq G$ a subgroup. Then $|H|$ divides $|G|$*

Proof. We have a partition $G = \bigcup_{x \in G} xH$ where each of the partition sets xH has equal size $|xH| = |H|$ so $|G| = k|H|$ where k is the number of distinct cosets xH. □

We saw in Corollary 3.22 that for any group G and a subgroup $H \leq G$, we have a quotient set G/H. It is natural to ask: when is G/H a group?

Proposition 3.24. *For G an abelian group and $H \leq G$ a subgroup, the quotient $G/H = \{xH | x \in G\}$ is a group under the operation $(xH) \cdot (yH) := (xy)H$ with unit $e_{G/H} := e \cdot H = H$ and the inverse of xH is given by $x^{-1}H$.*

Proof. Note that it could be that $xH = x'H$ with $x \neq x' \in G$ so we need to check that multiplication in G/H is well-defined (i.e. does not depend on representatives). So suppose $xH = x'H$ and $yH = y'H$. By Lemma 3.21 we have $xx'^{-1}, yy'^{-1} \in H$. In order to show that multiplication is well-defined, we need to show that $(xy)H = (x'y')H$ which by Lemma 3.21 is equivalent to show that $(xy)(x'y')^{-1} \in H$. But we have

$$(xy)(x'y')^{-1} = xy(x')^{-1}(y')^{-1} = xx'^{-1} \cdot yy'^{-1} \in H$$

since H is abelian. Associativity, unitary and existence of inverse in G/H follow from the corresponding axioms in G. □

Definition 3.25. For an abelian group G and a subgroup $H \leq G$ we call G/H with the multiplication defined above the **quotient group**. The map $q : G \longrightarrow G/H$ given by $x \mapsto xH$ is a group homomorphism and called the **quotient map**.

Examples 3.26.

1. Let $H = \{0, 5\} \leq \mathbb{Z}_{10} = G$. We claim that there is an isomorphism $\mathbb{Z}_{10}/H \cong \mathbb{Z}_5$. To see this, observe that we can write the (left) cosets as $G/H = \{0 + H, 1 + H, ..., 4 + H\}$ – if we take, for example, the coset $8 + H$, then $8-3 = 5 \in H$ so that $3+H = 8+H$. In light of this we define $f : \mathbb{Z}_{10}/H \longrightarrow \mathbb{Z}_5$ by $f(x + H) = x$ for $x = 0, 1, ..., 5$. This is clearly a homomorphism and an isomorphism of the underlying sets, hence an isomorphism of groups.

2. Let G, H be groups. We saw that the Cartesian product $G \times H$ has a subgroup of the form $G \times \{e_H\}$. Clearly, $G \times \{e_H\} \cong G$. If G, H are abelian, we get that the quotient $G \times H/G \times \{e_H\}$ is isomorphic to H. Similarly, we have $G \times H/\{e_G\} \times H \cong G$.

Definition 3.27. Let $f : G \longrightarrow H$ be a group homomorphism. The **kernel** of f is the set $\ker(f) = \{x \in G | f(x) = e_H\} \subseteq G$. The **image** of f is the set $\mathrm{Im}\,(f) = \{f(x) | x \in G\}$.

Exercise 3.28. Prove that a group homomorphism $f : G \longrightarrow H$ is a monomorphism iff $\ker f = \{e\}$

Proposition 3.29. *For any group homomorphism $f : G \longrightarrow H$ we have*

1. $\ker(f) \leq G$.
2. $Im\,(f) \leq H$.

Proof.

1. Clearly, $e \in \ker(f)$. If $x, y \in \ker(f)$ then $f(xy) = f(x)f(y) = e$ so $xy \in \ker(f)$ and for any $x \in \ker(f)$, $f(x^{-1}) = f(x)^{-1} = e$ so $x^{-1} \in \ker(f)$.

2. Clearly, $e = f(e) \in \mathrm{Im}\,(f)$. If $x, y \in \mathrm{Im}\,(f)$ then there are $a, b \in G$ such that $x = f(a)$ and $y = f(b)$. Thus, $xy = f(a)f(b) = f(ab)$ so that $xy \in \mathrm{Im}\,(f)$. Similarly, if $x = f(a) \in \mathrm{Im}\,(f)$ then $x^{-1} = f(a^{-1})$ so that $x^{-1} \in \mathrm{Im}\,(f)$.

□

Suppose $f : G \longrightarrow H$ is a homomorphism of abelian groups. By Proposition 3.29 we can consider the quotient groups $G/\ker(f)$.

Theorem 3.30 (The first isomorphism theorem)**.** *Let $f : G \longrightarrow H$ be a homomorphism of abelian groups and denote $K = \ker f$. Then there is an isomorphism of groups $G/K \cong Im\,(f)$.*

Proof. Define $f' : G/K \longrightarrow \mathrm{Im}\,(f)$ by $f'(xK) = f(x)$. To see that f' is well-defined, suppose $x, y \in G$ are such that $xK = yK$. Then $xy^{-1} \in K$ so that $f(xy^{-1}) = f(x)f(y)^{-1} = e$ and thus $f(x) = f(y)$ i.e. $f'(xK) = f'(yK)$ and f' is well-defined. Clearly, f' is a group homomorphism since f is and it remains to check that f' is an isomorphism. First, if $a = f(x) \in \mathrm{Im}\, f$ then $f'(xK) = f(x) = a$ so that f' is an epimorphism. Second, if xK, yK are such that $f'(xK) = f'(yK)$ then $f(x) = f(y)$ so that $f(xy^{-1}) = e$ and we get that $xy^{-1} \in K$ which by Lemma 3.21 means that $xK = yK$. Thus f' is also a monomorphism and we're done. □

Remark 3.31. Theorem 3.30 is true also without the assumption that G, H are abelian. We chose to restrict attention to the version above since defining the quotient group for non-abelian groups requires complication we wish to avoid at the moment.

Examples 3.32.

1. Let $f : \mathbb{Z} \longrightarrow \mathbb{Z}_n$ be the *mod n* homomorphism $x \mapsto x(\ mod\ n)$. Then clearly $\mathrm{Im}\, f = \mathbb{Z}_n$. On the other hand,

$$\ker f = \{nx | x \in \mathbb{Z}\} = n\mathbb{Z}.$$

   It follows from Theorem 3.30 that $\mathbb{Z}_n \cong \mathbb{Z}/n\mathbb{Z}$.

2. Let $f : \mathbb{Z}_{10} \longrightarrow \mathbb{Z}_2$ be the homomorphism given by $f(x) = x(\ mod\ 2)$. Then $\mathrm{Im}\, f = \mathbb{Z}_2$ and $\ker f = \{0, 2, 4, 6, 8\}$. It follows that $\mathbb{Z}_{10}/\ker f \cong \mathbb{Z}_2$. Note that $\ker f \cong \mathbb{Z}_5$ via the isomorphism $\ker(f) \longrightarrow \mathbb{Z}_5$ given by $x \mapsto \frac{x}{2}$.

3.3 Cyclic groups and structure theorems for abelian groups

Much of our attention in cryptography will be devoted to finite (abelian) groups. In such a context, it's natural to consider the following

Definition 3.33. Let G be a group and $g \in G$ an element.

1. The minimal number $n \in \mathbb{N}$ such that $g^n = e$ is called the **order** of g and denoted $o(g)$. If no such number exists, we set $o(g) = \infty$.

2. The group **generated** by g is the set $\langle g \rangle := \{g^n | n \in \mathbb{Z}\}$ which inherits a group structure from G. Note that by definition, $g^{-n} := \left(g^{-1}\right)^n$.

3. The group G is called **cyclic** if there exists $\gamma \in G$ such that $\langle \gamma \rangle = G$.

Examples 3.34. 1. Let $n \in \mathbb{N}$. The group $\mathbb{Z}_n$ is cyclic: as a generator we can choose $\gamma = 1$.

2. The group $\mathbb{Z}$ is cyclic. As a generator we can choose either 1 or -1.

3. The group $\mathbb{Z}_2 \times \mathbb{Z}_3$ is cyclic. As a generator we can choose $(1, 1)$.

Cyclic groups have a rather rigid structure. For example:

Observation 3.35. Let $G = \langle \gamma \rangle$ be a cyclic group. Then G is abelian: $\gamma^n \gamma^k = \gamma^{n+k} = \gamma^k \gamma^n$.

Suppose $G = \langle \gamma \rangle$ is a cyclic group of order n. The elements $\{e, \gamma, \gamma^2, ..., \gamma^{n-1}\}$ are all distinct since if there are $1 \leq k < l \leq n-1$ such that $\gamma^l = \gamma^k$ then $\gamma^{l-k} = e$ with $l - k < n - 1$ and so γ cannot be a generator. Thus $G = \{e, \gamma, \gamma^2, ..., \gamma^{n-1}\}$ and $\gamma^n = e$.

Proposition 3.36. *For any $n \in \mathbb{N}$, there exists a cyclic group of order n, and it is unique up to an isomorphism.*

Proof. As we saw, $\mathbb{Z}_n$ is a cyclic group of order n. Suppose $G = \langle e, \gamma, \gamma^2, ..., \gamma^{n-1} \rangle$ is a cyclic group of order n. Then the function $f : \mathbb{Z}_n \longrightarrow G$ by $f(k) = \gamma^k$ is an isomorphism of groups. □

As we saw in Example 3.34 (2), a cyclic group can have more than one generator. In order to account for all generators in $\mathbb{Z}_n$, we will need the following

Definition 3.37. For $n \in \mathbb{N}$, $\varphi(n) = |\{k | 1 \leq k \leq n \wedge \gcd(n, k) = 1\}|$ is the **Euler totient function**.

Proposition 3.38. *An element $k \in \mathbb{Z}_n$ is a generator iff* $\gcd(n, k) = 1$.

Proof. If $1 \leq k \leq n - 1$ is a generator then $\mathbb{Z}_n = \{0, k, 2k, ..., (n-1)k\}$ so if $l = \gcd(n, k) > 1$ we can write $n = ls$ and $k = lt$. Since $s < n - 1$ we get that $sk = slt = nt$ so $sk = 0$ (mod n) contradiction. Conversely, if $\gcd(n, k) = 1$ we claim that the numbers $\{0, k, 2k, ..., (n-1)k\}$ are all distinct ($mod\, n$): otherwise, if $sk = tk$ (mod n) for $1 \leq s < t \leq n - 1$ then $(t - s)k = nx$ for some x and since $k \nmid n$ we have $k \mid x$. But then, $(t - s)k = nky$ for some y so $t - s = ny$ which is a contradiction since $1 \leq (t - s) \leq n - 1$. □

Corollary 3.39. *Let G be a cyclic group of order n. Then G has exactly $\varphi(n)$ generators.*

Proof. This is the case for $G = \mathbb{Z}_n$ and the number of generators stays the same for isomorphic groups. □

For future reference, it is useful to record:

Theorem 3.40. *For $n \in \mathbb{N}$, the Euler's totient function can be calculated as $\varphi(n) = n\prod_{p|n}(1-\frac{1}{p})$ where the product is taken over all primes p that divide n. It follows that*

$$n = \sum_{d:d|n} \varphi(d).$$

The next result is useful in the study of cyclic groups

Theorem 3.41. *A subgroup of a finite cyclic group is cyclic.*

Proof. Suppose $G = \langle g \rangle$ is cyclic of order n and $H \leq G$ a subgroup. Denote $H = \{a_0, ..., a_{d-1}\} = \{1, g^{k_1}, ..., g^{k_{d-1}}\}$ where $d \mid n$ with $n = kd$. Since H is of order d, $\left(g^{k_i}\right)^d = g^{k_i d} = 1$ for each i so that $k_i d = nm = kdm \iff k_i = km$ for some m. Thus, each a_i is a power of g^k so that $H \subseteq \{1, g^k, g^{2k}, ..., g^{k(d-1)}\}$ and it follows that $H = \{1, g^k, g^{2k}, ..., g^{k(d-1)}\}$ which is cyclic with generator g^k. □

Exercise 3.42.

1. Let G be a group of prime order. Prove that G is cyclic.
2. Show that $\mathbb{Z}_2 \times \mathbb{Z}_3$ is cyclic.
3. Let G, H be groups such that $|G| = |H| = 3$. Is necessarily $G \cong H$?
4. Suppose G, H are groups with $|G| = |H| = 4$. Is necessarily $G \cong H$?

Recall from Proposition 3.15 that given two groups G, H we have their Cartesian product $G \times H$ with coordinate-wise group structure. The next lemma demonstrates how to calculate the order of elements in the Cartesian product

Lemma 3.43. *Let G, H be groups. If $g \in G$ has order n and $h \in H$ has order m then $(g, h) \in G \times H$ has order $\operatorname{lcm}(n, m)$.*

Proof. Note that $(g,h)^N = (e,e) \iff g^N = e \wedge h^N = e \iff o(g)|N \wedge o(h)|N$ so N must be a common multiple of n and m. Thus, the order of (g, h) is the least common multiple of n and m. □

We showed in Exercise 3.42 that $\mathbb{Z}_2 \times \mathbb{Z}_3$ is cyclic. Lemma 3.43 enables us to prove the general case. First, we need a small result from elementary number theory:

Lemma 3.44. *Let n, m be positive integers. Then $\operatorname{lcm}(n, m) = \frac{nm}{\gcd(n,m)}$.*

Proof. Left as an exercise: prove by contradiction. □

Corollary 3.45. *Let n, m be positive integers. Then $\mathbb{Z}_n \times \mathbb{Z}_m$ is cyclic iff $\gcd(n, m) = 1$.*

Proof. Suppose $\gcd(nm) = 1$ By Lemma 3.43 the order of $(1,1)$ is $\text{lcm}(n,m)$ which by Lemma 3.44 is $\frac{nm}{\gcd(n,m)} = nm$. Since $|\mathbb{Z}_n \times \mathbb{Z}_m| = nm$ we get that $\mathbb{Z}_n \times \mathbb{Z}_m = \langle(1,1)\rangle$ is cyclic.

Conversely, if $\gcd(n,m) > 1$ then the order of any element in $\mathbb{Z}_n$ divides n and the order of any element in $\mathbb{Z}_m$ divides m so that by Lemma 3.43 the order of any element in $G \times H$ divides $\text{lcm}(nm) = \frac{nm}{\gcd(nm)}$ which is strictly smaller than nm. Thus in that case $\mathbb{Z}_n \times \mathbb{Z}_m$ is not cyclic. □

Corollary 3.45 in turn leads to an insight on solving a system of equations modulo different integers:

Theorem 3.46. *Let $n, m \in \mathbb{N}$ such that $\gcd(n,m) = 1$. Then for all $r, s \in \mathbb{N}$, the system of equations*

$$\begin{aligned} x &= r \bmod n \\ x &= s \bmod m \end{aligned} \tag{1}$$

has a unique solution in $\mathbb{Z}_{nm}$.

Proof. Consider the map $f : \mathbb{Z}_{nm} \longrightarrow \mathbb{Z}_n \times \mathbb{Z}_m$ given by $a \mapsto (a \bmod n, a \bmod m)$. If $a, b \in \mathbb{Z}_{nm}$ then since $n \mid nm$ and $m \mid nm$, we have

$$(a+b) \bmod nm \mapsto (a \bmod n + b \bmod n, a \bmod m + b \bmod m)$$

so f is a group homomorphism. Clearly, $f(1) = (1,1)$ and since $(1,1)$ has order nm, f is an isomorphism. Let $[r]_n = r \bmod n \in \mathbb{Z}_n$ and $[s]_n = s \bmod m \in \mathbb{Z}_m$. Since f is an isomorphism, the element $([r]_n, [s]_m)$ has a unique $x \in \mathbb{Z}_{nm}$ such that $f(x) = (x \bmod n, x \bmod m) = ([r]_n, [s]_m)$. □

Theorem 3.46 has, in fact, a more general version:

Corollary 3.47 (The Chinese remainder Theorem)**.** *Let $n1, ..., n_k \in \mathbb{N}$ be such that $\forall i \neq j : \gcd(n_i, n_j) = 1$ and denote $N = n_1 \cdot ... \cdot n_k$. Then for all $a_1, ..., a_k \in \mathbb{N}$, the system of equations*

$$\begin{aligned} x &= a_1 \bmod n_1 \\ &\vdots \\ x &= a_k \bmod n_k \end{aligned} \tag{2}$$

has a unique solution in $\mathbb{Z}_N$.

Proof. By induction using Theorem 3.46. □

What happens when we want to solve a system of integral equations with non-coprime moduli? We state here a result for the sake of completeness.

Theorem 3.48. *Let* $n1, ..., n_k \in \mathbb{N}$ *and denote* $N = n_1 \cdot ... \cdot n_k$. *Then for all* $a_1, ..., a_k \in \mathbb{N}$, *the system of equations*

$$\begin{aligned} x &= a_1 \ mod \ n_1 \\ &\vdots \\ x &= a_k \ mod \ n_k \end{aligned} \tag{3}$$

has a solution in $\mathbb{Z}$ *if* $\forall i \neq j$, $a_i = a_j \ mod \ \gcd(n_i, n_j)$ *and that solution is unique modulo* $\text{lcm}(n_1, ..., n_k)$.

We shall finish this section with two structure theorems for abelian groups. Let $G_1, ..., G_n$ be groups. We define the nth Cartesian product $G = G_1 \times ... \times G_n$ to be the set of all n-tuples $(g_1, ..., g_n)$ with $g_i \in G_i$ for all i. Just like with the two-fold Cartesian product, we can define a binary operation on G by using the operation of G_i in the ith coordinate and this makes G into a group. Another common notation in this case is $G = G_1 \oplus ... \oplus G_n$. Keeping that in mind, we can now state:

Theorem 3.49 (structure theorem for finite abelian groups)**.** *Let* G *be a finite abelian group. Then*

$$G \cong \mathbb{Z}_{n_1} \oplus ... \oplus \mathbb{Z}_{n_k}$$

where $n_i \in \mathbb{N}$ *and for all* i $n_i \mid n_{i+1}$.

In case the abelian group in question is not finite, the 'closest' property for being finite is given by the following

Definition 3.50. An abelian group G (written in additive form) is said to be **finitely generated** if there is a set of elements $\{g_1, ..., g_k\} \subseteq G$, called a **set of generators**, such that for each $x \in G$ there are integers $n_1, ..., n_k \in \mathbb{Z}$ such that

$$x = n_1 g_1 + ... n_k g_k.$$

Examples 3.51.

1. $\mathbb{Z}$ is finitely generated since $\langle 1 \rangle = \mathbb{Z}$.
2. $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$ is finitely generated by $(1, 0)$ and $(0, 1)$: indeed, if $x = (a, b) \in \mathbb{Z} \times \mathbb{Z}$ then $x = a(1, 0) + b(0, 1)$.
3. more generally, the group $\mathbb{Z}^r = \underbrace{\mathbb{Z} \times ... \times \mathbb{Z}}_{r-times}$ is finitely generated and a set of generators is given by
$$\{(1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1)\}.$$
4. The group $(\mathbb{Q}, +)$ is not finitely generated. If $\mathcal{A} = \{\frac{p_1}{q_1}, ..., \frac{p_n}{q_n}\}$ is a finite set then the additive subgroup it generates is all fractions whose denominator is the least common multiple of $q_1, ..., q_n$ hence $\mathcal{A}$ cannot be a set of generators.

It turns out that finitely generated abelian groups have a similar structure theorem.

Theorem 3.52 (structure theorem for finitely generated abelian groups)**.** *Let G be a finitely generated abelian group. Then there are positive integers $n_1, ..., n_k, r$ and an isomorphism*

$$G \cong \mathbb{Z}_{n_1} \oplus ... \oplus \mathbb{Z}_{n_k} \oplus \mathbb{Z}^r.$$

*The number r is called the **rank** of G.*

4 Fields

4.1 Fields and field homomorphisms

The notion of a group is considered one of the most basic ones in modern Mathematics. In a group we are dealing with a set and a binary operation, that abstracts the properties of multiplication (or addition). However, going back to real numbers, we have in fact **two binary operations**, namely addition and multiplication and these operations are compatible in some sense. In order to abstract these properties for more general cases, we have the following

Definition 4.1. A **field** is the data of a set $\mathbb{F}$ together with two binary operations $+ : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F}$ and $\cdot : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F}$ called addition and multiplication and two specified elements $0, 1 \in \mathbb{F}$ subject to the following conditions:

1. (associativity of + and $\cdot$) for every $x, y, z \in \mathbb{F}$, $(x + y) + z = x + (y + z)$ and $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
2. (commutativity of + and $\cdot$) for every $x, y \in \mathbb{F}$, $x + y = y + x$ and $x \cdot y = y \cdot x$.
3. (distributivity) for every $x, y, z \in \mathbb{F}$, $x \cdot (y + z) = x \cdot y + x \cdot z$.
4. (neutral elements wrt + and $\cdot$) for every $x \in \mathbb{F}$, $x + 0 = x$, $x \cdot 0 = 0$ and $x \cdot 1 = x$.
5. (existence of inverses wrt +) for every $x \in \mathbb{F}$ there is an element $-x$ such that $x + (-x) = 0$.
6. (existence of inverses wrt $\cdot$) if $x \neq 0$ there exists an element $x^{-1} \in \mathbb{F}$ such that $x \cdot x^{-1} = 1$.

Remark 4.2. As with groups, we will usually ommit $\cdot$ when multiplying elements in a field, ie write xy instead of $x \cdot y$.

Examples 4.3.

1. The real numbers $\mathbb{R}$ with usual addition and multiplication are a field.
2. The rational numbers $\mathbb{Q}$ with usual addition and multiplication are a field. Note that we need to make sure that addition and multiplication of two rational numbers result in a rational number.

3. The complex numbers $\mathbb{C}$ with addition and multiplication of complex numbers are a field. The non-trivial condition to check is existence of multiplicative inverse: if $z = r(\cos\theta + i\sin\theta) \neq 0$, then $z^{-1} = \frac{1}{r}(\cos(-\theta) + i\sin(-\theta))$.

Exercise 4.4.

1. Prove that $\sqrt{2} \notin \mathbb{Q}$. Hint: suppose by contradiction that $\sqrt{2} \in \mathbb{Q}$ and let $a, b \in \mathbb{Z} \setminus \{0\}$ be such that $\sqrt{2} = \frac{a}{b}$ and $\gcd(a, b) = 1$. Then $a^2 = 2b^2$. Derive a contradiction by dividing to cases where a, b are odd or even.

2. Let $\mathbb{Q}(\sqrt{2}) = \{x + y\sqrt{2} | x, y \in \mathbb{Q}\}$ (note the analogy with $\mathbb{C}$) with addition and multiplication induced from $\mathbb{R}$. Prove that $\mathbb{Q}(\sqrt{2})$ is a field. Write an explicit formula for the inverse of $x + y\sqrt{2}$.

Lemma 4.5. *If $\mathbb{F}$ is a field and $x, y \in \mathbb{F}$ such that $xy = 0$ then either $x = 0$ or $y = 0$.*

Proof. Otherwise, we could multiply by x^{-1} and get $x^{-1}xy = 1 \cdot y = y = 0 \cdot x^{-1} = 0$ – a contradiction. □

Proposition 4.6. *For $n \in \mathbb{N}$, $\mathbb{Z}_n$ with addition and multiplication modolu n is a field iff $n = p$ is prime, and denoted $\mathbb{F}_p$*

Proof. The associativity, commutativity distributivity and neutral elements axioms hold for any $n \in \mathbb{N}$. In addition, inverses for + exist in $\mathbb{Z}_n$ as it is a group. The remaining axiom to check is existence of multiplicative inverses. If n is not prime, then $n = k \cdot l$ for some $1 \leq k, l \leq n - 1$ but then $k \cdot l = 0$ in $\mathbb{Z}_n$ whereas $k, l \neq 0$ – in contradiction to Lemma 4.5. Thus in that case, $\mathbb{Z}_n$ cannot be a field. conversely, suppose $n = p$ is prime. For $0 \neq x \in \mathbb{Z}_p$ we have, by Fermat's little theorem $x^{p-1} = 1$ ($mod\, p$). Thus $x^{p-2} = x^{-1}$ and we're done. □

What sets apart a field like $\mathbb{Q}$ from the field $\mathbb{F}_p$? The following definition gives such a suggestion:

Definition 4.7. Let $\mathbb{F}$ be a field. If there is a minimal number $n \in \mathbb{N}$ such that $\underbrace{1 + 1 + ... + 1}_{n-times} = 0$ is called the **characteristic** of $\mathbb{F}$ and we denote $\operatorname{char}\mathbb{F} = n$. If no such number exists, we say that $\mathbb{F}$ is of **characteristic 0** and denote $\operatorname{char}\mathbb{F} = 0$.

Proposition 4.8. *The characteristic of a field $\mathbb{F}$ must be 0 or prime number p.*

Proof. Let $n \in \mathbb{N}$ be the characteristic of $\mathbb{F}$. If n is not prime, then $n = kl$ for some $k, l > 1$. But then $\underbrace{1 + 1 + ... + 1}_{n-times} = (\underbrace{1 + 1 + ... + 1}_{k-times})(\underbrace{1 + 1 + ... + 1}_{l-times}) = 0$ in contradiction to Lemma 4.5. □

As with groups, we would like to have a notion of a map of fields.

Definition 4.9. Let $\mathbb{F}, \mathbb{K}$ be fields and $f : \mathbb{F} \longrightarrow \mathbb{K}$ a function. We say that f is a **field homomorphism** if:

1. $f(0) = 0$ and $f(1) = 1$.
2. for any $x, y \in \mathbb{F}$, $f(x + y) = f(x) + f(y)$.
3. for any $x, y \in \mathbb{F}$, $f(xy) = f(x)f(y)$.

We say that f is a **field isomorphism** if it is furthermore an isomorphism of sets.

Examples 4.10.
1. The inclusions $\mathbb{Q} \hookrightarrow \mathbb{R}$ and $\mathbb{R} \hookrightarrow \mathbb{C}$ are field homomorphisms.
2. There is no field homomorphism $f : \mathbb{F}_p \longrightarrow \mathbb{Q}$. If there were, then $0 = f(0) = f(\underbrace{1 + 1 + ... + 1}_{p-times}) = \underbrace{f(1) + f(1) + ... + f(1)}_{p-times} = \underbrace{1 + 1 + ... + 1}_{p-times}$ since f is a homomorphism but $\mathbb{Q}$ has characteristic 0 which is a contradiction.

Exercise 4.11. Prove that if $\operatorname{char} \mathbb{F} = p$ and $\operatorname{char} \mathbb{K} = p'$ for $p \neq p'$ then there is no field homomorphism $f : \mathbb{F} \longrightarrow \mathbb{K}$.

Our main basis in cryptography is that of finite fields. Note that a field of characteristic 0 must be infinite since the elements $1, 1 + 1, 1 + 1 + 1, ...$ must all be different. Thus, by Proposition 4.8 a finite field must have characteristic p for some prime number. We saw that $\mathbb{F}_p$ is such a field, and we may ask:

Question 4.12. What are the fields of characteristic p up to isomorphism?

Let $\mathbb{F}$ be a field with p elements (p prime). We can view $(\mathbb{F}_p, +)$ as an abelian group and consider the cyclic group generated by 1: $\langle 1 \rangle = \{1, 1 + 1, ...\}$. By Lagrange's Theorem 3.23, $|\langle 1 \rangle|$ must divide p. Since p is prime and $\langle 1 \rangle$ has at least two elements $1 \neq 1 + 1$, it follows that $|\langle 1 \rangle| = p$ so that $\langle 1 \rangle = \mathbb{F}$ and by Proposition 3.36 we have an isomorphism of groups $\langle 1 \rangle \cong \mathbb{Z}_p$. Define a map $f : \mathbb{F}_p \longrightarrow \mathbb{F}$ by $f(0) = 0$ and $f(n) = \underbrace{1 + ... + 1}_{n-times}$. Then it is easy to check that f is an isomorphism of fields so that $\mathbb{F} \cong \mathbb{F}_p$. We have just proved:

Corollary 4.13. *Let p be prime. Then any field with p elements is isomorphic to $\mathbb{F}_p$*

To begin answering Question 4.12 it will be useful to consider the following

Definition 4.14. Let $\mathbb{F}$ be a field. A subset $\mathbb{K} \subseteq \mathbb{F}$ is called a **subfield** if

1. $0_{\mathbb{F}}, 1_{\mathbb{F}} \in \mathbb{K}$.
2. $\mathbb{K}$ is a field under the addition and multiplication defined in $\mathbb{F}$

Examples 4.15.

1. $\mathbb{Q} \subseteq \mathbb{R}$ is a subfield inclusion.
2. $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2})$ is a subfield inclusion.
3. $\mathbb{F}_p \subseteq \mathbb{Q}$ is not a subfield inclusion since $\mathbb{F}_p$ is not a field under the addition and multiplication of $\mathbb{Q}$. This is because, e.g., $\underbrace{1+1+...+1}_{p-times} = 0$ in $\mathbb{F}_p$ and $\underbrace{1+1+...+1}_{p-times} = p$ in $\mathbb{Q}$.

Using Definition 4.14 we can identify 'minimal' fields of characteristic p and characteristic 0 as the following propositions show.

Proposition 4.16. *Let p be prime and $\mathbb{F}$ a field of characteristic p. Then $\mathbb{F}$ contains a subfield which is isomorphic to $\mathbb{F}_p$*

Proof. Consider the set

$$\mathbb{K} = \{1, 1+1, ..., \underbrace{1+1+...+1}_{(p)-times} = 0\}.$$

then $\mathbb{K} \subseteq \mathbb{F}$ is a subfield and we can define an isomorphism of fields $f : \mathbb{K} \longrightarrow \mathbb{F}_p$ by $\underbrace{1+1+...+1}_{(k)-times} \mapsto k$. □

Proposition 4.17. *Let $\mathbb{F}$ be a field of characteristic 0. Then $\mathbb{F}$ contains a subfield which is isomorphic to $\mathbb{Q}$.*

Proof. First, note that $\mathbb{F}$ contains the set $\mathbb{F}_{\mathbb{Z}} = \{..., -(1+1), -1, 0, 1, 1+1, ...\}$ which can be thought of as a copy of the integers. Second, $\mathbb{F}$ must contain inverses to all non-zero elements in $\mathbb{F}_{\mathbb{Z}}$ i.e. elements of the form $\frac{1}{b}$ with $b \in \mathbb{F}_{\mathbb{Z}} \setminus \{0\}$. Third $\mathbb{F}$ is closed to multiplication hence must contain all elements of the form $a \cdot \frac{1}{b} = \frac{a}{b}$ where $a \in \mathbb{F}_{\mathbb{Z}}$ and $b \in \mathbb{F}_{\mathbb{Z}} \setminus \{0\}$. The last set is a subfield of $\mathbb{F}$ isomorphic to $\mathbb{Q}$. □

4.2 Vector spaces over fields

To give a partial answer to Question 4.12 in full we will need to take a small digression to the theory of vector spaces.

Definition 4.18. Let $\mathbb{F}$ be a field. A **vector space** V over $\mathbb{F}$ is a set V together with a distinguised element $0_V \in V$ and two binary operations $+ : V \times V \longrightarrow V$, $\cdot : \mathbb{F} \times V \longrightarrow V$ called addition of vectors and multiplication of a vector by a scalar, satisfying the following conditions:

1. (associativity of addition) for all $u, v, w \in V$, $(u+v)+w = u+(v+w)$.
2. (commutativity of addition) for all $v, w \in V$, $v+w = w+v$.
3. (associativity of scalar multiplication) for every $\alpha, \beta \in \mathbb{F}$ and every $v \in V$, $(\alpha\beta)v = \alpha(\beta \cdot v)$.

4. (distributivity) for every $\alpha \in \mathbb{F}$ and $v, w \in V$, $\alpha \cdot (v + w) = \alpha v + \alpha w$.

5. (neutral element) for every $v \in V$, $0_V + v = v$ and $0_{\mathbb{F}} \cdot v = 0_V$.

6. (existence of inverse wrt addition) for every $v \in V$, there is an element $-v \in V$ such that $v + (-v) = 0_V$.

Examples 4.19.

1. Every field $\mathbb{F}$ is a vector space over itself.

2. For any field $\mathbb{F}$, the Cartesian product

$$\mathbb{F}^n = \underbrace{\mathbb{F} \times ... \times \mathbb{F}}_{n-times}$$

is a vector space over $\mathbb{F}$: addition of vectors is defined by

$$(x_1, ..., x_n) + (y_1, ..., y_n) := (x_1 + y_1, ..., x_n + y_n)$$

and scalar multiplication is defined by

$$\alpha \cdot (x_1, ..., x_n) := (\alpha x_1, ..., \alpha x_n).$$

The inverse for addition of vectors is given by

$$-(x_1, ..., x_n) := (-x_1, ..., -x_n)$$

where $-x_i$ is the inverse of x_i with respect to addition in $\mathbb{F}$.

3. The complex numbers $\mathbb{C}$ are a vector space over $\mathbb{R}$: for $\alpha \in \mathbb{R}$ and $z = x + iy \in \mathbb{C}$ we define scalar multiplication as $\alpha z := \alpha x + \alpha y i$.

As usual, we would like to have a notion of a map between vector spaces:

Definition 4.20. Let V, W be vector spaces over (the same) field $\mathbb{F}$. A function $f : V \longrightarrow W$ is called a **linear map** (synonyms: vector space homomorphism, linear transformation) if:

1. for every $\alpha \in \mathbb{F}$ and every $v \in V$, $f(\alpha v) = \alpha f(v)$.

2. for every $v, v' \in V$, $f(v + v') = f(v) + f(v')$.

A linear map $f : V \longrightarrow W$ is called a **vector space isomorphism** if f is an isomorphism of the underlying sets of V and W.

Examples 4.21.

1. the function $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ given by $f(x, y, z) = (2x + 4y - z, 5x + z)$ is a linear map.

2. the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by $f(x, y) = xy$ is not a linear map: in $\mathbb{R}^2$ we have $(1, 1) + (1, 1) = (2, 2)$ but $f(2, 2) = 4$ whereas $f(1, 1) + f(1, 1) = 2$.

3. if V, W are vector spaces over a field $\mathbb{F}$, the map $V \longrightarrow W$ given by $v \mapsto 0_W$ is a linear map.

4. Consider $\mathbb{C}$ as a vector space over $\mathbb{R}$ and let $f : \mathbb{C} \longrightarrow \mathbb{R}^2$ be defined by $f(x + yi) = (x, y)$. Then f is a linear isomorphism. Note that when $\mathbb{C}$ is viewed as a vector space over $\mathbb{R}$, we have $\mathbb{C} \cong \mathbb{R}^2$ but $\mathbb{C}$ can also be viewed as field whereas $\mathbb{R}^2$ has no apparent structure of a field. Furthermore, we can view $\mathbb{C}$ as a vector space over the field $\mathbb{C}$ and there is no apparent way to view $\mathbb{R}^2$ as a vector space over $\mathbb{C}$.

We will be mostly interested in a class of vector spaces that are particularly simple:

Definition 4.22. Let V be a vector space over a field $\mathbb{F}$. We say that V has **finite dimension** if there are vectors $\mathcal{B} = \{v_1, ..., v_n\} \subseteq V$ such that for any $v \in V$ there are scalars $\alpha_1, ..., \alpha_n \in \mathbb{F}$ that satisfy $v = \alpha_1 v_1 + ... + \alpha_n$.

For example, let $\mathbb{F}$ be a field and consider $\mathbb{F}^n$ as a vector space over $\mathbb{F}$. Take $\mathcal{B} = \{(1, 0, ..., 0), (0, 1, 0, ..., 0), ..., (0, ..., 0, 1)\} \in \mathbb{F}^n$. If $v = (x_1, ..., x_n) \in \mathbb{F}^n$ we can write $v = x_1 \cdot (1, 0, ..., 0) + x_2 \cdot (0, 1, 0, ..., 0) + ... + x_n \cdot (0, ..., 0, 1)$ where for all i, $x_i \in \mathbb{F}$. Thus $\mathbb{F}^n$ is a finite dimensional vector space over $\mathbb{F}$.

Theorem 4.23 (structure theorem for finite dimensional vector spaces)**.** *Let V be a finite dimensional vector space over a field $\mathbb{F}$. Then there is an isomorphism of vector spaces $V \cong \mathbb{F}^n$.*

Proof. See [Gol]. □

Let us come back to fields of positive characteristic. We saw that one such field is $\mathbb{F}_p$ and asked if there are other examples. Using Theorem 4.23 we can partially answer this question. Suppose $\mathbb{F}$ is a finite field of characteristic p. If $\alpha \in \mathbb{F}_p$ and $v \in \mathbb{F}$ we can define $\alpha \cdot v := \underbrace{v + v + ... + v}_{\alpha - times} \in \mathbb{F}$. It is easy to verify that this definition, makes $\mathbb{F}$ a vector space over $\mathbb{F}_p$ where addition of vectors is given by addition in $\mathbb{F}$. The vector space $\mathbb{F}$ is finite hence clearly finite dimensional and by Theorem 4.23 we have an isomorphism of vector spaces $\mathbb{F} \cong \mathbb{F}_p^n$. We thus get:

Corollary 4.24. *Let $\mathbb{F}$ be a finite field of positive characteristic p. Then there is an isomorphism of vector spaces $\mathbb{F} \cong \mathbb{F}_p^n$ for some $n \in \mathbb{N}$. In particular $|\mathbb{F}| = p^n$.*

Note that Corollary 4.24 is only a partial answer to Question 4.12 since we only have an isomorphism of vector spaces and not of fields.

4.3 Polynomials over fields

Corollary 4.24 tells us that a finite field of characteristic p must have p^n elements for some $n \in \mathbb{N}$. However, for $n > 1$, it is not clear that such a field actually exists and, if it does, what its explicit structure is. To answer this question, we

will need to digress to a discussion on polynomials over a field. Polynomials are also a fundamental object in Cryptography so the material in this section is of more general interest.

Definition 4.25. Let $\mathbb{F}$ be a field. A **polynomial** (in one variable) over $\mathbb{F}$ is a formal expression of the form

$$f(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n = \sum_{i=0}^{n} a_ix^i$$

where $a_0, ..., a_n \in \mathbb{F}$ are called the **coefficients** of f. We assume that $a_n \neq 0$ and write $\deg f = n$. The set of all polynomials over $\mathbb{F}$ is denoted

$$\mathbb{F}[x] := \{f(x) = a_0 + a_1x + ... + a_nx^n \mid 0 \leq n,\ a_0, ..., a_n \in \mathbb{F},\ a_n \neq 0\} \cup \{0\}.$$

The element $0 \in \mathbb{F}[x]$ is called the **zero polynomial** and by convention $\deg(0) = -\infty$.

Remark 4.26. By definition, two polynomials

$$f(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n$$

and

$$g(x) = b_0 + b_1x + b_2x^2 + ... + b_kx^k$$

are equal iff $n = k$ and for all i, $a_i = b_i$.

Example 4.27. Let $\mathbb{F} = \mathbb{F}_2 = \{0, 1\}$ be the field of two elements. Consider $f(x) = x^2$ and $g(x) = x$. Then $f(x) \neq g(x)$ although they define the same function $\mathbb{F}_2 \longrightarrow \mathbb{F}_2$: $0 \mapsto 0$, $1 \mapsto 1$.

If

$$f(x) = \sum_{i=0}^{n} a_ix^i$$

and

$$g(x) = \sum_{i=0}^{k} b_ix^i$$

are two polynomials in $\mathbb{F}[x]$ where (without loss of generality) $\deg g = k \leq n = \deg f$ we can define their sum as

$$f(x) + g(x) = \sum_{i=0}^{k}(a_i + b_i)x^i + \sum_{i=k+1}^{n} a_ix^i$$

Example 4.28. Over $\mathbb{F} = \mathbb{R}$,

$$(4x^3 + 2x) + (5x^4 + 3x^3 + 2) = 5x^4 + 7x^3 + 2x + 2.$$

Similarly we can define

$$f(x) \cdot g(x) = \left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{i=0}^{k} b_i x^i\right) = \sum_{i=0}^{n+k} c_i x^i$$

where

$$c_k = \sum_{i=0}^{k} a_i b_{k-i}$$

Example 4.29. Let $\mathbb{F} = \mathbb{F}_2$, $f(x) = x^2 - 1$ and $g(x) = x + 1$. Then

$$f(x)g(x) = (x^2 + x - 1)(x + 1) = x^3 + x^2 + x^2 + x - x - 1 = x^3 + 2x^2 - 1 = x^3 - 1.$$

where the last equality comes from the fact that $2 = 0$ in $\mathbb{F}_2$.

Although a polynomial $f(x) = \sum_i a_i x^i$ over a field $\mathbb{F}$ is a formal expression, every such polynomial defines a function $\mathbb{F} \longrightarrow \mathbb{F}$ given by $t \mapsto \sum_i a_i t^i$. We will be particularly interested in values t that are mapped to $0_{\mathbb{F}}$:

Definition 4.30. Let $\mathbb{F}$ be a field and $f = f(x) = \sum_i a_i x^i \in \mathbb{F}[x]$ a polynomial. An element $t \in \mathbb{F}$ is a called a **zero** (or root) of f if $\sum a_i t^i = 0$ (in $\mathbb{F}$).

Example 4.31. Let $f(x) = x^n - 1 \in \mathbb{C}[x]$ be a polynomial over the complex numbers. Then the set of all zeros of f is precisely the group μ_n of nth roots of unity. Note that when we think of f over the real numbers, i.e. $f(x) \in \mathbb{R}[x]$ the numbers of zeros changes: for example when $n = 3$ we have only one zero of f over $\mathbb{R}$ and 3 zeros of f over $\mathbb{C}$.

We have defined two binary operations

$$+ : \mathbb{F}[x] \times \mathbb{F}[x] \longrightarrow \mathbb{F}[x]$$

and

$$\cdot : \mathbb{F}[x] \times \mathbb{F}[x] \longrightarrow \mathbb{F}[x].$$

We note that the zero polynomial 0 can be viewed as a unit element with respect to + and the polynomial 1 can be viewed as a unit element with respect to $\cdot$. If $f(x) = \sum_i a^i x^i$ then $-f(x) := \sum_i -a_i x^i$ is clearly an additive inverse to $f(x)$ i.e. $f(x) + (-f(x)) = 0_{\mathbb{F}[x]}$. However, in general, there is no multiplicative inverse to $f(x)$: if $\deg f \geq 1$ then for any $g(x) \in \mathbb{F}[x]$ we have $\deg(f \cdot g) = \deg f + \deg g \geq 1$ whereas $\deg(1) = 0$ so there is no $g(x)$ such that $f(x)g(x) = 1$. Note that the data of $\mathbb{F}[x]$, the elements $0, 1 \in \mathbb{F}[x]$ and the binary operations $+, \cdot$ **almost form a field**: the only axiom not satisfied is that of a multiplicative inverse.

Remark 4.32. As can be easily seen, a non-zero polynomial $f(x) \in \mathbb{F}[x]$ is invertible iff $\deg f = 0$ (i.e. $f(x) = c$ for $c \in \mathbb{F}^\times$) and in that case we call $f(x)$ a **unit** in $\mathbb{F}[x]$: if $f(x) \in \mathbb{F}[x]$ is a polynomial of degree $1 \leq n$, then for any polynomial $g(x) \in \mathbb{F}[x]$, $\deg(f(x)g(x)) \geq n$ hence we cannot have $f(x)g(x) = 1$ for degree reasons.

Definition 4.33. A set R together with elements $0, 1 \in R$ and two binary operations

$$+ : R \times R \longrightarrow R$$

and

$$\cdot : R \times R \longrightarrow R$$

is called a **ring** if it satisfies all the axioms of a field in Definition 4.1 except (possibly) axiom (6) – existence of inverses wrt $\cdot$.

Examples 4.34.

1. Any field $\mathbb{F}$ is also a ring.
2. For $\mathbb{F}$ a field, $\mathbb{F}[x]$ is a ring.
3. The integers $\mathbb{Z}$ with ordinary addition and multiplication are a ring.
4. The integers modolu n, $\mathbb{Z}_n$ with addition and multiplication modolu n are a ring.

4.4 Euclidean Algorithm for Polynomials

Let $\mathbb{F}$ be a field and $f(x), g(x) \in \mathbb{F}[x]$ be non-zero polynomials with $n = \deg f \geq \deg g = k$. Write

$$f(x) = \sum_{i=0}^{n} a_i x^i$$

and

$$g(x) = \sum_{i=0}^{k} b_i x^i$$

with $a_n, b_k \in \mathbb{F}^\times$.

Set $n_0 = n$. Then for $c_0 := \frac{a_n}{b_k}$ we have

$$f(x) - c_0 x^{n_0 - k} g(x) = r_1(x) = (a_{n-1} - b_{k-1} c_1) x^{n-1} + \ldots$$

where $r_1(x)$ is a polynomial with $n_1 = \deg r_1 < n = n_0$. If $n_1 \geq k$ we can similarly write

$$r_1(x) - c_1 x^{n_1 - k} g(x) = r_2(x)$$

where $n_2 = \deg r_2 < n_1$ and $c_1 \in \mathbb{F}^\times$. Since the degrees of r_i's strictly decrease in each iteration, this process can last only finitely many steps s and we get a set of polynomials $\{r_i(x)\}_{i=1}^{s}$ and coefficients $\{c_i\}_{i=1}^{s-1} \subseteq \mathbb{F}^\times$ such that:

$$f(x) = c_0 x^{n_0 - k} g(x) + r_1(x),$$

$$r_1(x) = c_1 x^{n_1 - k} g(x) + r_2(x),$$

$$\ldots$$

$$r_{s-1}(x) = c_{s-1} x^{n_{s-1} - k} g(x) + r_s(x)$$

where $\deg r_s < k$. Combining the above equations we get

$$f(x) = g(x)\left(c_0 x^{n_0-k} + c_1 x^{n_1-k} + ... + c_{s-1}x^{n_{s-1}-k}\right) + r_s(x).$$

Let us denote

$$q(x) = c_0 x^{n_0-k} + c_1 x^{n_1-k} + ... + c_{s-1}x^{n_{s-1}-k}$$

and $r(x) = r_s(x)$. We are ready to state:

Theorem 4.35. *Let $\mathbb{F}$ be a field and $f(x), g(x) \in \mathbb{F}[x]$ be non-zero polynomials with $\deg f \geq \deg g$. Then there exist unique polynomials $q(x), r(x) \in \mathbb{F}[x]$ such that*

$$f(x) = g(x)q(x) + r(x)$$

and with $\deg r < \deg g$.

Proof. The discussion above proved existence so it remains to prove uniqueness. Suppose $q'(x), r'(x)$ also satisfy

$$f(x) = g(x)q'(x) - r'(x)$$

with $\deg r' < \deg g$. Then

$$g(x)q(x) - r(x) = g(x)q'(x) - r'(x)$$

so that

$$g(x)\left(q(x) - q'(x)\right) = r'(x) - r(x)$$

But

$$\deg(r - r') < \deg g \leq \deg(g(q - q'))$$

and to avoid contradiction we must have

$$q(x) - q'(x) = 0$$

after which it follows that $r'(x) - r(x) = 0$ and we are done. □

Terminology 1. *Under the notation of 4.35 we call q the **quotient** of f by g and call r the **remainder**. We also denote $f\%g = r$.*

The above terminology is meant to sharpen the analogy between modular arithmetic of integers and polynomials. Before diving into 'mod-$g(x)$ arithmetic', let us draw an important

Corollary 4.36. *Let $\mathbb{F}$ be a field and $f(x) \in \mathbb{F}[x]$. If $\alpha \in \mathbb{F}$ is a zero of $f(x)$ then there exists a polynomial $q(x) \in \mathbb{F}[x]$ such that $f(x) = (x - \alpha)q(x)$.*

Proof. Apply Theorem 4.35 with $g(x) = (x - \alpha)$ to get $q(x), r(x) \in \mathbb{F}[x]$ such that

$$f(x) = (x - \alpha)q(x) + r(x)$$

and $\deg r < \deg(x - \alpha) = 1$ so that $r(x) = c$ is a constant. Since α is a zero of f we get $r(\alpha) = 0$ so $r(x) = 0$ is the zero polynomial. □

Corollary 4.37. *Let $\mathbb{F}$ be a field. Then a polynomial $f(x) \in \mathbb{F}[x]$ of degree n has at most n roots.*

Proof. If f has $n+1$ roots (or more) then a repeated application of Corollary 4.36 yields a decomposition

$$f(x) = \prod_{i=1}^{n+1}(x - \alpha_i)g(x)$$

and the right-hand-side has degree $\geq n+1$ in contradiction to the fact that $\deg f = n$. □

We know that for every two distinct points in $\mathbb{R}^2$, we have a unique polynomial of degree 1 that passes through them. It is not hard to see that the formula for such a line is valid in any field. In fact, any $n+1$ distinct points in $\mathbb{F} \times \mathbb{F}$ determine a unique polynomial in $\mathbb{F}[x]$ that 'passes' through them as the next result shows:

Theorem 4.38 (Lagrange interpolation polynomial)**.** *Let $\mathbb{F}$ be a field and $(x_1, y_1), ..., (x_{n+1}, y_{n+1}) \in \mathbb{F} \times \mathbb{F}$ a set of $n+1$ pairs of elements of $\mathbb{F}$ such that $\forall i \neq j : x_i \neq x_j$. Then there exists a unique polynomial $f(x) \in \mathbb{F}[x]$ of degree at most n such that $f(x_i) = y_i$ for $i = 1, ..., n+1$.*

Proof. Fix $1 \leq i \leq n+1$. We have $\prod_{j \neq i}(x_i - x_j) \neq 0$ since all factors are non-zero. Define a polynomial

$$D_i(x) = \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)}.$$

Then $D_i(x_i) = 0$ and $D_i(x_j) = 1$ for every $j \neq i$.

We define

$$f(x) = \sum_{i=1}^{n+1} y_i D_i(x).$$

Then $f(x_i) = y_i \cdot 1 = y_i$ and f is of degree $\leq n$ as a sum of degree n polynomials. For uniqueness, suppose by contradiction that $g(x)$ is another polynomial of degree $\leq n$ such that $g(x_i) = y_i$ for $i = 1, ..., n+1$. Then $h(x) = f(x) - g(x)$ is a polynomial of degree $\leq n$ (as a sum of such) but for every i, $g(x_i) = f(x_i) - g(x_i) = y_i - y_i = 0$ so f has $n+1$ distinct roots in contradiction to Corollary 4.37. □

The discussion (or proof) preceding Theorem 4.35 gives in fact an algorithm to perform long division of polynomials.

Examples 4.39. Let $\mathbb{F} = \mathbb{Q}$.

1. Take $f(x) = 6x^3 - 2x^2 + x + 3$ and $g(x) = x^2 - x + 1$. Then long division $f(x)\%g(x)$ yields the following

$$
\begin{array}{r}
\multicolumn{2}{r}{6x+4.} \\
\cline{2-2}
x^2-x+1\big) & 6x^3-2x^2+x+3 \\
& -6x^3+6x^2-6x \\
\cline{2-2}
& 4x^2-5x+3 \\
& -4x^2+4x-4 \\
\cline{2-2}
& -x-1
\end{array}
$$

Thus

$$6x^3-2x^2+x+3=(x^2-x+1)(6x+4)-x-1.$$

2. Take $f(x)=2x^4-x^2+x+3$ and $g(x)=x^2-1$. Then long division $f(x)\%g(x)$ yields the following

$$
\begin{array}{r}
\multicolumn{2}{r}{2x^2+1.} \\
\cline{2-2}
x^2-1\big) & 2x^4-x^2+x+3 \\
& -2x^4+2x^2 \\
\cline{2-2}
& x^2+x+3 \\
& -x^2+1 \\
\cline{2-2}
& x+4
\end{array}
$$

Thus

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

Definition 4.40. Let $g(x)\in\mathbb{F}[x]$ be a polynomial of degree n. The set of possible **remainders** of mod-$g(x)$ division can be described as

$$\mathcal{R}_{\mathbb{F},n}=\left\{a_0+a_1x+...+a_{n-1}x^{n-1}|a_0,...,a_{n-1}\in\mathbb{F}\right\}$$

(note that here we allow all coefficients, including a_{n-1} to be 0).

Mod-$g(x)$ arithmetic is derived from polynomial arithmetic as follows. Let

$$r(x)=f(x)\ mod\ g(x)$$

and

$$s(x)=h(x)\ mod\ g(x).$$

Then there are quotient polynomials $q(x),t(x)$ such that

$$r(x)=f(x)-q(x)g(x)$$

and

$$s(x)=h(x)-t(x)g(x).$$

Thus

$$f(x)+h(x)=r(x)+s(x)-(q(x)+t(x))g(x)$$

$$f(x)h(x)=r(x)s(x)-(q(x)s(x)+t(x)r(x))g(x)+q(x)t(x)g(x)g(x).$$

It follows that

$$f(x)+h(x)=(r(x)+s(x))\ mod\ g(x)$$

and

$$f(x)h(x) = (r(x)s(x)) \ mod\ g(x)$$

Observe that mod-$g(x)$ addition on $\mathcal{R}_{\mathbb{F},n}$ for

$$r(x) = a_0 + a_1x + ... + a_{n-1}x^{n-1}$$

and

$$s(x) = b_0 + b_1x + ... + b_{n-1}x^{n-1}$$

is given by adding coefficients 'component-wise':

$$(r(x) + s(x)) \ mod\ g(x) = (a_0 + b_0) + (a_1 + b_1)x + ... + (a_{n-1} + b_{n-1})x^{n-1}$$

When $\mathbb{F} = \mathbb{F}_p$ is the field of p elements, we have $|\mathcal{R}_{\mathbb{F},n}| = p^n$ and the discussion above endows $\mathcal{R}_{\mathbb{F}_p,n}$ with the structure of a vector space that is isomorphic to $\mathbb{F}_p^n$. The addition of vectors arises from the addition of polynomials $mod\ g(x)$ and scalar multiplication arises from multiplying a polynomial by a scalar and taking $mod\ g(x)$ of the result.

Let $\mathbb{F}$ be a field and $g(x) \in \mathbb{F}[x]$ with $\deg g = n$. We will denote by $\mathbb{F}_{g(x)}$ the set $\mathcal{R}_{\mathbb{F},n}$ with mod-$g(x)$ arithmetic. From the discussion above, $\mathbb{F}_{g(x)}$ is a ring which is isomorphic as a vector space over $\mathbb{F}$ to $\mathbb{F}^n$. We are interested in polynomials $g(x)$ for which $\mathbb{F}_{g(x)}$ is a field.

Definition 4.41. Let $\mathbb{F}$ be a field. A polynomial $g(x)$ is **prime** (or irreducible) if it cannot be decomposed as $g(x) = a(x)b(x)$ for non-constant polynomials $a(x), b(x) \in \mathbb{F}[x]$.

Examples 4.42.

1. Clearly, any degree 1 polynomial $f(x) = x + a$ for $a \in \mathbb{F}$ is prime.
2. For $\mathbb{F} = \mathbb{F}_3$, the polynomial $g(x) = x^2 + 1$ is prime. If it wasn't we would have $g(x) = x^2 + 1 = a(x)b(x)$ where $\deg a = \deg b = 1$, i.e. $a(x) = \alpha_1x + \alpha_0$ and $b(x) = \beta_1x + \beta_0$ with $\alpha_1, \beta_1 \in \mathbb{F}_p^\times$. But then $a = -\frac{\alpha_0}{\alpha_1}$ is a root of g. However, we can easily see that $g(x)$ has no root over $\mathbb{F}$.
3. the polynomial $g(x) = x^2 + 1$ of the previous example is not prime over the field $\mathbb{F}_5$: since $3^2 + 1 = 0$ ($mod\ 5$) and $2^2 + 1 = 0$ ($mod\ 5$) we have $x^2 + 1 = (x - 3)(x - 2)$.

The terminology of Definition 4.41 is meant to suggest an analogy between numbers and polynomials. To flesh out the analogy, it will be convenient to restrict attention to **monic** polynomials i.e. polynomials of the form $x^n + a_{n-1}x^{n-1} + ... + a_1x + a_0$. Of course, any polynomial is a product of a monic polynomial and a non-zero field element. Let us start with

Proposition 4.43 (prime factorisation of polynomials)**.** *Let $f(x) \in \mathbb{F}[x]$ be a monic polynomial. Then there is $k \geq 1$ and monic prime polynomials*

$$a_1(x), ..., a_k(x) \in \mathbb{F}[x]$$

such that

$$f(x) = a_1(x) \cdot \ldots \cdot a_k(x).$$

Furthermore, this factorisation is unique up to re-ordering of the factors $a_i(x)$.

Proof. Clearly such factorisation must exist: if f is not prime we can write $f(x) = g(x)h(x)$ with $\deg g, \deg h < \deg f$ and continue factoring until we get monic prime factors. It remains to prove uniqueness. Suppose by contradiction that there is a polynomial with non-unique (monic) prime factorisation and let n be the minimal degree for which such polynomial $f(x)$ exist. Thus we can write $f(x)$ in two ways

$$a_1(x) \cdot \ldots \cdot a_k(x) = b_1(x) \cdot \ldots \cdot b_l(x) \tag{4}$$

where a_i, b_j are (monic) prime and $k, l \geq 1$. Now, $a_1(x)$ cannot appear on the right hand side of 4 since we could then cancel it from both sides and obtain contradiction to the minimality of n. Similarly, $b_1(x)$ cannot appear on the left hand side. Without loss of generality, assume $\deg b_1 \leq \deg a_1$. By Theorem 4.35 we can write $a_1(x) = b_1(x)q(x) + r(x)$. Since $a_1(x)$ is prime, $r(x) \neq 0$ and $\deg r < \deg b_1 \leq \deg a_1$. Now, $r(x)$ has a prime factorisation $r(x) = \alpha \cdot r_1(x) \cdot \ldots \cdot r_m(x)$ with $\alpha \in \mathbb{F}^\times$ and $b_1(x)$ cannot divide any of the $r_i(x)$ since it has a higher degree. Substituting that to 4 we get

$$\left(q(x)b_1(x) + \alpha r_1(x)...r_m(x)\right) a_2(x)...a_k(x) = b_1(x)...b_l(x).$$

Define $f'(x) = r_1(x)...r_m(x)a_2(x)...a_k(x)$. Then f' is monic as a product of such and $\deg f' < \deg f$. However, we can write

$$f'(x) = r_1(x)...r_m(x)a_2(x)...a_k(x) = \alpha^{-1}b_1(x)\left(b_2(x)...b_l(x) - q(x)a_2(x)...a_k(x)\right)$$

and these are two factorisations of f' in which the prime polynomial b_1 appear in one but not the other. This is a contradiction to the minimality of n. □

In light of Proposition 4.43 we can extend the analogy between polynomials and numbers with the following

Definition 4.44. Let $f(x), g(x)$ be monic polynomials. The **greatest common divisor** of f and g is the polynomial of maximal degree $\gamma(x)$ that divides both $f(x)$ and $g(x)$. We denote $\gamma = \gcd(f, g)$.

Remark 4.45. Note that the polynomial $\gamma = \gcd(f, g)$ of Definition 4.44 is unique by the uniqueness of factorisation to prime polynomials (Proposition 4.43).

Proposition 4.46. *For any field* $\mathbb{F}$, $\gcd(x^n - 1, x^m - 1) = x^{\gcd(n,m)} - 1$.

Proof. Without loss of generality, assume $n \leq m$ and use induction on m. If $m = 1$ then $n = 1$ and the claim is trivial. Suppose the claim is true for all $m' < m$. The case $n = m$ is trivial so we can further assume $n < m$. Then

$$x^m - 1 - x^{m-n}(x^n - 1) = x^{m-n} - 1$$

so that a polynomial g dividing both $x^m - 1$ and $x^n - 1$ must divide $x^{m-n} - 1$. By induction,

$$\gcd(x^{m-n} - 1, x^n - 1) = x^{\gcd(m-n,n)} - 1$$

but $\gcd(m - n, n) = \gcd(m, n)$ and

$$x^m - 1 = x^{m-n}(x^n - 1) + x^{m-n} - 1$$

so

$$\gcd(x^m - 1, x^n - 1) = \gcd(x^{m-n} - 1, x^n - 1) = x^{\gcd(m-n,n)} - 1 = x^{\gcd(m,n)} - 1.$$

□

4.5 Classification of finite fields

We are now ready to state

Theorem 4.47. *Let $\mathbb{F}$ be a field and $g(x)$ a prime polynomial in $\mathbb{F}[x]$ with $\deg g = n$. Then $\mathbb{F}_{g(x)}$ together with mod-$g(x)$ arithmetic is a field.*

The proof of Theorem 4.47 is remarkably similar to the proof that $\mathbb{F}_p$ is a field: using the extended Euclidean algorithm. We thus need a version of the extended Euclidean algorithm for polynomials:

Theorem 4.48. *Let $\mathbb{F}$ be a field and $a(x), b(x) \in \mathbb{F}[x]$ be polynomials. Then there exist polynomials $s(x), t(x), f(x)$ such that*

$$a(x)s(x) + b(x)t(x) = f(x)$$

where $f(x) = \gcd(a(x), b(x))$.

The proof of Theorem 4.48 in turn, is remarkably similar to that of the Extended Euclidean Algorithm for integers, using Theorem 4.35 and we will omit it.

Proof of Theorem 4.47. As discussed above, we only need to show multiplicative inverse. Let $r(x) \in \mathbb{F}_{g(x)}$. Using Theorem 4.48 with $a(x) = r(x)$ and $b(x) = g(x)$ we get polynomials s, t, γ such that

$$r(x)s(x) + g(x)t(x) = \gamma(x).$$

Since $g(x)$ is prime, $1 = \gamma(x) = \gcd(r(x), g(x))$ so

$$r(x)s(x) = 1 \ mod \ g(x).$$

□

Corollary 4.49. *Let $\mathbb{F} = \mathbb{F}_p$. Then every prime polynomial $g(x) \in \mathbb{F}_p$ of degree n gives rise to a field with p^n elements.*

Proof. The set $\mathbb{F}_{g(x)}$ with mod-$g(x)$ arithmetic is a field by Theorem 4.47. By construction, the set $\mathbb{F}_{g(x)}$ is in one-to-one correspondence with the set of remainder polynomials $\mathcal{R}_{\mathbb{F},n} = \{r(x) = a_0 + a_1x + ... + a_{n-1}x^{n-1} | a_0, ..., a_{n-1} \in \mathbb{F}_p\}$ whose size is p^n. □

Example 4.50. Let us construct a field with $2^2 = 4$ elements. We first observe that the polynomial $g(x) = x^2+x+1$ is prime over $\mathbb{F}_2$ since it has no roots. Thus, $\mathbb{F}_{g(x)}$ is a field with $2^{\deg g} = 4$ elements. There are four possible remainder polynomials $\{0, 1, x, x+1\}$. Addition is componentwise $mod\ 2$. For multiplication, note that $x*x = x^2 = x+1$ ($mod\ x^2+x+1$). Also, $x*x*x = x^3 = 1$ ($mod\ x^2+x+1$) so the three non-zero elements $\{1, x, x+1\}$ form a cyclic group under mod-$g(x)$ multiplication.

The complete mod-$g(x)$ addition and multiplication tables are given as follows.

+	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

×	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

In order to prove the existence of finite fields of order p^n we need to show that there exist a prime polynomial over $\mathbb{F}_p$ of degree n. It is easy to show a slightly weaker statement:

Proposition 4.51. *Let p be a prime. Then for any $N \in \mathbb{N}$ there exist $n > N$ and a prime polynomial over $\mathbb{F}_p$ of degree n.*

Proof. Assume by contradiction that there is $N \in \mathbb{N}$ for which the statement is false, i.e there is no prime polynomial of degree $\geq N$. Note that the set of polynomials of degree $\leq N$ in $\mathbb{F}_p[x]$ is finite since $\mathbb{F}_p$ has only finitely many elements. It follows that the set of all (monic) prime polynomials over $\mathbb{F}_p$ is finite, say, $\{p_1(x), ..., p_k(x)\}$. Consider the polynomial $p(x) = 1 + p_1(x)...p_k(x)$. Clearly $p_i(x) \nmid p(x)$ for all $1 \leq i \leq k$ (since the remainder is 1) but this is a contradiction to the prime factorisation of Proposition 4.43. □

Exercise 4.52. Describe a field with $3^2 = 9$ elements together with its addition and multiplication tables.

We have proved via Theorem 4.47 and Proposition 4.51 that for any prime p there exist finite fields of characteristic p of arbitrary size. For $n = 1$ we showed that such a field is unique up to isomorphism. Can we say the same for $n > 1$? The answer lies in the following

Theorem 4.53 (Structure theorem for finite fields)**.** *Let p be prime. Then for any $n \in \mathbb{N}$ there exist a field of characteristic p with exactly $q = p^n$ elements. Moreover, such a field is unique up to isomorphism and denoted $\mathbb{F}_q$.*

The proof of Theorem 4.53 is beyond the scope of these notes. Let us finish this section with a result that may be useful for future considerations.

Theorem 4.54. *Let p be prime and $q = p^k$. The multiplicative group of the field of q elements $G = \mathbb{F}_q^\times$ is cyclic.*

Proof. We denote $n = q - 1$ so that $|G| = n$. By Lagrange Theorem 3.23, the order d of an element $a \in G$ must divide n since $< a >$ is a subgroup of G of order d. Let d be a divisor of n and denote by $\psi(d)$ the number of elements in G of order d. If $a \in G$ is of order d, then $H =< a >= \{1, a, ..., a^{d-1}\} \leq G$ is a cyclic group of order d so that every $x \in H$ satisfies $x^d = 1$. Since $\mathbb{F} = \mathbb{F}_q$ is a field, the polynomial $f(x) = x^d - 1$ can have at most d roots and it follows that it has exactly d roots – the elements of H. Note that not all elements in H have order d. Rather, the number of elements of H of order d are precisely its generators. By Corollary 3.39 the number of such generators is precisely $\varphi(d)$ and it follows that $\psi(d) = 0$ or $\psi(d) = \varphi(n)$. By Theorem 3.40 we have

$$n = \sum_{d:d|n} \varphi(d)$$

and since G is a finite group, we have

$$n = \sum_{d:d|n} \psi(d).$$

Thus,

$$\sum_{d:d|n} \varphi(d) = \sum_{d:d|n} \psi(d)$$

and it follows that $\psi(n) = \varphi(n)$ so G has at least one generator of order n. □

4.6 Algebraic Closure

Let p be prime, $q = p^n$. If $\beta \in \mathbb{F}_q^\times$ then $\langle\beta\rangle \leq \mathbb{F}_q^\times$ is a subgroup of the multiplicative field of $\mathbb{F}_q$ so that $|\langle\beta\rangle| \mid q - 1$. It follows that $\beta^{q-1} = 1$ in $\mathbb{F}_q$ ie β is a root of the polynomial $f(x) = x^q - x$ over $\mathbb{F}_q$. Since f can have at most q roots, it follows that the elements of $\mathbb{F}_q$ are precisely the roots of f and we can thus write f as a product of distinct linear factors

$$f(x) = \prod_{\beta \in \mathbb{F}_q} (x - \beta). \tag{5}$$

Let p be a prime, $q = p^n$ and consider the field $\mathbb{F}_q$ given by Theorem 4.53. We revoke a small result from elementary Number Theory

Lemma 4.55. *If p is prime and $m \mid n$ then $p^m - 1 \mid p^n - 1$.*

Theorem 4.56. *Let $q = p^n$. Then for every $m \mid n$ there is a unique subfield of $\mathbb{F}_q$ given by the roots of $x^{p^m} - x$ over $\mathbb{F}_q$. Moreover, every subfield of $\mathbb{F}_q$ is obtained in this form.*

Proof. Let us start with the last statement. If $K \subseteq \mathbb{F}_q$ is a subfield then $\mathbb{F}_q$ is a vector space over K which is finite dimensional (since $\mathbb{F}_q$ is finite). Hence, by Theorem 4.23 we have an isomorphism of vector spaces over K, $\mathbb{F}_q \cong K^d$. It follows that $|K| \mid |\mathbb{F}_q| = p^n$ so that $|K| = p^m$ with $m \mid n$. Let us turn to uniqueness. If $K \subseteq \mathbb{F}_q$ is a subfield of order p^m then $K^\times$ is a cyclic group of order $p^m - 1$ hence its elements are roots of $x^{p^m-1} - 1$ over $\mathbb{F}_q$. Thus the elements of K must be precisely the roots of $x^{p^m} - x$. Lastly, we prove existence. Let $m \mid n$. We know from Equation 5 that the elements of $\mathbb{F}_q^\times$ can be identified with the roots of the polynomial $x^{p^n-1} - 1$. By Lemma 4.55 $p^m - 1 \mid p^n - 1$, and by Proposition 4.46 the polynomial $x^{p^m-1} - 1$ divides $x^{p^n-1} - 1$ in $\mathbb{F}_p[x]$ hence $x^{p^m-1} - 1$ decomposes to linear factors over $\mathbb{F}_q$. Let K be the set of roots of $x^{p^m} - x$ over $\mathbb{F}_q$, or the set of roots of $x^{p^m-1} - 1$ together with 0, so that $|K| = p^m$. Clearly, the elements of K are closed under multiplication since $(\alpha\beta)^{p^m-1} = \alpha^{p^m-1}\beta^{p^m-1} = 1 \cdot 1 = 1$. Similarly, the elements of K are closed under inverse. To see clossness to addition, observe that $(\alpha + \beta)^{p^m} = (\alpha + \beta)$ for any field of characteristic p by Lemma 3.7. Thus, $K \subseteq \mathbb{F}_q$ is indeed a subfield. □

Theorem 4.56 allows us to identify $\mathbb{F}_{p^m}$ as a subfield of $\mathbb{F}_{p^n}$ whenever $m \mid n$. In particular, we have a sequence of subfield inclusions

$$\mathbb{F}_p = \mathbb{F}_{p^{1!}} \subseteq \mathbb{F}_{p^{2!}} \subseteq \ldots \subseteq \mathbb{F}_{p^{n!}} \subseteq \ldots$$

This in turn, enables us to give the following

Definition 4.57. Let $\mathbb{F}$ be a field with $\operatorname{char}\mathbb{F} = p$. The **algebraic closure** of $\mathbb{F}$, denoted $\overline{\mathbb{F}}$ is the union

$$\overline{\mathbb{F}} = \bigcup_{n \in \mathbb{N}} \mathbb{F}_{p^{n!}}$$

of the subfield inclusions described above.

Proposition 4.58. *Let $\mathbb{F}$ be a field of characteristic p. Then $\overline{\mathbb{F}}$ admits a structure of a field.*

Proof. Note first that by Theorem 4.56, any field $\mathbb{F}_{p^n}$ can be identified as a subfield of $\mathbb{F}_{p^{n!}}$ (since $n \mid n!$) and hence is contained in $\overline{\mathbb{F}}$. In addition for every n, $\mathbb{F}_p$ is a subfield of $\mathbb{F}_{p^{n!}}$. Thus we define the neutral elements by $0, 1 \in \mathbb{F}_p$. To define addition and multiplication let $x, y \in \overline{\mathbb{F}}$. Then there are n, k such that $x \in \mathbb{F}_{p^{n!}}$ and $y \in \mathbb{F}_{p^{k!}}$. But then, $x, y \in \mathbb{F}_{p^{n!k!}}$ and we define their addition and multiplication to be the one in $\mathbb{F}_{p^{n!k!}}$. If $0 \neq x \in \overline{\mathbb{F}}$ then $x \in \mathbb{F}_{p^{n!}}$ and thus $-x$ and x^{-1} are defined in $\overline{\mathbb{F}}$ as they are in $\mathbb{F}_{p^{n!}}$. Associativity, commutativity and distributivity are satisfied in all $\mathbb{F}_{p^{n!}}$ hence also in $\overline{\mathbb{F}}$. □

Let $\mathbb{F}$ be a field and $g(x) \in \mathbb{F}[x]$ a prime polynomial. The set $\mathbb{F}_{g(x)}$ with mod-$g(x)$ addition and multiplication is a field by Theorem 4.47 and contains $\mathbb{F}$ as the set of constant remainder polynomials. Thus, we can consider g as a polynomial over $\mathbb{F}_{g(x)}$. Consider the quotient map $\pi : \mathbb{F}[x] \longrightarrow \mathbb{F}_{g(x)}$ given by $\pi(f(x)) = f(x) \ mod \ g(x)$. Note that when g is considered as a polynomial over $\mathbb{F}_{g(x)}$, $g(\pi(x)) = \pi(g(x)) = 0 \ mod \ g(x)$ so that $\pi(x)$ is a root of g. In light of this, we are ready to prove the main property of the algebraic closure:

Theorem 4.59. *Let $g(x) \in \overline{\mathbb{F}}_p[x]$ be a polynomial of degree n. Then $g(x)$ decomposes to linear factors*

$$g(x) = \prod_{i=1}^{n}(x - \alpha_i)$$

for $\alpha_i \in \overline{\mathbb{F}}_p$.

Proof. It is enough to prove the theorem for prime polynomials, so assume g is prime. Since g has finitely many coefficients, they all must lie in some field $K = \mathbb{F}_q$. By the discussion above, when we consider g as a polynomial over $K_{g(x)}$ we get $g(x) = (x - \alpha)g_1(x)$ for some polynomial $g_1(x)$ over $K_{g(x)}$ with $\deg g_1 = \deg g - 1$. But $K_{g(x)} \cong \mathbb{F}_{q^n}$ hence is contained in $\overline{\mathbb{F}}_p$. Thus, as polynomials over $\overline{\mathbb{F}}_p$ we get a decomposition $g(x) = (x - \alpha)g_1(x)$. Repeat this procedure for $g_1, g_2, ...$ to decompose $g(x)$ to linear factors. □

In light of Theorem 4.59, we can decompose the polynomial $x^n - 1$ into linear factors $x^n - 1 = \prod_{i=1}^{n}(x - \alpha_i)$ over $\overline{\mathbb{F}}_p$. The considerations in Example 3.3 (5), apply in this case as well so that the collection of all distinct roots of $x^n - 1$ forms a group under multiplication.

Definition 4.60. Let p be prime. The group of n**th roots of unity** over $\overline{\mathbb{F}}_p$, denoted $\mu_n = \mu_n(\mathbb{F}_p)$ is the set of roots of the polynomial $x^n - 1$ in $\overline{\mathbb{F}}_p$.

Since we are in positive characteristic, the order of $\mu_n(\overline{\mathbb{F}}_p)$ may be different than n.

Proposition 4.61. *If $p \nmid n$, then the group of nth roots of unity μ_n is cyclic of order n.*

Proof. The case $n = 1$ is trivial so suppose $n \geq 2$. Since $p \nmid n$, the polynomial $x^n - 1$ and its derivative nx^{n-1} have no common roots so the linear factors in the decomposition $x^n - 1 = \prod_{i=1}^{n}(x - \alpha_i)$ are all distinct so that $|\mu_n(\overline{\mathbb{F}}_p)| = n$. Moreover, $\mu_n(\overline{\mathbb{F}}_p)$ is a subgroup of $\mathbb{F}_q^\times$ for some q and the latter is cyclic by Theorem 4.54. By Theorem 3.41, a subgroup of a cyclic group must by cyclic and the result follow. □

Throughout this section, we developed the algebraic closure of finite fields. In fact, algebraic closure is a construction that exists for any field and is characterized by the property of Theorem 4.59, ie. that over an algebraic closure any polynomial splits to linear factors. For the sake of completeness, we record

Theorem 4.62 (fundamental theorem of algebra)**.** *The algebraic closure of the real numbers $\mathbb{R}$ are the complex numbers $\mathbb{C}$, i.e. $\overline{\mathbb{R}} = \mathbb{C}$. In particular, every polynomial $f(x) \in \mathbb{C}[x]$ admits a decomposition*

$$f(x) = \prod_i (x - \alpha_i).$$

Proof. We will not use this theorem much since we mostly work over finite fields. A proof can be found, for example, in [FR]. □

5 Elliptic Curves

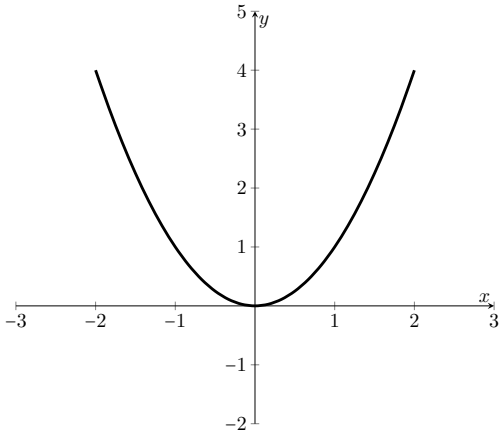
Let $\Bbbk$ be a field. We saw that a polynomial in one variable $f(x) \in \Bbbk[x]$ must have finitely many roots. We can however consider polynomials in more variables

Definition 5.1. A **polynomial in two variables** x, y over a field $\Bbbk$ is a formal expression of the form

$$f(x,y) = a_{0,0} + a_{1,0}x + a_{0,1}y + a_{1,1}xy + a_{2,1}x^2y + \ldots + a_{n,k}x^n y^k = \sum_{0 \le i \le n, 0 \le j \le k} a_{i,j} x^i y^j$$

where $a_{n,k} \neq 0$. Here we define $\deg f = n + k$. We will denote the set of all such polynomials by $\Bbbk[x,y]$. A **root** (or **zero**) of $f(x,y)$ is a point $(x_0, y_0) \in \Bbbk^2$ such that $f(x_0, y_0) = 0$. The set of all such roots is called the **solution set** of f.

Example 5.2. For $\Bbbk = \mathbb{R}$ we can take $f(x,y) = y - x^2$. The set of roots of f can be depicted as the parabola in the $X - Y$ plane $y = x^2$:



Example 5.2 indicates a general phenomenon. When our base field $\Bbbk$ has some geometry (as is the case with $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$), the solution set of a polynomial in two variables has a geometry as well – that of a one dimensional surface, aka a curve. Thus we can expect that the solution set of an n-dimensional polynomial will have a geometry of an $(n-1)$-dimensional surface. It turns out we

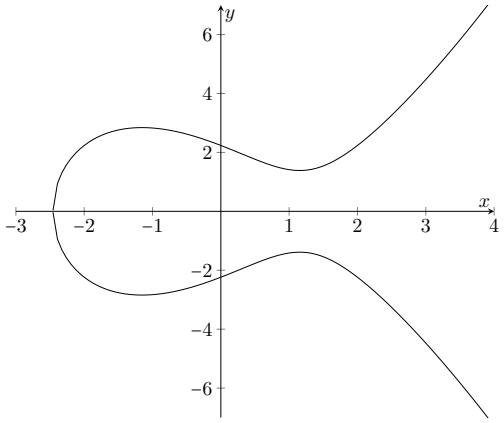


Figure 1: An elliptic curve defined over $\mathbb{R}$ by the equation $y^2 = x^3 - 4x + 5$

can use the geometry over $\mathbb{R}$ and $\mathbb{C}$ to get intuition on solution sets of polynomials over finite fields! In modern Mathematics, this approach is usually referred to as **Algebraic Geometry**. We will be restricting attention to solution sets of polynomials in two variables. More specifically

Definition 5.3. Let $\Bbbk$ be a field of characteristic $\neq 2, 3$. An **Elliptic curve** E defined over $\Bbbk$ (denoted $E/\Bbbk$) is a polynomial of the form

$$E : y^2 - x^3 - Ax - B \tag{6}$$

where $A, B \in \Bbbk$ satisfy

$$\Delta(E) := 4A^3 + 27B^2 \neq 0.$$

Notation 5.4. We will sometime write $E_{A,B}$ to denote an Elliptic curve of the form $E/\Bbbk : y^2 = x^3 + Ax + B$.

Remark 5.5. Henceforth, unless explicitly mentioned otherwise, we will assume that our base field $\Bbbk$ has $\operatorname{char} \Bbbk \neq 2, 3$.

Remark 5.6. The form appearing in 6 is called the **short Weierstrass form**. When $\operatorname{char} \Bbbk = 2$ or $\operatorname{char} \Bbbk = 3$ the definition of an Elliptic Curve over $\Bbbk$ is expressed in the **long Weierstrass form**:

$$E : y^2 + a_1xy + a_2y = x^3 + a_3x + a_4$$

such that $\Delta(E) \neq 0$. Note that here $\Delta(E)$ is the discriminant of E which has a slightly different formula than in the case of the short Weirstrass form.

The condition

$$4A^3 + 27B^2 \neq 0$$

of Definition 5.3 might seem odd but in fact is equivalent to a geometric property:

Proposition 5.7. *Let $E : y^2 = x^3 + Ax + B$ be a polynomial equation defined over $\mathbb{R}$. Then E has a well-defined and unique tangent line at every point iff E is an Elliptic curve, ie $4A^3 - 27B^2 \neq 0$.*

Proof. The solution set of E can be depicted as a curve in $\mathbb{R}$ and this curve in turn can be expressed as the union of graphs of two functions $y = \pm\sqrt{x^3 + Ax + B}$. When $y \neq 0$, a tangent line with respect to E must have a slope given by the condition $\frac{dy}{dx} = 0$ ie

$$\pm\frac{3x^2 + A}{2\sqrt{x^3 + Ax + B}} = 0.$$

Thus, we have a well-defined tangent line when $y \neq 0$ and $x^3 + Ax + B > 0$. If $x^3 + Ax + B < 0$ then there is no value y such that (x, y) is in the solution set; this leaves us with the case $x^3 + Ax + B = 0$. For this case we take a different method, namely that of implicit differentiation $\frac{d}{dx}$ of both sides:

$$\frac{d(y^2)}{dx} = 2y\frac{dy}{dx} = 3x^2 + A \iff \frac{dy}{dx} = \frac{3x^2 + A}{2y}.$$

Suppose $3x^2 + A \neq 0$ but $y = 0$. Then we can reflect our curve along the line $y = x \subseteq \mathbb{R}^2$ which means, algebraically, that we interchange $y \leftrightarrow x$, ie the nominator being zero and the denominator being non-zero – resulting in a tangent line of slope 0. Thus, flipping x, y again we get a tangent line parallel to the y-axis. Lastly, suppose $y = 0$ and $3x^2 + A = 0$. Then necessarily $A < 0$ and we can substitute $A = -A$ to rewrite the equation as $y^2 = x^3 - Ax + B$ with $A > 0$. In that case, $3x^2 - A = 0$ means that $x = \pm\sqrt{\frac{A}{3}}$. Since $x^3 - Ax + B = 0$ we get

$$(\sqrt{A/3})^3 - A\sqrt{A/3} + B = 0$$

$$\iff \frac{2A^{\frac{3}{2}}}{3\sqrt{3}} = B$$

$$\iff \frac{4A^3}{27} = B^2 \iff 4A^3 - 27B^2 = 0.$$

Substituting back $A = -A$ in the last equation shows us that we have a well-defined tangent line in the remaining case iff $4A^3 + 27B^2 \neq 0$ and this finishes the proof. □

In other words, Proposition 5.7 says that for a curve $E : y^2 = x^3 + Ax + B$ over $\mathbb{R}$ the condition $\Delta(E) = 4A^3 + 27B^2 \neq 0$ amount to a certain **smoothness**. As alluded earlier, in Algebraic Geometry one can extend the notion of smoothness to curves (or varieties) over arbitrary fields and in particular finite fields. We will not flesh out the definition of smoothness over finite fields here but merely point out that it remains equivalent to the condition $\Delta(E) = 4A^3 + 27B^2 \neq 0$ for all fields $\Bbbk$ with $\operatorname{char}\Bbbk \neq 2, 3$.

Let us examine the condition $\Delta(E) = 4A^3 + 27B^2 \neq 0$ further. Given a polynomial $f(x)$ of degree n over a field $\Bbbk$ we can ask when does $f(x)$ admit roots of multiplicity higher than 1 in the algebraic closure $\overline{\Bbbk}$.

Definition 5.8. Let $f(x) \in \Bbbk[x]$ be a polynomial of the form $f(x) = a_n x^n + ... + a_1 x + a_0$ and let $\alpha_1, ..., \alpha_n$ be its (not necessarily distinct) roots in the algebraic closure $\overline{\Bbbk}$. The **discriminant** of f is defined to be

$$\Delta(f) = a_n^{2n-2} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

What is relevant to us is the observation that f has roots of multiple degree iff $\Delta(f) = 0$. Since $f(x) = a_n \prod_{i=1}^{n} (x - \alpha_i)$ it follows that we can write the discriminant of f in terms of the coefficients $a_0, ..., a_n$.

Example 5.9.

1. A quadratic $f(x) = ax^2 + bx + c$ has the well-known discriminant

$$\Delta(f) = b^2 - 4ac$$

2. A cubic $f(x) = ax^3 + bx^2 + cx + d$ has discriminant

$$\Delta(f) = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd$$

3. A quartic $f(x) = ax^4 + bx^3 + cx^2 + dx + e$ has discriminant

$$\begin{aligned} \Delta(f) =& 256a^3e^3 - 192a^2bde^2 - 128a^2c^2e^2 + 144a^2cd^2e \\ & - 27a^2d^4 + 144ab^2ce^2 - 6ab^2d^2e - 80abc^2de \\ & + 18abcd^3 + 16ac^4e - 4ac^3d^2 - 27b^4e^2 + 18b^3cde \\ & - 4b^3d^3 - 4b^2c^3e + b^2c^2d^2 . \end{aligned} \tag{7}$$

Remark 5.10. For a polynomial $f(x) = a_n x^n + ... + a_1 x + a_0$, there is a formula for $\Delta(f)$ given by a determinant of a $2n \times 2n$ matrix built from the coefficients $a_0, ..., a_n$.

It follows from Example 5.9 that in the case of a cubic of the form $f(x) = x^3 + Ax + B$ we have $\Delta(f) = -(4A^3 + 27B^2)$.

Thus, that the condition $\Delta(E) = 4A^3 + 27B^2 \neq 0$ is equivalent to the condition that the polynomial $x^3 + Ax + B$ has no roots of multiplicity > 1.

5.1 The group law on an Elliptic curve

Let $E : y^2 = x^3 + Ax + B$ be an Elliptic curve over a field $\Bbbk$.

Definition 5.11. The **$\Bbbk$-rational** points of E are the set

$$E(\Bbbk) = \{(x, y) \in \Bbbk^2 | y^2 = x^3 + Ax + B\} \cup \{\mathcal{O}\}$$

where $\mathcal{O} = \infty$ is considered as a 'point at infinity' of E. If $\Bbbk \subseteq \mathbb{L}$ is any field inclusion, then the $\mathbb{L}$-rational points are simply

$$E(\mathbb{L}) = \{(x, y) \in \mathbb{L}^2 | y^2 = x^3 + Ax + B\} \cup \{\mathcal{O}\}.$$

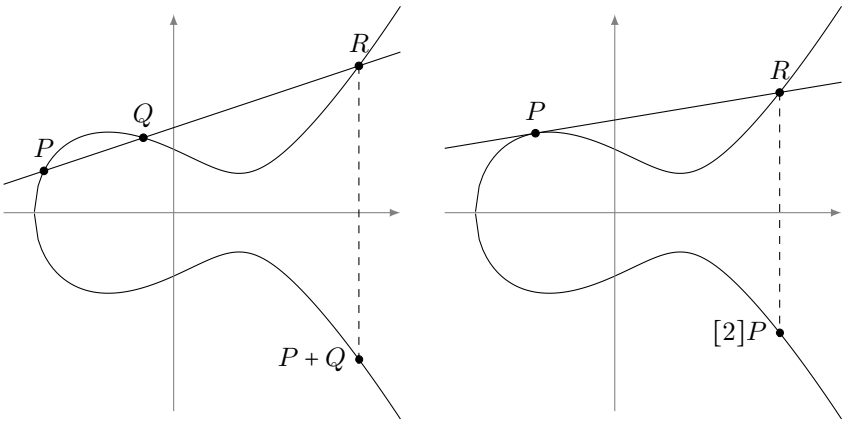


Figure 2: An elliptic curve defined over $\mathbb{R}$, and the geometric representation of its group law.

The reason for including $\mathcal{O}$ will become clear later, but for now it is useful to regard it as a point sitting simultaneously at the top and bottom of the y-axis so that lines parallel to the y-axis pass through $\mathcal{O}$.

Let $E : y^2 = x^3 + Ax + B$ be an Elliptic curve defined over $\mathbb{R}$. Start with two points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ on E as depicted in Figure 5.1. Draw the line L that intersects P and Q. We will see below (since E is a cubic) that L intersects E in a third point $R = (x_3, y_3)$. Since the graph of E is symmetric around the x-axis, the point $R' = (x_3, -y_3)$ must also lie on the curve E and we define $P + Q := R'$.

Assume first that $P \neq Q$ and that none of them is the point at infinity $\mathcal{O}$. Then the slope of the line L is

$$m = \frac{y_2 - y_1}{x_2 - x_1}.$$

If $x1 = x_2$ then L is vertical and we'll treat this case later, so suppose $x_1 \neq x_2$. Then L is given by $L : y = m(x - x_1) + y_1$. To find the intersection with E, we substitute that to get

$$(m(x - x_1) + y_1)^2 = x^3 + Ax + B$$

which can be re-arranged in the form

$$0 = x^3 - m^2x^2 + ...$$

the roots of the last cubic are the x-coordinates of the intersection points of L with E and we know that x_1, x_2 are two such roots. If we decompose this cubic to linear factors over $\overline{\mathbb{R}} = \mathbb{C}$ (Theorem 4.62) as $x^3 - m^2x^2 + ... = (x - x_1)(x - x_2)(x - x_3)$ then since $x_1, x_2 \in \mathbb{R}$, x_3 must also be a real number. Note that for any monic cubic with three roots,

$$x^3 + ax^2 + bx + c = (x - r)(x - s)(x - t) = x^3 - (r + s + t)x^2 + ...$$

so in our case $x_3 = m^2 - x_1 - x_2$ and thus $y_3 = m(x_3 - x_1) + y_1$. Now reflect along the x-axis to get

$$P + Q = R' = (m^2 - x_1 - x_2, m(x_1 - x_3) - y_1).$$

In the case $x_1 = x_2$ but $y_1 \neq y_2$, the line L is vertical hence intersects E at $\mathcal{O}$ and reflecting $\mathcal{O}$ along the x-axis yields $\mathcal{O}$ again so we define $P + Q = \mathcal{O}$.

Lastly, if $P = Q$, the line that passes through P, Q can be thought of the limit line when P, Q get closer to each other, i.e. the tangent line to E at $P = Q$ (recall that we showed such a line always exist on Elliptic curve). To find the slope of the tangent line, we use implicit differentiation

$$\frac{d}{dx}(y^2) = 2y\frac{dy}{dx} = 3x^2 + A \Rightarrow m = \frac{dy}{dx} = \frac{3x_1^2 + A}{2y_1}.$$

If $y_1 = 0$ the line is vertical and we set $P + Q = \mathcal{O}$ as before. Therefore, assume $y_1 \neq 0$. The equation for L is $L : y = m(x - x_1) + y_1$ and as before we obtain a cubic equation $0 = x^3 - m^2x^2 + ...$ this time, we know only one root, but it's a double root so we proceed as before to get $x_3 = m^2 - 2x_1$ and

$$P + P = (m^2 - 2x_1, m(x_1 - x_3) - y_1). \tag{8}$$

Finally, suppose $Q = \mathcal{O}$. The line through P and $\mathcal{O}$ is the vertical line that intersects E at P. The third intersection point of this line with E is the reflection of P along the x-axis and when we reflect this point along the x-axis we get P back. Thus, we define $P + \mathcal{O} = P$ for all points on E, so in particular $\mathcal{O} + \mathcal{O} := \mathcal{O}$.

Observation 5.12. The formulas for addition of points on E described above make sense for an Elliptic curve $E/\Bbbk$ defined over any field $\Bbbk$. The only amendment we need to do is to replace $\overline{\mathbb{R}} = \mathbb{C}$ with $\overline{\Bbbk}$ in order to decompose a polynomial into linear factors (by Theorem 4.59). We thus extend our definition of addition of points to this more general case.

Theorem 5.13. *Let $E : y^2 = x^3 + Ax + B$ be an Elliptic curve defined over a field $\Bbbk$ with $\operatorname{char}\Bbbk \neq 2, 3$. Then the $\Bbbk$-rational points*

$$E(\Bbbk) = \{(x, y) \in \Bbbk \times \Bbbk | y^2 = x^3 + Ax + B\} \cup \{\mathcal{O}\},$$

together with addition of points defined above, form an abelian group. The neutral element is $\mathcal{O}$ and the inverse of a point $P = (x_1, y_1)$ is given by $-P = (x_1, -y_1)$.

Proof sketch. It is easy to see from the definition of addition of points that $P + (-P) = \mathcal{O}$. However, proving that addition of points on E is associative is a tedious chase of equations and we will omit it from the current notes. A full proof may be found in [Wash]. See also Terrance Tau's blog for an intuitive explanation. □

Remark 5.14. The definition of addition of points on an Elliptic curve over a general field illustrates a typical reasoning in Algebraic Geometry: one first makes construction over $\mathbb{R}$ using geometric insights, and then extend it to arbitrary fields by analogy.

We finish this section with the following

Observation 5.15. Let $E/\Bbbk = E_{A,B}$ be an Elliptic curve defined over a field $\Bbbk$. If $\Bbbk \subseteq \Bbbk'$ is a subfield inclusion, then $E(\Bbbk) \subseteq E(\Bbbk')$ is a subgroup inclusion.

Proof. We have

$$\begin{aligned} E(\Bbbk) &= \{(x,y) \in \Bbbk \times \Bbbk \,|y^2 = x^3 + Ax + B\} \cup \{\mathcal{O}\} \\ &\subseteq \{(x,y) \in \Bbbk' \times \Bbbk' \,|y^2 = x^3 + Ax + B\} \cup \{\mathcal{O}\} = E(\Bbbk') \end{aligned} \tag{9}$$

since $\Bbbk \subseteq \Bbbk'$. The group operation in both sides agrees since it is defined in terms of the field operations and $\Bbbk \subseteq \Bbbk'$ is a subfield inclusion. □

5.2 Projective coordinates

We said that the $\Bbbk$-rational points of an Elliptic curve $E(\Bbbk)$ include a 'point at infinity' $\mathcal{O}$. In this section we will formalize that matter.

Let $\Bbbk$ be a field and consider the set $\Bbbk\times\Bbbk\times\Bbbk \setminus \{(0,0,0)\}$ of triples of elements in $\Bbbk$ with the origin removed. Define an equivalence relation on this set by setting for each $(x,y,z) \in \Bbbk^3$ and nonzero scalar $\lambda \in \Bbbk^\times$:

$$(x,y,z) \sim (\lambda x, \lambda y, \lambda z).$$

To see why this is an equivalence relation, recall Definition 2.23. Note that for reflexivity we can take $\lambda = 1$, and get $(x,y,z) \sim (x,y,z)$.

For transitivity, if

$$(x_1,y_1,z_1) \sim (x_2,y_2,z_2) \text{ and } (x_2,y_2,z_2) \sim (x_3,y_3,z_3)$$

then

$$(x_2,y_2,z_2) = \lambda(x_1,y_1,z_1) \text{ and } (x_3,y_3,z_3) = \lambda'(x_2,y_2,z_2)$$

so that

$$(x_3,y_3,z_3) = \lambda'\lambda(x_1,y_1,z_1)$$

and we get $(x_1,y_1,z_1) \sim (x_3,y_3,z_3)$ hence $\sim$ is transitive (symmetry is left for the reader). Informally speaking, an equivalence class of $\sim$ can be viewed as a line in $\Bbbk^3$ that passes through the origin.

Definition 5.16. The **two-dimensional projective plane** over $\Bbbk$ is the quotient set

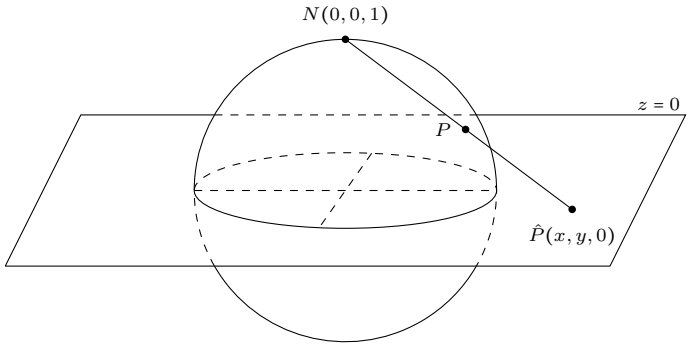
$$\mathbb{P}^2_\Bbbk = \Bbbk^3 \setminus \{(0,0,0)\}/\sim$$

by the equivalence relation described above.

We denote the equivalence class of a point (x,y,z) by $(x:y:z)$.

As an intuition for the definition of $\mathbb{P}^2_{\Bbbk}$ consider $\Bbbk = \mathbb{R}$. A point $(x : y : z) \in \mathbb{P}^2_{\mathbb{R}}$ corresponds to the collection $\{(\lambda x, \lambda y, \lambda z) | \lambda \in \mathbb{R}\}$ which can be considered as a line through the origin in $\mathbb{R}^3$. The lines on the x-y plane are the "points at infinity" which correspond to directions. A representative of this line may be taken to be a unit vector on this line i.e. a point on the 2-dimensional sphere S^2. Antipodal points on the sphere are identified with the same point on the projective plane.

Thus, we can identify $\mathbb{P}^2_{\mathbb{R}} \simeq S^2/\sim$ where $\sim$ identifies antipodal points. Under this identification, we can view $\mathbb{R}^2$ as points in $S^2/\sim$ via the **stereographic projection** shown in picture below. Here each point $\hat{P}$ on the plane is identified with two points on the sphere $P, -P$ by drawing a line that passes through $\hat{P}$ and either the north or south pole, and taking the intersection points $P, -P$ of this line with the sphere. One can find explicit formulas for this projection in terms of "spherical coordinates" ie in terms of Sine and Cosine.



A polynomial in three variables $F(x, y, z)$ over $\Bbbk$ is a sum of terms $a_{ijk}x^i y^j z^k$, called **monomials**, where $a_{ijk} \in \Bbbk$. Such polynomial is called **homogeneuos** of degree n if its monomials are all of the form $a_{ijk}x^i y^j z^k$ with $i + j + k = n$. If $F(x, y, z)$ is a homogeneous polynomial of degree n, then for any $\lambda \in \Bbbk^\times$, $F(\lambda x, \lambda y, \lambda z) = \lambda^n F(x, y, z)$. Thus, for such polynomials, $F(x, y, z) = 0$ if and only if for any $\lambda \in \Bbbk^\times$, $F(\lambda x, \lambda y, \lambda z) = 0$. It follows that if $F(x, y, z)$ is homogeneous of some degree and $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ then $F(x_1, y_1, z_1) = 0$ if and only if $F(x_2, y_2, z_2) = 0$.

Therefore, a zero of such polynomial F in $\mathbb{P}^2_{\Bbbk}$ is well-defined as it does not depend on the representative of the equivalence class.

Remark 5.17. The polynomial $F(x, y, z) = x^2 + 2y - 3z$ is not homogeneous and thus the considerations described above fail. For example, $F(1, 1, 1) = 0$ so we might be tempted to say that F has a zero at $(1 : 1 : 1)$, but $F(2, 2, 2) = 2 \neq 0$ whereas $(1 : 1 : 1) = (2 : 2 : 2)$.

Definition 5.18. If $f(x, y)$ is any polynomial in two variables of degree n, then

$$F(x, y, z) = z^n f(\frac{x}{z}, \frac{y}{z})$$

is called the **homogenization** of f.

Note the homogenization $F(x,y,z)$ of a polynomial $f(x,y)$ of degree n is a homogeneous polynomial of degree n and $F(x,y,1) = f(x,y)$. For example, if

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

then

$$F(x,y,z) = y^2z - x^3 - Axz^2 - Bz^3$$

is a homogeneous polynomial of degree 3.

Example 5.19. We can now understand what it formally means that two parallel lines intersect at infinity. Let

$$\ell_1 : y = mx + b_1$$

$$\ell_2 : y = mx + b_2$$

be two non-vertical parallel lines with $b_1 \neq b_2$. Their homogenization is

$$y = mx + b_1 z$$

$$y = mx + b_2 z$$

and to check their intersection in $\mathbb{P}^2_{\Bbbk}$ we look for zeros of the homogeneous polynomial

$$y - mx - b_1 z - (y - mx - b_2 z) = (b_2 - b_1)z$$

and see that $z = 0$. This means that $y = mx$ and since we cannot have $x = y = 0$ we get that a representative of the intersection of ℓ_1 and ℓ_2 in $\mathbb{P}^2_{\Bbbk}$ is $(x : mx : 0) = (1 : m : 0)$ which is a point at infinity.

Similarly, if $\ell_1 : x = b_1$ and $\ell_2 : x = b_1$ are two distinct parallel vertical lines, their intersection in $\mathbb{P}^2_{\Bbbk}$ is $(0 : 1 : 0)$ which is also one of the points at infinity.

We now look at an Elliptic curve $E : y^2 = x^3 + Ax + B$ defined over a field $\Bbbk$. Its homogeneous form is $y^2z = x^3 + Axz^2 + Bz^3$. The points (x,y) on the original curve in $\Bbbk \times \Bbbk$ correspond to points $(x : y : 1)$ in $\mathbb{P}^2_{\Bbbk}$. To see what points on the projective version of E lie at infinity, set $z = 0$ and obtain $x^3 = 0$ ie. $x = 0$. Thus, the coordinate y can be any non-zero element of $\Bbbk$. In other words, the point $(0 : y : 0) = (0 : 1 : 0)$ is the only point at infinity lying on the projective version of E.

5.3 Rational functions on an Elliptic curve

Throughout this section, we let $\Bbbk$ be a field with $\operatorname{char} \Bbbk \neq 2, 3$. Recall our notation $E_{A,B}$ for an Elliptic curve $E : y^2 = x^3 + Ax + B$ defined over $\Bbbk$.

Recall form section 4.4 that for a polynomial $g(x) \in \Bbbk[x]$ we can define mod-g arithmetic. A standard notation for mod-g arithmetic is $\Bbbk[x]/(g = 0)$. Similarly, if $g(x,y) \in \Bbbk[x,y]$ we can the analog of mod-g arithmetic is the set $\Bbbk[x,y]/(g = 0) = \Bbbk[x,y]/\sim$ where $a(x,y) \sim b(x,y)$ iff $a(x,y) - b(x,y) = f(x,y)g(x,y)$ for some $f(x,y) \in \Bbbk[x,y]$.

Definition 5.20. For an Elliptic curve $E/\Bbbk : y^2 = x^3 + Ax + B$ we define the set of **polynomials** over E to be

$$\Bbbk[E] := \Bbbk[x, y]/(y^2 - x^3 - Ax - B = 0).$$

Remark 5.21. It can be shown that the defining polynomial of an Elliptic curve $g(x, y) = y^2 - x^3 - Ax - B$ is irreducible in $\Bbbk[x, y]$ ie cannot be written as $g = a \cdot b$ for non-constant polynomials $a, b \in \Bbbk[x, y]$. However, unlike in $\Bbbk[x]$, $\Bbbk[x, y]$ does not admit a well-defined gcd. For example, if $f(x, y) = x$ and $g(x, y) = y$ we would want to say that $\gcd(f, g) = 1$, implying (by Bezout identity) that there are polynomials $a(x, y), b(x, y)$ such that $xa(x, y)+yb(x, y) = 1$ (as polynomials). However, if we substitute $(x, y) = (0, 0)$ we get $0 = 1$, a contradiction. Thus, $\Bbbk[E]$ is not a field.

By definition, we can replace every term y^2 in a polynomial $f \in \Bbbk[E]$ with $x^3 + Ax + B$ without changing the equivalence class of f. Thus, f can be written in a **canonical form** as $f(x, y) = v(x) + yw(x)$ for some $v, w \in \Bbbk[x]$.

Exercise 5.22. Show that the canonical form is unique.

Definition 5.23. Let $f \in \Bbbk[E]$ be given in canonical form $f(x, y) = v(x) + yw(x)$. The **conjugate** of f is $\overline{f} = v(x) - yw(x)$ and the **norm** of f is

$$N_f = f \cdot \overline{f} = v(x)^2 - y^2 w(x)^2 = v(x)^2 - (x^3 + Ax + B)w(x)^2 \in \Bbbk[x] \subseteq \Bbbk[E].$$

Exercise 5.24. Show that $N_{fg} = N_f N_g$ for any $f, g \in \Bbbk[E]$.

As pointed out in Remark 5.21, $\Bbbk[E]$ is not a field. In order to make it such, we have the following

Definition 5.25. For an Elliptic curve $E/\Bbbk$ the set of **rational functions** on E is the quotient set

$$\Bbbk(E) := \Bbbk[E] \times \Bbbk[E]/\sim$$

where $(f, g) \sim (h, k) \iff f \cdot k = h \cdot g \in \Bbbk[E]$. To check if equality holds, we can write $f \cdot k$ and $h \cdot g$ in canonical forms and compare coefficients. We denote the equivalence class of (f, g) by $\frac{f}{g}$. For $r \in \Bbbk(E)$ and a finite point $P \in E(\Bbbk)$ we say that r is **finite** at P if there exists a representation $r = \frac{f}{g}$ with $f, g \in \Bbbk[E]$ such that $g(P) \neq 0$. In this case, we define $r(P) = \frac{f(P)}{g(P)}$. Otherwise, we write $r(P) = \infty$.

Remark 5.26. For $r = \frac{f}{g} \in \Bbbk(E)$ we can write

$$\frac{f}{g} = \frac{f\overline{g}}{g\overline{g}} = \frac{f\overline{g}}{N_g}$$

and write $f\overline{g}$ in canonical form $(f\overline{g})(x, y) = v(x) + yw(x)$. We get

$$r(x, y) = \frac{f(x, y)}{g(x, y)} = \frac{(f\overline{g})(x, y)}{N_g(x)} = \frac{v(x)}{N_g(x)} + y\frac{w(x)}{N_g(x)}$$

which we will refer to as the **canonical form** for r.

Our next matter is defining the value of a rational function r at the point at infinity ie to give meaning to the expression $r(\mathcal{O})$. In the situation of a rational function in one variable, ie an expression of the form $r(x) = \frac{f(x)}{g(x)}$ with $f(x), g(x) \in \Bbbk[x]$ we typically (as in calculus) compare the degrees of f and g in order to get a meaningful value $r(\infty)$. For example, if $r(x) = \frac{x}{x^2+1}$ we would say that $r(\infty) = 0$ whereas if $r(x) = \frac{x^2}{x+1}$ we would say that $r(\infty) = \infty$. The situation in $\Bbbk[E]$ is more subtle since we have $y^2 = x^3 + Ax + B$ which means that the degree of y should be $\frac{2}{3}$ of the degree of x. Since we want to keep degrees as integers, we set $\deg(y) = 3$ and $\deg(x) = 2$ in $\Bbbk[E]$. The classical degree of a polynomial $f \in \Bbbk[x]$ will be denoted $\deg_x(f)$. This motivates the following

Definition 5.27. Let $f \in \Bbbk[E]$ and write it in canonical form $f(x, y) = v(x) + yw(x)$. The **degree** of f is

$$\deg(f) := \max\{2 \cdot \deg_x(v), 3 + 2 \cdot \deg_x(w)\}.$$

Remark 5.28. Recall that $\deg_x(0) = -\infty$ and $\deg_x(c) = 0, \ \ \forall c \in \Bbbk^\times$.

Proposition 5.29. *Let $E = E_{A,B}$ be an Elliptic curve defined over $\Bbbk$ and denote $s(x) = x^3 + Ax + B$. For $f, g \in \Bbbk[E]$:*

1. $\deg(f) = \deg_x(N_f)$.
2. $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Proof.

1. Write f in canonical form $f(x, y) = v(x) + yw(x)$, then $N_f = v(x)^2 - s(x)w(x)^2$. Since $\deg_x(v^2)$ and $\deg_x(w^2)$ are even and $\deg_x(s)$ is odd, it follows that

$$\begin{aligned} \deg_x(N_f) = \deg_x(v^2 - sw^2) &= \max\{\deg_x(v^2), \deg_x(s) + \deg_x(w^2)\} \\ &= \max\{2\deg_x(v), 3 + 2\deg_x(w)\} = \deg(f). \end{aligned} \tag{10}$$

2. We can easily calculate

$$\begin{aligned} \deg(fg) = \deg_x(N_{fg}) = \deg_x(N_f N_g) = \deg_x(N_f) + \deg_x(N_g) \\ = \deg(f) + \deg(g). \end{aligned} \tag{11}$$

□

It makes no sense to talk about the degree of the nominator of a rational function $r \in \Bbbk(E)$ since that depends on the representation $r = \frac{f}{g} = \frac{h}{k}$. However, if $r = \frac{f}{g} \in \Bbbk[E]$, the quantity $\deg(f) - \deg(g)$ does not depend on the representation since if $\frac{f}{g} = \frac{h}{k}$ we have $fk = hg$ and by Proposition 5.29, $\deg(f) - \deg(g) = \deg(h) - \deg(k)$.

Definition 5.30. Let $r = \frac{f}{g} \in \Bbbk(E)$ be a rational function and distinguish the following cases:

1. If $\deg(f) < \deg(g)$: set $r(\mathcal{O}) = 0$.
2. If $\deg(f) > \deg(g)$: say that r is not finite at $\mathcal{O}$.
3. If $\deg(f) = \deg(g)$ and $\deg(f)$ is **even**: write both f and g in canonical form, so that they both have leading terms ax^d and bx^d (respectively) with $a, b \in \Bbbk^\times$ and $d = \frac{\deg(f)}{2}$, and we set $r(\mathcal{O}) = \frac{a}{b}$.
4. If $\deg(f) = \deg(g)$ and $\deg(f)$ is **odd**: write both f and g in canonical form, so that they both have leading terms $ax^d y$ and $bx^d y$ (respectively), $a, b \in \Bbbk^\times$ and $\deg(f) = \deg(g) = 3 + 2d$, and we set $r(\mathcal{O}) = \frac{a}{b}$.

Remark 5.31. For $r = \frac{f}{g} \in \Bbbk(E)$, it may seem natural to define $\deg(r) = \deg(f) - \deg(g)$ so that the value $r(\mathcal{O})$ would depend on the sign of $\deg(r)$. However, this differs from the usual definition of a degree of a rational function in Algebraic Geometry so we avoid defining the degree of a rational function altogether.

Example 5.32. Consider $E = E_{A,B}$ and $\Bbbk(E)$. For

$$r(x, y) = \frac{x^3 + 2x + y + 2x^4 y}{x + x^2 + 5xy^3}$$

one can write

$$r(x, y) = \frac{x^3 + 2x + y + 2x^4 y}{x + x^2 + 5xy(x^3 + Ax + B)} = \frac{(x^3 + 2x) + y(1 + 2x^4)}{(x + x^2) + y(5x^4 + 5Ax^2 + 5Bx)}.$$

The last representative has nominator of degree $\max\{2 \cdot 3, 3 + 2 \cdot 4\} = 11$, and denominator of degree $\max\{2 \cdot 3, 3 + 2 \cdot 4\} = 11$ which are both odd. Thus $r(\mathcal{O}) = \frac{2}{5}$.

Exercise 5.33. For $r, s \in \Bbbk(E)$ with $r(\mathcal{O}), s(\mathcal{O})$ finite, we have $(rs)(\mathcal{O}) = r(\mathcal{O})s(\mathcal{O})$ and $(r + s)(\mathcal{O}) = r(\mathcal{O}) + s(\mathcal{O})$.

5.3.1 Zeros and poles

Definition 5.34. Let $E/\Bbbk$ be an Elliptic curve and let $r \in \Bbbk(E)$ be a rational function. We say that r has a **zero** in $P \in E$ if $r(P) = 0$ and that r has a **pole** in P if $r(P)$ is not finite.

The goal of this section is to define the **multiplicity** of zeros and poles. The motivation comes from functions in one variable. Consider $E = E_{1,0} : y^2 = x^3 + x$ and $P = (0, 0) \in E$. Then P is a zero of the functions x and y. However, between these two functions there is a relation: $x = y^2 - x^3$. In the analytic sense, when $x \longrightarrow 0$, the term x^3 can be neglected so we would like to say that the function x has a zero at P whose multiplicity is twice that of the function y at P. This is formalised in the following

Definition 5.35. Let $E/\Bbbk$ be an Elliptic curve and $P \in E$. A rational function $u \in \Bbbk(E)$ with $u(P) = 0$ is called a **uniformizer** at P if:

$\forall r \in \Bbbk(E) \setminus \{0\}$, $\exists d \in \mathbb{Z}, s \in \Bbbk(E)$ finite at P with $s(P) \neq 0$ such that

$$r = u^d \cdot s.$$

Remark 5.36. As we will see soon, uniformizers exist for all points in $E(\Bbbk)$ (though different points may require different uniformizers). However, even for a fixed point P there is usually more than one uniformizer at P. The uniformizers we present below are simply one common choice.

Proposition 5.37. *Let $E/\Bbbk$ be an Elliptic curve and $P = (a, b) \in E$ a finite point with $2P \neq \mathcal{O}$. Then the function $u(x, y) = x - a$ is a uniformizer at P.*

Proof. First note that $u(P) = 0$. Now let $r \in \Bbbk(E) \setminus \{0\}$ be arbitrary. If r has neither zero nor pole at P, we can take $d = 0$ and $r = s$.

Suppose r has a zero at P, ie $r(P) = 0$. Note that if we have proved that u is a uniformizer in the case P is a zero, then for $r \in \Bbbk(E)$ with a pole at P, $\frac{1}{r}$ has a zero at P so that there exists $d \in \mathbb{Z}$ and $s \in \Bbbk(E)$ which is finite and non-zero at P such that $\frac{1}{r} = u^d s$. But then $r = u^{-d}\frac{1}{s}$ shows that u is a uniformizer for $\frac{1}{r}$.

Thus, we assume $r(P) = 0$ so we can write $r\frac{f}{g}$ with $f(P) = 0$ and $g(P) \neq 0$. If we can decompose $f = u^d s$ as above then

$$r = \frac{f}{g} = \frac{u^d s}{g} = u^d \frac{s}{g}$$

with $\frac{s}{g}(P) \neq 0$ and finite so we are done.

Set $s_0(x, y) = f(x, y)$ and repeat the following process (starting from $i = 0$) while $s_i(P) = 0$:

Write $s_i(x, y) = v_i(x) + yw_i(x)$ in canonical form. Distinguish the cases $\overline{s}_i(P) = 0$ and $\overline{s}_i(P) \neq 0$ (recall that $\overline{s}_i = v_i(x) - yw_i(x)$ is the conjugate of s_i.)

$\underline{\overline{s}_i(P) = 0}$: Since $y(P) = b \neq 0$, the system of linear equations

$$\begin{aligned} v_i(a) + bw_i(a) &= 0 \\ v_i(a) - bw_i(a) &= 0 \end{aligned} \tag{12}$$

has a unique solution (e.g. its rank equals 2) of the form $v_i(a) = w_i(a) = 0$.

Thus, we can write

$$s_i(x, y) = v_i(x) + yw_i(x) = (x - a)v_{i+1}(x) + (x - a)yw_{i+1}(x) = (x - a)s_{i+1}(x) \tag{13}$$

for $s_{i+1}(x) := v_{i+1}(x) + yw_{i+1}(x)$ with some polynomials $v_{i+1}(x), w_{i+1}(x) \in \Bbbk[x]$.

$\underline{\overline{s}_i(P) \neq 0}$: multiply s_i by $1 = \frac{\overline{s}_i}{\overline{s}_i}$ to get

$$s_i(x, y) = \frac{N_{s_i}}{\overline{s}_i}.$$

Now, $s_i(P) = 0$ and $\overline{s}_i(P) \neq 0$ implies that $N_{s_i}(a) = 0$ so we can write

$$N_{s_i}(x) = (x-a)n(x)$$

for some $n(x) \in \Bbbk[x]$.

We now set

$$s_{i+1}(x) = \frac{n(x)}{\overline{s}_i(x,y)}$$

(which is finite at P), and we again get

$$s_i(x,y) = \frac{N_{s_i(x)}}{\overline{s}_i(x)} = \frac{(x-a)n(x)}{\overline{s}_i(x,y)} = (x-a)s_{i+1}(x,y).$$

If the process terminates, we get $f(x,y) = (x-a)^i s_i(x,y)$ where $s := s_i$ is finite and non-zero with $u(x,y) = x-a$ and $d = i$ so we are done.

Since s_i is a rational function and not a polynomial, it is not clear that this process indeed terminates. Let us show it anyhow.

$$\begin{aligned} N_f(x) &= N_{u^i s_i}(x) \\ &= ((x-a)^i v_i(x))^2 - y^2((x-a)^i w_i(x))^2 \\ &= (x-a)^{2i}(v_i(x)^2 - y^2 w_i(x)^2) \\ &= (x-a)^{2i} N_{s_i} \end{aligned} \tag{14}$$

and we see that i is bounded since $\deg(N_f) = 2i + \deg(N_{s_i})$ and since $\deg(N_{s_i}) > 0$. Thus, there can only be finitely many iterations i and this finishes the proof. □

Lemma 5.38. *Let $E/\Bbbk = E_{A,B}$ be an Elliptic curve over a field $\Bbbk$ such that $E(\Bbbk)$ contains all point of order two (e.g. $\Bbbk$ algebraically closed). Let $P \in E$ such that $2P = \mathcal{O}$. Then the rational function $u_P(x,y) = u(x,y) = y$ is a uniformizer at P.*

Proof. Since E is an Elliptic curve, $s_E(x) = x^3 + Ax + B$ has three distinct roots $\alpha_1, \alpha_2, \alpha_3$ ie

$$s_E(x) = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3).$$

We saw that if P is of order 2 then (without loss of generality) $P = (\alpha_1, 0)$ so that $u(P) = 0$. Let $r = \frac{f}{g} \in \Bbbk(E) \setminus \{0\}$ be such that $r(P) = 0$. Then we can assume that the presentation $r = \frac{f}{g}$ is such that $f(P) = 0$ and $g(P) \neq 0$. Write f in canonical form $f(x,y) = v(x) + yw(x)$ which means $v(\alpha_1) = 0$ so $v(x) = (x-\alpha_1)v_1(x)$ for some $v_1(x)$ with $\deg v_1 < \deg v$. We can thus write

$$\begin{aligned} f(x,y) = (x-\alpha_1)v_1(x) + yw(x) &= \frac{(x-\alpha_1)(x-\alpha_2)(x-\alpha_3)v_1(x) + yw_1(x)}{(x-\alpha_2)(x-\alpha_3)} \\ &= \frac{y^2 v_1(x) + yw_1(x)}{(x-\alpha_2)(x-\alpha_3)} = y\frac{yv_1(x) + w_1(x)}{(x-\alpha_2)(x-\alpha_3)} = u(x,y)W(x,y) \end{aligned} \tag{15}$$

where $w_1(x) = w(x)(x-\alpha_2)(x-\alpha_3)$ and $W(x,y) = \frac{yv_1(x)+w_1(x)}{(x-\alpha_2)(x-\alpha_3)}$. Note that $W(P)$ is finite . If $W(P) \neq 0$ we are done since we can take

$$s(x,y) = W(x,y)/g(x,y)$$

and write $r(x,y) = u(x,y)^1 \cdot s(x,y)$. Otherwise, we repeat the process with W/g instead of r. More specifically, we take

$$r'(x,y) = \frac{W(x,y)}{g(x,y)} = \frac{yv_1(x)+w_1(x)}{g(x,y)(x-\alpha_2)(x-\alpha_3)}$$

and by assumption $r'(P) = 0$. This means that $w_1(\alpha_1) = 0$ so that

$$w_1(x) = (x-\alpha_2)(x-\alpha_3)w(x) = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3)w_2(x)$$

for some $w_2(x)$ such that $\deg w_2 < \deg w$. We can thus write

$$\begin{aligned} W(x,y) &= \frac{yv_1(x)+w_1(x)}{(x-\alpha_2)(x-\alpha_3)} \\ &= \frac{yv_1(x)+(x-\alpha_1)(x-\alpha_2)(x-\alpha_3)w_2(x)}{(x-\alpha_2)(x-\alpha_3)} \\ &= \frac{yv_1(x)+y^2w_2(x)}{(x-\alpha_2)(x-\alpha_3)} \\ &= y\frac{v_1(x)+yw_2(x)}{(x-\alpha_2)(x-\alpha_3)} =: yW_1(x,y) \end{aligned} \tag{16}$$

where

$$W_1(x,y) = \frac{v_1(x)+yw_2(x)}{(x-\alpha_2)(x-\alpha_3)}.$$

As before, $W_1(P)$ is finite and if $W_1(P) \neq 0$ we are done since we can take

$$s'(x,y) = W_1(x,y)/g(x,y)$$

and get $r' = u^1 s'$ with $s'(P) \neq 0$ so that $r = u^2 s'$. As can be seen form the above, repeating the process produces a sequence of polynomials $v_1, v_2, \ldots$ and $w_1, w_2, \ldots$ such that for n odd $\deg v_n < \deg v_{n-2}$ and for n even $\deg w_n < \deg w_{n-2}$. It follows that the process must terminate since v and w have only finitely many roots.

Lastly, if $r(P) = \infty$, i.e. r has a pole at P , then $\frac{1}{r}(P) = 0$ so that by the argument above, $\frac{1}{r} = u^d w$ with $w(P) \neq 0$ and thus $r = u^{-d}\frac{1}{w}$ as required. □

Lemma 5.39. *Let $E/\Bbbk$ be an Elliptic curve. Then the function $u(x,y) = \frac{x}{y}$ is a uniformizer at $\mathcal{O}$.*

Proof. Since $\deg(y) = 3 > \deg(x) = 2$, it follows from Definition 5.30 that $u(\mathcal{O}) = 0$. Let $r = \frac{f}{g} \in \Bbbk(E) \setminus \{0\}$ be such that $r(\mathcal{O}) = 0$ or $r(\mathcal{O})$ is not finite. This

means, by Definition 5.30 that $d = \deg(f) - \deg(g) \neq 0$. We would like to take $s(x,y) = \left(\frac{y}{x}\right)^d r(x,y)$ since then we have

$$u(x,y)^d s(x,y) = \left(\frac{x}{y}\right)^d \left(\frac{y}{x}\right)^d r(x,y) = r(x,y)$$

However, in order for this to work we need to show that $s(\mathcal{O})$ is finite and non-zero. We have

$$s(x,y) = \frac{y^d f(x,y)}{x^d g(x,y)}$$

and because

$$\begin{aligned} \deg(y^d f(x,y)) - \deg(x^d g(x,y)) &= \deg(y^d) + \deg(f) - (\deg(x^d) + \deg(g)) \\ &= 3d + \deg(f) - 2d - \deg(g) = 0 \end{aligned} \tag{17}$$

we get from Definition 5.30 that $s(P)$ is finite and non-zero.

□

Theorem 5.40. *Let $E/\Bbbk$ be an Elliptic curve. Then any point on $E(\Bbbk)$ has a uniformizer and the number d of Definition 5.35 does not depend on its choice.*

Proof. The previous claims ensure the existence of a uniformizer for every point $P \in E(\Bbbk)$. It is left to show that the integer d does not depend on the choice of uniformizer. Let $P \in E(\Bbbk)$ and let $u, u' \in \Bbbk(E)$ be uniformizers at P. Then we can write $u = u'^a p$ and $u' = u^b q$ for $p, q \in \Bbbk(E)$ such that $p(P), q(P) \neq 0, \infty$. We thus get

$$u = u'^a p = (u^b q)^a p = u^{ab} q^a p \iff 1 = u^{ab-1} q^a p.$$

If $ab \neq 1$ then evaluating at P gives $1 = 0 \cdot q^a(P) p(P) = 0$ which is a contradiction. Thus $ab = 1$ ie. $a = b = \pm 1$. If $a = b = -1$ we get

$$u = u'^{-1} p \iff uu' = p$$

which cannot be true since $p(P) \neq 0$ whereas $u(P) = u'(P) = 0$ so we must have, $a = b = 1$.

If $r \in \Bbbk(E) \setminus \{0\}$ then since u and u' are uniformizers at P we get $r = u^d s$ and $r = u^{d'} s'$ for some $d, d' \in \mathbb{Z}$ and $s, s' \in \Bbbk(E)$ such that $s(P), s'(P) \neq 0, \infty$. But then

$$u^d s = u'^{d'} s' = (uq)^{d'} s' = u^{d'} q^{d'} s'$$

which yields

$$u^{d-d'} = \frac{q^{d'} s'}{s}.$$

If $d \neq d'$ we get a contradiction since the LHS evaluated at P is zero while the RHS evaluated at P is non-zero. Thus, $d = d'$ as desired.

□

Definition 5.41. Let $E/\Bbbk$ be an Elliptic curve, $P \in E(\Bbbk)$ and $u \in \Bbbk(E)$ a uniformizer at P. For $r \in \Bbbk(E) \setminus \{0\}$ a rational function with $r = u^d \cdot s$ with $s(P) \neq 0, \infty$, we say that r has **order** d at P and write

$$\mathrm{ord}_P(r) = d.$$

The **multiplicity of a zero** of r is the order of r at that point and the **multiplicity of a pole** of r is the order of r at that point.

Observation 5.42. Let $E/\Bbbk$ be an Elliptic curve and $r \in \Bbbk(E)$ a rational function. If $P \in E(\Bbbk)$ a point which is neither a zero or a pole of r, then $\mathrm{ord}_P(r) = 0$.

Proof. Pick a uniformizer u and set $s(x,y) = r(x,y)$. Then $s(P)$ is finite and non-zero and $r = u^0 s$. □

Example 5.43. Let $E = E_{A,B}$ be an Elliptic curve and $P = (a,b) \in E(\Bbbk)$, finite and not of order 2. We want to calculate the orders of $r(x,y) = x - a$ at all points $Q \in E(\Bbbk)$ where $r(Q)$ is zero or not finite (otherwise, $\mathrm{ord}_Q(r) = 0$ by Observation 5.42). Note that $Q = P = (a,b)$ and $Q = -P = (a,-b)$ are both zeros of r. Since in this case r itself is a uniformizer (with $s(x,y) = 1$) we get that $\mathrm{ord}_Q(r) = 1$.

When $Q = \mathcal{O}$, we have a pole for r. We take as a uniformizer $u(x,y) = \frac{x}{y}$ and $s(x,y) = \frac{x^3 - ax^2}{y^2}$ (note that $s(Q) = 1$) and get

$$u(x,y)^{-2} s(x,y) = \frac{y^2}{x^2} \cdot \frac{x^3 - ax^2}{y^2} = x - a = r(x,y)$$

so that $\mathrm{ord}_{\mathcal{O}}(r) = -2$.

Example 5.44. Let $E/\Bbbk : y^2 = x^3 + Ax + B$ be an Elliptic curve and let $r(x,y) = y \in \Bbbk(E)$. From Example 5.76, we know that a point $P \in E(\Bbbk)$ is of order two iff $P = (\alpha, 0)$ where α is a root of $x^3 + Ax + B$. According to Lemma 5.38, the rational function $u(x,y) = y$ is a uniformizer for points of order two, and we thus get $r = u^1 \cdot 1$ deducing that at point P of order two, r has a zero of multiplicity 1. According to Lemma 5.39, $u(x,y) = \frac{x}{y}$ is a uniformizer at $\mathcal{O}$. The rational function $s(x,y) = \frac{x^3 y}{y^3}$ satisfy $s(\mathcal{O}) \neq 0, \infty$ by Definition 5.30 since $\deg(x^3 y) - \deg(y^3) = 6 + 3 - 3 \cdot 3 = 0$. But then we get

$$u(x,y)^{-3} s(x,y) = \left(\frac{x}{y}\right)^{-3} \frac{x^3 y}{y^3} = y = r(x,y)$$

so we see that $\mathrm{ord}_{\mathcal{O}}(r) = -3$.

5.4 Divisors

Examples 5.44 shows an interesting phenomenon: if $\Bbbk$ is **algebraically closed** then $x^3 + Ax + B$ has 3 distinct roots $\alpha_1, \alpha_2, \alpha_3$ and thus $E = E_{A,B}$ has 3 points of order 2:

$$\{P_1, P_2, P_3\} = \{(\alpha_1, 0), (\alpha_2, 0), (\alpha_3, 0)\} \subseteq E(\Bbbk).$$

The function $r(x, y) = y \in \Bbbk(E)$ has 3 zeros, namely P_1, P_2, P_3, each of order 1 and a pole at $\mathcal{O}$ of order 3. In other words if we sum the orders of all points (recall that order of a point which is not a zero or a pole is 0 by Observation 5.42), we get $\sum_{P \in E(\Bbbk)} \operatorname{ord}_P(r) = 1 + 1 + 1 - 3 = 0$. In fact, the same is true for Example 5.43 since in that case $\sum_{Q \in E(\Bbbk)} \operatorname{ord}_Q(r) = 1 + 1 - 2 = 0$. This motivates the following

Definition 5.45. Let $E/\Bbbk$ be an Elliptic curve. A **divisor** on E is an expression

$$D = \sum_{P \in E(\Bbbk)} n_P[P]$$

where $\forall P,\ n_P \in \mathbb{Z}$ and only finitely many n_P's are non-zero. The **degree** of a divisor D is

$$\deg(D) = \sum_{P \in E(\Bbbk)} n_P.$$

The **sum** of a divisor D is

$$\operatorname{sum}(D) = \sum_{P \in E(\Bbbk)} n_P \cdot P \in E(\Bbbk).$$

Observation 5.46. The set of all divisors on E forms a group: If $D = \sum_{P \in E(\Bbbk)} n_P[P]$ and $D' = \sum_{P \in E(\Bbbk)} m_P[P]$ then

$$D + D' \coloneqq \sum_{P \in E(\Bbbk)} (n_P + m_P)[P].$$

The unit element is the divisor $\sum_{P \in E(\Bbbk)} 0 \cdot [P]$ and the inverse of a divisor $D = \sum_{P \in E(\Bbbk)} n_P[P]$ is the divisor $-D = \sum_{P \in E(\Bbbk)} -n_P[P]$.

Definition 5.47. Let $E/\Bbbk$ be an Elliptic curve and $r \in \Bbbk(E) \setminus \{0\}$ a rational function. The **associated divisor** of r is defined to be

$$\operatorname{div}(r) \coloneqq \sum_{P \in E(\Bbbk)} \operatorname{ord}_P(r)[P].$$

Lemma 5.48. *Let $E/\Bbbk$ be an Elliptic curve and $r, s \in \Bbbk(E) \setminus \{\mathcal{O}\}$ rational functions.*

1. $\operatorname{div}(rs) = \operatorname{div}(r) + \operatorname{div}(s)$.
2. $\operatorname{div}(\frac{r}{s}) = \operatorname{div}(r) - \operatorname{div}(s)$.

Proof. Both claims follow from the claim that for any $P \in E(\Bbbk)$, $\mathrm{ord}_P(rs) = \mathrm{ord}_P(r) + \mathrm{ord}_P(s)$ which in turn follows from the fact that is u is a uniformizer at P with degrees d_r, d_s for r, s respectively, then the degree of u for rs is $d_r + d_s$. Analogously, $\mathrm{ord}_P(\frac{r}{s}) = \mathrm{ord}_P(r) - \mathrm{ord}_P(s)$ □

There is another astonishing fact emerging from Examples 5.43 and 5.44: for a rational functions $r \in \Bbbk(E)$, we have in the group $E(\Bbbk)$:

$$\sum_{P \in E(\Bbbk)} \mathrm{ord}_P(r) \cdot P = \mathcal{O}.$$

In Example 5.43 we have

$$\sum_{Q \in E(\Bbbk)} \mathrm{ord}_Q(r) \cdot Q = 1 \cdot P + 1 \cdot (-P) + (-2) \cdot \mathcal{O} = \mathcal{O}$$

where as in Example 5.44 we have

$$\sum_{P \in E(\Bbbk)} \mathrm{ord}_P(r) \cdot P = P_1 + P_2 + P_3 + (-3) \cdot \mathcal{O} = P_1 + P_2 + P_3 = \mathcal{O}$$

Note that the last equality comes from the fact that $P_1 + P_2 + P_3$ must have order at most 2 in $E(\Bbbk)$ since $2(P_1 + P_2 + P_3) = 2P_1 + 2P_2 + 2P_3 = \mathcal{O}$ but if, for example, $P_1 + P_2 + P_3 = P_1$ we get $P_2 + P_3 = \mathcal{O} \iff P_2 = -P_3$ and we know that this is not true (since $-P_3 = P_3$), hence $P_1 + P_2 + P_3 = \mathcal{O}$. This leads us to a key

Theorem 5.49. *Let $E/\Bbbk$ be an Elliptic curve over an algebraically closed field $\Bbbk$.*

1. *Let r and r' be rational functions on E. If $\mathrm{div}(r) = \mathrm{div}(r')$ there exists a non-zero constant $c \in \Bbbk$ such that $r = cr'$.*
2. *Let $D = \sum_{P \in E(\Bbbk)} n_P[P]$ be a divisor. Then there exists a rational function $r \in \Bbbk(E)$ such that $\mathrm{div}(r) = D$ if and only if:*
   - $\deg(D) = 0$.
   - $\mathrm{sum}(D) = 0$.

In particular, if a rational function $r \in \Bbbk(E)$ has no zeros and no poles, it is constant.

Example 5.50. Let $E = E_{A,B}$ be an Elliptic curve. Suppose $P \in E(\Bbbk)$ has order m, ie $mP = \mathcal{O}$. By Theorem 5.49 there exists a rational function f_P such that $\mathrm{div}(f_P) = m[P] - m[\mathcal{O}]$. The case $m = 2$ is particularly simple. We saw that points of order 2 are of the form $P = (\alpha, 0)$ where α is a root of $x^3 + Ax + B$. As we saw in Example 5.43, $\mathrm{div}(x - \alpha) = 2[P] - 2[\mathcal{O}]$.

For our last result in this section, let us define the **support** of a divisor $\sum_P n_P[P]$ to be the points $P \in E(\Bbbk)$ such that $n_P \neq 0$. We will need the following

Definition 5.51. Let $r \in \Bbbk(E) \setminus \{0\}$ be a rational function and $D = \sum_P n_P[P]$ be a divisor whose support does not include zeros or poles of r. Then the function r **evaluated** at D is

$$r(D) := \prod_{P \in E(\Bbbk)} r(P)^{n_P}.$$

5.5 From rational functions to rational maps

Let $E/\Bbbk$ be an Elliptic curve. We discussed the notion of rational functions on E and saw that they give rise to functions $E(\Bbbk) \longrightarrow \mathbb{P}^1_\Bbbk = \Bbbk \cup \{\infty\}$. Since $\mathbb{P}^1$ is a curve, one may wonder if it is possible to extend the notion of rational function to a map $E(\Bbbk) \longrightarrow E(\Bbbk)$ between the Elliptic curve to itself. This is indeed possible as we will see below.

Definition 5.52. Let

$$E/\Bbbk = E_{A,B} : y^2 = x^3 + Ax + B$$

be an Elliptic curve. A **rational map** $\rho : E \longrightarrow E$ is a pair $\rho = (r, s)$ where $r, s \in \Bbbk(E)$ are rational functions on E such that for all $P \in E(\Bbbk)$,

$$s(P)^2 = r(P)^3 + Ar(P) + B.$$

In particular, $r(P) = \infty$ if and only if $s(P) = \infty$.

A rational map $\rho = (r, s) : E \longrightarrow E$ induces a map $\rho : E(\Bbbk) \longrightarrow E(\Bbbk)$ given by $P \mapsto (r(P), s(P))$ if $r(P), s(P) \neq \infty$ and $P \mapsto \mathcal{O}$ if $r(P) = s(P) = \infty$.

Example 5.53. Let $E/\Bbbk$ be an Elliptic curve and let $1 \leq n$. The map $[n] : E(\Bbbk) \longrightarrow E(\Bbbk)$ given by $[n](P) := nP$ is a rational map. The construction of Section 5.1, gives rational functions $r, s \in \Bbbk(E)$ such that $[n] = (r, s)$.

Example 5.54. Let E be an Elliptic curve and $Q \in E(\Bbbk)$. The map $\tau_Q : E \longrightarrow E$ given by $\tau_Q(P) = P + Q$ is a rational map.

As usual, we have:

Proposition 5.55. *Let $E/\Bbbk$ be an Elliptic curve and $\rho, \tau : E \dashrightarrow E$ be two rational maps. Then $\tau \circ \rho : E \longrightarrow E$ is a rational map.*

Proof. Suppose $\rho = (r, s)$ and $\tau = (u, v)$. Then

$$(\tau \circ \rho)(x, y) = \Big(u\big(r(x, y), s(x, y)\big), v\big(r(x, y), s(x, y)\big)\Big)$$

which is a pair of rational functions since each coordinate is a substitution of rational functions into rational functions. Since for every $P \in E(\Bbbk)$ we have $\big(r(P), s(P)\big) \in E(\Bbbk)$ (and similarly for u, v), we have $(\tau \circ \rho)(P) \in E(\Bbbk)$, so that the pair of rational functions representing $\tau \circ \rho$ satisfies the equation of E, ie $\tau \circ \rho$ is again a rational map. □

Rational maps have a rather rigid structure, as the following proposition shows:

Proposition 5.56. *Let $E/\Bbbk$ be an Elliptic curve over field $\Bbbk$ and $\rho = (r, s) : E \longrightarrow E$ be a rational map. If ρ is non-constant, then it induces a surjective map $E(\overline{\Bbbk}) \longrightarrow E(\overline{\Bbbk})$.*

Proof. We first show the analogous assertion for rational functions $E \longrightarrow \mathbb{P}^1$. If $r \in \Bbbk(E)$ is a non-constant rational function, then it must have a zero by Theorem 5.49. For $x_0 \in \Bbbk$, the same argument as above, applied to $r - x_0$, implies that $r - x_0$ has a zero, so there is $P \in E(\Bbbk)$ such that $r(P) = x_0$. Thus, r is surjective.

Now consider the rational map $\rho = (r, s) : E(\Bbbk) \longrightarrow E(\Bbbk)$. If r is constant, then since $\forall P \in E(\Bbbk)$

$$s(P)^2 = r(P)^3 + Ar(P) + B,$$

we get that s can take at most two values, namely the roots of $r(P)^3 + Ar(P) + B$, hence must be constant by the argument above.

Otherwise, r is surjective. It has a zero hence a pole by Theorem 5.49. In particular there is $P \in E(\Bbbk)$ such that $r(P) = \infty$ so that $s(P) = \infty$ as well and $\rho(P) = \mathcal{O}$. Let $Q \in E(\Bbbk)$. The map $\tau_{-Q} : E \longrightarrow E$ is a rational map by Example 5.54, and the map $\tau_{-Q} \circ \rho$ is a rational map as a composition of such (Proposition 5.55). The same argument as above, shows that there is $P' \in E(\Bbbk)$ such that $(\tau_{-Q} \circ \rho)(P') = \mathcal{O}$. But then, $\rho(P') = Q$ so that ρ is surjective. □

We now proceed to study a specific rational map that plays a key role in the study of Elliptic curves over finite fields. Throughout the rest of this section, let $\mathbb{F}_q$ be a finite field of characteristic p, so that $q = p^k$ for some $1 \leq k$.

Definition 5.57. The map $\Phi_q : \overline{\mathbb{F}}_q \longrightarrow \overline{\mathbb{F}}_q$ given by $\Phi_q(x) = x^q$ is called the **Frobenius endomorphism**. If $E/\mathbb{F}_q$ is an Elliptic curve defined over $\mathbb{F}_q$, Φ_q act on the coordinates of points in $E(\overline{\mathbb{F}}_q)$ by

$$\begin{aligned} \Phi_q(x, y) &= (x^q, y^q) \\ \Phi_q(\mathcal{O}) &= \mathcal{O} \end{aligned} \tag{18}$$

Proposition 5.58. *Let $E/\mathbb{F}_q$ be an Elliptic curve defined over $\mathbb{F}_q$ and $(x, y) \in E(\overline{\mathbb{F}}_q)$. Then*

1. *$\Phi_q(x, y) = E(\overline{\mathbb{F}}_q)$ so that Φ_q defines a rational map $\Phi_q : E \longrightarrow E$ that we call (with slight abuse) the Frobenius endomorphism.*
2. *$(x, y) \in E(\mathbb{F}_q) \iff \Phi_q(x, y) = (x, y)$.*

Proof. Note first that $\overline{\mathbb{F}}_q$ is a field of characteristic p.

1. Recall from field theory that in a field of characteristic p, for any a, b, $(a + b)^p = a^p + b^p$. Thus, $(a + b)^{p^2} = ((a + b)^p)^p = (a^p + b^p)^p = a^{p^2} + b^{p^2}$.

By trivial induction on k, we get that for any $a, b \in \overline{\mathbb{F}}_q$, $(a+b)^q = a^q + b^q$. Recall also that for any $a \in \mathbb{F}^q$ $a^q = a$ since the multiplicative group $\mathbb{F}_q^\times$ is cyclic group of order $q-1$. Since $(x,y) \in E(\overline{\mathbb{F}}_q)$,

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

where $A, B \in \mathbb{F}_q$ since E is defined over $\mathbb{F}_q$. Thus

$$\begin{aligned} (y^2)^q &= (x^3 + Ax + B)^q \\ \iff (y^q)^2 &= (x^q)^3 + A^q x^q + B^q \\ \iff (y^q)^2 &= (x^q)^3 + Ax^q + B \\ \iff (x^q, y^q) &\in E(\overline{\mathbb{F}}_q). \end{aligned} \tag{19}$$

2. Note that for $a \in \overline{\mathbb{F}}_q$, $a^q = a \iff a \in \mathbb{F}_q$ since this all elements $a \in \mathbb{F}_q$ satisfy it and the polynomial $x^q - x$ can have at most q distinct roots in $\overline{\mathbb{F}}_q$. Thus,

$$(x,y) \in E(\mathbb{F}_q) \iff x, y \in \mathbb{F}_q \iff \Phi(x,y) = (x,y).$$

□

Corollary 5.59. *Let $E/\mathbb{F}_q$ be an Elliptic curve and $\Phi_q : E \longrightarrow E$ the Frobenius endomorphism. Then for any rational map $g : E(\mathbb{F}_q) \longrightarrow \mathbb{P}^1_{\mathbb{F}_q}$ we have a commutative diagram*

$$\begin{array}{ccc} E(\mathbb{F}_q) & \xrightarrow{\Phi_q} & E(\mathbb{F}_q) \\ \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle g} \\ \mathbb{P}^1_{\mathbb{F}_q} & \xrightarrow{\Phi_q} & \mathbb{P}^1_{\mathbb{F}_q} \end{array}$$

where the bottom Φ_q sends ∞ to ∞.

Proof. Follows immediately from Proposition 5.58 □

5.6 Weil Reciprocity

Our final piece of theory before going to pairings is a result used by Weil to construct what is now known as the Weil Pairing. The statement is as follows.

Theorem 5.60 (Weil reciprocity)**.** *Let $E/\Bbbk$ be an Elliptic curve defined over an algebraically closed field. If $r, s \in \Bbbk(E) \setminus \{0\}$ are rational functions whose divisors have disjoint support, then*

$$r(\operatorname{div}(s)) = s(\operatorname{div}(r))$$

Weil Reciprocity is in fact a general property of "projective" curves (ie ones with an additional point at infinity), not just Elliptic curves.

The proof of Weil reciprocity for Elliptic curves is carried out in two stages. In the first stage, one proves Weil reciprocity for the projective line $\mathbb{P}^1_{\Bbbk}$ (see Definition 5.61 below). In the second stage, one uses a formal argument for projective curves to "transfer" the proof from the projective line to a general Elliptic curve.

We will devote the remainder of this section to formulate and prove the first stage of Theorem 5.60, ie in the case of the projective line. We omit the second stage since it requires to develop general theory for curves, which we believe is too big of a digression to take in these notes. We hope that the proof of Weil reciprocity for the projective line will give the reader a feeling why it should be true for Elliptic curves as well.

We start with

Definition 5.61. Let $\Bbbk$ be a field. The **projective line** over $\Bbbk$, denoted $\mathbb{P}^1 = \mathbb{P}^1_{\Bbbk}$ is the set

$$\mathbb{P}^1 = \Bbbk \cup \{\infty\}.$$

Just like an Elliptic curve E is the set of all solutions of a polynomial $f_E(x, y)$ + a point at infinity, the projective line can be viewed as the set of all solutions of the zero polynomial + a point at infinity. Thus, the projective line is another example of a curve.

Warning 5.62. Unlike an Elliptic curve, $\mathbb{P}^1_{\Bbbk}$ does not have a group structure.

Definition 5.63. Let $\Bbbk$ be a field. A **rational function** on $\mathbb{P}^1 = \mathbb{P}^1_{\Bbbk}$ is a quotient of polynomials $r(x) = \frac{u(x)}{v(x)}$. We denote the set of all rational functions on $\mathbb{P}^1$ by $\Bbbk(\mathbb{P}^1)$.

As before, we can evaluate a rational function at a point:

Definition 5.64. Let $r \in \Bbbk(\mathbb{P}^1)$. The **induced function** from r is the function $r : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ defined as follows. If $a \neq \infty$, $r(a) = \frac{u(a)}{v(a)}$ except in the case a is a root of $v(x)$, in which case $r(a) : \infty$. If $a = \infty$, $r(a) := \infty$ if $\deg u > \deg v$, $r(a) = 0$ if $\deg u < \deg v$ and if $\deg u = \deg v$, with $u(x) = \sum_{i=1}^n a_i x^i$ and $v(x) = \sum_{i=1}^n b_i x^i$ then $r(a) := \frac{a_n}{b_n}$.

The notion of zeros and poles of a rational function on $\mathbb{P}^1$ is the same as before:

Definition 5.65. Let $r \in \Bbbk(\mathbb{P}^1)$. A point $a \in \mathbb{P}^1$ is called a **zero** of r if $r(a) = 0$ and is called a **pole** of r if $r(a) = \infty$.

We now want to define the multiplicity/order of zeros and poles. Since we are dealing with polynomials in one variable the situation is simpler than with Elliptic curves. The only subtlety is the multiplicity of the point at infinity ∞.

Definition 5.66. Let $r \in \Bbbk(\mathbb{P}^1)$ be a rational function and $a \in \mathbb{P}^1$ a point. Suppose $r(x) = \frac{u(x)}{v(x)}$ where u, v do not have common factors (so in particular have a disjoint set of roots).

1. If a is not a zero or a pole of r, we set $\mathrm{ord}_a(r) = 0$.
2. If $a \neq \infty$:
   - If a is a zero of r, its order $\mathrm{ord}_a(r)$ is the largest positive integer n such that $(x-a)^n | u(x)$.
   - If a is a pole of r, its order $\mathrm{ord}_a(r)$ is $-n$ where n is the largest positive integer such that $(x-a)^n | v(x)$.
3. If $a = \infty$: we define $\mathrm{ord}_\infty(r) = \deg v - \deg u$.

Remark 5.67. Note that by definition, we always have $\mathrm{ord}_\infty(r) = \deg v - \deg u$, since when ∞ is not a zero or a pole of r, $\deg u = \deg v$ so that $\mathrm{ord}_\infty(r) = \deg v - \deg u = 0$.

Next, we define divisors on $\mathbb{P}^1$:

Definition 5.68. A **divisor** on $\mathbb{P}^1$ is a formal sum $D = \sum_{a \in \mathbb{P}^1} n_a [a]$ where the n_a's are integers, and only finitely many of them are non-zero.

Similarly to the case of Elliptic curves, we can associate a divisor to every rational function:

Definition 5.69. Let $r \in \Bbbk(\mathbb{P}^1)$ be a rational function. The **associated divisor** of r is

$$\mathrm{div}(r) = \sum_{a \in \mathbb{P}^1} \mathrm{ord}_a(r)[a].$$

We now have a simple version of Theorem 5.49:

Proposition 5.70. *Let $r = \frac{u}{v} \neq 0 \in \Bbbk(\mathbb{P}^1)$ be a rational function such that both u and v decompose to linear factors over $\Bbbk$ (eg $\Bbbk$ is algebraically closed). Write $\mathrm{div}(r) = \sum_{a \in \mathbb{P}^1} \mathrm{ord}_a(r)[a]$. Then*

$$\sum_{a \in \mathbb{P}^1} n_a = 0.$$

Proof. Since u and v decompose to linear factors, we can write

$$u(x) = \prod_{i=1}^{m} (x - a_i)^{n_{a_i}}$$

and

$$v(x) = \prod_{j=1}^{m'} (x - b_j)^{n_{b_j}}$$

where

$$d_u = \deg u = \sum_{i=1}^{m} n_{a_i}$$

and

$$d_v = \deg v = \sum_{j=1}^{m'} n_{b_j}.$$

Clearly, each a_i is a zero of r with multiplicity/order n_{a_i} and each b_j is a pole of r with multiplicity/order $-n_{b_j}$. Thus, the associated divisor of r can be written as

$$\operatorname{div}(r) = \sum_{i=1}^{m} n_{a_i}[a_i] - \sum_{j=1}^{m'} n_{b_j}[b_j] + \operatorname{ord}_\infty(r)[\infty]$$

and the sum of the coefficients can be written as

$$S = \sum_{i=1}^{m} n_{a_i} - \sum_{j=1}^{m'} n_{b_j} + \operatorname{ord}_\infty(r) = d_u - d_v + \operatorname{ord}_\infty(r).$$

If $d_u = \deg u = \deg v d_v$ then by definition ∞ is neither a zero nor a pole of r so $\operatorname{ord}_\infty(r) = 0$ and in this case $S = d_u - d_v = 0$.

If $d_u > d_v$ then by definition ∞ is a pole and $\operatorname{ord}_\infty(r) = -(d_u - d_v)$ so that again $S = 0$.

Lastly, if $d_u < d_v$, then by definition ∞ is a zero and

$$\operatorname{ord}_\infty(r) = d_v - d_u = -(d_u - d_v)$$

so that $S = 0$ as well.

□

As in the section on divisors that the **support** of a divisor $D = \sum_{a\in\mathbb{P}^1} n_a[a]$ is the set of points $a \in \mathbb{P}^1$ for which $n_a \neq 0$. If $r \in \Bbbk(\mathbb{P}^1)$ is a rational function and $D = \sum_{a\in\mathbb{P}^1} n_a[a]$ a divisor whose support does not contain zeros or poles of r, the **evaluation** of r at D is

$$r(D) = \prod_{a\in\mathbb{P}^1} r(a)^{n_a}.$$

We are now ready to prove Weil Reciprocity for the projective line.

Theorem 5.71 (Weil Reciuprocity for the projective line)**.** *Let $\Bbbk$ be algebraically closed and $r, s \in \Bbbk(\mathbb{P}^1)$ two rational functions with disjoint support. Then*

$$r(\operatorname{div}(s)) = s(\operatorname{div}(r)).$$

Proof. Write $r = u/v$ and $s = u'/v'$. Since $\Bbbk$ is algebraically closed, the polynomials $u, v, u'.v'$ decompose to linear factors. Thus, we can write

$$r(x) = \prod_{i=1}^{m} (x - a_i)^{n_{a_i}}$$

and

$$s(x) = \prod_{j=1}^{m'} (x - b_j)^{n_{b_j}}$$

where $n_{a_i}, n_{b_j} \in \mathbb{Z}$ are the orders of a_i, b_j respectively.

Write

$$\operatorname{div}(r) = \sum_{i=1}^{m} n_{a_i}[a_i] - n_\infty^r[\infty]$$

and

$$\operatorname{div}(s) = \sum_{j=1}^{m'} n_{b_j}[b_j] - n_\infty^s[\infty].$$

By Proposition 5.70, $n_\infty^r = \sum_{i=1}^{m} n_{a_i}$ and $n_\infty^s = \sum_{j=1}^{m'} n_{b_j}$.

Suppose first that the support of both r and s do not contain the point at infinity ∞, ie $\forall i, j,\ a_i \neq b_j$ and

$$\sum_{i=1}^{m} n_{a_i} = 0 = \sum_{j=1}^{m'} n_{b_j}.$$

Then

$$\begin{aligned} r(\operatorname{div}(s)) &= \prod_{j=1}^{m'} \left[\prod_{i=1}^{m} (b_j - a_i)^{n_{a_i}} \right]^{n_{b_j}} \\ &= \prod_{j=1}^{m'} \prod_{i=1}^{m} (b_j - a_i)^{n_{a_i} n_{b_j}} \\ &= (-1)^{\sum_{i=1}^{m} \sum_{j=1}^{m'} n_{a_i} n_{b_j}} \prod_{i=1}^{m} \prod_{j=1}^{m'} (a_i - b_j)^{n_{a_i} n_{b_j}} \\ &= s(\operatorname{div}(r)) \end{aligned} \tag{20}$$

where the sign equals 1 since

$$\sum_{i=1}^{m} \sum_{j=1}^{m'} n_{a_i} n_{b_j} = \left(\sum_{i=1}^{m} n_{a_i} \right) \left(\sum_{j=1}^{m'} n_{b_j} \right) = 0 \cdot 0 = 0.$$

If, without loss of generality, ∞ is in the support of $r(x) = \frac{u(x)}{v(x)}$ (so by assumption, ∞ is not in the support of s), then

$$\operatorname{div}(r) = \sum_{i=1}^{m} n_{a_i}[a_i] - n_\infty[\infty]$$

with

$$n_\infty = n_\infty^r = \sum_{i=1}^{m} n_{a_i} = \deg u - \deg v,$$

and

$$\operatorname{div}(s) = \sum_{j=1}^{m'} n_{b_j}[b_j]$$

with

$$\sum_{j=1}^{m'} n_{b_j} = 0$$

as before.

Then,

$$\begin{aligned} r(\mathrm{div}(s)) &= \left[\prod_{j=1}^{m'}\prod_{i=1}^{m}(b_j - a_i)^{n_{a_i}n_{b_j}}\right] \cdot \prod_{j=1}^{m'}(\infty - b_j)^{n_\infty n_{b_j}} = 1 \\ &= \left[(-1)^{\sum_{i=1}^{m}\sum_{j=1}^{m'} n_{a_i}n_{b_j}} \prod_{i=1}^{m}\prod_{j=1}^{m'}(a_i - b_j)^{n_{a_i}n_{b_j}}\right] \cdot \prod_{j=1}^{m'}(\infty - b_j)^{n_\infty n_{b_j}} \\ &= \left[\prod_{i=1}^{m}\prod_{j=1}^{m'}(a_i - b_j)^{n_{a_i}n_{b_j}}\right] \cdot \prod_{j=1}^{m'}(\infty - b_j)^{n_\infty n_{b_j}} \\ &= \prod_{i=1}^{m}\prod_{j=1}^{m'}(a_i - b_j)^{n_{a_i}n_{b_j}} \\ &= s(\mathrm{div}(r)) \end{aligned} \tag{21}$$

where similarly as before, the sign equals 1 since

$$\sum_{i=1}^{m}\sum_{j=1}^{m'} n_{a_i}n_{b_j} = \left(\sum_{i=1}^{m} n_{a_i}\right)\left(\sum_{j=1}^{m'} n_{b_j}\right) = n_\infty \cdot 0 = 0$$

and in addition

$$\prod_{j=1}^{m'}(\infty - b_j)^{n_\infty n_{b_j}} = 1,$$

because

$$\sum_{j=1}^{m'} n_\infty n_{b_j} = n_\infty \sum_{j=1}^{m'} n_{b_j} = n_\infty \cdot 0 = 0$$

and by our convention, for every pair of **monic** polynomials $f(x), g(x)$ of the same degree, $\frac{f(\infty)}{g(\infty)} = 1$ (here, the polynomial $f(x)$ is given by

$$f(x) = \prod_{n_{b_j}>0} (x - b_j)^{n_{b_j}}$$

and the polynomial $g(x)$ is given by

$$g(x) = \prod_{n_{b_j}<0} (x - b_j)^{n_{b_j}}$$

) This completes the proof. □

We finish this section by a generalised form of Weil reciprocity. First, consider the following

Definition 5.72. Let $E/\Bbbk$ be an Elliptic curve, $f, g \in \Bbbk(E)$ rational functions and $P \in E(\overline{\Bbbk})$. The **tame symbol** of f and g is

$$\langle f, g\rangle_P = (-1)^{\mathrm{ord}_P(f)\,\mathrm{ord}_P(g)} \left(\frac{f^{\mathrm{ord}_P(g)}}{g^{\mathrm{ord}_P(f)}}\right)(P).$$

Theorem 5.73. *Let $E/\Bbbk$ be an Elliptic curve and $f, g \in \Bbbk(E)$ be rational functions. Then*

$$\prod_{P \in E(\overline{\Bbbk})} \langle f, g\rangle_P = 1.$$

Remark 5.74. Note that $\langle f, g\rangle_P = 1$ if P is not a zero nor a pole of both f and g. T Thus, for fixed rational functions $f, g \in \Bbbk(E)$, in case we know f and g do not admit zeros or pole over $\overline{\Bbbk}$ that did not exist in $\Bbbk$, Theorem 5.73 remains valid with the product on the RHS taken over $E(\Bbbk)$.

5.7 Torsion points

Definition 5.75. Let $E : y^2 = x^3 + Ax + B$ be an Elliptic curve defined over $\Bbbk$. The n**-torision** points of E are

$$E[n] = E(\Bbbk)[n] = \{P \in E(\Bbbk) \mid nP = \mathcal{O}\} \subseteq E(\overline{\Bbbk}).$$

If $\Bbbk \subseteq \Bbbk'$ is a subfield inclusion, we also define

$$E(\Bbbk')[n] = \{P \in E(\Bbbk') \mid nP = \mathcal{O}\} \subseteq E(\overline{\Bbbk})$$

that by observation 5.15 gives an inclusion

$$E(\Bbbk)[n] \subseteq E(\Bbbk')[n].$$

The set $E(\Bbbk)[n]$ has an evident abelian group structure since for $P, Q \in E(\Bbbk)[n]$, $n(P + Q) = nP + nQ$.

Example 5.76. It is easy to determine $E[2]$: over $\overline{\Bbbk}$ we can write E as follows, with distinct α_i:

$$y^2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$$

A point $P \in E(\overline{\Bbbk})$ satisfies $2P = \mathcal{O}$ iff the tangent line to P is vertical which means that $y = 0$. It follows that $E[2] = \{\mathcal{O}, (\alpha_1, 0), (\alpha_2, 0), (\alpha_3, 0)\}$. As an abelian group, $E[2] \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$.

Let us examine $E(\overline{\Bbbk})[3]$. $3P = \mathcal{O} \iff 2P = -P$ and this means that the x-coordinate of $2P$ equals to that of P and the y-coordinates of $2P$ and P differ by a sign. In equations:

$m^2 - 2x = x$, where $m = \frac{3x^2+A}{2y}$ (see Equation 8). Substituting that in the equation of E we get

$$\begin{aligned} \frac{(3x^2+A)^2}{4m^2} &= x^3 + Ax + B \iff \\ (3x^2+A)^2 &= 12x(x^3+Ax+B) \iff \\ 3x^4 + 6Ax^2 + 12Bx - A^2 &= 0. \end{aligned} \tag{22}$$

The discriminant of the resulting polynomial is $-6912(4A^3+27B^2)^2$ (as can be verified by the formula in Example 5.9) which is non-zero since E is smooth. Thus, this polynomial has no multiple roots. There are 4 different roots, ie. values of x, over $\overline{\Bbbk}$ and each of them yields 2 values of y. Together with $\mathcal{O}$ we get $|E[3]| = 9$ where each non-zero element has order 3. It follows that

$$E(\overline{\Bbbk})[3] \cong \mathbb{Z}_3 \times \mathbb{Z}_3.$$

The general situation is given by the following

Theorem 5.77. *Let E be an Elliptic curve defined over $\mathbb{F}_p$ and let n be a positive integer. If $p \nmid n$, there exists k such that for any $k < N$:*

$$E(\mathbb{F}_{p^N})[n] \cong \mathbb{Z}_n \times \mathbb{Z}_n.$$

Corollary 5.78. *Let E be an Elliptic curve over $\mathbb{F}_{p^k}$. Then*

$$E(\mathbb{F}_{p^k}) \cong \begin{cases} \mathbb{Z}_n \\ \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \ \text{ where } \ n_1 \mid n_2 \end{cases}$$

Proof. By the structure theorem of finite abelian groups 3.52,

$$E(\mathbb{F}_{p^k}) \cong \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_r}$$

where for every $1 \le i < r$, $n_i \mid n_{i+1}$ and r is the "rank". Each component $\mathbb{Z}_{n_i}$ contains n_1 elements of order dividing n_1 so that $E(\mathbb{F}_{p^k})$ contains n_1^r elements of order dividing n_1. However,

$$E(\mathbb{F}_{p^k})[n_1] \subseteq E[n_1] \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_1}$$

so $E(\mathbb{F}_{p^k})$ has at most n_1^2 elements of order dividing n_1. It follows that $r \le 2$. □

As an aside let us state the case where $p \mid n$:

Theorem 5.79. *Let E be an Elliptic curve defined over $\mathbb{F}_p$ and let n be a positive integer. If $p \mid n$ where $n = p^r n'$ with $p \nmid n'$ then there exists k such that for any $k < N$:*

$$E(\mathbb{F}_{p^N})[n] = \mathbb{Z}_{n'} \oplus \mathbb{Z}_{n'} \ \text{ or } \ \mathbb{Z}_n \oplus \mathbb{Z}_{n'} \tag{23}$$

Observation 5.80. For an Elliptic curve $E = E_{A,B}$, the group $E(\overline{\mathbb{F}}_p)$ is always infinite: given any point $y_0 \in \overline{\mathbb{F}}_p$, the polynomial $x^3 + Ax + B - y_0$ must have a root x_0 in $\overline{\mathbb{F}}_p$ so that $(x_0, y_0) \in E(\overline{\mathbb{F}_p})$.

For future use, let us record the following

Definition 5.81. An Elliptic curve E defined over $\Bbbk = \overline{\mathbb{F}}_p$ is called **ordinary** if $E(\Bbbk)[p] \cong \mathbb{Z}_p$ and is called **supersingular** if $E[p] \cong 0$.

Remark 5.82. Note that the notion supersingular is unrelated to the notion of singular curves appears in the literature.

5.8 Weil pairing and its properties

Recall our setting: E is an Elliptic curve defined over $\mathbb{F}_p$, $1 \leq n$ an integer with $p = \operatorname{char} \Bbbk \nmid n$ and $\mathbb{F}_p \subseteq \Bbbk$ a field such that

$$E[n] := E(\Bbbk)[n] \cong \mathbb{Z}_n \oplus \mathbb{Z}_n.$$

Let us start with a construction of Weil pairing.

Construction 5.83. Let $Q \in E[n]$ and $f_Q \in \Bbbk(E)$ a rational function such that

$$\operatorname{div}(f_Q) = n[Q] - n[\mathcal{O}].$$

By Proposition 5.56, the map $[n] : E(\Bbbk) \longrightarrow E(\Bbbk)$ given by $P \mapsto nP$ is surjective so let $Q' \in E(\Bbbk)[n^2]$ be such that $nQ' = Q$.

Consider the divisor

$$\sum_{R \in E[n]} \left([Q' + R] - [R]\right).$$

Since $|E[n]| = n^2$,

$$\sum_{R \in E[n]} (Q' + R - R) = n^2 Q' = \mathcal{O}$$

so that there exists a rational function $g = g_Q \in \Bbbk(E)$ such that

$$\operatorname{div}(g) = \sum_{R \in E[n]} \left([Q' + R] - [R]\right)$$

Note that g does not depend on the choice of Q': if $Q'' \in E[n^2]$ is such that $nQ'' = Q$, then $Q' - Q'' \in E[n]$ so that

$$\sum_{R \in E[n]} \left([Q'' + R] - [R]\right)$$

is unchanged.

Consider the rational function $f_Q \circ [n] \in \Bbbk(E)$. The points $R \in E[n]$ are poles of $f_Q \circ [n]$ of order n each. The points $X = Q' + R$ for $R \in E[n]$ are those points X for which $nX = Q$, hence the zeros of $f_Q \circ [n]$, each with order n as before. It follows that

$$\operatorname{div}(f_Q \circ [n]) = n\left(\sum_{R \in E[n]} [Q' + R]\right) - n\left(\sum_{R \in E[n]} [R]\right) = \operatorname{div}(g^n).$$

Thus, $f_Q \circ [n]$ is a constant multiple of g^n and wlog, we may choose f_Q such that $f_Q \circ [n] = g^n$.

Let $P \in E[n]$ and $S \in E(\Bbbk)$. Then

$$g(S+P)^n = f_Q\left(n(S+P)\right) = f_Q(nS) = g(S)^n$$

so that

$$\frac{g(S+P)^n}{g(S)^n} = 1.$$

We define the (abstract) **Weil pairing** to be

$$e_n(P,Q) = \frac{g(S+P)}{g(S)}$$

and thus get a function

$$e_n : E(\Bbbk)[n] \times E(\Bbbk)[n] \longrightarrow \Bbbk$$

where for every $P, Q \in E[n]$, $e_n(P,Q) \in \mu_n(\overline{\mathbb{F}_p})$.

Before going to the main proof of this section, we need an auxiliary result:

Lemma 5.84. *Let $E/\Bbbk$ be an Elliptic curve, and $1 \leq n$ such that $p = \mathrm{char}(\Bbbk) \nmid n$ and $E(\Bbbk)[n] \cong \mathbb{Z}_n \oplus \mathbb{Z}_n$. Suppose $g \in \Bbbk(E)$ is a rational function such that*

$$g(S+P) = g(S)$$

for all $S \in E(\Bbbk)$ and $P \in E(\Bbbk)[n]$. Then there is a rational function

$$h \in \Bbbk(E)$$

such that for all $S \in E(\Bbbk)$,

$$g(S) = h(nS).$$

In other words, if we consider the rational map $[n] : E(\Bbbk) \longrightarrow E(\Bbbk)$ given by $S \mapsto nS$ then

$$g = h \circ [n].$$

Proof. The proof involves Galois Theory and thus omitted from these notes. □

We are ready to for the main result of this section.

Theorem 5.85. *The Weil pairing of Construction 5.83 has the following properties:*

1. *$\forall P, Q \in E[n]$, $e_n(P,Q)^n = 1$ ie $e_n(P,Q) \in \mu_n(\overline{\mathbb{F}_p})$.*
2. *$\forall P, Q \in E[n]$, $e_n(P,Q)$ is independent of the choice of function $g = g_Q$ and the point S.*

3. e_n *is* ***bilinear*** *in each variable, ie for all* $P_1, P_2, Q \in E(\Bbbk)[n]$

$$e_n(P_1 + P_2, Q) = e_n(P_1, Q)e_n(P_2, Q)$$

and

$$e_n(Q, P_1 + P_2) = e_n(Q, P_1)e_n(Q, P_2).$$

4. *The Weil pairing is* ***skew-symmetric*** *(or: alternating) ie.*

$$e_n(Q, Q) = 1 \ \forall Q \in E(\mathbb{K})[n]$$

and

$$e_n(P, Q) = e_n(Q, P)^{-1} \ \forall P, Q \in E(\mathbb{K})[n].$$

5. e_n *is* ***non-degenerate*** *in each variable, ie if* $e_n(P, Q) = 1$ *for all* $Q \in E(\Bbbk)[n]$ *then* $P = \mathcal{O}$ *and if* $e_n(P, Q) = 1$ *for all* $P \in E(\Bbbk)[n]$ *then* $Q = \mathcal{O}$.

6. *The set of maps* $\{e_n\}$ *is* ***compatible*** *in that for any* m, n *with* $p \nmid n, m$, *for any* $P \in E[mn]$ *and for any* $Q \in E[n] \subseteq E[mn]$,

$$e_{mn}(P, Q) = e_n(mP, Q).$$

Proof.

1. This was proved in Construction 5.83.

2. Fix $P, Q \in E(\Bbbk)[n]$. Then

$$e_n(P, Q) = \frac{g_Q(P + S)}{g_Q(S)}$$

can be viewed as a function of $S \in E(\Bbbk)$ ie as a rational function $E(\Bbbk) \longrightarrow \Bbbk$. Enlarging $\Bbbk$ if necessary, we invoke Proposition 5.56 that says this function is either constant or surjective. However, since we know $e_n(P, Q) \in \mu_n(\overline{\mathbb{F}_p})$, it cannot be surjective hence constant ie does not depend on S.

3. Let us prove linearity in the first variable. Since e_n is independent of the choice of S, we may replace S by $S + P_1$ to get

$$\begin{aligned} e_n(P_1, Q)e_n(P_2, Q) &= \frac{g(P_1 + S)}{g(S)} \frac{g(P_2 + P_1 + S)}{g(P_1 + S)} \\ &= \frac{g(P_1 + P_2 + S)}{g(S)} = e_n(P_1 + P_2, Q). \end{aligned} \tag{24}$$

For linearity in the second variable, suppose $Q_1, Q_2, Q_3 \in E[n]$ such that $Q_1 + Q_2 = Q_3$. For $1 \leq i \leq 3$, let f_{Q_i}, g_{Q_i} be the functions used to define $e_n(P, Q_i)$ in Construction 5.83, and let $h \in \Bbbk(E)$ be a rational function such that

$$\operatorname{div}(h) = [Q_3] - [Q_2] - [Q_1] + [\mathcal{O}].$$

Then we have

$$\operatorname{div}\left(\frac{f_{Q_3}}{f_{Q_1}f_{Q_2}}\right) = n\operatorname{div}(h) = \operatorname{div}(h^n).$$

Thus, there exists a constant $c \in \Bbbk$ such that $f_{Q_3} = cf_{Q_1}f_{Q_2}h^n$ and this means that

$$\begin{aligned} g_{Q_3}^n = f_{Q_3} \circ [n] &= (c \cdot f_{Q_1} \cdot f_{Q_2} \cdot h^n) \circ [n] \\ &= c \cdot (f_{Q_1} \circ [n]) \cdot (f_{Q_2} \circ [n])(h \circ [n])^n. \end{aligned} \tag{25}$$

Taking nth root, we get

$$g_{Q_3} = c^{\frac{1}{n}}(g_{Q_1})(g_{Q_2})(h \circ [n]).$$

The definition of e_n now yields

$$\begin{aligned} e_n(P, Q_1 + Q_2) &= \frac{g_{Q_3}(P+S)}{g_{Q_3}(S)} \\ = \frac{g_{Q_1}(P+S)}{g_{Q_1}(S)}\frac{g_{Q_2}(P+S)}{g_{Q_2}(S)}\frac{h(n(P+S))}{h(nS)} &= e_n(P, Q_1)e_n(P, Q_2) \end{aligned} \tag{26}$$

where the last equality follows since $nP = \mathcal{O}$ so that $h(n(P+S)) = h(nS)$.

4. Let $\tau_{jQ} : E(\Bbbk) \longrightarrow E(\Bbbk)$ be the translation by jQ so $f \circ \tau_{jQ}$ is the function $P \mapsto f(P + jQ)$. A direct inspection shows that

$$\operatorname{div}(f_Q \circ \tau_{jQ}) = n[Q - jQ] - n[-jQ].$$

Thus,

$$\operatorname{div}\left(\prod_{j=0}^{n-1} f_Q \circ \tau_{jQ}\right) = \sum_{j=0}^{n-1}(n[(1-j)Q] - n[-jQ]) = 0,$$

so that $\prod_{j=0}^{n-1} f_Q \circ \tau_{jQ}$ is constant. We now have

$$\begin{aligned} \left(\prod_{j=0}^{n-1} g \circ \tau_{jQ'}\right)^n &= \prod_{j=0}^{n-1} f_Q \circ [n] \circ \tau_{jQ'} \\ = \prod_{j=0}^{n-1} f_Q \circ \tau_{jQ} \circ [n] \quad (since\ \ nQ' = Q) \\ &= \left(\prod_{j=0}^{n-1} f_Q \circ \tau_{jQ}\right) \circ [n]. \end{aligned} \tag{27}$$

so that $\left(\prod_{j=0}^{n-1} g \circ \tau_{jQ'}\right)^n$ is constant. We now invoke Proposition 5.56 to deduce that there is some field extension K of $\Bbbk$ for which $\prod_{j=0}^{n-1} g \circ \tau_{jQ'}$ is

constant, hence must be constant already over $\Bbbk$. Thus, $\prod_{j=0}^{n-1} g \circ \tau_{jQ'}$ has the same value at S and $S+Q'$ so that

$$\prod_{j=0}^{n-1} g(S+Q'+jQ') = \prod_{j=0}^{n-1} g(S+jQ').$$

Canceling all common terms (we assume S is chosen such that all terms are finite and non-zero), we get

$$g(S+nQ') = g(S).$$

Since $nQ' = Q$ we get

$$e_n(Q,Q) = \frac{g(S+Q)}{g(S)} = 1.$$

As for the second part, bilinearity yields

$$1 = e_n(P+Q,P+Q) = e_n(P,P)e_n(P,Q)e_n(Q,P)e_n(Q,Q)$$

and since we have just showed that $e_n(P,P) = e_n(Q,Q) = 1$ we get

$$e_n(P,Q) = e_n(Q,P)^{-1}.$$

5. Suppose $Q \in E[n]$ is such that $e_n(P,Q) = 1$ for all $P \in E[n]$. By Theorem 5.88, this means that $e_n(P,Q) = 1$ so that $g(S+P) = g(S)$ for all $P \in E[n]$ and $S \in E(\Bbbk)$. By Lemma 5.84, there is a rational function $h \in \Bbbk(E)$ such that $g = h \circ [n]$. Then,

$$(h \circ [n])^n = g^n = f \circ [n].$$

Note that $(h \circ [n])^n = h^n \circ [n]$ so that $h^n \circ [n] = f \circ [n]$. By Proposition 5.56, $[n] : E(\Bbbk) \longrightarrow E(\Bbbk)$ is surjective, so by Lemma 2.16, $h^n = f$. Thus, we have

$$n \operatorname{div}(h) = \operatorname{div}(f) = n[Q] - n[\mathcal{O}],$$

so that

$$\operatorname{div}(h) = [Q] - [\mathcal{O}].$$

Since h is a rational function, $Q = \mathcal{O}$, and this proves half of Theorem 5.85 (5). The second half follows from the first half in conjunction with (4).

6. The proof requires a use of another definition of the Weil pairing. We thus deffer it to the end of section 5.9.

□

Corollary 5.86. *In the setting of Construction 5.83, $\mu_n(\overline{\mathbb{F}_p}) \subseteq \Bbbk$ so that the Weil pairing can be viewed as a map*

$$e_n : E[n] \times E[n] \longrightarrow \mu_n(\overline{\mathbb{F}_p}).$$

Proof. Let $P_0, Q_0 \in E[n]$ be such that $\langle (P_0, Q_0) \rangle = E[n] = \mathbb{Z}_n \times \mathbb{Z}_n$. We claim that $e_n(P_0, Q_0)$ is a generator of $\mu_n(\overline{\mathbb{F}_p})$ which means that e_n is surjective onto $\mu_n(\overline{\mathbb{F}_p})$. Since by definition $e_n(P, Q) \in \Bbbk$ for all P, Q, we get that $\mu_n(\overline{\mathbb{F}_p}) \subseteq \Bbbk$.

Suppose by contradiction that $e_n(P_0, Q_0)$ is not a generator of μ_n. Then there exists $m < n$ such that $e_n(P_0, Q_0)^m = 1$. By bilinearity we get $e_n(mP_0, Q_0) = 1$. If $X \in E[n]$ then there are $a, b \in \mathbb{Z}$ such that $X = aP_0 + bQ_0$. Then

$$e_n(mP_0, X) = e_n(mP_0, P_0)^a e_n(mP_0, Q_0)^b = e_n(P_0, P_0)^{ma} e_n(P_0, Q_0)^{mb} = 1$$

and thus by non-degeneracy, $mP_0 = \mathcal{O}$ – contradiction. □

5.9 Equivalence of definitions of Weil pairing

The Weil pairing admits a few equivalent definitions in the literature, each useful for different purposes. Our goal in this section is to two other definitions of the Weil pairing and prove their equivalence.

Definition 5.87. Let $E/\Bbbk$ be an Elliptic curve. Suppose $P, Q \in E(\Bbbk)[n]$ such that $P \neq Q$ and $P, Q \neq \mathcal{O}$. Let $f_P, f_Q \in \Bbbk(E)$ be monic rational functions such that

$$\operatorname{div}(f_P) = n[P] - n[\mathcal{O}], \quad \operatorname{div}(f_Q) = n[Q] - n[\mathcal{O}].$$

The **efficient Weil pairing** of P, Q is defined to be

$$e_n^{\text{eff}}(P, Q) = (-1)^n \frac{f_P(Q)}{f_Q(P)}.$$

If $P = Q$ or either $P = \mathcal{O}$ or $Q = \mathcal{O}$, we set

$$e_n^{\text{eff}}(P, Q) = 1.$$

The following Theorem is stated here for convenience, but we deffer its proof to a later part in this section.

Theorem 5.88. *Under the setup of Construction 5.83, for every $P, Q \in E[n]$, $e_n(P, Q) = e_n^{\text{eff}}(P, Q)$.*

Remark 5.89. In Definition 5.87, since f_P, f_Q are monic and have a pole of order n at $\mathcal{O}$, it follows that $\frac{f_Q(\mathcal{O})}{f_P(\mathcal{O})} = 1$ so that

$$e_n^{\text{eff}}(P, Q) = (-1)^n \frac{f_P(Q)}{f_Q(P)} \cdot \frac{f_Q(\mathcal{O})}{f_P(\mathcal{O})}.$$

There is yet another definition of the Weil pairing which is convenient for "theoretical" purposes as it is expressed in terms of general divisors. First, let us introduce a bit of terminology: we say that a divisor D on an Elliptic curve $E/\Bbbk$ is **principal** if there is a rational function $f \in \Bbbk(E)$ such that $\operatorname{div}(f) = D$. Recall that according to Theorem 5.49, a divisor D is principal if and only if:

$$\operatorname{sum} D = \mathcal{O}, \quad \deg D = 0.$$

Next, let us define

Definition 5.90. Let $E/\Bbbk$ be an Elliptic curve and D, D' divisors on E. We say that D is equivalent to D' and write $D \sim D'$ if $D - D'$ is principal.

Exercise 5.91. Prove that equivalence of divisors is an equivalence relation on the set of all divisors.

We are ready to define the generic version of the Weil pairing:

Definition 5.92. Let $E/\Bbbk$ be an Elliptic curve. For a degree zero divisor D, such that $nD \sim 0$ we let f_D to denote a monic rational function in $\Bbbk(E)$ such that

$$\operatorname{div}(f_D) = nD.$$

(e.g. $f_{[P]-[\mathcal{O}]} = f_P$ in the notation of Construction 5.83).

Let $P, Q \in E(\Bbbk)[n]$. Choose divisors D_P, D_Q with disjoint support such that

$$D_P \sim [P] - [\mathcal{O}], \quad D_Q \sim [Q] - [\mathcal{O}].$$

The **generic Weil pairing** is defined to be

$$e_n^{\text{gen}}(P, Q) = \frac{f_{D_P}(D_Q)}{f_{D_Q}(D_P)}.$$

Our first order of business is to verify Definition 5.92 is well-defined.

Lemma 5.93. *Definition 5.92 does not depend on the choice of divisors D_P, D_Q or rational functions f_{D_P}, f_{D_Q}.*

Proof. The choice of rational functions with a prescribed divisor is unique up to a constant and the requirement they are monic means they are unique. Let us show independence of the choice of D_Q; independence of the choice of D_P is proven analogously. Suppose D'_Q is another divisor such that $D'_Q \sim [Q] - [\mathcal{O}]$. Then $D'_Q = D_Q + \operatorname{div}(h)$ for some $h \in \Bbbk(E)$ with support disjoint from D_P. Then $f_{D'_Q} = f_{D_Q} h^n$ and thus

$$\begin{aligned} \frac{f_{D_P}(D'_Q)}{f_{D'_Q}(D_P)} &= \frac{f_{D_P}(D_Q) f_{D_P}(\operatorname{div} h)}{f_{D_Q}(D_P) h(D_P)^n} \\ &= \frac{f_{D_P}(D_Q) f_{D_P}(\operatorname{div} h)}{f_{D_Q}(D_P) h(\operatorname{div}(f_{D_P}))} \\ &= \frac{f_{D_P}(D_Q)}{f_{D_Q}(D_P)}. \end{aligned} \tag{28}$$

□

We now wish to show that the Definitions of generic Weil pairing and efficient Weil pairing are equivalent and thus, by Remark 5.89 are equivalent to the (computational) definition of the Weil pairing we originally defined.

Proposition 5.94. *Let $E/\Bbbk$ be an Elliptic curve and $1 \leq n$ such that $p = \operatorname{char} \Bbbk \nmid n$ and $E(\Bbbk)[n] \cong \mathbb{Z}_n \times \mathbb{Z}_n$. For any $P, Q \in E(\Bbbk)[n]$,*

$$e_n^{\text{eff}}(P,Q) = e_n^{\text{gen}}(P,Q).$$

Proof. Assume first that $P \neq Q$. Let $D_P = [P] - [\mathcal{O}]$ so that $f_{D_P} = f_P$. Let $S \in E(\Bbbk)$ be such that $D_Q := [Q+S] - [S]$ has disjoint support with D_P (i.e. $S \notin \{\mathcal{O}, P, -Q, P-Q\}$).

Then, $D_Q = [Q] - [\mathcal{O}] + \operatorname{div}(h)$ for some monic $h \in \Bbbk(E)$ such that

$$\operatorname{div} h = [Q+S] - [S] - [Q] + [\mathcal{O}]$$

and thus

$$f_{D_Q} = f_Q h^n.$$

By generalised Weil reciprocity 5.73,

$$\begin{aligned} prod_{A \in E(\bar{\Bbbk})} \langle f_P, h \rangle_A &= prod_{A \in E(\Bbbk)} \langle f_P, h \rangle_A \\ &= (-1)^n \frac{f_P(Q+S) f_P(\mathcal{O})}{f_P(S) f_P(Q) h^n(P) h^{-n}(\mathcal{O})} \\ &= \frac{f_P(Q+S)}{f_P(S) f_P(Q) h^n(P)} \cdot (-1)^n (f_P h^n)(\mathcal{O}) \\ &= \frac{f_P(Q+S)}{f_P(S) f_P(Q) h^n(P)} \cdot (-1)^n. \end{aligned} \tag{29}$$

Thus,

$$\begin{aligned} e_n^{\text{gen}}(P,Q) = \frac{f_{D_P}(D_Q)}{f_{D_Q}(D_P)} &= \frac{(f_Q h^n)(\mathcal{O})}{(f_Q h^n)(P)} \cdot \frac{f_P(Q+S)}{f_P(S)} \\ &= \frac{f_P(Q)}{f_Q(P)} \cdot \frac{f_P(Q+S)}{f_P(Q) h^n(P) f_P(S)} \\ &= (-1)^n \frac{f_P(Q)}{f_Q(P)} = e_n^{\text{eff}}(P,Q) \end{aligned} \tag{30}$$

If $P = Q$, a similar calculation shows that

$$e_n^{\text{gen}}(P,Q) = 1.$$

□

In light of Proposition 5.94, the proof of Theorem 5.9 can be reduced to the following Theorem, whose proof will occupy the remain of this section.

Theorem 5.95. *Let $E/\Bbbk$ be an Elliptic curve and let $1 \leq n$ be an integer such that $p = \operatorname{char} \Bbbk \nmid n$ and $E(\Bbbk)[n] \cong \mathbb{Z}_n \times \mathbb{Z}_n$. Then for any $P, Q \in E(\Bbbk)[n]$,*

$$e_n(P,Q) = e_n^{\text{gen}}(P,Q).$$

Let $V, W \in E(\Bbbk)[n^2]$ and let f_{nV}, g_{nV} be as in Construction 5.83 ie such that

$$\operatorname{div}(f_{nV}) = n[nV] - n[\mathcal{O}],$$
$$g_{nV}^n = f_{nV} \circ [n].$$

Define

$$c(nV, nW) = \frac{f_{nV+nW}(X)}{f_{nV}(X) f_{nW}(X - nV)}$$

and

$$d(V, W) = \frac{g_{nV+nW}(X)}{g_{nV}(X) g_{nW}(X - V)}.$$

where $X \in E(\Bbbk)$. The notation on the left does not include X because

Lemma 5.96. *$c(nV, nW)$ and $d(V, W)$ are constants and*

$$d(V, W)^n = c(nV, nW).$$

Proof. Using $\operatorname{div}(f_{nV}) = n[nV] - n[\mathcal{O}]$ we get that $\operatorname{div} c(nV, nW) = 0$ and thus $c(nV, nW)$ is constant. Since $g_{nV}^n = f_{nV} \circ [n]$ we get that

$$d(V, W)^n = \frac{f_{nV+nW}(X)}{f_{nV}(X) f_{nW}(nX - nV)} = c(nV, nW)$$

where the last equality holds since $c(nV, nW)$ does not depend on X. Thus, $d(V, W)$ is constant as well.

□

The next few Lemmas tie c and d to e_n.

Lemma 5.97. *Let $U, V, W \in E[n^2]$. Then*

$$d(V, W + nU) = d(V, W)$$

and

$$d(V + nU, W) = d(V, W) e_n(nU, nW).$$

Proof. Since $n(W + nU) = nW$, we have $g_{nV+n(W+nU)} = g_{nV+nW}$. Thus,

$$d(V, W + nU) = \frac{g_{nV+nW}(X)}{g_{nV}(X) g_{nW}(X - V)} = d(V, W).$$

Similarly,

$$\begin{aligned} d(V + nU, W) &= \frac{g_{nV+nW}(X)}{g_{nV}(X) g_{nW}(X - V - nU)} \\ &= \frac{g_{nV+nW}(X)}{g_{nV}(X) g_{nW}(X - V)} \frac{g_{nW}(X - V)}{g_{nW}(X - V - nU)} \\ &= d(V, W) \frac{g_{nW}\left((X - V - nU) + nU\right)}{g_{nW}(X - V - nU)} \\ &= d(V, W) e_n(nU, nW) \end{aligned} \tag{31}$$

where the last equality follows from taking $S = X - V - nU$ in the definition of e_n. □

Lemma 5.98. *For $U, V, W \in E[n^2]$,*

$$\frac{d(U,V)}{d(V,U)} = \frac{d(V,W)d(U+W,V)}{d(V,U+W)d(W,V)}.$$

Proof. From the definition of d we get:

$$\begin{aligned} g_{nU+(nV+nW)}(X) &= d(U,V+W)g_{nU}(X)g_{nV+nW}(X-U) \\ &= d(U,V+W)g_{nU}(X)d(V,W)g_{nV}(X-U)g_{nW}(X-U-V). \end{aligned} \tag{32}$$

Similarly,

$$\begin{aligned} g_{(nU+nV)+nW}(X) &= d(U+V,W)g_{nU+nV}(X)g_{nW}(X-U-V) \\ &= d(U+V,W)d(U,V)g_{nU}(X)g_{nV}(X-U)g_{nW}(X-U-V). \end{aligned} \tag{33}$$

Since

$$g_{nU+(nV+nW)} = g_{(nU+nV)+nW},$$

we can cancel common terms and obtain

$$d(U,V+W)d(V,W) = d(U+V,W)d(U,V). \tag{34}$$

Interchange the roles of U, V in Equation 34 and divide to obtain

$$\frac{d(U,V)}{d(V,U)} = \frac{d(U,V+W)d(V,W)}{d(V,U+W)d(U,W)}. \tag{35}$$

Now, swap the roles of V, W in Equation 34, solve for $d(U,W)$ and substitute in 35 to obtain the result. □

Lemma 5.99. *Let $P, Q \in E[n]$. Then*

$$e_n(P,Q) = \frac{c(P,Q)}{c(Q,P)}.$$

Proof. By Proposition 5.56, the map $[n] : E(\Bbbk) \longrightarrow E(\Bbbk)$ is surjective so we may choose $U, V \in E[n^2]$ such that $nU = P$ and $nV = Q$. The left-hand side of the formula of Lemma 5.98 does not depend on W so we may substitute $W = jU$ for $0 \le j < n$ and multiply the results to obtain:

$$\frac{c(P,Q)}{c(Q,P)} = \left(\frac{d(U,V)}{d(V,U)}\right)^n = \prod_{j=0}^{n-1} \frac{d(V,jU)d(U+jU,V)}{d(V,U+jU)d(jU,V)}. \tag{36}$$

Most terms on the right-hand side of Equation 36 cancel except those for $j = 0$ and $j = n-1$ so that

$$\frac{c(P,Q)}{c(Q,P)} = \frac{d(V,\mathcal{O})d(nU,V)}{d(V,nU)d(\mathcal{O},V)}.$$

In the first equation of Lemma 5.97 we substitute $W = \mathcal{O}$ to obtain $d(V, nU) = d(V, \mathcal{O})$. In the second equation of Lemma 5.97 we substitute $V = \mathcal{O}$ and $W = V$ to get

$$d(nU, V) = d(\mathcal{O}, V)e_n(nU, nV) = d(\mathcal{O}, V)e_n(P, Q),$$

and the result follows. □

We are ready for the

Proof of Theorem 5.95. Lemma 5.99 and the definition of c shows that

$$e(P, Q) = \frac{c(P, Q)}{c(Q, P)} = \frac{f_Q(X)f_P(X - Q)}{f_P(X)f_Q(X - P)},$$

which is independent of X.

Let $D_P = [P] - [\mathcal{O}]$ and $D_Q = [S] - [S - Q]$ where S is chosen such that $\mathrm{supp}(D_P) \cap \mathrm{supp}(D_Q) = \varnothing$ ie $S \notin \{P, \mathcal{O}, Q, P + Q\}$.

Let $F_P(X) = f_P(X)$ and $F_Q(X) = \frac{1}{f_Q(S-X)}$.

Then

$$\mathrm{div}(F_P) = n[P] - n[\mathcal{O}] = nD_P$$

and

$$\mathrm{div}(F_Q) = n[S] - n[S - Q] = nD_Q.$$

We therefore have

$$e_n(P, Q) = \frac{F_Q(D_P)}{F_P(D_Q)} = e_n^{\mathrm{gen}}(P, Q)$$

and this completes the proof.

□

Corollary 5.100. *In the setup of Construction 5.83, for any $P, Q \in E[n]$,*

$$e_n(P, Q) = e_n^{\mathrm{eff}}(P, Q) = e_n^{\mathrm{gen}}(P, Q).$$

Let us finish this section with the proof of the last property of Weil pairing.

Proof of Theorem 5.85(6) . In light of Corollary 5.100, let us write e_n for e_n^{gen} in this proof. We have $mnP = \mathcal{O}$ and $nQ = \mathcal{O}$. Let

$$\begin{aligned} f_1 &: \ \mathrm{div}(f_1) = mn([P] - [\mathcal{O}]), \\ f_2 &: \ \mathrm{div}(f_2) = n([Q + T] - [T]), \\ f_3 &: \ \mathrm{div}(f_3) = n([mP] - [\mathcal{O}]), \end{aligned} \tag{37}$$

where $T \notin \{P, -Q, P - Q, \mathcal{O}, mP, mP - Q\}$. Then, for $i \neq j$, $\mathrm{div}(f_i), \mathrm{div}(f_j)$ have disjoint support. Let $D_P = [P] - [\mathcal{O}]$ and $D_Q = [Q + T] - [T]$. Then D_P, D_Q have disjoint support and

$$\mathrm{div}\, f_1 = (mn)D_P$$

and

$$\operatorname{div}(f_2^m) = (mn)D_Q.$$

Thus,

$$e_{mn}(P,Q) = \frac{f_1([Q+T]-[T])}{f_2^m([P]-[\mathcal{O}])}.$$

Similarly, let $D_P = [mP]-[\mathcal{O}]$ and $D_Q = [Q+T]-[T]$. Then, by the choice of T, D_P, D_Q have disjoint support. Moreover

$$\operatorname{div} f_3 = nD_P$$

and

$$\operatorname{div} f_2 = nD_Q.$$

Thus,

$$e_n(mP,Q) = \frac{f_3([Q+T]-[T])}{f_2([mP]-[\mathcal{O}])}.$$

Observe that

$$\begin{aligned} \operatorname{div}(f_3) &= n([mP]-[\mathcal{O}]) \\ &= n([mP]+(m-1)[\mathcal{O}]-m[P])+mn([P]-[\mathcal{O}]) \\ &= \operatorname{div}(f_4^n \cdot f_1) \end{aligned} \tag{38}$$

where $\operatorname{div}(f_4) = [mP]+(m-1)[\mathcal{O}]-m[P]$. Thus,

$$\begin{aligned} e_{mn}(P,Q) &= \frac{f_3 f_4^{-n}([Q+T]-[T])}{f_2^m([P]-[\mathcal{O}])} \\ &= \frac{f_3([Q+T]-[T])f_4(-\operatorname{div}(f_2))}{f_2^m([P]-[\mathcal{O}])} \\ &= \frac{f_3([Q+T]-[T])}{f_2(\operatorname{div}(f_4)+m([P]-[\mathcal{O}]))} \\ &= \frac{f_3([Q+T]-[T])}{f_2([mP]-[\mathcal{O}])} \\ &= e_n(mP,Q). \end{aligned} \tag{39}$$

□

5.10 Miller's algorithm

In order to use the Weil pairing e_n in practice, we need an efficient algorithm that given a point $P \in E(\Bbbk)[n]$, allows us to evaluate a rational function f_P such that $\operatorname{div}(f_P) \sim n[P]-n[\mathcal{O}]$ at various points. The main algorithm for such a calculation is called **Miller's algorithm** and originally appeared in [Mil].

Henceforth, we work in the setup of the Weil pairing appearing in Construction 5.83: $E/\mathbb{F}_p$ an Elliptic curve, $1 \leq n$ s.th. $p \nmid n$ and $\mathbb{F}_p \subseteq \Bbbk$ a finite field extension such that

$$E(\Bbbk)[n] = E[n] = \mathbb{Z}_n \times \mathbb{Z}_n.$$

Suppose $U, V \in E(\Bbbk)$.

Let $\mathcal{L}_{U,V}$ be the **(horizontal) line** through U, V. For $U \neq V$ and $U, V \neq \mathcal{O}$,

$$\operatorname{div} \mathcal{L}_{U,V} = [U] + [V] + [-(U+V)] - 3[\mathcal{O}].$$

For $U \neq V$ and $V = \mathcal{O}$,

$$\operatorname{div} \mathcal{L}_{U,V} = [U] + [-U] - 2[U].$$

Let $\mathcal{T}_U$ be the **tangent line** through U. Then

$$\operatorname{div} \mathcal{T}_U = 2[U] + [-2U] - 3[\mathcal{O}].$$

Let $\mathcal{V}_U$ be the **vertical line** through U. Then

$$\operatorname{div} \mathcal{V}_U = [U] + [-U] - 2[\mathcal{O}].$$

Observation 5.101. For any $U \in E(\Bbbk)$, $\mathcal{T}_U = \mathcal{L}_{U,U}$ and $\mathcal{V}_U = \mathcal{L}_{U,-U}$.

Let $P, Q \in E[n]$ and $R, S \in E(\Bbbk)$ such that

$$S \notin \{R, P+R, P+R-Q, R-Q\}.$$

Consider the divisors

$$D_P = [P+R] - [R]$$

and

$$D_Q = [Q+S] - [S]$$

that have disjoint support by the choice of R, S. Clearly, $D_P \sim [P] - [\mathcal{O}]$ and $D_Q \sim [Q] - [\mathcal{O}]$. Let

$$f_P \text{ monic} : \operatorname{div} f_P = n[P+R] - n[R],$$

$$f_Q \text{ monic} : \operatorname{div} f_Q = n[Q+S] - [S]$$

Then $\operatorname{div} f_P = nD_P$ and $\operatorname{div} f_Q = nD_Q$ so by Definition 5.92 and Theorem 5.88 we can write

$$e_n(P,Q) = \frac{f_P(D_Q)}{f_Q(D_P)} = \frac{f_P(Q+S)}{f_P(S)} \Big/ \frac{f_Q(P+R)}{f_Q(R)}.$$

Henceforth, we aim to describe a way to evaluate $f_P(X)$ for various $X \in E(\Bbbk)$. One can then use this to evaluate $f_Q(Y)$ for various $Y \in E(\Bbbk)$ in a similar way.

For $k \in \mathbb{Z}$, let

$$f_k \text{ monic} : \operatorname{div} f_k = k[P+R] - k[R] - [kP] + [\mathcal{O}]$$

so that $f_n = f_P$.

We now have the following

Lemma 5.102. *For $a, b \in \mathbb{Z}$,*

$$\operatorname{div}\left(f_a \cdot f_b \cdot \frac{\mathcal{L}_{aP,bP}}{\mathcal{V}_{(a+b)P}}\right) = \operatorname{div}(f_{a+b}).$$

Proof. WLOG, $P \neq \mathcal{O}$. Then

$$\begin{aligned}
\text{LHS} = a[P+R] - a[R] - [aP] + [\mathcal{O}] \\
+b[P+R] - b[R] - [bP] + [\mathcal{O}] \\
+[aP] + [bP] + [-(a+b)P] - 3[\mathcal{O}] \\
-([(a+b)P] + [-(a+b)P] - 2[\mathcal{O}]) \\
= (a+b)[P+R] - a[R] - b[R] - [(a+b)P] + [\mathcal{O}] \\
= (a+b)[P+R] - (a+b)[R] - [(a+b)P] + [\mathcal{O}] \\
= \text{RHS}.
\end{aligned} \tag{40}$$

□

Corollary 5.103.

$$\operatorname{div}\left(f_k^2 \cdot \frac{\mathcal{T}_{kP}}{\mathcal{V}_{2kP}}\right) = \operatorname{div}(f_{2k}).$$

Furthermore,

Lemma 5.104. *We have*

$$\operatorname{div}(f_1) = \operatorname{div}\left(\frac{\mathcal{V}_{P+R}}{\mathcal{L}_{P,R}}\right).$$

Proof.

$$\text{LHS} = [P+R] - [R] - [P] + [\mathcal{O}] \tag{41}$$

whereas

$$\begin{aligned}
\text{RHS} = [P+R] + [-(P+R)] - 2[\mathcal{O}] - ([P] + [R] + [-(P+R)] - 3[\mathcal{O}]) \\
= \text{LHS}.
\end{aligned} \tag{42}$$

□

To describe Miller's Algorithm, recall that

$$f_P : \operatorname{div}(f_P) = n[P+R] - n[R]$$

and let us write

$$n = \sum_{i=0}^{t} 2^i \cdot n_i$$

where $n_i \in \{0, 1\}$.

Algorithm 1 Miller's algorithm to compute $f_P(Q)$. Output $x = f_P(Q)$.

1: $x_1 := \frac{\mathcal{V}_{P+R}(Q)}{\mathcal{L}_{P,R}(Q)}$.
2: $x := x_1$.
3: $Z := P$.
4: **for** $i := t-1, ..., 0$ **do**
5: $\quad x := x^2 \cdot \frac{\mathcal{T}_Z(Q)}{\mathcal{V}_{2Z}(Q)}$.
6: $\quad Z := 2Z$.
7: $\quad$ **if** $n_i = 1$ **then**
8: $\quad\quad x := x \cdot x_1 \cdot \frac{\mathcal{L}_{Z,P}(Q)}{\mathcal{V}_{Z+P}(Q)}$.
9: $\quad\quad Z := Z + P$.
10: $\quad$ **end if**
11: **end for**

Proposition 5.105. *Miller's algorithm has computational complexity of* $\mathcal{O}(\log n)$ *points addition. On termination, we have* $Z = nP = \mathcal{O}$ *and* $x = f_P(Q)$.

Proof. The complexity statement follows from the fact that we run a loop of order $\log n$, in which a constant number of points addition needs to be calculated. The rest of the computation admits closed formulas hence can be done in constant time. It is trivial to check the Algorithm works when $t = 0$, so suppose $t \geq 1$. By Lemma 5.104, the pre-for loop setup is:

$$\begin{aligned} x_1 &= f_1(Q), \\ x &= f_1(Q), \\ Z &= P. \end{aligned} \tag{43}$$

For $i = t-1$, we have

$$x = f_1^2(Q) \cdot \frac{\mathcal{T}_P(Q)}{\mathcal{V}_{2P}(Q)} = f_2(Q)$$

and we then set $Z = 2P$. If $n_{t-1} = 1$ we get by Step 8 & 9:

$$x = f_1(Q) \cdot f_2(Q) \cdot \frac{\mathcal{L}_{2P,P}(Q)}{\mathcal{V}_{3P}(Q)} = f_3(Q)$$

(where the last equality holds by Lemma 5.102) and $Z = 3P$. The proof proceeds by a straightforward induction on t. □

5.11 The Tate Pairing

In this section we define the Tate pairing and prove it's basic properties. Tate pairing is an example of an **asymmetric** bilinear pairing and its implementation is considered to have lower computational complexity.

As we will see below, the Tate pairing can be expressed in terms of the Weil pairing. Thus, the effort we spent on proving the properties of the Weil pairing helps in proving the properties of the Tate pairing.

Our setup throughout this section is identical to that of the Weil pairing, but we repeat it for convenience. Let $E/\mathbb{F}_p$ be an Elliptic curve defined over $\mathbb{F}_p$ and $1 \leq n$ an integer such that $p \nmid n$. We let $\mathbb{F}_p \subseteq \Bbbk$ be finite field extension such that $E(\Bbbk)[n] = E(\overline{\Bbbk})[n] \cong \mathbb{Z}_n \times \mathbb{Z}_n$. Recall from Theorem 5.85 that this setup implies that the group of nth roots of unity μ_n satisfies $\mu_n \subseteq \Bbbk = \mathbb{F}_{p^k}$.

Our setup implies, in particular, $\Bbbk = \mathbb{F}_{p^k}$ for some k. We start with a general result that expresses μ_n as a quotient of $\Bbbk^\times$.

Lemma 5.106. *Let $1 \leq n$ be an integer and suppose k is such that $\mu_n \subseteq \mathbb{F}_{p^k}$. Then*

$$\mathbb{F}_{p^k}^\times / \left(\mathbb{F}_{p^k}^\times\right)^n \cong \mu_n.$$

Proof. Define a map

$$\Phi : \mathbb{F}_{p^k}^\times / \left(\mathbb{F}_{p^k}^\times\right)^n \longrightarrow \mu_n$$

by $\Phi(\gamma) = \gamma^{\frac{p^k-1}{n}}$. The map Φ has range μ_n since $\Phi(\gamma)^n = \gamma^{p^k-1} = 1$ (given that $\mathbb{F}_{p^k}^\times$ is a cyclic group of order $p^k - 1$) and one can easily check it is a group homomorphism. Recall that $\left(\mathbb{F}_{p^k}^\times\right)^n = \{\alpha^n | \alpha \in \mathbb{F}_{p^k}^\times\}$ is a subgroup of $\mathbb{F}_{p^k}^\times$ since $\alpha^n\beta^n = (\alpha\beta)^n$.

Let us evaluate the order of $\left(\mathbb{F}_{p^k}^\times\right)^n$. If $\alpha, \beta \in \mathbb{F}_{p^k}^\times$, then

$$\alpha^n = \beta^n \iff \left(\frac{\alpha}{\beta}\right)^n = 1 \iff \frac{\alpha}{\beta} \in \mu_n.$$

Since $\mu_n \subseteq \mathbb{F}_{p^k}^\times$, it follows that $|\left(\mathbb{F}_{p^k}^\times\right)^n| = \frac{p^k-1}{n}| = \frac{p^k-1}{n}$. By the first Isomorphism theorem (Theorem 3.30), it follows that

$$|\mathrm{Im}\,(\Phi)| = |\mathbb{F}_{p^k}^\times| / |\left(\mathbb{F}_{p^k}^\times\right)^n| = p^k - 1 / \frac{p^k-1}{n} = n = |\mu_n|$$

and thus Φ is an isomorphism □

We are ready to phrase

Construction 5.107. Let $E/\mathbb{F}_p$ be an Elliptic curve, n a prime with $n \mid \#E(\mathbb{F}_p)$ and k the minimal integer such that $n \mid p^k - 1$. Note that by Theorem 6.1, our assumptions mean that $E[n] \subseteq E(\mathbb{F}_{p^k})$. Let $P \in E[n]$ and $Q \in E(\mathbb{F}_{p^k})$.

Let D_P, D_Q be divisors with disjoint support such that

$$D_P \sim [P] - [\mathcal{O}]$$
$$D_Q \sim [Q] - [\mathcal{O}],$$

and let f_{nD_P} be a function such that

$$\mathrm{div}(f_{nD_P}) = nD_P.$$

We define the (reduced) **Tate pairing** to be

$$\tau_n(P,Q) = \left(f_{nD_P}(D_Q)\right)^{\frac{p^k-1}{n}}.$$

Theorem 5.108. *The Tate pairing of Construction 5.107 defines a bilinear non-degenerate pairing*

$$\tau_n : E[n] \times E(\mathbb{F}_{p^k})/nE(\mathbb{F}_{p^k}) \longrightarrow \mu_n.$$

Proof. We first show that $\tau_n(P,Q)$ does not depend on the choices of D_P and D_Q (since we are not in a symmetric setting, both cases require a proof).

Let $D'_P = D_P + \operatorname{div}(g)$. If f_{nD_P} is a rational function corresponding to nD_P then

$$\operatorname{div}(f_{nD_P} g^n) = nD_P + n\operatorname{div}(g) = nD'_P$$

and thus:

$$\begin{aligned} f_{nD'_P}(D_Q)^{\frac{p^k-1}{n}} &= f_{nD_P}(D_Q)^{\frac{p^k-1}{n}} \cdot g(D_Q)^{p^k-1} \\ &= \left(f_{nD_P}(D_Q)\right)^{\frac{p^k-1}{n}}. \end{aligned}$$

Similarly, if $D'_Q = D_Q + \operatorname{div}(h)$ s.th. D_P and $\operatorname{div}(h)$ have disjoint support, then

$$f_{nD_P}(\operatorname{div} h) = h(\operatorname{div} f_{nD_P}) = h([P] - [\mathcal{O}])^n$$

and thus

$$\begin{aligned} f_{nD_P}(D'_Q)^{\frac{p^k-1}{n}} &= f_{nD_P}(D_Q)^{\frac{p^k-1}{n}} f_{nD_P}(\operatorname{div} h)^{\frac{p^k-1}{n}} \\ = f_{nD_P}(D_Q)^{\frac{p^k-1}{n}} \cdot h([P] - [\mathcal{O}])^{p^k-1} &= f_{nD_P}(D_Q)^{\frac{p^k-1}{n}}. \end{aligned}$$

Note that a-priori, Construction 5.107 only defines a map

$$\tau_n : E[n] \times E(\mathbb{F}_{p^k}) \longrightarrow \mu_n.$$

To prove that we have a map

$$\tau_n : E[n] \times E(\mathbb{F}_{p^k})/nE(\mathbb{F}_{p^k}) \longrightarrow \mu_n,$$

fix $P \in E[n]$ and consider the map

$$\tau_n(P,-) : E(\mathbb{F}_{p^k}) \longrightarrow \mu_n$$

evaluated at $Q + nR$ for $Q, R \in E(\mathbb{F}_{p^k})$. Choose a divisor $D_P \sim [P] - [\mathcal{O}]$ whose support is disjoint from $\{\mathcal{O}, R, nR, Q\}$. Let $D_Q = [Q] - [\mathcal{O}]$ and $D_{Q+nR} = [Q + nR] - [\mathcal{O}]$. Then

$$D_{Q+nR} \sim D_Q + n[R] - n[\mathcal{O}].$$

Since τ_n is independent of the divisor class of D_{Q+nR} we get:

$$\begin{aligned}
\tau_n(P, Q+nR) &= f_{nD_P}(D_{Q+nR})^{\frac{p^k-1}{n}} \\
&= f_{nD_P}(D_Q + n[R] - n[\mathcal{O}])^{\frac{p^k-1}{n}} \\
&= f_{nD_P}(D_Q)^{\frac{p^k-1}{n}} \cdot f_{nD_P}(n[R] - n[\mathcal{O}])^{\frac{p^k-1}{n}} \\
&= f_{nD_P}(D_Q)^{\frac{p^k-1}{n}} \cdot f_{nD_P}([R] - [\mathcal{O}])^{p^k-1} \\
&= f_{nD_P}(D_Q)^{\frac{p^k-1}{n}} = \tau_n(P, Q)
\end{aligned}$$

as required.

To prove that τ_n is a non-degenerate bilinear pairing, let us give an alternative definition of τ_n.

Let $P \in E[n]$ and $Q \in E(\mathbb{F}_{p^k})$. Choose $Q_0 \in E(\overline{\mathbb{F}}_{p^k})$ s.th. $Q = nQ_0$. Let

$$\Phi : E(\overline{\mathbb{F}}_p) \longrightarrow E(\overline{\mathbb{F}}_p)$$

be the Frobenius map over $\mathbb{F}_p$ (given by $\Phi(x, y) = (x^p, y^p)$). Then

$$\Phi^k : E(\overline{\mathbb{F}}_{p^k}) \longrightarrow E(\overline{\mathbb{F}}_{p^k})$$

is the Frobenius map over $\mathbb{F}_{p^k}$. Denote $Q_1 = (\Phi^k - 1)(Q_0) = \Phi^k Q_0 - Q_0$. Then

$$nQ_1 = \Phi^k(nQ_0) - nQ_0 = \Phi^k Q - Q = 0$$

where the last equality follows since Φ^k acts as identity on points $Q \in E(\mathbb{F}_{p^k})$. Thus $Q_1 \in E[n]$.

Observe – Q_1 does not depend on the choice of Q_0: if $T \in E[n]$ then $n(Q_0 + T) = Q_1$ and

$$(\Phi^k - 1)(Q_0 + T) = Q_1 + (\Phi^k - 1)(T) = Q_1,$$

where the last equality holds since $T \in E[n] \subseteq E(\mathbb{F}_{p^k})$.

We thus get a well-defined map

$$\frac{\Phi^k - 1}{n} : E(\mathbb{F}_{p^k}) \longrightarrow E[n] \subseteq E(\mathbb{F}_{p^k})$$

given by

$$Q \mapsto (\frac{\Phi^k - 1}{n})(Q) = (\frac{\Phi^k - 1}{n})(nQ_0) = Q_1.$$

Furthermore, it is easy to check that $\frac{\Phi^k-1}{n}$ is an endomorphism of $E(\mathbb{F}_{p^k})$.

Recall from Construction 5.83 that

$$e_n(P, Q_1) = \frac{g_P(S + Q_1)}{g_P(S)}$$

for an arbitrary S. Taking $S = Q_0$, we get

$$e_n(P, Q_1) = e_n(P, \Phi^k Q_0 - Q_0) = \frac{g_P(\Phi^k Q_0)}{g_P(Q_0)}.$$

Now, $P \in E[n] \subseteq E(\mathbb{F}_{p^k})$ so that $g_P \in \mathbb{F}_{p^k}(E)$.
By Corollary 5.59, we get:

$$\begin{aligned} e_n(P, Q_1) &= \frac{g_P(\Phi^k Q_0)}{g_P(Q_0)} \\ &= \frac{\Phi^k(g_P(Q_0))}{g_P(Q_0)} = g_P(Q_0)^{p^k - 1} \\ = (g_P^n(Q_0))^{\frac{p^k-1}{n}} &= \left(f_{n[P]-n[\mathcal{O}]}(nQ_0)\right)^{\frac{p^k-1}{n}} \\ &= f_{nD_P}(Q)^{\frac{p^k-1}{n}} = \tau_n(P, Q) \end{aligned} \tag{44}$$

where $D_P = [P] - [\mathcal{O}]$.

Using equation 44 we deduce bilinearity of the Tate pairing from bilinearity of the Weil pairing.

For non-degeneracy, it is enough to show that

$$\frac{\Phi^k - 1}{n} : E(\mathbb{F}_{p^k}) \longrightarrow E[n]$$

is surjective since we could then use non-degeneracy of the Weil pairing. We have

$$\ker\left(\frac{\Phi^k - 1}{n}\right) = nE(\mathbb{F}_{p^k})$$

and so by the first isomorphism theorem 3.30,

$$\mathrm{Im}\left(\frac{\Phi^k - 1}{n} : E(\mathbb{F}_{p^k}) \longrightarrow E[n]\right) \cong E(\mathbb{F}_{p^k})/nE(\mathbb{F}_{p^k}).$$

Now, by Corollary 5.78 , $E(\mathbb{F}_{p^k}) \cong \mathbb{Z}_a \times \mathbb{Z}_b$ for some $a \mid b$ and since $\mathbb{Z}_n \times \mathbb{Z}_n \leq E(\mathbb{F}_{p^k}) \cong \mathbb{Z}_a \times \mathbb{Z}_b$, we deduce from [Toth, Theorem 4.5] that $n \mid a$ and $n \mid b$.

Write $a = \alpha n$ and $b = \beta n$. Then the map

$$\varphi : \mathbb{Z}_a \times \mathbb{Z}_b \longrightarrow \mathbb{Z}_n \times \mathbb{Z}_n$$

given by

$$\varphi(x, y) = (x \bmod n, y \bmod n)$$

is a homomorphism whose kernel satisfies

$$\ker\varphi = \{(x, y) | n \mid x \wedge n \mid y\} = n(\mathbb{Z}_a \times \mathbb{Z}_b).$$

Since φ is clearly surjective, we deduce from the first isomorphism theorem that

$$E(\mathbb{F}_{p^k})/nE(\mathbb{F}_{p^k}) \cong \mathbb{Z}_a \times \mathbb{Z}_b/n(\mathbb{Z}_a \times \mathbb{Z}_b) \cong \mathbb{Z}_n \times \mathbb{Z}_n.$$

Since $E[n] \cong \mathbb{Z}_n \times \mathbb{Z}_n$, we deduce that the image of $\frac{\Phi^k-1}{n}$ is $E[n]$, ie that it is surjective.

□

6 Pairing-Friendly Curves

6.1 Embedding degree

Theorem 6.1 (Balasubramanian-Koblitz)**.** *[BK] Let $E/\mathbb{F}_p$ be an Elliptic curve and suppose that $E(\mathbb{F}_p)$ has a subgroup $G = \langle P \rangle$ of order n with $\gcd(n, p-1) = 1$. Then $E[n] \subseteq E(\mathbb{F}_{p^k})$ iff $n \mid p^k - 1$.*

Definition 6.2. Let $E/\mathbb{F}_p$ be an Elliptic curve and suppose that $E(\mathbb{F}_p)$ has a subgroup $G = \langle P \rangle$ of order r. The **embedding degree** of G is the smallest integer k such that $E[r] \subseteq E(\mathbb{F}_{p^k})$.

Remark 6.3. If $\gcd(r, p-1) = 1$ then the embedding degree is simply the smallest k such that $r \mid p^k - 1$.

References

[BK] Balasubramanian, R. and Koblitz, N., 1998. The improbability that an elliptic curve has subexponential discrete log problem under the Menezes—Okamoto—Vanstone algorithm. Journal of cryptology, 11(2), pp.141-145.

[FR] Fine, B. and Rosenberger, G., 1997. The fundamental theorem of algebra. Springer Science & Business Media.

[Gol] Golan, J.S., 2013. Foundations of linear algebra (Vol. 11). Springer Science & Business Media.

[Lyn] Lynn, B., 2007. On the implementation of pairing-based cryptosystems (Doctoral dissertation, Stanford University).

[Mil] Miller, V.S., 2004. The Weil pairing, and its efficient calculation. Journal of cryptology, 17(4), pp.235-261.

[Sil] Silverman, J.H., 2009. The arithmetic of elliptic curves (Vol. 106, pp. xx+-513). New York: Springer.

[Toth] Tóth, L., 2013. Subgroups of finite abelian groups having rank two via Goursat's lemma. arXiv preprint arXiv:1312.1485.

[Wash] Washington, L.C., 2008. Elliptic curves: number theory and cryptography. Chapman and Hall/CRC.

List of symbols

$E(\Bbbk)$	points on an Elliptic curve over a field $\Bbbk$
$E/\Bbbk$	Elliptic curve defined over a field $\Bbbk$
$E[n]$	n torsion points
G, H	group
$\overline{\Bbbk}, \overline{\mathbb{F}}$	algebraic closure over a field
$\circ$	composition
$\varnothing$	empty set
$\mathbb{C}$	complex numbers
$\mathbb{F}, \Bbbk$	field
$\mathbb{F}[x]$, $\Bbbk[x]$	polynomials over a field with one indeterminate
$\mathbb{F}_{p^n}$	field with p^n elements
$\mathbb{N}$	natural numbers
$\mathbb{Q}$	rational numbers
$\mathbb{R}$	real numbers
$\mathbb{Z}_n$	additive group of integers modulo n
$\mathbb{Z}$	integers
$\mathcal{O}$	point at infinity
$\mu_n(\mathbb{F})$	nth roots of unity over a field $\mathbb{F}$
$\prod$	product over an indexed set
$\sim$	equivalence relation
$\sum$	sum over an indexed set
τ_n	Tate pairing
div	divisor
e_n	Weil pairing