Finite Fields

Sophie Huczynska (with changes by Max Neunhöffer)

Semester 2, Academic Year 2012/13

Contents

Chapter 1

Introduction

Finite fields is a branch of mathematics which has come to the fore in the last 50 years due to its numerous applications, from combinatorics to coding theory. In this course, we will study the properties of finite fields, and gain experience in working with them.

In the first two chapters, we explore the theory of fields in general. Throughout, we emphasize results particularly important to finite fields, but allow fields to be arbitrary unless otherwise stated.

1 Group theory: a brief summary

We begin by recalling the definition of a group.

Definition 1.1
A *group* is a set G, together with a binary operation $*$, such that the following axioms hold:

Closure: G is closed under the operation $*$: $x, y \in G \Longrightarrow x * y \in G$;

Associativity: $(x * y) * z = x * (y * z)$ for all $x, y, z \in G$;

Identity: there exists an element $e \in G$ (called the identity of G) such that $x * e = e * x = x$ for all $x \in G$;

Inverses: for every element $x \in G$ there exists an element $x^{-1} \in G$ (called the inverse of x) such that $x * x^{-1} = x^{-1} * x = e$.

Note: We often write $\cdot$ instead of $*$ or leave it out completely.

Definition 1.2
A group G is said to be *abelian* if the binary operation $*$ is commutative, i.e. if $x * y = y * x$ for all $x, y \in G$. The operation $*$ is often replaced by $+$ for abelian groups, i.e. $x * y$ is written $x + y$
We then say that the group is "written additively" (as opposed to being "written multiplicatively").

Example 1.3
The following are examples of groups:

- The set $\mathbb{Z}$ of integers with the operation of addition (this is abelian);
- the set $\mathrm{GL}_n(\mathbb{R})$ of invertible $(n \times n)$-matrices with real entries, with the operation of matrix multiplication, forms a group (for $n > 1$ this is not abelian).
- Let G be the set of remainders of all the integers on division by n, e.g. $G = \{0, 1, \ldots, n-1\}$. Let $a * b$ be the operation of taking the integer sum $a + b$ and reducing it modulo n. Then $(G, *)$ is a group (this is abelian).

Definition 1.4
A multiplicative group G is said to be *cyclic* if there is an element $a \in G$ such that for any $b \in G$ there is some integer j with $b = a^j$. Such an element is called a generator of the cyclic group, and we write $G = \langle a \rangle$.

Note we may have more than one generator, e.g. either 1 or -1 can be used to generate the additive group $\mathbb{Z}$.

Definition 1.5
For a set S, a subset R of $S \times S$ is called an *equivalence relation* on S if it satisfies:

- $(s, s) \in R$ for all $s \in S$ (reflexive)
- If $(s, t) \in R$ then $(t, s) \in R$ (symmetric)
- If $(s, t), (t, u) \in R$ then $(s, u) \in R$ (transitive).

An equivalence relation R on S induces a *partition* of S. If we collect all elements of S equivalent to a fixed $s \in S$, we obtain the *equivalence class* of s, denoted by

$$[s] = \{t \in S : (s, t) \in R\}.$$

The collection of all equivalence classes forms a partition of S, and $[s] = [t] \Leftrightarrow (s, t) \in R$.

Definition 1.6
For arbitrary integers a,b and positive integer n, we say that a is *congruent* to b modulo n if the difference $a - b$ is a multiple of n, i.e. $a = b + kn$ for some integer k. We write $a \equiv b \bmod n$ for this.

It is easily checked that "congruence modulo n" is an equivalence relation on the set $\mathbb{Z}$ of integers. Consider the equivalence classes into which the relation partitions $\mathbb{Z}$. These are the sets:

$$\begin{aligned} [a] &= \{m \in \mathbb{Z} : m \equiv a \bmod n\} \\ &= \{m \in \mathbb{Z} : m = a + kn \text{ for some } k \in \mathbb{Z}\}. \end{aligned}$$

E.g. for $n = 4$ we have:
$[0] = \{\ldots, -8, -4, 0, 4, 8, \ldots\}$;
$[1] = \{\ldots, -7, -3, 1, 5, 9, \ldots\}$;
$[2] = \{\ldots, -6, -2, 2, 6, 10, \ldots\}$;
$[3] = \{\ldots, -5, -1, 3, 7, 11, \ldots\}$.
We may define on the set $\{[0], [1], \ldots, [n-1]\}$ a binary operation, which we shall write as $+$ (though it is not ordinary addition) by

$$[a] + [b] := [a + b]$$

where a and b are any elements of the sets $[a]$ and $[b]$ respectively, and $a + b$ is the ordinary sum of a and b. Can show (exercise) that this is well-defined, i.e. does not depend on choice of representatives.

Theorem 1.7
Let $n \in \mathbb{N}$. The set $\{[0], [1], \ldots, [n-1]\}$ of equivalence classes modulo n forms a group under the operation $+$ given by $[a] + [b] := [a + b]$. It is called the group of integers modulo n and is denoted $\mathbb{Z}_n$. It is cyclic with $[1]$ as a generator.

Proof. See Exercise Sheet 1. ■

Definition 1.8
A group is called *finite* (respectively, *infinite*) if it contains finitely (respectively, infinitely) many elements. The number of elements of a finite group G is called its order, written $|G|$.

Definition 1.9
A subset H of the group G is a *subgroup* of G if H is itself a group with respect to the operation of G, this is written $H \leq G$. The (cyclic) subgroup consisting of all powers of some element $a \in G$ is denoted $\langle a \rangle$ and called the subgroup *generated by* a. If $|\langle a \rangle|$ is finite, it is called the *order of* a, it is the smallest natural number i such that $a^i = e$.

Next, we generalize the notion of congruence, as follows.

Theorem 1.10
If H is a subgroup of G, then the relation R_H on G defined by $(a, b) \in R_H$ if and only if $a = bh$ for some $h \in H$ (additively, $a = b + h$ for some $h \in H$) is an equivalence relation. The relation is called left congruence modulo H.

The equivalence classes are called the *left cosets of H in G*; each has size $|H|$. Right congruence and right cosets are defined analogously.

Note that when $(G, *) = (\mathbb{Z}, +)$ and $H = \langle n \rangle$, we get back our previous definition of congruence, since $a \equiv b \bmod n \Leftrightarrow a = b + h$ for some $h \in \langle n \rangle$.

Definition 1.11
The *index* of H in G (denoted by $[G : H]$) is the number of left cosets of H in G, and is equal to the number of right cosets of H in G.

Theorem 1.12
The order of a finite group G is equal to the product of the order of any subgroup H and the index of H in G. In particular, the order of H divides the order of G and the order of any element $a \in G$ divides the order of G.

Proof. Exercise ■

We can easily describe subgroups and orders for cyclic groups. In what follows, ϕ is Euler's function; i.e. $\phi(n) :=$ the number of integers k with $1 \leq k \leq n$ which are relatively prime to n. If the integer n has the prime factorization $p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}$, then

$$\phi(n) = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_r}).$$

So, for example, $\phi(7) = 6$ and $\phi(30) = 2 \cdot 3 \cdot 5 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} = 8$. See Example Sheet for more details.

Theorem 1.13

(i) Every subgroup S of a cyclic group $G = \langle a \rangle$ is cyclic.

(ii) In a finite cyclic group $\langle a \rangle$ of order m, the element a^k generates a subgroup of order $\frac{m}{\gcd(k,m)}$.

(iii) For any positive divisor d of m, $\langle a \rangle$ contains precisely one subgroup of order d and precisely one subgroup of index d.

(iv) Let f be a positive divisor of m. Then $\langle a \rangle$ contains $\phi(f)$ elements of order f.

(v) A finite cyclic group $\langle a \rangle$ of order m contains $\phi(m)$ generators, namely the powers a^r with $\gcd(r, m) = 1$.

Proof.

(i) If $S = \{e\}$, then S is cyclic with generator e. Otherwise, let k be the least positive integer for which $a^k \in S$. We will show: $S = \langle a^k \rangle$. Clearly $\langle a^k \rangle \subseteq S$. Now, take an arbitrary $s \in S$, then $s = a^n$ for some $n \in \mathbb{Z}$. By the division algorithm for integers, there exist $q, r \in \mathbb{Z}$ with $0 \leq r < k$ such that $n = qk + r$. Then $a^n = a^{qk+r} = (a^k)^q \cdot a^r$, implying $a^r \in S$. If $r > 0$, this contradicts the minimality of k, so we must have $r = 0$ and hence $s = a^n = (a^k)^q \in \langle a^k \rangle$.

(ii) Set $d := \gcd(k, m)$. The order of a^k is the least positive integer n such that $a^{kn} = e$. This identity holds if and only if m divides kn, i.e. if and only if $\frac{m}{d}$ divides n. The least positive n with this property is $n = \frac{m}{d}$.

(iii) Exercise: see Exercise Sheet 1.

(iv) Let $|\langle a \rangle| = m$ and $m = df$. By (ii), the element a^k is of order f if and only if $\gcd(k, m) = d$. So the number of elements of order f is equal to the number of integers k with $1 \leq k \leq m$ and $\gcd(k, m) = d$. Equivalently, writing $k = dh$ with $1 \leq h \leq f$, the condition becomes $\gcd(h, f) = 1$. There are precisely $\phi(f)$ such h.

(v) The first part follows from (iv), since the generators of $\langle a \rangle$ are precisely the elements of order m. The second part follows from (ii).

■

Definition 1.14

- A subgroup H of G is *normal* $\Leftrightarrow$ its left and right cosets coincide.

  We write $H \triangleleft G$ in that case.

- For a normal subgroup H, the set of (left) cosets of H in G forms a group, denoted G/H. The operation is
$$(aH)(bH) := (ab)H.$$

Definition 1.15

A mapping $f : G \to H$ of the group G into the group H is called a *homomorphism* of G into H if f preserves the operation of G, i.e. $(gk)f = (gf) \cdot (kf)$ for all $g, k \in G$. If f is a bijective homomorphism it is called an *isomorphism* and we say G and H are isomorphic and write $G \cong H$. An isomorphism of G onto itself is called an *automorphism* of G.

Definition 1.16

The *kernel* of the homomorphism $f : G \to H$ of the group G into the group H is the set (actually, normal subgroup)
$$\ker(f) := \{a \in G : af = e_H\}.$$
The image of f is the set (actually, subgroup)
$$\mathrm{im}(f) := \{af : a \in G\}.$$

Theorem 1.17

[First Isomorphism Theorem] Let $f : G \to H$ be a homomorphism of groups. Then $\ker f$ is a normal subgroup of G and
$$G/\ker f \cong \mathrm{im} f \quad \textit{by the isomorphism } g \ker(f) \mapsto gf.$$

Proof. Omitted. ■

Example 1.18

Take $G := \mathbb{Z}$, $H := \mathbb{Z}_n$ and $f : a \mapsto [a]$. Then f is a homomorphism with $\ker(f) = \langle n \rangle$ and $\mathrm{im}(f) = \mathbb{Z}_n$, and so the First Isomorphism Theorem says that $\mathbb{Z}/\langle n \rangle$ and $\mathbb{Z}_n$ are isomorphic as groups.

2 Rings and fields

Definition 2.1
A *ring* $(R, +, *)$ is a set R, together with two binary operations, denoted by $+$ and $*$, such that

- R is an *abelian group* with respect to $+$;
- R is closed under $*$;
- $*$ is *associative*, that is $(a * b) * c = a * (b * c)$ for all $a, b, c \in R$;
- the *distributive laws* hold, that is, for all $a, b, c \in R$ we have $a * (b + c) = (a * b) + (a * c)$ and $(b + c) * a = (b * a) + (c * a)$.

Typically, we use 0 to denote the identity element of the abelian group R with respect to addition, and $-a$ to denote the additive inverse of $a \in R$.

Definition 2.2
- A ring is called a *ring with identity* if the ring has a multiplicative identity (usually denoted e or 1).
- A ring is called *commutative* if $*$ is commutative.
- A ring is called an *integral domain* if it is a commutative ring with identity $e \neq 0$ in which $ab = 0$ implies $a = 0$ or $b = 0$ (i.e. *no zero divisors*).
- A ring is called a *division ring* (or skew field) if the non-zero elements form a group under $*$.
- A commutative division ring is called a *field*.

Example 2.3
- the integers $(\mathbb{Z}, +, *)$ form an integral domain but not a field;
- the rationals $(\mathbb{Q}, +, *)$, reals $(\mathbb{R}, +, *)$ and complex numbers $(\mathbb{C}, +, *)$ form fields;
- the set of 2×2 matrices with real entries forms a non-commutative ring with identity w.r.t. matrix addition and multiplication.
- the group $\mathbb{Z}_n$ with addition as before and multiplication defined by $[a][b] := [ab]$ is a commutative ring with identity $[1]$.

So, in summary: a field is a set F on which two binary operations, called addition and multiplication, are defined, and which contains two distinguished elements e and 0 with $0 \neq e$. Moreover, F is an abelian group with respect to addition, having 0 as the identity element, and the non-zero elements of F (often written F^*) form an abelian group with respect to multiplication having e as the identity element. The two operations are linked by the distributive laws.

Theorem 2.4
Every finite integral domain is a field.

Proof. Let R be a finite integral domain, and let its elements be $r_1, r_2, \ldots, r_n$. Consider a fixed non-zero element $r \in R$. Then the products $rr_1, rr_2, \ldots, rr_n$ must be distinct, since $rr_i = rr_j$ implies $r(r_i - r_j) = 0$, and since $r \neq 0$ we must have $r_i - r_j = 0$, i.e. $r_i = r_j$. Thus, these products are precisely the n elements of R. Each element of R is of the form rr_i; in particular, the identity $e = rr_i$ for some $1 \leq i \leq n$. Since R is commutative, we also have $r_i r = e$, and so r_i is the multiplicative inverse of r. Thus the non-zero elements of R form a commutative group, and R is a field. ■

Definition 2.5

- A subset S of a ring R is called a *subring* of R if S is closed under $+$ and $*$ and forms a ring under these operations.
- A subset J of a ring R is called an *ideal* if J is a subring of R and for all $a \in J$ and $r \in R$ we have $ar \in J$ and $ra \in J$.
- Let R be a commutative ring with an identity. Then the smallest ideal containing an element $a \in R$ is $(a) := \{ra : r \in R\}$. We call (a) the *principal ideal* generated by a.

Definition 2.6
An integral domain in which every ideal is principal is called a *principal ideal domain* (PID).

Example 2.7
$\mathbb{Z}$ is a PID.

An ideal J of R defines a partition of R into disjoint cosets (with respect to $+$), *residue classes* modulo J. These form a ring w.r.t. the following operations:

$$(a + J) + (b + J) = (a + b) + J,$$

$$(a + J)(b + J) = ab + J.$$

This ring is called the *residue class ring* and is denoted R/J.

Example 2.8
The residue class ring $\mathbb{Z}/(n)$
Here, (n) is the principal ideal generated by the integer n (same set $n\mathbb{Z}$ as the subgroup $\langle n\rangle$ but now with two operations). As in the group case, we denote the residue class of a modulo n by $[a]$, as well as by $a + (n)$. The elements of $\mathbb{Z}/(n)$ are $[0] = 0 + (n), [1] = 1 + (n), \ldots, [n - 1] = n - 1 + (n)$.

Theorem 2.9
$\mathbb{Z}/(p)$, the ring of residue classes of the integers modulo the principal ideal generated by a prime p, is a field.

Proof. By Theorem 2.4, it is enough to show that $\mathbb{Z}/(p)$ is an integral domain. Now, $[a][b] = [ab] = [0]$ if and only if $ab = kp$ for some $k \in \mathbb{Z}$. Since p is prime, p divides ab if and only if p divides one of the factors. So, either $[a] = [0]$ or $[b] = [0]$, so $\mathbb{Z}/(p)$ contains no zero divisors. ■

These are our first examples of *finite fields*!

Example 2.10
Here are the addition and multiplication tables for the field $\mathbb{Z}/(3)$:

$+$	$0+(3)$	$1+(3)$	$2+(3)$
$0+(3)$	$0+(3)$	$1+(3)$	$2+(3)$
$1+(3)$	$1+(3)$	$2+(3)$	$0+(3)$
$2+(3)$	$2+(3)$	$0+(3)$	$1+(3)$

,

$*$	$0+(3)$	$1+(3)$	$2+(3)$
$0+(3)$	$0+(3)$	$0+(3)$	$0+(3)$
$1+(3)$	$0+(3)$	$1+(3)$	$2+(3)$
$2+(3)$	$0+(3)$	$2+(3)$	$1+(3)$

.

Remark 2.11
As you will prove in Exercise Sheet 1, the above result does not hold if p is replaced by a composite n.

Definition 2.12
A mapping $\phi : R \to S$ (R,S rings) is called a *ring homomorphism* if for any $a, b \in R$ we have

$$\phi(a + b) = \phi(a) + \phi(b) \text{ and } \phi(ab) = \phi(a)\phi(b).$$

A ring homomorphism preserves both $+$ and $*$ and induces a homomorphism of the additive group of R into that of S. Concepts such as kernel and image are defined analogously to the groups case. We have a ring version of the First Isomorphism Theorem:

Theorem 2.13 (First Isomorphism Theorem for Rings)
If ϕ is a ring homomorphism from a ring R onto a ring S then the factor ring $R/\ker\phi$ and the ring S are isomorphic by the map

$$r + \ker\phi \mapsto \phi(r).$$

We can use mappings to transfer a structure from an algebraic system to a set without structure. Given a ring R, a set S and a bijective map $\phi : R \to S$, we can use ϕ to define a ring structure on S that converts ϕ into an isomorphism. Specifically, for $s_1 = \phi(r_1)$ and $s_2 = \phi(r_2)$, define

$$s_1 + s_2 \text{ to be } \phi(r_1 + r_2), \text{ and } s_1 s_2 \text{ to be } \phi(r_1)\phi(r_2).$$

This is called the ring structure *induced by* ϕ; any extra properties of R are inherited by S.

This idea allows us to obtain a more convenient representation for the finite fields $\mathbb{Z}/(p)$.

Definition 2.14
For a prime p, let $\mathbb{F}_p$ be the set $\{0, 1, \ldots, p-1\}$ of integers, and let $\phi : \mathbb{Z}/(p) \to \mathbb{F}_p$ be the mapping defined by $\phi([a]) = a$ for $a = 0, 1, \ldots, p-1$. Then $\mathbb{F}_p$ endowed with the field structure induced by ϕ is a finite field, called the *Galois field of order p*.

From above, the mapping ϕ becomes an isomorphism, so $\phi([a] + [b]) = \phi([a]) + \phi([b])$ and $\phi([a][b]) = \phi([a])\phi([b])$. The finite field $\mathbb{F}_p$ has zero element 0, identity element 1 and its structure is that of $\mathbb{Z}/(p)$. So, computing with elements of $\mathbb{F}_p$ now means ordinary arithmetic of integers with reduction modulo p.

Example 2.15

- $\mathbb{F}_2$: the elements of this field are 0 and 1. The operation tables are:

$+$	0	1
0	0	1
1	1	0

,

$*$	0	1
0	0	0
1	0	1

- We have $\mathbb{Z}/(5)$, isomorphic to $\mathbb{F}_5 = \{0, 1, 2, 3, 4\}$, where the isomorphism is given by $[0] \mapsto 0, \ldots, [4] \mapsto 4$. The operation tables are:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

,

$*$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Definition 2.16
If R is an arbitrary ring and there exists a positive integer n such that $nr = 0$ for every $r \in R$ (i.e. r added to itself n times is the zero element) then the least such positive integer n is called the *characteristic* of R, and R is said to have positive characteristic. If no such positive integer n exists, R is said to have characteristic 0.

Example 2.17

- $\mathbb{F}_2$ and $\mathbb{F}_5$ have characteristic 2 and 5 respectively.

- $\mathbb{Q}$ and $\mathbb{R}$ have characteristic 0.

Theorem 2.18
A ring $R \neq \{0\}$ of positive characteristic with an identity and no zero divisors must have prime characteristic.

Proof. Since R contains non-zero elements, R has characteristic $n \geq 2$. If n were not prime, we could write $n = km$ with $k, m \in \mathbb{Z}$, $1 < k, m < n$. Then $0 = ne = (km)e = (ke)(me)$, so either $ke = 0$ or $me = 0$, since R has no zero divisors. Hence either $kr = (ke)r = 0$ for all $r \in R$ or $mr = (me)r = 0$ for all $r \in R$, contradicting the definition of n as the characteristic. ■

Corollary 2.19
A finite field has prime characteristic.

Proof. From Theorem 2.18, we need only show that a finite field F has a positive characteristic. Consider the multiples $e, 2e, 3e, \ldots$ of the identity. Since F contains only finitely many elements, there must exist integers k and m with $1 \leq k < m$ such that $ke = me$, i.e. $(k - m)e = 0$, and thus $(k - m)f = (k - m)ef = 0f = 0$ for all $f \in F$ so F has a positive characteristic. ■

Example 2.20
The field $\mathbb{Z}/(p)$ (equivalently, $\mathbb{F}_p$) has characteristic p.

Theorem 2.21 (Freshmen's Exponentiation)
Let R be a commutative ring of prime characteristic p. then

$$(a + b)^{p^n} = a^{p^n} + b^{p^n} \text{ and } (a - b)^{p^n} = a^{p^n} - b^{p^n}$$

for $a, b \in R$ and $n \in \mathbb{N}$.

Proof. *It can be shown (see Exercise Sheet 1) that*

$$\binom{p}{i} = \frac{p(p-1)\cdots(p-i+1)}{1 \cdot 2 \cdots i} \equiv 0 \bmod p$$

for all $i \in \mathbb{Z}$ with $0 < i < p$. By the Binomial Theorem,

$$(a + b)^p = a^p + \binom{p}{1}a^{p-1}b + \cdots + \binom{p}{p-1}ab^{p-1} + b^p = a^p + b^p$$

and induction on n establishes the first identity. The second identity follows since

$$a^{p^n} = ((a - b) + b)^{p^n} = (a - b)^{p^n} + b^{p^n}.$$

■

3 Polynomials

Let R be an arbitrary ring. A *polynomial over R* is an expression of the form

$$f = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n,$$

where n is a non-negative integer, the *coefficients* a_i ($0 \leq i \leq n$) are elements of R, and x is a symbol not belonging to R, called an *indeterminate* over R.

Definition 3.1
Let $f = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + \cdots + a_n x^n$ be a polynomial over R which is not the zero polynomial, so we can suppose $a_n \neq 0$. Then n is called the *degree* of f. By convention, $\deg(0) = -\infty$. Polynomials of degree 0 are called *constant polynomials*. If the leading coefficient of f is 1 (the identity of R) then f is called a *monic* polynomial.

Given two polynomials f and g, we can write $f = \sum_{i=0}^{n} a_i x^i$ and $g = \sum_{i=0}^{n} b_i x^i$ (taking coefficients zero if necessary to ensure the same n). We define their sum to be

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

and their product to be

$$fg = \sum_{k=0}^{2n} c_k x^k, \text{ where } c_k = \sum_{i+j=k, 0 \leq i \leq n, 0 \leq j \leq n} a_i b_j.$$

Note that the degree of the product of two non-zero polynomials f and g is equal to the sum of the degrees of f and g.

Theorem 3.2
With the above operations, the set of polynomials over R forms a ring. It is called the polynomial ring over R *and denoted by $R[x]$. Its zero element is the* zero polynomial*, all of whose coefficients are zero.*

Proof. Exercise. ■

Let F denote a (not necessarily finite) field. From now on, we consider polynomials over fields.

We say that the polynomial $g \in F[x]$ *divides* $f \in F[x]$ if there exists a polynomial $h \in F[x]$ such that $f = gh$.

Theorem 3.3 (Division Algorithm)
Let $g \neq 0$ be a polynomial in $F[x]$. Then for any $f \in F[x]$, there exist polynomials $q, r \in F[x]$ such that

$$f = qg + r, \text{ where } \deg(r) < \deg(g).$$

Using the division algorithm, we can show that every ideal of $F[x]$ is principal:

Theorem 3.4
$F[x]$ is a principal ideal domain. In fact, for every ideal $J \neq (0)$ of $F[x]$ there is a uniquely determined monic polynomial $g \in F[x]$ such that $J = (g)$.

Proof. Let I be an ideal in $F[x]$. If $I = \{0\}$, then $I = (0)$. If $I \neq \{0\}$, choose a non-zero polynomial $k \in I$ of smallest degree. Let b be the leading coefficient of k, and set $m = b^{-1}k$. Then $m \in I$ and m is monic. We will show: $I = (m)$. Clearly, $(m) \subseteq I$. Now take $f \in I$; by the division algorithm there are polynomials q, r with $f = qm + r$ where either $r = 0$ or $\deg(r) < \deg(m)$. Now, $r = f - qm \in I$. If $r \neq 0$, we contradict the minimality of m; so we must have $r = 0$, i.e. f is a multiple of m and $I = (m)$.

We now show uniqueness: if $m_1 \in F[x]$ is another monic polynomial with $I = (m_1)$, then $m = c_1 m_1$ and $m_1 = c_2 m$ with $c_1, c_2 \in F[x]$. Then $m = c_1 c_2 m$, i.e. $c_1 c_2 = 1$, and so c_1, c_2 are constant polynomials. Since both m and m_1 are monic, we must have $m = m_1$. ■

We next introduce an important type of polynomial.

Definition 3.5
A polynomial $p \in F[x]$ is said to be *irreducible over F* if p has positive degree and $p = bc$ with $b, c \in F[x]$ implies that either b or c is a constant polynomial. A polynomial which does allow a non-trivial factorization over F is called *reducible over* F.

Note that the field F under consideration is all-important here, e.g. the polynomial $x^2 + 1$ is irreducible in $\mathbb{R}[x]$ but reducible in $\mathbb{C}[x]$, where it factors as $(x + i)(x - i)$.

Theorem 3.6 (Unique Factorization)
Any polynomial $f \in F[x]$ of positive degree can be written in the form

$$f = ap_1^{e_1} \dots p_k^{e_k}$$

where $a \in F$, $p_1, \dots, p_k$ are distinct monic irreducibles in $F[x]$ and $e_1, \dots, e_k$ are positive integers. This factorization is unique up to the order in which the factors occur; it is called the canonical factorization *of f in $F[x]$.*

Proof. Omitted. ■

Example 3.7
Find all irreducible polynomials over $\mathbb{F}_2$ of degree 3.

First, note that a non-zero polynomial in $\mathbb{F}_2[x]$ must be monic. The degree 3 polynomials are of the form $x^3 + ax^2 + bx + c$, where each coefficient is 0 or 1, i.e. there are $2^3 = 8$ of them. Such a polynomial is reducible over $\mathbb{F}_2$ precisely if it has a divisor of degree 1. Compute all products $(x + a_0)(x^2 + b_1x + b_0)$ to obtain all reducible degree 3 polynomials over $\mathbb{F}_2$. There are 6 of these, leaving 2 irreducibles: $x^3 + x + 1$ and $x^3 + x^2 + 1$.

Theorem 3.8
For $f \in F[x]$, the residue class ring $F[x]/(f)$ is a field if and only if f is irreducible over F.

Proof. *Details omitted. For those who know some ring theory this is immediate since, for a PID S, $S/(c)$ is a field if and only if c is a prime element of S. Here, the prime elements of the PID $R[x]$ are precisely the irreducible polynomials.* ■

We will be very interested in the structure of the residue class ring $F[x]/(f)$, for arbitrary non-zero polynomial $f \in F[x]$. To summarize,

- $F[x]/(f)$ consists of residue classes $g + (f)$ (also denoted $[g]$) with $g \in F[x]$.
- Two residue classes $g+(f)$ and $h+(f)$ are identical if and only if $g \equiv h \bmod f$, i.e. precisely if $g - h$ is divisible by f. This is equivalent to: g and h have the same remainder on division by f.
- Each residue class $g + (f)$ contains a unique representative $r \in F[x]$ with $\deg(r) < \deg(f)$, namely the remainder when g is divided by f. The process of moving from g to r is called *reduction mod f*. (Exercise: uniqueness?)
- Hence the distinct residue classes comprising $F[x]/(f)$ are precisely the residue classes $r + (f)$, where r runs through all polynomials in $F[x]$ with $\deg(r) < \deg(f)$.
- In particular, if $F = \mathbb{F}_p$ and $\deg(f) = n$, then the number of elements of $\mathbb{F}_p/(f)$ is equal to the number of polynomials in $\mathbb{F}_p/(f)$ of degree $< n$, namely p^n.

Example 3.9
- Let $f = x \in \mathbb{F}_2[x]$. The field $\mathbb{F}_2[x]/(x)$ has $2^1 = 2$ elements, namely $0 + (x)$ and $1 + (x)$. This field is isomorphic to $\mathbb{F}_2$.

- Let $f = x^2 + x + 1 \in \mathbb{F}_2[x]$. Then $\mathbb{F}_2[x]/(f)$ is a finite field of $2^2 = 4$ elements: $\{0 + (f), 1 + (f), x + (f), x + 1 + (f)\}$. Its behaviour under addition and multiplication is shown below (remember our field has characteristic 2). When performing field operations note that, since we replace each occurrence of f by 0, the polynomial representative for each residue class has degree less than 2.

$+$	$0+(f)$	$1+(f)$	$x+(f)$	$x+1+(f)$
$0+(f)$	$0+(f)$	$1+(f)$	$x+(f)$	$x+1+(f)$
$1+(f)$	$1+(f)$	$0+(f)$	$x+1+(f)$	$0+(f)$
$x+(f)$	$x+(f)$	$x+1+(f)$	$0+(f)$	$1+(f)$
$x+1+(f)$	$x+1+(f)$	$x+(f)$	$1+(f)$	$0+(f)$

,

$*$	$0+(f)$	$1+(f)$	$x+(f)$	$x+1+(f)$
$0+(f)$	$0+(f)$	$0+(f)$	$0+(f)$	$0+(f)$
$1+(f)$	$0+(f)$	$1+(f)$	$x+(f)$	$x+1+(f)$
$x+(f)$	$0+(f)$	$x+(f)$	$x+1+(f)$	$1+(f)$
$x+1+(f)$	$0+(f)$	$x+1+(f)$	$1+(f)$	$x+(f)$

.

Note that, in the multiplication table,

$$(x+(f))(x+(f)) = x^2 + (f) = f - x - 1 + (f) = x + 1 + (f),$$

$$(x+(f))(x+1+(f)) = x^2 + x + (f) = f - 1 + (f) = 1 + (f),$$

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

Comparing these tables to those of $\mathbb{Z}_4$ we see that the field $\mathbb{F}_2[x]/(f)$ is *not* isomorphic to $\mathbb{Z}_4$, which is not a field since in $\mathbb{Z}_4$ we have $2 \cdot 2 = 0$.

What is the multiplicative order of $x+(f)$ in $\mathbb{F}_2[x]/(f)$? The multiplicative group of this field has order $2^2-1 = 3$, so the order must be 1 or 3. Clearly $x+(f) \neq 1+(f)$, so the order must be 3. Check: $(x+(f))^3 = (x+(f))(x^2+(f)) = x(x+1)+(f) = x^2+x+(f) = 1+(f)$.

- Let $f = x^2 + 2 \in \mathbb{F}_3[x]$. We find that $\mathbb{F}_3[x]/(f)$ is a ring of 9 elements which is not even an integral domain, let alone a field. Its elements are $\{0 + (f), 1 + (f), 2 + (f), x + (f), x + 1 + (f), x + 2 + (f), 2x + (f), 2x + 1 + (f), 2x + 2 + (f)\}$. To see that it is not an integral domain, note that $(x + 1 + (f))(x - 1 + (f)) = x^2 - 1 + (f) = x^2 + 2 + (f) = 0 + (f)$, but neither $x + 1 + (f)$ nor $x - 1 + (f)$ are zero.

Definition 3.10
An element $a \in F$ is called a *root* (or *zero*) of the polynomial $f \in F[x]$ if $f(a) = 0$.

Example 3.11
(i) The elements $2, 3 \in \mathbb{Q}$ are roots of $x^2 - 5x + 6 \in \mathbb{Q}[x]$.
(ii) The polynomial $x^2 + 1 \in \mathbb{Q}[x]$ has no roots in $\mathbb{Q}$, but two roots $\pm i \in \mathbb{C}$.

Definition 3.12
If $f = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \in F[x]$, then the *derivative* f' of f is defined by $f' = a_1 + 2a_2x + \cdots + na_nx^{n-1} \in F[x]$.

This obeys the familiar rules:

$$(f + g)' = f' + g'$$

and

$$(fg)' = fg' + f'g.$$

Theorem 3.13
An element $a \in F$ is a root of the polynomial $f \in F[x]$ if and only if $x - a$ divides f.

Proof. Using the Division Algorithm, we can write

$$f = q \cdot (x - a) + c$$

with $q \in F[x]$ and $c \in F$. Substituting $x = a$, we get $f(a) = c$, hence $f = q \cdot (x - a) + f(a)$. The theorem follows from this identity. ■

Definition 3.14
Let $a \in F$ be a root of $f \in F[x]$. If k is a positive integer such that f is divisible by $(x - a)^k$ but not $(x - a)^{k+1}$, then k is called the *multiplicity* of a. If $k \geq 2$ then a is called a *multiple root* of f.

Theorem 3.15
An element $a \in F$ is a multiple root of $f \in F[x]$ if and only if it is a root of both f and its derivative f'.

Proof. Exercise ■

Example 3.16
Consider the polynomial $f = x^3 - 7x^2 + 16x - 12 \in \mathbb{Q}[x]$. It factors as $(x-2)^2(x-3)$, so its roots are 2 (with multiplicity 2) and 3 (with multiplicity 1). Here, $f' = 3x^2 - 14x + 16$ which factors as $(x - 2)(3x - 8)$, so we can verify that 2 is also a root of f'.

The following observation is very important.

Theorem 3.17
If F is a field and $f \in F[x]$ has degree n, then F contains at most n roots of f.

Proof. Outline: Suppose F contains $n + 1$ distinct roots $a_1, \ldots, a_{n+1}$ of f. By Theorem 3.13, we can show that this implies $f = (x - a_1)(x - a_2) \cdots (x - a_{n+1})g$ for some polynomial g, contradicting $\deg(f) = n$. ■

Chapter 2

Some field theory

4 Field Extensions

Definition 4.1
Let F be a field. A subset K that is itself a field under the operations of F is called a *subfield* of F. The field F is called an *extension field* of K. If $K \neq F$, K is called a *proper* subfield of F.

Definition 4.2
A field containing no proper subfields is called a *prime field*.

For example, $\mathbb{F}_p$ is a prime field, since any subfield must contain the elements 0 and 1, and since it is closed under addition it must contain all other elements, i.e. it must be the whole field.

Definition 4.3
The intersection of all subfields of a field F is itself a field, called the *prime subfield* of F.

Remark 4.4
The prime subfield of F is a prime field, as defined above (see Exercise sheet).

Theorem 4.5
The prime subfield of a field F is isomorphic to $\mathbb{Q}$ if F has characteristic 0 and is isomorphic to $\mathbb{F}_p$ if F has characteristic p.

Proof. Denote by $P(F)$ the prime subfield of F. Let F be a field of characteristic 0; then the elements $n1_F$ ($n \in \mathbb{Z}$) are all distinct, and form a subring of F isomorphic to $\mathbb{Z}$. The set

$$Q(F) = \{m1_F/n1_F : m, n \in \mathbb{Z}, n \neq 0\}$$

is a subfield of F isomorphic to $\mathbb{Q}$. Any subfield of F must contain 1 and 0 and so must contain $Q(F)$, so $Q(F) \subseteq P(F)$. Since $Q(F)$ is itself a subfield of F, we also have $P(F) \subseteq Q(F)$, so in fact $Q(F)$ is the prime subfield of F. If F has characteristic p, a similar argument holds with the set

$$Q(F) = \{0 \cdot (1_F), 1 \cdot (1_F), 2 \cdot (1_F), \ldots, (p-1) \cdot 1_F\},$$

and this is isomorphic to $\mathbb{F}_p$. ■

Definition 4.6

- Let K be a subfield of the field F and M any subset of F. Then the field $K(M)$ is defined to be the intersection of all subfields of F containing both K and M; i.e. it is the smallest subfield of F containing both K and M. It is called the extension field obtained by *adjoining* the elements of M.

- For finite $M = \{\alpha_1, \ldots, \alpha_n\}$, we write $K(M) = K(\alpha_1, \ldots, \alpha_n)$.
- If $M = \{\alpha\}$, then $L = K(\alpha)$ is called a *simple extension* of K and α is called a *defining element* of L over K.

The following type of extension is very important in the theory of fields in general.

Definition 4.7

- Let K be a subfield of F and $\alpha \in F$. If α satisfies a nontrivial polynomial equation with coefficients in K, i.e. if

$$a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha_1 + a_0 = 0$$

for some $a_i \in K$ not all zero, then α is *algebraic over* K.

- An extension L of K is called *algebraic over* K (or an *algebraic extension of* K) if every element in L is algebraic over K.

Example 4.8

- The element $\sqrt[3]{3} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$, since it is a root of the polynomial $x^3 - 3 \in \mathbb{Q}[x]$.
- The element $i \in \mathbb{C}$ is algebraic over $\mathbb{R}$, since it is a root of $x^2 + 1 \in \mathbb{R}[x]$.
- The element $\pi \in \mathbb{R}$ is not algebraic over $\mathbb{Q}$. An element which is not algebraic over a field F is said to be *transcendental* over F.

Given $\alpha \in F$ which is algebraic over some subfield K of F, it can be checked (exercise!) that the set $J = \{f \in K[x] : f(\alpha) = 0\}$ is an ideal of $F[x]$ and $J \neq (0)$. By Theorem 3.4, it follows that there exists a uniquely determined monic polynomial $g \in K[x]$ which generates J, i.e. $J = (g)$.

Definition 4.9
If α is algebraic over K, then the uniquely determined monic polynomial $g \in K[x]$ generating the ideal $J = \{f \in K[x] : f(\alpha) = 0\}$ of $K[x]$ is called the *minimal polynomial* of α over K. We refer to the degree of g as the *degree of* α *over* K.

The key properties of the minimal polynomial are summarised in the next theorem. The third property is the one most useful in practice.

Theorem 4.10
Let $\alpha \in F$ be algebraic over a subfield K of F, and let g be the minimal polynomial of α. Then

(i) g is irreducible in $K[x]$;

(ii) For $f \in K[x]$, we have $f(\alpha) = 0$ if and only if g divides f;

(iii) g is the monic polynomial of least degree having α as a root.

Proof. (i) Since g has the root α, it has positive degree. Suppose $g = h_1h_2$ in $K[x]$ with $1 \leq \deg(h_i) < \deg(g)$ $(i = 1, 2)$. This implies $0 = g(\alpha) = h_1(\alpha)h_2(\alpha)$, and so one of h_1 or h_2 must lie in J and hence is divisible by g, a contradiction.
(ii) Immediate from the definition of g.
(iii) Any monic polynomial in $K[x]$ having α as a root must be a multiple of g by (ii), and so is either equal to g or has larger degree than g. ■

Example 4.11

- The element $\sqrt[3]{3} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$ since it is a root of $x^3 - 3 \in \mathbb{Q}[x]$. Since $x^3 - 3$ is irreducible over $\mathbb{Q}$, it is the minimal polynomial of $\sqrt[3]{3}$ over $\mathbb{Q}$, and hence $\sqrt[3]{3}$ has degree 3 over $\mathbb{Q}$.

- The element $i = \sqrt{-1} \in \mathbb{C}$ is algebraic over the subfield $\mathbb{R}$ of $\mathbb{C}$, since it is a root of the polynomial $x^2 + 1 \in \mathbb{R}[x]$. Since $x^2 + 1$ is irreducible over $\mathbb{R}$, it is the minimal polynomial of i over $\mathbb{R}$, and hence i has degree 2 over $\mathbb{R}$.

5 Field extensions as vector spaces

Let L be an extension field of K. An important observation is that L may be viewed as a vector space over K. The elements of L are the "vectors" and the elements of K are the "scalars".

We briefly recall the main properties of a vector space.

Definition 5.1
A *vector space* V over F is a non-empty set of objects (called vectors) upon which two operations are defined

- addition: there is some rule which produces, from any two objects in V, another object in V (denote this operation by $+$)
- scalar multiplication: there is some rule which produces, from an element of F (a scalar) and an object in V, another object in V

and these objects and operations obey the Vector Space Axioms:

1. $x + y = y + x$ for all $x, y \in V$
2. $(x + y) + z = x + (y + z)$ for all $x, y, z \in V$
3. there exists an object $0 \in V$ such that $x + 0 = x$ for all $x \in V$
4. for every $x \in V$ there exists an object $-x$ such that $x + (-x) = 0$
5. $\lambda(x + y) = \lambda x + \lambda y$ for all $x, y \in V$ and all scalars $\lambda \in F$
6. $(\lambda + \mu)x = \lambda x + \mu x$ for all $x \in V$ and all scalars $\lambda, \mu \in F$
7. $(\lambda\mu)x = \lambda(\mu x)$ for all $x \in V$ and all scalars $\lambda, \mu \in F$
8. $1x = x$ for all $x \in V$

Definition 5.2

- A basis of a vector space V over F is defined as a subset $\{v_1, \ldots, v_n\}$ of vectors in V that are linearly independent and span V. If $v_1, \ldots, v_n$ is a list of vectors in V, then these vectors form a basis if and only if every $v \in V$ can be *uniquely* written as

  $$v = a_1 v_1 + \cdots + a_n v_n$$

  where $a_1, \ldots, a_n$ are elements of the base field F.

- A vector space will have many different bases, but there are always the same number of basis vectors in each. The number of basis vectors in any basis is called the *dimension* of V over F.

- Suppose V has dimension n over F. Then any sequence of more than n vectors in V is linearly dependent.

To see that the vector space axioms hold for a field L over a subfield K, note that the elements of L form an abelian group under addition, and that any "vector" $\alpha \in L$ may be multiplied by an $r \in K$ (a "scalar") to get $r\alpha \in L$ (this is just multiplication in L). Finally, the laws for multiplication by scalars hold since, for $r, s \in L$ and $\alpha, \beta \in K$ we have $r(\alpha + \beta) = r\alpha + r\beta$, $(r + s)\alpha = r\alpha + s\alpha$, $(rs)\alpha = r(s\alpha)$ and $1\alpha = \alpha$.

Example 5.3
Take $L = \mathbb{C}$ and let K be its subfield $\mathbb{R}$. Then we can easily check that $\mathbb{C}$ is a vector space over $\mathbb{R}$. Since we know from school that $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$, it is clear that a basis is given by $\{1, i\}$.

Definition 5.4
Let L be an extension field of K. If L is finite-dimensional as a vector space over K, then L is said to be a *finite extension* of K. The dimension of the vector space L over K is called the *degree* of L over K and written $[L : K]$.

Example 5.5
From above, $\mathbb{C}$ is a finite extension of $\mathbb{R}$ of degree 2.

Theorem 5.6
If L is a finite extension of K and M is a finite extension of L, then M is a finite extension of K with

$$[M : K] = [M : L][L : K].$$

Proof. Let $[M : L] = m$, $[L : K] = n$; let $\{\alpha_1, \ldots, \alpha_m\}$ be a basis of M over L and let $\{\beta_1, \ldots, \beta_n\}$ be a basis of L over K. We shall use them to form a basis of M over K of appropriate cardinality.

Every $\alpha \in M$ can be expressed as a linear combination $\alpha = \gamma_1\alpha_1 + \cdots + \gamma_m\alpha_m$ for some $\gamma_1, \ldots, \gamma_m \in L$. Writing each γ_i as a linear combination of the β_j's we get

$$\alpha = \sum_{i=1}^{m} \gamma_i\alpha_i = \sum_{i=1}^{m}(\sum_{j=1}^{n} r_{ij}\beta_j)\alpha_i = \sum_{i=1}^{m}\sum_{j=1}^{n} r_{ij}\beta_j\alpha_i$$

with coefficients $r_{ij} \in K$. We claim that the mn elements $\beta_j\alpha_i$ form a basis of M over K. Clearly they span M; it suffices to show that they are linearly independent over K.

Suppose we have

$$\sum_{i=1}^{m}\sum_{j=1}^{n} s_{ij}\beta_j\alpha_i = 0$$

where the coefficients $s_{ij} \in K$. Then

$$\sum_{i=1}^{m}(\sum_{j=1}^{n} s_{ij}\beta_j)\alpha_i = 0,$$

and since the α_i are linearly independent over L we must have

$$\sum_{j=1}^{n} s_{ij}\beta_j = 0$$

for $1 \leq i \leq m$. Now, since the β_j are linearly independent over K, it follows that all the s_{ij} are 0, as required. ■

Theorem 5.7
Every finite extension of K is algebraic over K.

Proof. Let L be a finite extension of K and let $[L : K] = m$. For $\alpha \in L$, the $m + 1$ elements $1, \alpha, \ldots, \alpha^m$ must be linearly dependent over K, i.e. must satisfy $a_0 + a_1\alpha + \cdots a_m\alpha^m = 0$ for some $a_i \in K$ (not all zero). Thus α is algebraic over K. ■

Remark 5.8
The converse of Theorem 5.7 is not true, however. See the Exercise sheet for an example of an algebraic extension of $\mathbb{Q}$ which is not a finite extension.

We now relate our new vector space viewpoint to the residue class rings considered previously.

Theorem 5.9
Let F be an extension field of K and $\alpha \in F$ be algebraic of degree n over K and let g be the minimal polynomial of α over K. Then

(i) $K(\alpha)$ is isomorphic to $K[x]/(g)$;

(ii) $[K(\alpha) : K] = n$ and $\{1, \alpha, \ldots, \alpha^{n-1}\}$ is a basis of $K(\alpha)$ over K;

(iii) Every $\beta \in K(\alpha)$ is algebraic over K and its degree over K is a divisor of n.

Proof. (i) Consider the "evaluation at α" mapping $\tau : K[x] \to K(\alpha)$, defined by

$$\tau(f) = f(\alpha) \text{ for } f \in K[x];$$

it is easily shown that this is a homomorphism. Then

$$\ker\tau = \{f \in K[x] : f(\alpha) = 0\} = (g)$$

by the definition of the minimal polynomial. Let S be the image of τ, i.e. the set of polynomial expressions in α with coefficients in K. By the First Isomorphism Theorem for rings we have $S \cong K[x]/(g)$. Since g is irreducible, by Theorem 3.8, $K[x]/(g)$ is a field and so S is a field. Since $K \subseteq S \subseteq K(\alpha)$ and $\alpha \in S$, we have $S = K(\alpha)$ by the definition of $K(\alpha)$, and (i) follows.
(ii) Spanning set: Since $S = K(\alpha)$, any $\beta \in K(\alpha)$ can be written in the form $\beta = f(\alpha)$ for some polynomial $f \in K[x]$. By the division algorithm, $f = qg + r$ for some $q, r \in K[x]$ and $\deg(r) < \deg(g) = n$. Then

$$\beta = f(\alpha) = q(\alpha)g(\alpha) + r(\alpha) = r(\alpha),$$

and so β is a linear combination of $1, \alpha, \ldots, \alpha^{n-1}$ with coefficients in K.
L.I.: if $a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} = 0$ for some $a_0, \ldots, a_{n-1} \in K$, then the polynomial $h = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} \in K[x]$ has α as a root, and is thus a multiple of its minimal polynomial g. Since $\deg(h) < n = \deg(g)$, this is possible only if $h = 0$, i.e. $a_0 = \cdots = a_{n-1} = 0$. Thus the elements $1, \alpha, \ldots, \alpha^{n-1}$ are linearly independent over K.
(iii) $K(\alpha)$ is a finite extension of K by (ii), and so $\beta \in K(\alpha)$ is algebraic over K by Theorem 5.7. Moreover, $K(\beta)$ is a subfield of $K(\alpha)$. If d is the degree of β over K, then $n = [K(\alpha) : K] = [K(\alpha) : K(\beta)][K(\beta) : K] = [K(\alpha) : K(\beta)]d$, i.e. d divides n. ■

Remark 5.10
This theorem tells us that the elements of the simple extension $K(\alpha)$ of K are polynomial expressions in α, and any $\beta \in K(\alpha)$ can be uniquely expressed in the form $\beta = a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}$ for some $a_i \in K$.

Example 5.11
Consider the simple extension $\mathbb{R}(i)$ of $\mathbb{R}$. We saw earlier that i has minimal polynomial $x^2 + 1$ over $\mathbb{R}$.

So $\mathbb{R}(i) \cong \mathbb{R}[x]/(x^2 + 1)$, and $\{1, i\}$ is a basis for $\mathbb{R}(i)$ over $\mathbb{R}$. So

$$\mathbb{R}(i) = \{a + bi : a, b \in \mathbb{R}\} = \mathbb{C}.$$

Example 5.12
Consider the simple extension $\mathbb{Q}(\sqrt[3]{3})$ of $\mathbb{Q}$. We saw earlier that $\sqrt[3]{3}$ has minimal polynomial x^3-3 over $\mathbb{Q}$.

So $\mathbb{Q}(\sqrt[3]{3}) \cong \mathbb{Q}[x]/(x^3-3)$, and $\{1, \sqrt[3]{3}, (\sqrt[3]{3})^2\}$ is a basis for $\mathbb{Q}(\sqrt[3]{3})$ over $\mathbb{Q}$. So

$$\mathbb{Q}(\sqrt[3]{3}) = \{a + b\sqrt[3]{3} + c(\sqrt[3]{3})^2 : a, b, c \in \mathbb{Q}\}.$$

Note that we have been assuming that both K and α are embedded in some larger field F. Next, we will consider constructing a simple algebraic extension without reference to a previously given larger field, i.e. "from the ground up".

The next result, due to Kronecker, is one of the most fundamental results in the theory of fields: it says that, given any non-constant polynomial over any field, there exists an extension field in which the polynomial has a root.

Theorem 5.13 (Kronecker)
Let $f \in K[x]$ be irreducible over the field K. Then there exists a simple algebraic extension of K with a root of f as a defining element.

Proof.

- Consider the residue class ring $L = K[x]/(f)$, which is a field since f is irreducible. Its elements are the residue classes $[h] = h + (f)$, with $h \in K[x]$.
- For any $a \in K$, think of a as a constant polynomial in $K[x]$ and form the residue class $[a]$. The mapping $a \mapsto [a]$ gives an isomorphism from K onto a subfield K' of L (exercise: check!), so K' may be identified with K. Thus we can view L as an extension of K.
- For every $h = a_0 + a_1x + \cdots + a_mx^m \in K[x]$, we have

$$\begin{aligned}[h] &= [a_0 + a_1x + \cdots + a_mx^m] \\ &= [a_0] + [a_1][x] + \cdots + [a_m][x]^m \\ &= a_0 + a_1[x] + \cdots + a_m[x]^m\end{aligned}$$

  (making the identification $[a_i] = a_i$). So, every element of L can be written as a polynomial in $[x]$ with coefficients in K. Since any field containing K and $[x]$ must contain these expressions, L is a simple extension of K obtained by adjoining $[x]$.
- If $f = b_0 + b_1x + \cdots + b_nx^n$, then

$$f([x]) = b_0 + b_1[x] + \cdots + b_n[x]^n = [f] = [0],$$

  i.e. $[x]$ is a root of f and L is a simple algebraic extension of K.

■

Example 5.14
Consider the prime field $\mathbb{F}_3$ and the polynomial $x^2 + x + 2 \in \mathbb{F}_3[x]$, irreducible over $\mathbb{F}_3$. Take θ to be a "root" of f, in the sense that θ is the residue class $[x] = x + (f) \in L = \mathbb{F}_3[x]/(f)$. Explicitly, we have:

$$\begin{aligned}f(\theta) = f([x]) &= f(x + (f)) \\ &= (x + (f))^2 + (x + (f)) + (2 + (f)) \\ &= x^2 + x + 2 + (f) \\ &= f + (f) \\ &= 0 + (f) \\ &= [0].\end{aligned}$$

The other root of f in L is $2\theta+2$, since $f(2\theta+2) = \theta^2+\theta+2 = 0$. By Theorem 5.9, the simple algebraic extension $L = \mathbb{F}_3(\theta)$ consists of the nine elements $0, 1, 2, \theta, \theta+1, \theta+2, 2\theta, 2\theta+1, 2\theta+2$.

Example 5.15
Consider the polynomial $f = x^2+x+1 \in \mathbb{F}_2[x]$, irreducible over $\mathbb{F}_2$. Let θ be the root $[x] = x+(f)$ of f; then the simple algebraic extension $L = \mathbb{F}_2(\theta)$ consists of the four elements $0, 1, \theta, \theta+1$. (The other root is $\theta + 1$). The tables for addition and multiplication are precisely those of Example 3.9, now appropriately relabelled. We give the addition table:

$+$	0	1	θ	$\theta+1$
0	0	1	θ	$\theta+1$
1	1	0	$\theta+1$	θ
θ	θ	$\theta+1$	0	1
$\theta+1$	$\theta+1$	θ	1	0

.

Note that, in the above examples, adjoining either of two roots of f would yield the same extension field.

Theorem 5.16
Let F be an extension field of the field K and $\alpha, \beta \in F$ be two roots of a polynomial $f \in K[x]$ that is irreducible over K. Then $K(\alpha)$ and $K(\beta)$ are isomorphic under an isomorphism mapping α to β and keeping the elements of K fixed.

Proof. By Theorem 5.9 both are isomorphic to the field $K[x]/(f)$ since the irreducible f is the minimal polynomial of both α and β. ■

Given a polynomial, we now want an extension field which contains all its roots.

Definition 5.17
Let $f \in K[x]$ be a polynomial of positive degree and F an extension field of K. Then we say that f *splits in* F if f can be written as a product of linear factors in $F[x]$, i.e. if there exist elements $\alpha_1, \ldots, \alpha_n \in F$ such that

$$f = a(x-\alpha_1)\cdots(x-\alpha_n)$$

where a is the leading coefficient of f. The field F is called a *splitting field* of f over K if it splits in F and if $F = K(\alpha_1, \ldots, \alpha_n)$.

So, a splitting field F of a polynomial f over K is an extension field containing all the roots of f, and is "smallest possible" in the sense that no subfield of F contains all roots of f. The following result answers the questions: can we always find a splitting field, and how many are there?

Theorem 5.18 (Existence and uniqueness of splitting field)

(i) If K is a field and f any polynomial of positive degree in $K[x]$, then there exists a splitting field of f over K.

(ii) Any two splitting fields of f over K are isomorphic under an isomorphism which keeps the elements of K fixed and maps roots of f into each other.

So, we may therefore talk of *the* splitting field of f over K. It is obtained by adjoining to K finitely many elements algebraic over K, and so we can show (exercise!) that it is a finite extension of K.

Example 5.19
Find the splitting field of the polynomial $f = x^2 + 2 \in \mathbb{Q}[x]$ over $\mathbb{Q}$.

The polynomial f splits in $\mathbb{C}$, where it factors as $(x - i\sqrt{2})(x + i\sqrt{2})$. However, $\mathbb{C}$ itself is not the splitting field for f. It turns out to be sufficient to adjoin just one of the complex roots of f to $\mathbb{Q}$. The field $K = \mathbb{Q}(i\sqrt{2})$ contains both of the roots of f, and no smaller subfield has this property, so K is the splitting field for F.

Splitting fields will be central to our characterization of finite fields, in the next chapter.

Chapter 3

Finite fields

We have seen, in the previous chapters, some examples of finite fields. For example, the residue class ring $\mathbb{Z}/p\mathbb{Z}$ (when p is a prime) forms a field with p elements which may be identified with the Galois field $\mathbb{F}_p$ of order p.

The fields $\mathbb{F}_p$ are important in field theory. From the previous chapter, every field of characteristic p contains a copy of $\mathbb{F}_p$ (its prime subfield) and can therefore be thought of as an extension of $\mathbb{F}_p$. Since every finite field must have characteristic p, this helps us to classify finite fields.

6 Characterizing finite fields

Lemma 6.1
Let F be a finite field containing a subfield K with q elements. Then F has q^m elements, where $m = [F : K]$.

Proof. F is a vector space over K, finite-dimensional since F is finite. Denote this dimension by m; then F has a basis over K consisting of m elements, say $b_1, \ldots, b_m$. Every element of F can be uniquely represented in the form $k_1 b_1 + \cdots k_m b_m$ (where $k_1, \ldots, k_m \in K$). Since each $k_i \in K$ can take q values, F must have exactly q^m elements. ■

We are now ready to answer the question: "What are the possible cardinalities for finite fields?"

Theorem 6.2
Let F be a finite field. Then F has p^n elements, where the prime p is the characteristic of F and n is the degree of F over its prime subfield.

Proof. Since F is finite, it must have characteristic p for some prime p (by Corollary 2.19). Thus the prime subfield K of F is isomorphic to $\mathbb{F}_p$, by Theorem 4.5, and so contains p elements. Applying Lemma 6.1 yields the result. ■

So, all finite fields must have prime power order - there is no finite field with 6 elements, for example.

We next ask: does there exist a finite field of order p^n for every prime power p^n? How can such fields be constructed?

We saw, in the previous chapter, that we can take the prime fields $\mathbb{F}_p$ and construct other finite fields from them by adjoining roots of polynomials. If $f \in \mathbb{F}_p[x]$ is irreducible of degree n over $\mathbb{F}_p$, then adjoining a root of f to $\mathbb{F}_p$ yields a finite field of p^n elements. However, it is not clear whether we can find an irreducible polynomial in $\mathbb{F}_p[x]$ of degree n, for *every* integer n.

The following two lemmas will help us to characterize fields using root adjunction.

Lemma 6.3
If F is a finite field with q elements, then every $a \in F$ satisfies $a^q = a$.

Proof. Clearly $a^q = a$ is satisfied for $a = 0$. The non-zero elements form a group of order $q - 1$ under multiplication. Using the fact that $a^{|G|} = 1_G$ for any element a of a finite group G, we have that all $0 \neq a \in F$ satisfy $a^{q-1} = 1$, i.e. $a^q = a$. ■

Lemma 6.4
If F is a finite field with q elements and K is a subfield of F, then the polynomial $x^q - x$ in $K[x]$ factors in $F[x]$ as

$$x^q - x = \prod_{a \in F} (x - a)$$

and F is a splitting field of $x^q - x$ over K.

Proof. Since the polynomial $x^q - x$ has degree q, it has at most q roots in F. By Lemma 6.3, all the elements of F are roots of the polynomial, and there are q of them. Thus the polynomial splits in F as claimed, and cannot split in any smaller field. ■

We are now ready to prove the main characterization theorem for finite fields.

Theorem 6.5 (Existence and Uniqueness of Finite Fields)
For every prime p and every positive integer n, there exists a finite field with p^n elements. Any finite field with $q = p^n$ elements is isomorphic to the splitting field of $x^q - x$ over $\mathbb{F}_p$.

Proof. (Existence) For $q = p^n$, consider $x^q - x$ in $\mathbb{F}_p[x]$, and let F be its splitting field over $\mathbb{F}_p$. Since its derivative is $qx^{q-1} - 1 = -1$ in $\mathbb{F}_p[x]$, it can have no common root with $x^q - x$ and so, by Theorem 3.15, $x^q - x$ has q distinct roots in F. Let $S = \{a \in F : a^q - a = 0\}$. Then S is a subfield of F since

- S contains 0;
- $a, b \in S$ implies (by Freshmen's Exponentiation) that $(a-b)^q = a^q - b^q = a - b$, so $a - b \in S$;
- for $a, b \in S$ and $b \neq 0$ we have $(ab^{-1})^q = a^q b^{-q} = ab^{-1}$,so $ab^{-1} \in S$.

On the other hand, $x^q - x$ must split in S since S contains all its roots, i.e its splitting field F is a subfield of S. Thus $F = S$ and, since S has q elements, F is a finite field with $q = p^n$ elements. (Uniqueness) Let F be a finite field with $q = p^n$ elements. Then F has characteristic p by Theorem 6.2, and so contains $\mathbb{F}_p$ as a subfield. So, by Lemma 6.4, F is a splitting field of $x^q - x$. The result now follows from the uniqueness (up to isomorphism) of splitting fields, from Theorem 5.18. ■

As a result of the uniqueness part of Theorem 6.5, we may speak of *the* finite field (or *the* Galois field) of q elements. We shall denote this field by $\mathbb{F}_q$, where q denotes a power of the prime characteristic p of $\mathbb{F}_q$.

Example 6.6

- In Example 5.14, we constructed a field $L = \mathbb{F}_3(\theta)$ of 9 elements, where θ is a root of the polynomial $x^2 + x + 2 \in \mathbb{F}_3[x]$. By Theorem 6.5, L is *the* field of 9 elements, i.e. $\mathbb{F}_9$.
- In Example 5.15, we constructed a field $L = \mathbb{F}_2(\theta)$ of 4 elements, where θ is a root of the polynomial $x^2 + x + 1 \in \mathbb{F}_2[x]$. By Theorem 6.5, L is *the* field of 4 elements, i.e. $\mathbb{F}_4$.

We can also completely describe the subfields of a finite field $\mathbb{F}_q$.

Theorem 6.7 (Subfield Criterion)
Let $\mathbb{F}_q$ be the finite field with $q = p^n$ elements. Then every subfield of $\mathbb{F}_q$ has order p^m, where m is a positive divisor of n. Conversely, if m is a positive divisor of n, then there is exactly one subfield of $\mathbb{F}_q$ with p^m elements.

Proof. Clearly, a subfield K of F must have order p^m for some positive integer $m \leq n$. By Lemma 6.1, $q = p^n$ must be a power of p^m, and so m must divide n.

Conversely, if m is a positive divisor of n, then $p^m - 1$ divides $p^n - 1$, and so $x^{p^m-1} - 1$ divides $x^{p^n-1} - 1$ in $\mathbb{F}_p[x]$. So, every root of $x^{p^m} - x$ is a root of $x^q - x$, and hence belongs to $\mathbb{F}_q$. It follows that $\mathbb{F}_q$ must contain a splitting field of $x^{p^m} - x$ over $\mathbb{F}_p$ as a subfield, and (from proof of Theorem 6.5) such a splitting field has order p^m. If there were two distinct subfields of order p^m in $\mathbb{F}_q$, they would together contain more than p^m roots of $x^{p^m} - x$ in $\mathbb{F}_q$, a contradiction. ■

So, the unique subfield of $\mathbb{F}_{p^n}$ of order p^m, where m is a positive divisor of n, consists precisely of the roots of $x^{p^m} - x$ in $\mathbb{F}_{p^n}$.

Example 6.8
Determine the subfields of the finite field $\mathbb{F}_{2^{30}}$. To do this, list all positive divisors of 30. The containment relations between subfields are equivalent to divisibility relations among the positive divisors of 30. (For diagram, see lectures!)

We can also completely characterize the multiplicative group of a finite field. For the finite field $\mathbb{F}_q$, we denote the multiplicative group of non-zero elements of $\mathbb{F}_q$ by $\mathbb{F}_q^*$.

Theorem 6.9
For every finite field $\mathbb{F}_q$, the multiplicative group $\mathbb{F}_q^*$ of nonzero elements of $\mathbb{F}_q$ is cyclic.

Proof. We may assume $q \geq 3$. Set $h = q - 1$, the order of $\mathbb{F}_q^*$, and let $h = p_1^{r_1} p_2^{r_2} \dots p_m^{r_m}$ be its prime factor decomposition. For each i, $1 \leq i \leq m$, the polynomial $x^{h/p_i} - 1$ has at most h/p_i roots in $\mathbb{F}_q$. Since $h/p_i < h$, it follows that there are nonzero elements of $\mathbb{F}_q$ which are not roots of this polynomial. Let a_i be such an element, and set $b_i = a_i^{h/p_i^{r_i}}$. Now, $b_i^{p_i^{r_i}} = 1$, so the order of b_i divides $p_i^{r_i}$ and so has the form $p_i^{s_i}$ for some $0 \leq s_i \leq r_i$. On the other hand,

$$b_i^{p_i^{r_i-1}} = a_i^{h/p_i} \neq 1,$$

so the order of b_i is precisely $p_i^{r_i}$.

Let $b = b_1 b_2 \dots b_m$. We claim: b has order $h(= q - 1)$, i.e. is a generator for the group. Suppose, on the contrary, that the order of b is a proper divisor of h. It is therefore a divisor of at least one of the m integers h/p_i, $1 \leq i \leq m$; wlog, say of h/p_1. Then

$$1 = b^{h/p_1} = b_1^{h/p_1} b_2^{h/p_1} \cdots b_m^{h/p_1}.$$

Now, if $2 \leq i \leq m$, then $p_i^{r_i}$ divides h/p_1, and so $b_i^{h/p_1} = 1$. This forces $b_1^{h/p_1} = 1$. Thus the order of b_1 must divide h/p_1, which is impossible since the order of b_1 is $p_1^{r_1}$. Thus $\mathbb{F}_q^*$ is a cyclic group with generator b. ■

Definition 6.10
A generator of the cyclic group $\mathbb{F}_q^*$ is called a *primitive element* of $\mathbb{F}_q$.

By Theorem 1.13, $\mathbb{F}_q$ contains $\phi(q-1)$ primitive elements, where ϕ is Euler's function: the number of integers less than and relatively prime to $q - 1$. Recall that, if the integer n has the prime factorization $p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}$, then

$$\phi(n) = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_r}).$$

Example 6.11

- $\mathbb{F}_5$ has $\phi(4) = 2$ primitive elements, namely 2 and 3.
- $\mathbb{F}_4$ has $\phi(3) = 2$ primitive elements. Expressing $\mathbb{F}_4$ as $\mathbb{F}_2(\theta) = \{0, 1, \theta, \theta + 1\}$, where $\theta^2 + \theta + 1 = 0$, we find that both θ and $\theta + 1$ are primitive elements.

We are now ready to prove an important result.

Theorem 6.12
Let $\mathbb{F}_q$ be a finite field and $\mathbb{F}_r$ a finite extension field. Then

- *$\mathbb{F}_r$ is a simple extension of $\mathbb{F}_q$, i.e $\mathbb{F}_r = \mathbb{F}_q(\beta)$ for some $\beta \in \mathbb{F}_r$;*
- *every primitive element of $\mathbb{F}_r$ can serve as a defining element β of $\mathbb{F}_r$ over $\mathbb{F}_q$.*

Proof. Let α be a primitive element of $\mathbb{F}_r$. Clearly, $\mathbb{F}_q(\alpha) \subseteq \mathbb{F}_r$. On the other hand, since $\mathbb{F}_q(\alpha)$ contains 0 and all powers of α, it contains all elements of $\mathbb{F}_r$. So $\mathbb{F}_r = \mathbb{F}_q(\alpha)$. ■

So, we can express *any* finite field K with subfield F, by adjoining to F a root β of an appropriate irreducible polynomial f, which of course must have degree $d = [K : F]$. Although the proof of Theorem 6.12 uses a β which is a primitive element of K, it is not in fact necessary for β to be a multiplicative generator of K^*, as the next example shows.

Example 6.13
Consider the finite field $\mathbb{F}_9$. We can express $\mathbb{F}_9$ in the form $\mathbb{F}_3(\beta)$, where β is a root of the polynomial $x^2 + 1$, irreducible over $\mathbb{F}_3$. However, since $\beta^4 = 1$, β does not generate the whole of $\mathbb{F}_9^*$, i.e. β is not a primitive element of $\mathbb{F}_9$.

Corollary 6.14
For every finite field $\mathbb{F}_q$ and every positive integer n, there exists an irreducible polynomial in $\mathbb{F}_q[x]$ of degree n.

Proof. Let $\mathbb{F}_r$ be the extension field of $\mathbb{F}_q$ of order q^n, so that $[\mathbb{F}_r : \mathbb{F}_q] = n$. By Theorem 6.12, $\mathbb{F}_r = \mathbb{F}_q(\alpha)$ for some $\alpha \in \mathbb{F}_r$. Then, by properties of minimal polynomials, the minimal polynomial of α over $\mathbb{F}_q$ is an irreducible polynomial in $\mathbb{F}_q[x]$ of degree n. ■

7 Irreducible polynomials

In this section, we investigate irreducible polynomial over finite fields.

Lemma 7.1
Let $f \in \mathbb{F}_q[x]$ be an irreducible polynomial over a finite field $\mathbb{F}_q$ and let α be a root of f in an extension field of $\mathbb{F}_q$. Then, for a polynomial $h \in \mathbb{F}_q[x]$, we have $h(\alpha) = 0$ if and only if f divides h.

Proof. The minimal polynomial of α over $\mathbb{F}_q$ is given by $a^{-1}f$, where a is the leading coefficient of f (since it is a monic irreducible polynomial in $\mathbb{F}_q[x]$ having α as a root). The proposition then follows from part (ii) of Theorem 4.10. ■

Lemma 7.2
Let $f \in \mathbb{F}_q[x]$ be an irreducible polynomial over $\mathbb{F}_q$ of degree m. Then f divides $x^{q^n} - x$ if and only if m divides n.

Proof. First, suppose f divides $x^{q^n} - x$. Let α be a root of f in the splitting field of f over $\mathbb{F}_q$. Then $\alpha^{q^n} = \alpha$, so $\alpha \in \mathbb{F}_{q^n}$. Thus $\mathbb{F}_q(\alpha)$ is a subfield of $\mathbb{F}_{q^n}$. Since $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$ and $[\mathbb{F}_{q^n} : \mathbb{F}_q] = n$, we have $n = [\mathbb{F}_{q^n} : \mathbb{F}_q(\alpha)]m$, so m divides n.

Conversely, suppose m divides n. Then by Theorem 6.7, $\mathbb{F}_{q^n}$ contains $\mathbb{F}_{q^m}$ as a subfield. Let α be a root of f in the splitting field of f over $\mathbb{F}_q$. Then $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$, and so $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$. Thus $\alpha \in \mathbb{F}_{q^n}$, hence $\alpha^{q^n} = \alpha$, and so α is a root of $x^{q^n} - x \in \mathbb{F}_q[x]$. Therefore, by Lemma 7.1, f divides $x^{q^n} - x$. ■

We are now ready to describe the set of roots of an irreducible polynomial.

Theorem 7.3
If f is an irreducible polynomial in $\mathbb{F}_q[x]$ of degree m, then f has a root α in $\mathbb{F}_{q^m}$. Moreover, all the roots of f are simple and are given by the m distinct elements $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{m-1}}$ of $\mathbb{F}_{q^m}$.

Proof. Let α be a root of f in the splitting field of f over $\mathbb{F}_q$. Then $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$, hence $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$, and so $\alpha \in \mathbb{F}_{q^m}$.

We now show that, if $\beta \in \mathbb{F}_{q^m}$ is a root of f, then β^q is also a root of f. Write $f = a_m x^m + \cdots + a_1 x + a_0$ ($a_i \in \mathbb{F}_q$). Then

$$\begin{aligned} f(\beta^q) &= a_m \beta^{qm} + \cdots + a_1 \beta^q + a_0 \\ &= a_m^q \beta^{qm} + \cdots + a_1^q \beta^q + a_0^q \\ &= (a_m \beta^m + \cdots + a_1 \beta + a_0)^q \\ &= f(\beta)^q = 0, \end{aligned}$$

using Lemma 6.3 and Freshmen's Exponentiation.

Thus, the elements $\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}$ are roots of f. We must check that they are all distinct. Suppose not, i.e. $\alpha^{q^j} = \alpha^{q^k}$ for some $0 \leq j < k \leq m-1$. Raising this to the power q^{m-k} , we get

$$\alpha^{q^{m-k+j}} = \alpha^{q^m} = \alpha.$$

It then follows from Lemma 7.1 that f divides $x^{q^{m-k+j}} - x$. By Lemma 7.2, this is possible only if m divides $m - k + j$, a contradiction since $0 < m - k + j < m$. ■

This result gives us two useful corollaries.

Corollary 7.4
Let f be an irreducible polynomial in $\mathbb{F}_q[x]$ of degree m. Then the splitting field of f over $\mathbb{F}_q$ is $\mathbb{F}_{q^m}$.

Proof. Theorem 7.3 shows that f splits in $\mathbb{F}_{q^m}$. To see that this is the splitting field, note that $\mathbb{F}_q(\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}) = \mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$. ■

Corollary 7.5
Any two irreducible polynomials in $\mathbb{F}_q[x]$ of the same degree have isomorphic splitting fields.

As we shall see later, sets of elements such as those in Theorem 7.3 appear often in the theory of fields.

Theorem 7.6
For every finite field $\mathbb{F}_q$ and every $n \in \mathbb{N}$, the product of all monic irreducible polynomials over $\mathbb{F}_q$ whose degrees divide n is equal to $x^{q^n} - x$.

Proof. By Lemma 7.2, the monic irreducible polynomials over $\mathbb{F}_q$ which occur in the canonical factorization of $g = x^{q^n} - x$ in $\mathbb{F}_q[x]$ are precisely those whose degrees divide n. Since $g' = -1$, by Theorem 3.15 g has no multiple roots in its splitting field over $\mathbb{F}_q$. Thus each monic irreducible polynomial over $\mathbb{F}_q$ whose degree divides n occurs exactly once in the canonical factorization of g in $\mathbb{F}_q[x]$. ■

Example 7.7
Take $q = n = 2$; the monic irreducible polynomials over $\mathbb{F}_2[x]$ whose degrees divide 2 are x, $x+1$ and x^2+x+1. It is easily seen that $x(x+1)(x^2+x+1) = x^4+x = x^4-x$.

Corollary 7.8
If $N_q(d)$ is the number of monic irreducible polynomials in $\mathbb{F}_q[x]$ of degree d, then

$$q^n = \sum_{d|n} dN_q(d) \text{ for all } n \in \mathbb{N},$$

where the sum is extended over all positive divisors d of n.

Proof. This follows immediately from Theorem 7.6, upon comparing the degree of $g = x^{q^n} - x$ with the total degree of the canonical factorization of g. ■

This corollary allows us to obtain an explicit formula for the number of monic irreducible polynomials in $\mathbb{F}_q[x]$ of a given degree. To do so, we need the following arithmetic function, which will also prove useful in the next chapter.

Definition 7.9
The *Moebius function* μ is the function on $\mathbb{N}$ defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1; \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes} \\ 0 & \text{if } n \text{ is divisible by the square of a prime.} \end{cases}$$

Example 7.10
(i)$\mu(5) = -1$; (ii)$\mu(35) = 1$; (iii)$\mu(50) = 0$.

Lemma 7.11
For $n \in \mathbb{N}$, the Moebius function satisfies

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1. \end{cases}$$

Proof. The $n = 1$ case is immediate. For $n > 1$ we need only consider the positive divisors d of n for which $\mu(d)$ is non-zero, namely those d for which $d = 1$ or d is a product of distinct primes. If $p_1, \ldots, p_k$ are the distinct prime divisors of n then

$$\begin{aligned} \sum_{d|n} \mu(d) &= \mu(1) + \sum_{i=1}^{k} \mu(p_i) + \sum_{1 \le i_1 < i_2 \le k} \mu(p_{i_1} p_{i_2}) + \cdots + \mu(p_1 p_2 \ldots p_k) \\ &= 1 + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + \cdots + \binom{k}{k}(-1)^k \\ &= (1 + (-1))^k = 0. \end{aligned}$$

■

Theorem 7.12 (Moebius Inversion Formula)

- *Additive version: let h and H be two functions from $\mathbb{N}$ into an additively written abelian group G. Then*

$$H(n) = \sum_{d|n} h(d) \text{ for all } n \in \mathbb{N} \tag{3.1}$$

if and only if

$$h(n) = \sum_{d|n} \mu(\frac{n}{d})H(d) = \sum_{d|n} \mu(d)H(\frac{n}{d}) \text{ for all } n \in \mathbb{N}. \tag{3.2}$$

- *Multiplicative version: let h and H be two functions from $\mathbb{N}$ into a multiplicatively written abelian group G. Then*

$$H(n) = \prod_{d|n} h(d) \text{ for all } n \in \mathbb{N} \tag{3.3}$$

if and only if

$$h(n) = \prod_{d|n} H(d)^{\mu(\frac{n}{d})} = \prod_{d|n} H(\frac{n}{d})^{\mu(d)} \text{ for all } n \in \mathbb{N}. \tag{3.4}$$

Proof. Additive version: we prove the forward implication; the converse is similar and is left as an exercise. Assume the first identity holds. Using Lemma 7.11, we get

$$\begin{aligned} \sum_{d|n} \mu(\frac{n}{d})H(d) &= \sum_{d|n} \mu(d)H(\frac{n}{d}) = \sum_{d|n} \mu(d) \sum_{c|\frac{n}{d}} h(c) \\ &= \sum_{c|n} \sum_{d|\frac{n}{c}} \mu(d)h(c) = \sum_{c|n} h(c) \sum_{d|\frac{n}{c}} \mu(d) = h(n) \end{aligned}$$

for all $n \in \mathbb{N}$.

Multiplicative version: immediate upon replacing sums by products and multiples by powers. ■

We can now apply this result as follows.

Theorem 7.13

The number $N_q(n)$ of monic irreducible polynomials in $\mathbb{F}_q[x]$ of degree n is given by

$$N_q(n) = \frac{1}{n} \sum_{d|n} \mu(\frac{n}{d})q^d = \frac{1}{n} \sum_{d|n} \mu(d)q^{\frac{n}{d}}.$$

Proof. Apply the additive case of the Moebius Inversion Formula to the group $G = (\mathbb{Z}, +)$. Take $h(n) = nN_q(n)$ and $H(n) = q^n$ for all $n \in \mathbb{N}$. By Corollary 7.8, the identity (3.1) is satisfied, and so the result follows. ■

Remark 7.14

Since it is clear from this formula that $N_q(n)$ is greater than zero for all n, this gives an alternative proof of Theorem 6.14.

Example 7.15

The number of monic irreducibles in $\mathbb{F}_q[x]$ of degree 12 is given by

$$\begin{aligned} N_q(12) &= \frac{1}{12}(\mu(1)q^{12} + \mu(2)q^6 + \mu(3)q^4 + \mu(4)q^3 + \mu(6)q^2 + \mu(12)q) \\ &= \frac{1}{12}(1.q^{12} + (-1)q^6 + (-1)q^4 + 0.q^3 + 1.q^2 + 0.q) \\ &= \frac{1}{12}(q^{12} - q^6 - q^4 + q^2). \end{aligned}$$

We can also obtain a formula for the *product* of all monic irreducible polynomials in $\mathbb{F}_q[x]$ of fixed degree.

Theorem 7.16
The product $I(q,n;x)$ of all monic irreducible polynomials in $\mathbb{F}_q[x]$ of degree n is given by:

$$I(q,n;x) = \prod_{d|n}(x^{q^d} - x)^{\mu(\frac{n}{d})} = \prod_{d|n}(x^{q^{\frac{n}{d}}} - x)^{\mu(d)}.$$

Proof. From Theorem 7.6 we know that

$$x^{q^n} - x = \prod_{d|n} I(q,d;x).$$

Now apply Moebius Inversion in the multiplicative form to the multiplicative group G of non-zero rational functions over $\mathbb{F}_q$. Take $h(n) = I(q,n;x)$ and $H(n) = x^{q^n} - x$ to get the desired formula. ■

Example 7.17
Take $q = 2$ and $n = 4$. Then the product of all monic irreducible quartics in $\mathbb{F}_2[x]$ is:

$$\begin{aligned} I(2,4;x) &= (x^{16} - x)^{\mu(1)}(x^4 - x)^{\mu(2)}(x^2 - x)^{\mu(4)} \\ &= \frac{x^{16} - x}{x^4 - x} = \frac{x^{15} - 1}{x^3 - 1} \\ &= x^{12} + x^9 + x^6 + x^3 + 1 \end{aligned}$$

Chapter 4

Finite fields: further properties

8 Roots of unity in finite fields

In this section, we will generalize the concept of roots of unity (well-known for complex numbers) to the finite field setting, by considering the splitting field of the polynomial $x^n - 1$. This has links with irreducible polynomials, and provides an effective way of obtaining primitive elements and hence representing finite fields.

Definition 8.1
Let $n \in \mathbb{N}$. The splitting field of $x^n - 1$ over a field K is called the *nth cyclotomic field* over K and denoted by $K^{(n)}$. The roots of $x^n - 1$ in $K^{(n)}$ are called the nth roots of unity over K and the set of all these roots is denoted by $E^{(n)}$.

The following result, concerning the properties of $E^{(n)}$, holds for an arbitrary (not just a finite!) field K.

Theorem 8.2
Let $n \in \mathbb{N}$ and K a field of characteristic p (where p may take the value 0 in this theorem). Then

(i) If $p \nmid n$, then $E^{(n)}$ is a cyclic group of order n with respect to multiplication in $K^{(n)}$.

(ii) If $p \mid n$, write $n = mp^e$ with positive integers m and e and $p \nmid m$. Then $K^{(n)} = K^{(m)}$, $E^{(n)} = E^{(m)}$ and the roots of $x^n - 1$ are the m elements of $E^{(m)}$, each occurring with multiplicity p^e.

Proof.

(i) The $n = 1$ case is trivial. For $n \geq 2$, observe that $x^n - 1$ and its derivative nx^{n-1} have no common roots; thus $x^n - 1$ cannot have multiple roots and hence $E^{(n)}$ has n elements. To see that $E^{(n)}$ is a multiplicative group, take $\alpha, \beta \in E^{(n)}$: we have $(\alpha\beta^{-1})^n = \alpha^n(\beta^n)^{-1} = 1$ and so $\alpha\beta^{-1} \in E^{(n)}$. It remains to show that the group $E^{(n)}$ is cyclic; this can be proved by an analogous argument to the proof of Theorem 6.9 (exercise: fill in details).

(ii) Immediate from $x^n - 1 = x^{mp^e} - 1 = (x^m - 1)^{p^e}$ and part (i).

■

Definition 8.3
Let K be a field of characteristic p and n a positive integer not divisible by p. Then a generator of the cyclic group $E^{(n)}$ is called a *primitive nth root of unity over K*.

By Theorem 1.13, $E^{(n)}$ has $\phi(n)$ generators, i.e. there are $\phi(n)$ primitive nth roots of unity over K. Given one such, ζ say, the set of all primitive nth roots of unity over K is given by

$$\{\zeta^s : 1 \leq s \leq n,\ \gcd(s,n) = 1\}.$$

We now consider the polynomial whose roots are precisely this set.

Definition 8.4
Let K be a field of characteristic p, n a positive integer not divisible by p and ζ a primitive nth root of unity over K. Then the polynomial

$$Q_n(x) = \prod_{\substack{s=1 \\ (s,n)=1}}^{n} (x - \zeta^s)$$

is called the nth cyclotomic polynomial over K. It is clear that $Q_n(x)$ has degree $\phi(n)$.

Theorem 8.5
Let K be a field of characteristic p and n a positive integer not divisible by p. Then

(i) $x^n - 1 = \prod_{d|n} Q_d(x)$;

(ii) the coefficients of $Q_n(x)$ belong to the prime subfield of K (and in fact to $\mathbb{Z}$ if the prime subfield is $\mathbb{Q}$).

Proof. (i) Each nth root of unity over K is a primitive dth root of unity over K for exactly one positive divisor d of n. Specifically, if ζ is a primitive nth root of unity over K and ζ^s is an arbitrary nth root of unity over K, then $d = n/\gcd(s,n)$, i.e. d is the order of ζ^s in $E^{(n)}$. Since

$$x^n - 1 = \prod_{s=1}^{n} (x - \zeta^s).$$

we obtain the result by collecting together those factors $(x - \zeta^s)$ for which ζ^s is a primitive dth root of unity over K.
(ii) Proved by induction on n. It is clearly true for $Q_1(x) = x - 1$. Let $n > 1$ and suppose it is true for all $Q_d(x)$ where $1 \leq d < n$. By (i),

$$Q_n(x) = \frac{x^n - 1}{\prod_{d|n, d<n} Q_d(x)}.$$

By the induction hypothesis, the denominator is a polynomial with coefficients in the prime subfield of K (or $\mathbb{Z}$ if $\mathrm{char}K = 0$). Applying long division yields the result. ■

Example 8.6
Let $n = 3$, let K be any field with $\mathrm{char}K \neq 3$, and let ζ be a primitive cube root of unity over K. Then

$$Q_3(x) = (x - \zeta)(x - \zeta^2) = x^2 - (\zeta + \zeta^2)x + \zeta^3 = x^2 + x + 1.$$

Example 8.7
Let r be a prime and let $k \in \mathbb{N}$. Then

$$Q_{r^k}(x) = 1 + x^{r^{k-1}} + x^{2r^{k-1}} + \cdots + x^{(r-1)r^{k-1}}$$

since

$$Q_{r^k}(x) = \frac{x^{r^k} - 1}{Q_1(x)Q_r(x)\cdots Q_{r^{k-1}}(x)} = \frac{x^{r^k} - 1}{x^{r^{k-1}} - 1}$$

by Theorem 8.5 (i). When $k = 1$, we have $Q_r(x) = 1 + x + x^2 + \cdots + x^{r-1}$.

In fact, using the Moebius Inversion Formula, we can establish an explicit formula for the nth cyclotomic polynomial Q_n, for every $n \in \mathbb{N}$.

Theorem 8.8
For a field K of characteristic p and $n \in \mathbb{N}$ not divisible by p, the nth cyclotomic polynomial Q_n over K satisfies

$$Q_n(x) = \prod_{d|n}(x^d - 1)^{\mu(\frac{n}{d})} = \prod_{d|n}(x^{\frac{n}{d}} - 1)^{\mu(d)}.$$

Proof. Apply the multiplicative form of the Mobius Inversion Formula (Theorem 7.12) to the multiplicative group G of non-zero rational functions over K. Take $h(n) = Q_n(x)$ and $H(n) = x^n - 1$ for all $n \in \mathbb{N}$. By Theorem 8.5, the identity (3.3) is satisfied, and so applying Moebius Inversion yields the desired formula. ■

Example 8.9
Let $n = 12$, and let K be any field over which Q_{12} is defined. Then

$$\begin{aligned} Q_{12}(x) &= \prod_{d|12}(x^{\frac{12}{d}} - 1)^{\mu(d)} \\ &= (x^{12} - 1)^{\mu(1)}(x^6 - 1)^{\mu(2)}(x^4 - 1)^{\mu(3)}(x^3 - 1)^{\mu(4)}(x^2 - 1)^{\mu(6)}(x - 1)^{\mu(12)} \\ &= \frac{(x^{12} - 1)(x^2 - 1)}{(x^6 - 1)(x^4 - 1)} = x^4 - x^2 + 1. \end{aligned}$$

Before the next theorem, we make a definition.

Definition 8.10
Let n be a positive integer and b an integer relatively prime to n. Then the least positive integer k such that $n|b^k - 1$ (equivalently, $b^k \equiv 1 \bmod n$) is called the *multiplicative order* of b modulo n, and denoted $\mathrm{ord}_n(b)$.

Example 8.11
(i) $\mathrm{ord}_8(5) = 2$; (ii) $\mathrm{ord}_{31}(2) = 5$; (iii) $\mathrm{ord}_9(4) = 3$.

Theorem 8.12
The cyclotomic field $K^{(n)}$ is a simple algebraic extension of K. Moreover, if $K = \mathbb{F}_q$ with $\gcd(q, n) = 1$, and $d = \mathrm{ord}_n(q)$, then

- *Q_n factors into $\phi(n)/d$ distinct polynomials in $K[x]$ of the same degree d;*
- *$K^{(n)}$ is the splitting field of any such irreducible factor over K;*
- *$[K^{(n)} : K] = d$.*

Proof. If there exists a primitive nth root of unity ζ over K, then $K^{(n)} = K(\zeta)$. Otherwise, we have the situation of Theorem 8.2 (ii); here $K^{(n)} = K^{(m)}$ and the first result again holds.

Now let K be the finite field $\mathbb{F}_q$, assume $\gcd(q, n) = 1$, such that primitive nth roots of unity over $\mathbb{F}_q$ exist. Let η be one of them. Then

$$\eta \in \mathbb{F}_{q^k} \Leftrightarrow \eta^{q^k} = \eta \Leftrightarrow q^k \equiv 1 \bmod n.$$

The smallest positive integer for which this holds is $k = d$, so η is in $\mathbb{F}_{q^d}$ but not in any proper subfield. Thus the minimal polynomial of η over $\mathbb{F}_q$ has degree d. Since η was an arbitrary root of $Q_n(x)$, the result follows, because we can successively divide by the minimal polynomials of the roots of $Q_n(x)$. ■

Example 8.13
Take $q = 11$ and $n = 12$.

- From Example 8.9, we have $K = \mathbb{F}_{11}$ and $Q_{12}(x) = x^4 - x^2 + 1 \in \mathbb{F}_{11}[x]$. We are interested in $K^{(12)}$.
- Since $12 \nmid 11 - 1$ but $12 | 11^2 - 1$, the multiplicative order d of 11 modulo 12 is 2.
- So, $Q_{12}(x)$ factors into $\phi(12)/2 = 4/2 = 2$ monic quadratics, both irreducible over $\mathbb{F}_{11}[x]$, and the cyclotomic field $K^{(12)} = \mathbb{F}_{121}$.
- We can check that the factorization is in fact $Q_{12}(x) = (x^2 + 5x + 1)(x^2 - 5x + 1)$.

The following result, which ties together cyclotomic and finite fields, is very useful.

Theorem 8.14
The finite field $\mathbb{F}_q$ is the $(q-1)$st cyclotomic field over any one of its subfields.

Proof. Since the $q - 1$ non-zero elements of $\mathbb{F}_q$ are all the roots of the polynomial $x^{q-1} - 1$, this polynomial splits in $\mathbb{F}_q$. Clearly, it cannot split in any proper subfield of $\mathbb{F}_q$, so that $\mathbb{F}_q$ is the splitting field of $x^{q-1} - 1$ over any one of its subfields. ■

9 Using cyclotomic polynomials

Cyclotomic fields give us another way of expressing the elements of a finite field $\mathbb{F}_q$. Since $\mathbb{F}_q$ is the $(q-1)$st cyclotomic field over $\mathbb{F}_p$, we can construct it as follows:

- Find the decomposition of the $(q-1)$st cyclotomic polynomial $Q_{q-1} \in \mathbb{F}_p[x]$ into irreducible factors in $\mathbb{F}_p[x]$, which are all of the same degree.
- A root α of any of these factors is a primitive $(q-1)$st root of unity over $\mathbb{F}_p$, and hence a primitive element of $\mathbb{F}_q$.
- For such an α we have $\mathbb{F}_q = \{0, \alpha, \alpha^2, \ldots, \alpha^{q-2}, \alpha^{q-1} = 1\}$.

Example 9.1
Consider the field $\mathbb{F}_9$.

- $\mathbb{F}_9 = \mathbb{F}_3^{(8)}$, the eighth cyclotomic field over $\mathbb{F}_3$.
- As in Example 8.7,
$$Q_8(x) = \frac{x^8 - 1}{x^4 - 1} = x^4 + 1 \in \mathbb{F}_3[x].$$
Its decomposition into irreducible factors in $\mathbb{F}_3[x]$ is
$$Q_8(x) = (x^2 + x + 2)(x^2 + 2x + 2);$$
we have $\phi(8)/\mathrm{ord}_8(3) = 4/2 = 2$ factors of degree 2.
- Let ζ be a root of $x^2 + x + 2$; then ζ is a primitive eighth root of unity over $\mathbb{F}_3$. Hence $\mathbb{F}_9 = \{0, \zeta, \zeta^2, \ldots, \zeta^7, \zeta^8 = 1\}$.

We can now ask: how does this new representation for $\mathbb{F}_9$ correspond to our earlier viewpoint, where $\mathbb{F}_9$ was considered as a simple algebraic extension of $\mathbb{F}_3$ of degree 2, obtained by adjoining a root of an irreducible quadratic?

Example 9.2
Consider the polynomial $f(x) = x^2 + 1 \in \mathbb{F}_3[x]$. This quadratic is irreducible over $\mathbb{F}_3$. So we can build $\mathbb{F}_9$ by adjoining a root α of $f(x)$ to $\mathbb{F}_3$. Then $f(\alpha) = \alpha^2 + 1 = 0$ in $\mathbb{F}_9$, and the nine elements of $\mathbb{F}_9$ are given by $\{0, 1, 2, \alpha, \alpha + 1, \alpha + 2, 2\alpha, 2\alpha + 1, 2\alpha + 2\}$.

Now, note that the polynomial $x^2 + x + 2 \in \mathbb{F}_3[x]$, from Example 9.1, has $\zeta = 1 + \alpha$ as a root. So, the elements in the two representations of $\mathbb{F}_9$ correspond as in the following table

i	ζ^i
1	$1 + \alpha$
2	2α
3	$1 + 2\alpha$
4	2
5	$2 + 2\alpha$
6	α
7	$2 + \alpha$
8	1

Another use of cyclotomic polynomials is that they help us to determine irreducible polynomials.

Theorem 9.3
Let $I(q, n; x)$ be (as in Theorem 7.16) the product of all monic irreducible polynomials in $\mathbb{F}_q[x]$ of degree n. Then for $n > 1$ we have

$$I(q, n; x) = \prod_m Q_m(x),$$

where the product is extended over all positive divisors m of $q^n - 1$ for which n is the multiplicative order of q modulo m, and where $Q_m(x)$ is the mth cyclotomic polynomial over $\mathbb{F}_q$.

Proof.

- For $n > 1$, let S be the set of elements of $\mathbb{F}_{q^n}$ that are of degree n over $\mathbb{F}_q$. Then every $\alpha \in S$ has a minimal polynomial over $\mathbb{F}_q$ of degree n and is therefore a root of $I(q, n; x)$. Conversely, if β is a root of $I(q, n; x)$, then β is a root of some monic irreducible polynomial in $\mathbb{F}_q[x]$ of degree n, implying $\beta \in S$. Thus

$$I(q, n; x) = \prod_{\alpha \in S} (x - \alpha).$$

- If $\alpha \in S$, then $\alpha \in \mathbb{F}_{q^n}^*$, so the order of α in that multiplicative group is a divisor of $q^n - 1$. In fact, the order m of an element of S must be such that n is the least positive integer with $m | q^n - 1$, i.e. $n = \mathrm{ord}_m(q)$. This is because an element $\gamma \in \mathbb{F}_{q^n}^*$ lies in a proper subfield $\mathbb{F}_{q^d}$ if and only if $\gamma^{q^d} = \gamma$, i.e. if and only if the order of γ divides $q^d - 1$.

- For a positive divisor m of $q^n - 1$ which satisfies $n = \mathrm{ord}_m(q)$, let S_m be the set of elements of S of order m. Then S is the disjoint union of the subsets S_m, so we have

$$I(q, n; x) = \prod_m \prod_{\alpha \in S_m} (x - \alpha).$$

  Now, S_m contains precisely all elements of $\mathbb{F}_{q^n}^*$ of order m. So S_m is the set of primitive mth roots of unity over $\mathbb{F}_q$. From the definition of cyclotomic polynomials, we have

$$\prod_{\alpha \in S_m} (x - \alpha) = Q_m(x),$$

  and hence the result follows.

■

Example 9.4
We determine all monic irreducible polynomials in $\mathbb{F}_3[x]$ of degree 2.

- Here $q = 3$ and $n = 2$, so $q^n - 1 = 8$ and $2 = \mathrm{ord}_m(3)$ for divisors $m = 4$ and $m = 8$ of $q^n - 1$. Thus from Theorem 9.3 we have
$$I(3, 2; x) = Q_4(x)Q_8(x).$$
- From Theorem 8.12, we know that $Q_4(x)$ factors into $\phi(4)/2 = 1$ monic irreducible quadratic over $\mathbb{F}_3$, while $Q_8(x)$ factors into $\phi(8)/2 = 2$ monic irreducible quadratics over $\mathbb{F}_3$.
- By Theorem 8.8,
$$Q_4(x) = \prod_{d|4}(x^{\frac{4}{d}} - 1)^{\mu(d)} = \frac{x^4 - 1}{x^2 - 1} = x^2 + 1,$$
while
$$Q_8(x) = x^4 + 1 = (x^2 + x + 2)(x^2 + 2x + 2)$$
as in Example 9.1. Thus the irreducible polynomials in $\mathbb{F}_3[x]$ of degree 2 are x^2+1, x^2+x+2 and $x^2 + 2x + 2$.

Example 9.5
We determine all monic irreducible polynomials in $\mathbb{F}_2[x]$ of degree 4.

- Here $q = 2$ and $n = 4$, so $q^n - 1 = 15$ and $4 = \mathrm{ord}_m(2)$ for divisors $m = 5$ and $m = 15$ of $q^n - 1$. Thus from Theorem 9.3 we have
$$I(2, 4; x) = Q_5(x)Q_{15}(x).$$
- From Theorem 8.12, we know that $Q_5(x)$ factors into $\phi(5)/4 = 1$ monic irreducible quartic over $\mathbb{F}_2$, while $Q_{15}(x)$ factors into $\phi(15)/4 = 8/4 = 2$ monic irreducible quartics over $\mathbb{F}_2$.
- By Theorem 8.8,
$$Q_5(x) = \prod_{d|5}(x^{\frac{5}{d}} - 1)^{\mu(d)} = \frac{x^5 - 1}{x - 1} = x^4 + x^3 + x^2 + x + 1$$
and
$$\begin{aligned} Q_{15}(x) &= \prod_{d|15}(x^{\frac{15}{d}} - 1)^{\mu(d)} \\ &= \frac{(x^{15} - 1)(x - 1)}{(x^5 - 1)(x^3 - 1)} \\ &= \frac{x^{10} + x^5 + 1}{x^2 + x + 1} \\ &= x^8 + x^7 + x^5 + x^4 + x^3 + x + 1. \end{aligned}$$
We note that $Q_5(x+1) = x^4 + x^3 + 1$ is also irreducible in $\mathbb{F}_2[x]$ (since Q_5 is and shifting by a constant preserves irreducibility) and hence must divide $Q_{15}(x)$ (from the above we know that every irreducible polynomial of degree 4 over $\mathbb{F}_2$ either divides Q_5 or Q_{15}), leading to the factorization
$$Q_{15}(x) = (x^4 + x^3 + 1)(x^4 + x + 1).$$
Thus the irreducible polynomials in $\mathbb{F}_2[x]$ of degree 4 are $x^4 + x^3 + x^2 + x + 1$, $x^4 + x^3 + 1$ and $x^4 + x + 1$.

Chapter 5

Automorphisms and bases

10 Automorphisms

In this chapter, we will once again adopt the viewpoint that a finite extension $F = \mathbb{F}_{q^m}$ of a finite field $K = \mathbb{F}_q$ is a vector space of dimension m over K.

In Theorem 7.3 we saw that the set of roots of an irreducible polynomial $f \in \mathbb{F}_q[x]$ of degree m is the set of m distinct elements $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{m-1}}$ of $\mathbb{F}_{q^m}$.

Definition 10.1
Let $\mathbb{F}_{q^m}$ be an extension of $\mathbb{F}_q$ and let $\alpha \in \mathbb{F}_{q^m}$. The elements $\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}$ are called the *conjugates* of α with respect to $\mathbb{F}_q$.

Remark 10.2

- *The conjugates of $\alpha \in \mathbb{F}_{q^m}$ with respect to $\mathbb{F}_q$ are distinct if and only if the minimal polynomial g of α over $\mathbb{F}_q$ has degree m.*
- *Otherwise, the degree d of the minimal polynomial g of α over $\mathbb{F}_q$ is a proper divisor of m, and in this case the conjugates of α with respect to $\mathbb{F}_q$ are the distinct elements $\alpha, \alpha^q, \ldots, \alpha^{q^{d-1}}$, each repeated m/d times.*

Theorem 10.3
The conjugates of $\alpha \in \mathbb{F}_q^*$ with respect to any subfield of $\mathbb{F}_q$ have the same order in the group $\mathbb{F}_q^*$.

Proof. Apply Theorem 1.13 to the cyclic group $\mathbb{F}_q^*$, using the fact that every power of the characteristic of $\mathbb{F}_q$ is coprime to the order $q-1$ of $\mathbb{F}_q^*$. ■

This immediately implies the following observation.

Corollary 10.4
If α is a primitive element of $\mathbb{F}_{q^m}$, then so are all its conjugates with respect to $\mathbb{F}_q$.

Example 10.5
Expressing $\mathbb{F}_4$ as $\mathbb{F}_2(\theta) = \{0, 1, \theta, \theta+1\}$, where $\theta^2 + \theta + 1 = 0$, we saw in Example 6.11 that θ is a primitive element of $\mathbb{F}_4$. The conjugates of $\theta \in \mathbb{F}_4$ with respect to $\mathbb{F}_2$ are θ and θ^2; from Example 6.11, $\theta^2 = \theta + 1$ is also a primitive element.

Example 10.6
Let $\alpha \in \mathbb{F}_{16}$ be a root of $f = x^4 + x + 1 \in \mathbb{F}_2[x]$. Then the conjugates of α with respect to $\mathbb{F}_2$ are $\alpha, \alpha^2, \alpha^4 = \alpha + 1, \alpha^8 = \alpha^2 + 1$, and all of these are primitive elements of $\mathbb{F}_{16}$. The conjugates of α with respect to $\mathbb{F}_4$ are α and $\alpha^4 = \alpha + 1$.

We next explore the relationship between conjugate elements and certain automorphisms of a finite field.

Definition 10.7
An *automorphism of* $\mathbb{F}_{q^m}$ *over* $\mathbb{F}_q$ is an automorphism σ of $\mathbb{F}_{q^m}$ which fixes the elements of $\mathbb{F}_q$ pointwise. Thus, σ is a one-to-one mapping from $\mathbb{F}_{q^m}$ onto itself with

$$\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$$

and

$$\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$$

for all $\alpha, \beta \in \mathbb{F}_{q^m}$ and

$$\sigma(a) = a \text{ for all } a \in \mathbb{F}_q.$$

This definition may look familiar to anyone who has studied Galois theory!

Theorem 10.8
The distinct automorphisms of $\mathbb{F}_{q^m}$ *over* $\mathbb{F}_q$ *are precisely the mappings* $\sigma_0, \sigma_1, \ldots, \sigma_{m-1}$ *defined by*

$$\sigma_j(\alpha) = \alpha^{q^j}$$

for $\alpha \in \mathbb{F}_{q^m}$ *and* $0 \leq j \leq m-1$.

Proof. We first establish that the mappings σ_j are automorphisms of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$.

- For each σ_j and all $\alpha, \beta \in \mathbb{F}_{q^m}$, we have $\sigma_j(\alpha\beta) = \sigma_j(\alpha)\sigma_j(\beta)$ and $\sigma_j(\alpha + \beta) = \sigma_j(\alpha) + \sigma_j(\beta)$ by Freshmen's Exponentiation, so clearly σ_j is an endomorphism of $\mathbb{F}_{q^m}$.
- Since $\sigma_j(\alpha) = 0 \Leftrightarrow \alpha = 0$, σ_j is injective. Since $\mathbb{F}_{q^m}$ is a finite set, σ_j is also surjective, and hence is an automorphism of $\mathbb{F}_{q^m}$.
- We have $\sigma_j(a) = a$ for all $a \in \mathbb{F}_q$ by Lemma 6.3, and so each σ_j is an automorphism of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$.
- The mappings $\sigma_1, \ldots, \sigma_{m-1}$ are distinct as they return distinct values for a primitive element of $\mathbb{F}_{q^m}$.

Now, suppose σ is an arbitrary automorphism of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$; we show that it is in fact σ_j for some $0 \leq j \leq m-1$.

Let β be a primitive element of $\mathbb{F}_{q^m}$ and let $f = x^m + a_{m-1}x^{m-1} + \cdots + a_0 \in \mathbb{F}_q[x]$ be its minimal polynomial over $\mathbb{F}_q$. Then

$$\begin{aligned} 0 &= \sigma(\beta^m + a_{m-}\beta^{m-1} + \cdots + a_0) \\ &= \sigma(\beta)^m + a_{m-1}\sigma(\beta)^{m-1} + \cdots + a_0, \end{aligned}$$

so that $\sigma(\beta)$ is a root of f in $\mathbb{F}_{q^m}$. By Theorem 7.3, we must have $\sigma(\beta) = \beta^{q^j}$ for some j, $0 \leq j \leq m-1$. Since σ is a homomorphism and β primitive, we get that $\sigma(\alpha) = \alpha^{q^j}$ for all $\alpha \in \mathbb{F}_{q^m}$. ■

Hence the conjugates of $\alpha \in \mathbb{F}_{q^m}$ are obtained by applying all automorphisms of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ to the element α.

Remark 10.9
The automorphisms of $\mathbb{F}_{q^m}$ *over* $\mathbb{F}_q$ *form a group under composition of mappings, called the* Galois group *of* $\mathbb{F}_{q^m}$ *over* $\mathbb{F}_q$ *and denoted* $\mathrm{Gal}(\mathbb{F}_{q^m}/\mathbb{F}_q)$. *From Theorem 10.8, this group of automorphisms is a cyclic group of order* m, *generated by* σ_1.

11 Traces and Norms

Let $F = \mathbb{F}_{q^m}$ and $K = \mathbb{F}_q$. We introduce a mapping from F to K which turns out to be K-linear.

Definition 11.1
For $\alpha \in F$, the *trace* $\mathrm{Tr}_{F/K}(\alpha)$ of α over K is defined by

$$\begin{aligned}\mathrm{Tr}_{F/K}(\alpha) &= \text{sum of conjugates of } \alpha \text{ w.r.t. } K \\ &= \alpha + \alpha^q + \alpha^{q^2} + \cdots + \alpha^{q^{m-1}}\end{aligned}$$

If K is the prime subfield of F, i.e. $K = \mathbb{F}_p$ where p is the characteristic of F, then $\mathrm{Tr}_{F/K}(\alpha)$ is called the absolute trace of α and denoted simply by $\mathrm{Tr}(\alpha)$.

A useful alternative way to think of the trace is as follows.

Definition 11.2
Let $\alpha \in F$ and $f \in K[x]$ be the minimal polynomial of α over K; its degree d is a divisor of $m = [F : K]$. Then $g = f^{m/d} \in K[x]$ is called the *characteristic polynomial* of α over K.

By Theorem 7.3, the roots of f in F are $\alpha, \alpha^q, \ldots, \alpha^{q^{d-1}}$; from Remark 10.2, the roots of g in F are precisely the conjugates of α with respect to K. So

$$\begin{aligned}g &= x^m + a_{m-1}x^{m-1} + \cdots + a_0 \\ &= (x-\alpha)(x-\alpha^q)\cdots(x-\alpha^{q^{m-1}}).\end{aligned}$$

Comparing coefficients we see that

$$\mathrm{Tr}_{F/K}(\alpha) = -a_{m-1}.$$

In particular, $\mathrm{Tr}_{F/K}(\alpha)$ must be an element of K.

Theorem 11.3
Let $K = \mathbb{F}_q$ and let $F = \mathbb{F}_{q^m}$. Then the trace function $\mathrm{Tr}_{F/K}$ satisfies the following properties.

(i) $\mathrm{Tr}_{F/K}(\alpha+\beta) = \mathrm{Tr}_{F/K}(\alpha) + \mathrm{Tr}_{F/K}(\beta)$ *for all* $\alpha, \beta \in F$*;*

(ii) $\mathrm{Tr}_{F/K}(c\alpha) = c\mathrm{Tr}_{F/K}(\alpha)$ *for all* $c \in K$, $\alpha \in F$*;*

(iii) $\mathrm{Tr}_{F/K}$ *is a linear transformation from* F *onto* K *(both viewed as* K *vector spaces);*

(iv) $\mathrm{Tr}_{F/K}(a) = ma$ *for all* $a \in K$*;*

(v) $\mathrm{Tr}_{F/K}(\alpha^q) = \mathrm{Tr}_{F/K}(\alpha)$ *for all* $\alpha \in F$*.*

Proof. (i) For $\alpha, \beta \in F$, Freshmen's Exponentiation yields

$$\begin{aligned}\mathrm{Tr}_{F/K}(\alpha+\beta) &= \alpha + \beta + (\alpha+\beta)^q + \cdots + (\alpha+\beta)^{q^{m-1}} \\ &= \alpha + \beta + \alpha^q + \beta^q + \cdots + \alpha^{q^{m-1}} + \beta^{q^{m-1}} \\ &= \mathrm{Tr}_{F/K}(\alpha) + \mathrm{Tr}_{F/K}(\beta).\end{aligned}$$

(ii) By Lemma 6.3, for $c \in K$ we have $c^{q^i} = c$ for all $i \geq 0$. Then for $\alpha \in F$,

$$\begin{aligned}\mathrm{Tr}_{F/K}(c\alpha) &= c\alpha + c^q\alpha^q + \cdots + c^{q^{m-1}}\alpha^{q^{m-1}} \\ &= c\alpha + c\alpha^q + \cdots + c\alpha^{q^{m-1}} \\ &= c\mathrm{Tr}_{F/K}(\alpha).\end{aligned}$$

(iii) For all $\alpha \in F$ we have $\mathrm{Tr}_{F/K}(\alpha) \in K$; this follows from the discussion above, or immediately from

$$\begin{aligned}(\mathrm{Tr}_{F/K}(\alpha))^q &= (\alpha + \alpha^q + \cdots + \alpha^{q^{m-1}})^q \\ &= \alpha^q + \cdots + \alpha^{q^{m-1}} + \alpha \\ &= \mathrm{Tr}_{F/K}(\alpha).\end{aligned}$$

Combining this with (i) and (ii) shows that $\mathrm{Tr}_{F/K}$ is a K-linear transformation from F into K. To show that it is surjective, it suffices to demonstrate that there exists some $\alpha \in F$ with $\mathrm{Tr}_{F/K}(\alpha) \neq 0$. We have $\mathrm{Tr}_{F/K}(\alpha) = 0 \Leftrightarrow \alpha$ is a root of $x^{q^{m-1}} + \cdots + x^q + x \in K[x]$ in F; since this polynomial has at most q^{m-1} roots in F whereas F has q^m elements, the result follows.
(iv) By Lemma 7.3, $a^{q^i} = a$ for all $a \in K$ and $i \geq 0$, and the result follows.
(v) For $\alpha \in F$ we have $\alpha^{q^m} = \alpha$, and so

$$\mathrm{Tr}_{F/K}(\alpha^q) = \alpha^q + \alpha^{q^2} + \cdots + \alpha^{q^m} = \mathrm{Tr}_{F/K}(\alpha).$$

■

In fact, the trace function provides a description for all linear transformations from F into K, in the following sense.

Theorem 11.4
Let F be a finite extension of the finite field K (both viewed as vector spaces over K). Then the K-linear transformations from F into K are precisely the mappings L_β ($\beta \in F$) given by $L_\beta(\alpha) = \mathrm{Tr}_{F/K}(\beta\alpha)$ for all $\alpha \in F$. Moreover, if α, β are distinct elements of F then $L_\alpha \neq L_\beta$.

Proof. Omitted. Idea: $(\alpha, \beta) \mapsto \mathrm{Tr}_{F/K}(\alpha\beta)$ is a symmetric non-degenerate bilinear form on the K-vectorspace F. ■

For a chain of extensions, we have the following rule.

Theorem 11.5 (Transitivity of trace)
Let K be a finite field, let F be a finite extension of K and E a finite extension of F. Then

$$\mathrm{Tr}_{E/K}(\alpha) = \mathrm{Tr}_{F/K}(\mathrm{Tr}_{E/F}(\alpha))$$

for all $\alpha \in E$.

Proof. Let $K = \mathbb{F}_q$, let $[F : K] = m$ and let $[E : F] = n$, so that $[E : K] = mn$ by Theorem 5.6. For $\alpha \in E$,

$$\begin{aligned}\mathrm{Tr}_{F/K}(\mathrm{Tr}_{E/F}(\alpha)) &= \sum_{i=0}^{m-1} \mathrm{Tr}_{F/K}(\alpha)^{q^i} \\ &= \sum_{i=0}^{m-1} (\sum_{j=0}^{n-1} \alpha^{q^{jm}})^{q^i} \\ &= \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \alpha^{q^{jm+i}} \\ &= \sum_{k=0}^{mn-1} \alpha^{q^k} = \mathrm{Tr}_{E/K}(\alpha).\end{aligned}$$

■

The multiplicative analogue of the trace function is called the norm.

Definition 11.6
For $\alpha \in F = \mathbb{F}_{q^m}$ and $K = \mathbb{F}_q$, the *norm* $N_{F/K}(\alpha)$ of α over K is defined by

$$\begin{aligned} N_{F/K}(\alpha) &= \text{product of conjugates of } \alpha \text{ w.r.t. } K \\ &= \alpha \cdot \alpha^q \cdots \alpha^{q^{m-1}} \\ &= \alpha^{(q^m-1)/(q-1)}. \end{aligned}$$

Comparing this definition with the characteristic polynomial g of α over K, as before, we see that

$$N_{F/K}(\alpha) = (-1)^m a_0.$$

In particular, $N_{F/K}(\alpha)$ is always an element of K.

Theorem 11.7
Let $K = \mathbb{F}_q$, and F its degree m extension. The norm function $N_{F/K}$ satisfies the following properties:

(i) $N_{F/K}(\alpha\beta) = N_{F/K}(\alpha)N_{F/K}(\beta)$ *for all* $\alpha, \beta \in F$;

(ii) $N_{F/K}$ *maps* F *onto* K *and* F^* *onto* K^*;

(iii) $N_{F/K}(a) = a^m$ *for all* $a \in K$;

(iv) $N_{F/K}(\alpha^q) = N_{F/K}(\alpha)$ *for all* $\alpha \in F$.

Proof. (i) Immediate from definition of norm.
(ii) From above, $N_{F/K}$ maps F into K; since $N_{F/K}(\alpha) = 0 \Leftrightarrow \alpha = 0$, we have that $N_{F/K}$ maps F^* into K^*.

We must now show that $N_{F/K}$ is surjective. By (i), $N_{F/K}$ is a homomorphism between the multiplicative groups F^* and K^*. The elements of the kernel are the roots of $x^{\frac{q^m-1}{q-1}} - 1 \in K[x]$ in F; denoting the order of the kernel by d, we have $d \leq \frac{q^m-1}{q-1}$. By the First Isomorphism Theorem, the image has order $(q^m - 1)/d$, which is at least $q - 1$. So $N_{F/K}$ maps F^* onto K^* and hence F onto K.
(iii) Result is immediate upon noting that, for $a \in K$, all conjugates of a are equal to a.
(iv) By (i), $N_{F/K}(\alpha^q) = N_{F/K}(\alpha)^q$; by (ii), $N_{F/K}(\alpha) \in K$ and so $N_{F/K}(\alpha)^q = N_{F/K}(\alpha)$. ■

Theorem 11.8 (Transitivity of Norm)
Let K be a finite field, let F be a finite extension of K and let E be a finite extension of F. Then

$$N_{E/K}(\alpha) = N_{F/K}(N_{E/F}(\alpha))$$

for all $\alpha \in E$.

Proof. Let $[F : K] = m$ and $[E : F] = n$. Then for $\alpha \in E$,

$$\begin{aligned} N_{F/K}(N_{E/F}(\alpha)) &= N_{F/K}(\alpha^{\frac{q^{mn}-1}{q^m-1}}) \\ &= (\alpha^{\frac{q^{mn}-1}{q^m-1}})^{\frac{q^m-1}{q-1}} \\ &= \alpha^{\frac{q^{mn}-1}{q-1}} = N_{E/K}(\alpha). \end{aligned}$$

■

12 Bases and the Normal Basis Theorem

We first consider two important, and very natural, kinds of bases.

Recall that $\mathbb{F}_{q^m} = \mathbb{F}_q(\alpha) \cong \mathbb{F}_q[x]/(f)$, where f is an irreducible polynomial of degree m and α is a root of f in $\mathbb{F}_{q^m}$. So, every element of $\mathbb{F}_{q^m}$ can be uniquely expressed as a polynomial in α over $\mathbb{F}_q$ of degree less than m and hence, for any defining element α, the set $\{1, \alpha, \ldots, \alpha^{m-1}\}$ is a basis for $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$.

Definition 12.1
Let $K = \mathbb{F}_q$ and $F = \mathbb{F}_{q^m}$.

A *polynomial basis* of F over K is a basis of the form $\{1, \alpha, \alpha^2, \ldots, \alpha^{m-1}\}$, where α is a defining element of F over K.

We can always insist that the element α is a primitive element of F, since by Theorem 6.12, every primitive element of F can serve as a defining element of F over K.

Example 12.2
Let $K = \mathbb{F}_3$ and $F = \mathbb{F}_9$. Then F is a simple algebraic extension of K of degree 2, obtained by adjoining an appropriate θ to K. Let θ be a root of the irreducible polynomial $x^2 + 1 \in K[x]$; then $\{1, \theta\}$ is a polynomial basis for F over K. However, θ is not primitive since $\theta^4 = 1$. Now let α be a root of $x^2 + x + 2$; then $\{1, \alpha\}$ is another polynomial basis for F over K, and α is a primitive element of F.

Definition 12.3
Let $K = \mathbb{F}_q$ and $F = \mathbb{F}_{q^m}$. A *normal basis* of F over K is a basis of the form $\{\alpha, \alpha^q, \ldots, \alpha^{q^{m-1}}\}$, consisting of a suitable element $\alpha \in F$ and all its conjugates with respect to K. Such an α is called a *free* or *normal* element.

Example 12.4
Let $K = \mathbb{F}_2$ and $F = \mathbb{F}_8$. Let $\alpha \in \mathbb{F}_8$ be a root of the irreducible polynomial $x^3 + x^2 + 1$ in $\mathbb{F}_2[x]$. Then $B = \{\alpha, \alpha^2, 1 + \alpha + \alpha^2\}$ is a basis of $\mathbb{F}_8$ over $\mathbb{F}_2$. Since $\alpha^4 = 1 + \alpha + \alpha^2$, this is in fact a normal basis for F over K. To the contrary, let $\beta \in F$ be a root of the irreducible polynomial $x^3 + x + 1 \in K[x]$. Then the conjugates $\{\beta, \beta^2, \beta^2 + \beta\}$ of β do not form a basis of K.

We now ask: does a normal basis exist for every F and K?
We require two lemmas before proving the main result.

Lemma 12.5 (Artin Lemma)
Let $\chi_1, \ldots, \chi_m$ be distinct homomorphisms from a group G into the multiplicative group F^* of an arbitrary field F, and let $a_1, \ldots, a_m \in F$, not all zero. Then for some $g \in G$ we have

$$a_1\chi_1(g) + \ldots + a_m\chi_m(g) \neq 0.$$

Proof. The proof is by induction on m. We omit the details. ■

Next, we recall a few concepts from linear algebra.

Definition 12.6

- If T is a linear operator on a finite dimensional vector space V over an arbitrary field K, then a polynomial $f = a_n x^n + \cdots + a_1 x + a_0 \in K[x]$ is said to *annihilate* T if $a_n T^n + \cdots + a_1 T + a_0 I = 0$, where I and 0 are the identity and zero operator on V, respectively.
- The uniquely determined monic polynomial of least degree with this property is called the *minimal polynomial* for T. It divides any other polynomial in $K[x]$ which annihilates T.

- The *characteristic polynomial* g for T is given by $g := \det(xI - T)$. It is a monic polynomial of degree $n = \dim(V)$; by the Cayley-Hamilton theorem it annihilates T (and hence is divisible by the minimal polynomial). In fact, the roots of the two polynomials are the same up to multiplicity.
- A vector $\alpha \in V$ is called a *cyclic vector* for T if the vectors $T^k\alpha$, $k = 0, 1, \ldots$ span V.

We are now ready for the second lemma.

Lemma 12.7
Let T be a linear operator on the finite-dimensional vector space V. Then T has a cyclic vector if and only if the characteristic and minimal polynomials for T are identical.

Proof. Omitted. ■

Theorem 12.8 (Normal Basis Theorem)
For any finite field K and any finite extension F of K, there exists a normal basis of F over K.

Proof. Let $K = \mathbb{F}_q$ and $F = \mathbb{F}_{q^m}$ with $m \geq 2$.

- From Theorem 10.8, the distinct automorphisms of F over K are given by
$$\epsilon, \sigma, \sigma^2, \ldots, \sigma^{m-1},$$
where ϵ is the identity map on F, $\sigma(\alpha) = \alpha^q$ for $\alpha \in F$ and σ^i means composing σ with itself i times.
- Since $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$ and $\sigma(c\alpha) = c\sigma(\alpha)$ for $\alpha, \beta \in F$ and $c \in K$, we can think of σ as a linear operator on the vector space F over K.
- Since $\sigma^m = \epsilon$, the polynomial $x^m - 1 \in K[x]$ annihilates σ. Consider $\epsilon, \sigma, \sigma^2, \ldots, \sigma^{m-1}$ as endomorphisms of F^*, and apply the Artin Lemma; this tells us that no nonzero polynomial in $K[x]$ of degree less than m annihilates σ. Thus, $x^m - 1$ is the minimal polynomial for the linear operator σ.
- Since the characteristic polynomial for σ is a monic polynomial of degree m divisible by the minimal polynomial for σ, we must have that $x^m - 1$ is the characteristic polynomial also.
- By Lemma 12.7, there must exist a cyclic vector for V; i.e. there exists some $\alpha \in F$ such that $\alpha, \sigma(\alpha), \sigma^2(\alpha), \ldots$ span F.
- Dropping repeated elements, this says that $\alpha, \sigma(\alpha), \ldots, \sigma^{m-1}(\alpha)$ span F, and hence form a basis of F over K. Since this basis consists of an element and its conjugates with respect to K, it is a normal basis, as required!

■

In fact, it turns out that this result can be strengthened, in the following way.

Theorem 12.9 (Primitive Normal Basis Theorem)
For any finite extension F of a finite field K, there exists a normal basis of F over K that consists of primitive elements of F.

Proof. Beyond the scope of this course! ■

Example 12.10
Let $K = \mathbb{F}_2$ and $F = \mathbb{F}_8 = K(\alpha)$, where $\alpha^3 + \alpha^2 + 1 = 0$. We saw in Example 12.4 that the basis $B = \{\alpha, \alpha^2, 1 + \alpha + \alpha^2 = \alpha^4\}$ from Example 12.4 is a normal basis for F over K. In fact, α is a primitive element of F (F^* is the cyclic group of order 7 and hence any non-identity element is a generator). So this is a primitive normal basis for F over K.

Chapter 6

Applications to Cryptography

13 The Digital Signature Algorithm

In this section we present an application of finite fields to implement digital signatures.

Definition 13.1
Let G be a cyclic group and $x \in G$ a generator. The *discrete logarithm problem* (DLP) is the following: Given $y \in G$, find $k \in \mathbb{Z}$ with $y = x^k$ and $0 \leq k < |G|$.

Remark 13.2
Since x is a generator, such a k always exists and is uniquely defined.

Remark 13.3
The difficulty of the DLP depends on the representation *of G. For example if $G = \mathbb{Z}/m\mathbb{Z}$ with addition as operation, then for $x = 1$ the result k is just equal to y, so it can be read off readily from the input. Even for other x with $\gcd(x, m) = 1$ it is an easy calculation to find k for any given y. However, if $G = \mathbb{F}_q^*$ for some prime power q with multiplication as the operation, then the DLP is much more difficult.*

13.1 The scheme

Preparations:

- Choose a prime q (typically $\approx 2^{160}$).
- Choose another prime p with $p \equiv 1 \pmod{q}$ (typically $\approx 2^{1024}$).
- Choose $g \in \mathbb{F}_p$ of order q (possible since $q \mid p - 1$)
  (One can take a primitive root $h \in \mathbb{F}_p$ and set $g := h^{(p-1)/q} \pmod{p}$).
- Publish (p, q, g) to all participating parties.

Public and secret keys:

- Choose a random $x \in \mathbb{N}$ with $0 < x < q$.
- Compute $y := g^x \pmod{p}$.
- The public key is (p, q, g, y), the secret key is x.
- Note that computing x from y is an instance of the discrete logarithm problem.

Signing documents:

- Let m be a document, use a *secure hash function* H to compute $H(m) \in \{0, 1, \ldots, q-1\}$.
- Generate a random $k \in \mathbb{Z}$ with $0 < k < q$ (different for each document).
- Compute $r := (g^k \pmod p) \pmod q$.
- Compute $s := k^{-1}(H(m) + xr) \pmod q$.
- The pair (r, s) is the signature of m.

Verification of signatures:

- Given m, r and s (and the publicly available (p, q, g)):
- Compute $w := s^{-1} \pmod q$.
- Compute $u_1 := (H(m) \cdot w) \pmod q$.
- Compute $u_2 := (rw) \pmod q$.
- Compute $v := ((g^{u_1} \cdot y^{u_2}) \pmod p) \pmod q$.
- If $v = r$ then the signature is successfully verified, otherwise it is rejected.

Correctness: By the definition of w and s it follows that $k \equiv H(m)w + xrw \pmod q$. Thus

$$g^k \equiv g^{H(m)w+xrw} \equiv g^{H(m)w} \cdot y^{rw} \equiv g^{u_1} \cdot y^{u_2} \pmod p.$$

Therefore $r \equiv v \pmod q$.

Faking signatures with discrete logarithms:
Provided you can solve the DLP in $\mathbb{F}_p$ you can just compute x from the published p, q, y and sign happily. Note that you are actually just looking at the abelian group of order q here, realised as a subgroup of the multiplicative group of $\mathbb{F}_p$.

Important security measure:
Note that k must be kept secret under all circumstances! Namely, knowing k gives rise to the following simple attack:

$$x \equiv (s \cdot k - H(m)) \cdot r^{-1} \pmod q$$

Even knowing that the same k was used twice producing signatures (r, s) for m and (r', s') for m' leads to an attack:

$$\begin{aligned} x &\equiv (s \cdot k - H(m)) \cdot r^{-1} \pmod q \\ x &\equiv (s' \cdot k - H(m')) \cdot r'^{-1} \pmod q \end{aligned}$$

gives

$$(s \cdot k - H(m)) \cdot r^{-1} \equiv (s' \cdot k - H(m')) \cdot r'^{-1} \pmod q$$

and this allows to compute k and thus x.
Even knowing a few bits of the binary expansion of k in a few signatures could be enough to break the signature scheme.

13.2 Cryptographic hash functions

What properties should the cryptographic hash function H have?
(also called "digital fingerprints" or "digital checksums")

- H takes a block of data of arbitrary length and returns a fixed length result (e.g. 160 bit).
- H is necessarily **not injective**.
- $H(m)$ is easily (and efficiently) computable for any given message m.
- It is **infeasible** to find an m with a given $H(m)$ ("preimage resistance").
- Given m (and thus $H(m)$), it is **infeasible** to modify m to m' such that $H(m) = H(m')$ ("second preimage resistance").
- It is **infeasible** to find two different messages m_1 and m_2 with $H(m)$ ("collision resistance").

In summary: A cryptographic hash function should be as random as possible, while still being deterministic and efficiently computable.

Remark 13.4
Note that the length of the result gives an approximate upper bound for the strength of the preimage resistance and second preimage resistance. The half of the length of the result gives an approximate upper bound for the strength of the collision resistance due to the birthday paradox. "Difficulty" or "infeasibility" is usually measured by complexity theory (upper bound for the runtime in terms of the input size).

13.3 Algorithms to solve the DLP

Let $G = \langle x \rangle$ be a cyclic group of order n. What is known about solving the DLP?

Open Problem 13.5
For the discrete logarithm problem there is no general algorithm known whose runtime is bounded from above by a polynomial in the input size.

Algorithm 13.6
*The trivial method is **trial multiplication**: Given $y \in G$, simply try $k = 1, 2, \ldots, n$, compute x^k and compare with y. If $n = |G| \approx 2^\ell$, then storing a group element needs approximately ℓ bits. In the worse case, this needs $\approx n$ multiplications, which is exponential in ℓ.*

Algorithm 13.7
*A fairly simple extension of this is the **Baby-step Giant-step** algorithm: Let $m := \lceil \sqrt{n} \rceil$, imagine $k = im + j$. Compute and store all x^j for $1 \leq j \leq m$, if y is among them the solution is found. Otherwise, compute $y, yx^m, yx^{2m}, \ldots$ (each with one multiplication) and try to find it amongst the stored x^j. This finds the solution after at most $2m$ multiplications but uses memory for m group elements.*

Algorithm 13.8
*The memory requirements can be reduced by **Pollard's rho algorithm for discrete logarithms**: Assume we know $y = x^k$ and x and want to find k. We try to find $a, b, A, b \in \mathbb{Z}$ with $x^a y^b = x^A y^B$, then*

$$(b - B)k = (A - a) \pmod{n},$$

which can be computed easily.

To this end, let $N := \{0, 1, 2, \ldots, n-1\}$ *and define a function* $f : N \times N \to N \times N$ *that "jumps randomly". Let* $(a_{i+1}, b_{i+1}) := f(a_i, b_i)$ *for* $i > 1$ *and* $(a_1, b_1) \in N \times N$*. We try to find a cycle in the sequence* $(g_i := x^{a_i} y^{b_i})_{i \in \mathbb{N}}$ *of group elements. Provided* f *really behaves randomly, we expect a cycle after approximately* $\sqrt{\pi n/2}$ *steps. Let* ℓ *be the cycle length and* s *be the least integer with* $g_i = g_{i+\ell}$ *for all* $i \geq s$*.*

We look for ℓ *and* s *as follows (Floyd's cycle finding algorithm): We have* $g_i = g_{i+\ell j}$ *for all* $i \geq s$ *and all* $j \geq 0$*. In particular, if* $i = \ell j$ *we have* $g_i = g_{2i}$*. So we compute* g_i *and* g_{2i} *for* $i = 1, 2, \ldots$ *until we find equality. Note that we do not have to store the intermediate values, each step requires only* 3 *applications of* f*. Having found an* i *with* $g_i = g_{2i}$*, we know that* i *is a multiple of* λ*. Now start from the beginning and look for* $j \in \mathbb{N}$ *with* $g_j = g_{j+i}$*, this finds* s*. Then look for* $k \in \mathbb{N}$ *with* $g_s = g_{s+k}$ *to find* ℓ*.*

Remark 13.9
Both these methods are generic, they do not use any special representation of the cyclic group. Both have a runtime of $\approx \sqrt{n}$*, so they are still exponential in the input length* $\log n$*.*

Remark 13.10
There is an efficient procedure by Peter Shor that uses a (so for non-existing) ***quantum computer****.*