Mathematical preliminaries of crypto
Finite fields

Sven Laur

April 12, 2017

1 What are finite fields and what are they good for?

In the high school you probably have studied many arithmetic laws and how to apply them in order to compute or solve equations. In practice, most of them do not hold if you consider floating point numbers. For instance, $(a+b)+c \neq a+(b+c)$ when b and c are near the floating point precision and a is moderately large number. We also often treat a digital document as a raw sequence of bits and then perform some sort of computational algorithm on top of them. In many cases, we would like to organise computations in such way that the rules we learnt in the high school still hold. In particular, we need that the order of additions and multiplications should be irrelevant:

$$\begin{aligned} a+b &= b+a & \qquad a \cdot b &= b \cdot a \\ (a+b)+c &= a+(b+c) & \qquad (a \cdot b) \cdot c &= a \cdot (b \cdot c) \end{aligned}$$

and we could zero terms from sums and unit factors from products:

$$\begin{aligned} a+0 &= a & \qquad a \cdot 1 &= a \\ 0+a &= a & \qquad 1 \cdot a &= a \end{aligned}$$

and negative numbers and inverses behave as they should:

$$a+(-a) = 0 \qquad\qquad a \cdot (a^{-1}) = 1 \ .$$

Finally, we should be able to open brackets as usual:

$$\begin{aligned} (a+b) \cdot c &= a \cdot c + b \cdot c \ , \\ a \cdot (b+c) &= a \cdot c + a \cdot c \ . \end{aligned}$$

Assume that we somehow manage to define a data structure[1] $(\mathcal{R}; +, \cdot)$ together with addition and multiplication operations that all these assumptions are satisfied. Then we have defined a *ring*. To be precise, each element a should have an opposite element $-a$

[1] In object-oriented programming terms it is a class with operations $+$ and $\cdot$.

but not all elements should have inverses. In fact, zero can have inverse only if $0 = 1$, since the equation

$$1 = 0 \cdot 0^{-1} = (0 + 0) \cdot 0^{-1} = (1 \cdot 0 + 1 \cdot 0) \cdot 0^{-1} = (1 + 1) \cdot 0 \cdot 0^{-1} = 2$$

assures that

$$0 = 1 + (-1) = (1 + 1) - 1 = 1 \ .$$

The latter again implies that each element is zero:

$$a = a \cdot 1 = a \cdot 0 = 0$$

and thus $\mathcal{R} = \{0\}$ is rather degenerate data structure. As the ability to divide numbers is rather important, we could require that all non-zero elements should have inverses. A ring $(\mathcal{F}; +, \cdot)$ that satisfies this additional assumption is called *field*. A field is finite if it contains finite number of elements. Note that rational numbers $(\mathbb{Q}; +, \cdot)$ form a infinite field. Real and complex numbers are also examples of infinite fields.

1.1 Examples of finite fields

The simplest way to define a set that satisfies all restrictions specified above (*field axioms*) is to take some data structure with addition and multiplication defined over it and throw out all elements that violate some of the axioms. Of course, this approach can be unsuccessful, as we might have to throw out all elements. For instance, we cannot find any subsets of rational numbers that is both finite and still satisfies all field axioms. The element 1 is culprit. If we cannot drop 1 and thus elements

$$\mathbb{Z} = \{0, 1, -1, 1 + 1, -1 - 1, \ldots\}$$

must be in our set. However, as the set of all integers $\mathbb{Z}$ is infinite, our set must be infinite, as well. Hence, a field can be finite only if the set of all integers $\mathbb{Z}$ is finite.

This seems rather absurd requirement. However, you have already seen such a data structure with addition and multiplication defined over it so that $\mathbb{Z}$ maps to finite number of elements. Namely, we can consider the set residues modulo n, often denoted as $\mathbb{Z}_n$. Formally, the addition and multiplication is defined as follows:

$$a + b = a + b \mod n \qquad\qquad a \cdot b = a \cdot b \mod n \ .$$

Previous exercise sessions showed also that modular arithmetic follows all field axioms except not all elements have inverses. Hence, subsets of residue rings $\mathbb{Z}_n$ seems good candidates for finding finite fields.

Examples. Residue ring $\mathbb{Z}_2 = \{0, 1\}$ is a finite field as the inverse of 1 is 1. Residue ring $\mathbb{Z}_3 = \{0, 1, 2\}$ is also a finite field, as 2^{-1} is 1:

$$2 \cdot 2 = 4 \mod 3 = 1 \ .$$

With the ring $\mathbb{Z}_4$ we encounter some problems. Namely, the element 2 does not have an inverse: $2 \cdot 1 = 2$, $2 \cdot 2 = 0$ and $2 \cdot 3 = 2$. We cannot get a field form $\mathbb{Z}_4$ by dropping some elements, since we cannot drop 1 and its multiples generate all elements $\mathbb{Z}_4$. We can generalise this observation or other residue rings as well.

Lemma 1. *A residue ring $\mathbb{Z}_n$ is either a finite field or none of its subsets is a finite fields wrt modular addition and multiplication.*

Proof. Let $\mathcal{F} \subseteq \mathbb{Z}_n$. Then $1 \in \mathcal{F}$ and thus also $1 + 1 \in \mathcal{F}, 1 + 1 + 1 \in \mathcal{F}, \ldots$. As a result $\mathbb{Z}_n \subseteq \mathcal{F}$ and the claim follows. □

The next natural question what we can ask is when $\mathbb{Z}_n$ is a finite field. By the field axioms any non-zero element must have inverse. However, if n is a composite number then one of its factors p cannot gave an inverse. Indeed, let $px = 1 \mod n$ then

$$0 = px \mod p = (px \mod n) \mod p = 1 \mod p = 1$$

and we have derived a contradiction. Hence the following theorem holds.

Theorem 1. *A residue ring $\mathbb{Z}_n$ is a finite field if and only if n is a prime.*

Proof. We proved above that composite integer n cannot create a finite field. For the converse, recall that if two integers a, b are coprime then there exists $u, v \in \mathbb{Z}$ such that $ua + vb = 1$. If n is a prime number then any residue $a \in \{1, \ldots, n-1\}$ is coprime with n. Consequently, we can represent $1 = ua + vn$ and thus $ua = 1 \mod n$. □

Exercise 1. *Recall that according to Fermat Little theorem $a^p = a \mod p$ for any prime number. Describe an alternative inversion algorithm that is based on this idea. How many multiplications are needed to invert an element with this algorithm?*

2 Alternative ways to define finite fields

Another way to define finite fields is to start with a data structure for which the addition is already defined and then define multiplication so that the resulting structure would satisfy all field axioms. For instance, recall that bitwise xor operation satisfies all field axioms that are connected to addition. Hence, the set of two bit integers $\mathcal{B}_2 = \{00, 01, 10, 11\}$ would form a finite field if we could come up with a multiplication operation so that the remaining field axioms are satisfied. One potential candidate for multiplication operation is bitwise and operation, as $a \wedge b = b \wedge a$ and $(a \wedge b) \wedge c = a \wedge (b \wedge c)$. Unfortunately, bitwise and is not distributive, since

$$10 \wedge (01 \oplus 11) = 10 \wedge 01 = 00$$
$$10 \wedge 01 \oplus 10 \wedge 11 = 00 \oplus 10 = 10 \ .$$

However, if we define multiplication in the following way

$\times$	00	01	10	11
00	00	00	00	00
01	00	01	10	11
10	00	10	11	01
11	00	11	01	10

Then it is not hard though extremely tedious to verify that resulting data structure $(\mathcal{B}_1; \oplus, \times)$ satisfies all field axioms.

Exercise 2. *Verify that the data structure* $(\mathcal{B}_2; \oplus, \times)$ *satisfies all field axioms. Which of those axioms are automatically satisfied?*

Lets study this data structure a bit more to unravel the magic behind the definition of multiplication operation. Let $\alpha = 10$ and $1 = 01$ as 01 acts like the unit in multiplication. Then it is easy to express all field elements as linear combinations of α and 1.

Combination	0	1	α	$\alpha + 1$
Binary representation	00	01	10	11

The multiplication table forces following substitutions:

$$\alpha^2 = \alpha \oplus 1 \qquad (\alpha \oplus 1)^2 = \alpha \qquad \alpha \times (\alpha \oplus 1) = 1 \ .$$

If we open brackets and move all term to left-hand side[2], we get the same equation

$$\alpha^2 \oplus \alpha \oplus 1 = 0$$

after simplifications. To be punctual, note that $2 \times x = x \oplus x = 0$ and $x \oplus x = 0$ thus we can always drop even multiples and ignore signs.

From this insight it is easy to verify that the multiplication table captures arithmetic with polynomials where the result is reduced modulo $\alpha^2 \oplus \alpha \oplus 1$:

$$\begin{aligned} \alpha \times \alpha = \alpha^2 &= \alpha^2 \oplus (\alpha^2 \oplus \alpha \oplus 1) = \alpha \oplus 1 \\ \alpha \times (\alpha + 1) = \alpha^2 \oplus \alpha &= \alpha^2 \oplus \alpha \oplus (\alpha^2 \oplus \alpha \oplus 1) = 1 \\ (\alpha \oplus 1) \times (\alpha + 1) = \alpha^2 \oplus 1 &= \alpha^2 \oplus 1 \oplus (\alpha^2 \oplus \alpha \oplus 1) = \alpha \ . \end{aligned}$$

The latter allows us to declare that $(\mathcal{B}_2; \oplus, \times)$ is actually defined as a set of polynomials

$$\mathbb{Z}_2[\alpha]/(\alpha^2 \oplus \alpha \oplus 1) = \{0, 1, \alpha, \alpha + 1\}$$

where $\mathbb{Z}_2$ means that we consider all polynomial coefficients modulo 2 and $(\alpha^2 \oplus \alpha \oplus 1)$ under division slash means that whenever the degree of a polynomial is larger or equal to 2 we compute residue of the polynomial modulo $\alpha^2 \oplus \alpha \oplus 1$. Such a formalisation means that most of the field axioms are automatically satisfied.

[2]We can do high school math as all field axioms are satisfied.

Definition 1. *Let $\mathbb{Z}_n$ be a residue ring. Then $\mathbb{Z}_n[\alpha] = \{0, 1, \ldots, \alpha, 2\alpha, \ldots\}$ denotes the set of all polynomials with coefficients from $\mathbb{Z}_n$. Addition and multiplication with these polynomials is defined as usual, except all coefficients of resulting polynomials are reduced modulo n. Let $p(\alpha) \in \mathbb{Z}_n[\alpha]$. Then $\mathbb{Z}_2[\alpha]/(p(\alpha))$ is set of all polynomials over $\mathbb{Z}_n$ that are reduced modulo $p(\alpha)$. Similarly, addition and multiplication with these polynomials is defined as in $\mathbb{Z}_n[\alpha]$, except resulting polynomials are reduced modulo $p(\alpha)$.*

Lemma 2. *The set $\mathbb{Z}_2[\alpha]/(p(\alpha))$ is ring for any polynomial $p(\alpha) \in \mathbb{Z}_n[\alpha]$.*

In other words, the easiest way to define a field is to choose a integer modulus n and a polynomial $p(\alpha)$ and then check whether all non-zero polynomials in $\mathbb{Z}_n[\alpha]/(p(\alpha))$ are invertible or not. After that one has to choose how to represent these polynomials in binary. The easiest way is to store polynomials by their coefficient vectors.

Note that not all polynomials lead to a field. Let to exhaustive classification of first and second degree polynomials over $\mathbb{Z}_2$, i.e., addition is defined as bitwise xor. There are two linear polynomials α and $\alpha \oplus 1$. Corresponding rings consist of two elements

$$\mathbb{Z}_2[\alpha]/(\alpha) = \{0, 1\} = \mathbb{Z}_2 \qquad \mathbb{Z}_2[\alpha]/(\alpha + 1) = \{0, 1\} = \mathbb{Z}_2 \ ,$$

since reminder of first degree polynomial is free term. There are four quadratic polynomials over $\mathbb{Z}_2$:

$$\alpha^2 \ , \alpha^2 \oplus 1 \ \ \alpha^2 \oplus \alpha \ , \alpha^2 \oplus \alpha \oplus 1 \ .$$

Note that the first three can be easily factored:

$$\begin{aligned} \alpha^2 &= \alpha \times \alpha \\ \alpha^2 \oplus 1 &= (\alpha \oplus 1) \times (\alpha \oplus 1) \\ \alpha^2 \oplus \alpha &= \alpha \times (\alpha \oplus 1) \ . \end{aligned}$$

Note that if a polynomial $p(\alpha)$ can be expressed as non-trivial multiple of two other polynomials $p(\alpha) = u(\alpha) \times v(\alpha)$, then $u(\alpha)$ is non-invertible modulo $p(\alpha)$. Indeed, if

$$u(\alpha) \times q(\alpha) = 1 \mod p(\alpha)$$

then also

$$u(\alpha) \times q(\alpha) = 1 \mod u(\alpha) \ .$$

That is a contradiction as $u(\alpha) = 0 \mod u(\alpha)$. It is common to call polynomials *irreducible* if they cannot be represented as a non-trivial multiple of two polynomials. To be punctual, we require that the degree of both factors is at least one. This avoids trivial factorisations where one of factors is one or any other scalar. Now we are ready to state and prove the analog of Theorem 1.

Theorem 2. *A ring $\mathbb{Z}_n[\alpha]/(p(\alpha)))$ is a finite field if and only if n is a prime and $p(\alpha)$ is an irreducible polynomial.*

Proof. First note that any non-zero scalar $c \in \mathbb{Z}_n$ is invertible only if $\mathbb{Z}_n$ is a field and thus n must be a prime. We already saw that factors of a reducible polynomial cannot be inverted. To prove the theorem, we must prove that any non-zero polynomial has its inverse when $p(\alpha)$ is irreducible. For that recall that if two polynomials $a(\alpha), b(\alpha)$ are coprime over $\mathbb{Z}_n$ then there exists $u(\alpha), v(\alpha) \in \mathbb{Z}_n[\alpha]$ such that $u(\alpha)a(\alpha) + v(\alpha)b(\alpha) = 1$ over $\mathbb{Z}_n$. If $p(\alpha)$ is irreducible then any polynomial with smaller degree is coprime with $p(\alpha)$ over $\mathbb{Z}_n$. Consequently, we can represent $1 = u(\alpha)a(\alpha) + v(\alpha)p(\alpha)$ and thus $u(\alpha)a(\alpha) = 1 \mod p(\alpha)$ over $\mathbb{Z}_n$. □

More importantly, all alternative ways for defining finite fields are analogous and thus we can concentrate our attention to irreducible polynomials. We will later find out that all finite of the same size are identical up to element representation. Thus, finite fields are often denoted by the symbol $\mathbb{F}_{p^k}$ where p is a prime and p^k is the size of the field.

Theorem 3. *In any finite field there exists an element α and a polynomial $p(\alpha)$ such that addition and multiplication tables are defined according to $\mathbb{Z}_n[\alpha]/(p(\alpha))$.*

Exercise 3. *Find all irreducible polynomials of degree 3 and 4 over $\mathbb{Z}_2$. Construct the corresponding multiplication tables.*

Sieve of Eratosthenes. Irreducible polynomials are the analog of prime numbers. In particular, if a polynomial is not reducible then some smaller degree irreducible polynomial must divide it. More precisely if the degree of a polynomial is n then it is sufficient to consider all irreducible polynomials up to degree $\lfloor n/2 \rfloor$.

Example. To find all irreducible polynomials of degree five over $\mathbb{Z}_2$ we must first find all irreducible polynomials up to degree 2. These are

$$\alpha, \qquad \alpha \oplus 1, \qquad \alpha^2 \oplus \alpha \oplus 1 \ .$$

Now lets consider all 16 polynomials of degree 5. Half of them have the free term zero and thus divide by α and cannot be irreducible. Secondly, note that $\alpha \oplus 1$ divided polynomial $p(\alpha)$ if $p(1) = 0$. The latter can happen only if the number of nonzero monomials is even. As a result, we should consider only polynomials with odd number of monomials:

$$\begin{array}{ll}
\alpha^5 \oplus \alpha^4 \oplus 1 & \alpha^5 \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha \ \oplus 1 \\
\alpha^5 \oplus \alpha^3 \oplus 1 & \alpha^5 \oplus \alpha^4 \oplus \alpha^2 \oplus \alpha \ \oplus 1 \\
\alpha^5 \oplus \alpha^2 \oplus 1 & \alpha^5 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \ \oplus 1 \\
\alpha^5 \oplus \alpha \ \oplus 1 & \alpha^5 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha^2 \oplus 1
\end{array}$$

Now note that $\alpha^2 \oplus \alpha \oplus 1$ divides only two of them:

$$\begin{aligned}
\alpha^5 \oplus \alpha^4 \oplus 1 &= (\alpha^3 \oplus \alpha \oplus 1) \times (\alpha^2 \oplus \alpha \oplus 1) \\
\alpha^5 \oplus \alpha \oplus 1 &= (\alpha^3 \oplus \alpha^2 \oplus 1) \times (\alpha^2 \oplus \alpha \oplus 1)
\end{aligned}$$

and thus there are six irreducible polynomials:

$$\alpha^5 \oplus \alpha^3 \oplus 1 \qquad \alpha^5 \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha \ \oplus 1 \qquad \alpha^5 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \ \oplus 1$$
$$\alpha^5 \oplus \alpha^2 \oplus 1 \qquad \alpha^5 \oplus \alpha^4 \oplus \alpha^2 \oplus \alpha \ \oplus 1 \qquad \alpha^5 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha^2 \oplus 1$$

which all give a rise to a 32-element finite fields.

3 Computations in finite fields

3.1 Addition and multiplication

For clarity, let us consider 8-element finite field which is defined by the irreducible polynomial $\alpha^3 \oplus \alpha \oplus 1$. Then any computation is just arithmetic with polynomials with the exception that the end result must be reduced modulo $\alpha^3 \oplus \alpha \oplus 1$. For instance,

$$\begin{aligned}(\alpha^2 \oplus \alpha \oplus 1)(\alpha^2 \oplus 1) &= \alpha^4 \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha^2 \oplus \alpha \oplus 1 \\ &= \alpha^4 \oplus \alpha^3 \oplus \oplus\alpha \oplus 1 = \alpha^2 \oplus \alpha \ .\end{aligned}$$

3.2 Computing inverses

Extended Euclid algorithm. The easiest way to invert an element is to use extended Euclid algorithm. Let $a(\alpha)$ and $b(\alpha)$ be two polynomials. Then for the extended Euclid algorithm we must divide elements in the following manner:

$$\begin{aligned} a(\alpha) &= q_1(\alpha)b(\alpha) + r_1(\alpha) \\ b(\alpha) &= q_2(\alpha)r_1(\alpha) + r_2(\alpha) \\ r_1(\alpha) &= q_3(\alpha)r_2(\alpha) + r_3(\alpha) \\ &\cdots \\ r_n(\alpha) &= q_{n+2}(\alpha)r_{n+1}(\alpha) + r_{n+2}(\alpha) \\ r_{n+1}(\alpha) &= q_{n+3}(\alpha)r_{n+2}(\alpha) \end{aligned}$$

and then express $r_{n+2}(\alpha)$ as $r_{n+2}(\alpha) = r_n(\alpha) - q_{n+2}(\alpha)r_{n+1}(\alpha)$. In most cases, terms $r_n(\alpha)$ and $r_{n+1}(\alpha)$ are not $a(\alpha)$ and $b(\alpha)$. Hence, we must express them from the previous equations. As a result, we start to substitute one by one equations:

$$\begin{aligned} r_{n+1}(\alpha) &= r_{n-1}(\alpha) - q_{n+1}(\alpha)r_n(\alpha) \\ &\cdots \\ r_3(\alpha) &= r_1(\alpha) - q_3(\alpha)r_2(\alpha) \\ r_2(\alpha) &= b(\alpha) - q_2(\alpha)r_1(\alpha) \\ r_1(\alpha) &= a(\alpha) - q_1(\alpha)b(\alpha) \ . \end{aligned}$$

Although this process is very error prone to carry out with pen and pencil, it leads to a final expression

$$r_{n+1} = u(\alpha)a(\alpha) + v(\alpha)b(\alpha) \ .$$

Computing inverses. To invert a field element $a(\alpha)$ modulo $p(\alpha)$ you have to carry out the extended Euclid algorithm for $a(\alpha)$ and $b(\alpha) = p(\alpha)$. The inverse is $u(\alpha)$.

Example. Let us consider the specific finite field $\mathbb{F}_{2^8} = \mathbb{Z}_2[\alpha]/(\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1)$, which is used in the Advanced Encryption Standard. To find inverses of α and $\alpha^7 \oplus \alpha^3 \oplus \alpha$, we must find polynomials $u_1(\alpha), v_1(\alpha), u_2(\alpha), v_2(\alpha)$ over $\mathbb{Z}_2$ such that

$$\begin{aligned} \alpha \times u_1(\alpha) \oplus (\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1) \times v_1(\alpha) &= 1 \\ (\alpha^7 \oplus \alpha^3 \oplus \alpha) \times u_2(\alpha) \oplus (\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1) \times v_2(\alpha) &= 1 \end{aligned}$$

Let us use the Euclidean algorithm for that. Since

$$\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1 = \alpha(\alpha^7 \oplus \alpha^3 \oplus \alpha^2 \oplus 1) \oplus 1$$

over $\mathbb{Z}_2$, we can express

$$\alpha \times (\alpha^7 \oplus \alpha^3 \oplus \alpha^2 \oplus 1) \oplus (\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1) \times 1 = 1 \ .$$

The polynomial $u_1(\alpha) = \alpha^7 \oplus \alpha^3 \oplus \alpha^2 \oplus 1$ is the inverse of α, as

$$\alpha \times (\alpha^7 \oplus \alpha^3 \oplus \alpha^2 \oplus 1) = \alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha = 1 \ .$$

To find the second inverse note

$$\begin{aligned} \alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1 &= \alpha \cdot (\alpha^7 \oplus \alpha^3 \oplus \alpha) \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1 \\ \alpha^7 \oplus \alpha^3 \oplus \alpha &= (\alpha^4 \oplus \alpha^3) \cdot (\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1) \oplus \alpha \\ \alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1 &= (\alpha^2 \oplus \alpha \oplus 1) \times \alpha \oplus 1 \end{aligned}$$

and thus

$$\begin{aligned} 1 &= 1 \times (\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1) \oplus (\alpha^2 \oplus \alpha \oplus 1) \times \alpha \\ \alpha &= 1 \times (\alpha^7 \oplus \alpha^3 \oplus \alpha) \oplus (\alpha^4 \oplus \alpha^3) \times (\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1) \\ \alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1 &= 1 \times (\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1) \oplus \alpha \times (\alpha^7 \oplus \alpha^3 \oplus \alpha) \end{aligned}$$

Systematic substitution of terms allows us to represent 1 as a linear combination of terms $\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1$ and $\alpha^7 \oplus \alpha^3 \oplus \alpha$:

$$\begin{aligned} 1 &= 1 \times (\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1) \oplus (\alpha^2 \oplus \alpha \oplus 1) \times ((\alpha^7 \oplus \alpha^3 \oplus \alpha) \oplus (\alpha^4 \oplus \alpha^3) \times (\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1)) \\ &= (\alpha^2 \oplus \alpha \oplus 1) \times (\alpha^7 \oplus \alpha^3 \oplus \alpha) \oplus (\alpha^6 \oplus \alpha^3 \oplus 1)(\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1) \\ &= (\alpha^2 \oplus \alpha \oplus 1) \times (\alpha^7 \oplus \alpha^3 \oplus \alpha) \oplus (\alpha^6 \oplus \alpha^3 \oplus 1) \times ((\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1) \oplus \alpha(\alpha^7 \oplus \alpha^3 \oplus \alpha)) \\ &= (\alpha^6 \oplus \alpha^3 \oplus 1) \times (\alpha^8 \oplus \alpha^4 \oplus \alpha^3 \oplus \alpha \oplus 1) \oplus (\alpha^7 \oplus \alpha^4 \oplus \alpha^2 \oplus 1) \times (\alpha^7 \oplus \alpha^3 \oplus \alpha) \end{aligned}$$

and thus $(\alpha^7 \oplus \alpha^3 \oplus \alpha)^{-1} = \alpha^7 \oplus \alpha^4 \oplus \alpha^2 \oplus 1$.

Exercise 4. *Consider a finite field $\mathbb{F}_{32}$ specified by the polynomial $\alpha^5 \oplus \alpha^3 \oplus 1$. Find inverses of all elements that contain α^4 in their representation.*

4 Multiplicative group in finite fields

The set of invertible elements is closed under multiplication and under inverse, since multiple of invertible elements is invertible. Let us consider the set of invertible elements $\mathbb{F}_8^*$ when the field is specified by the polynomial $\alpha^3 \oplus \alpha \oplus 1$. Recall that order of any element must divide the group order. As there are 7 elements in $\mathbb{F}_8^*$ the order of any element must be either 1 or 7. For instance,

$$\begin{aligned} \alpha^3 &= \alpha \oplus 1 & \alpha^5 &= \alpha^2 \oplus \alpha \oplus 1 & \alpha^7 &= 1 \\ \alpha^4 &= \alpha^2 \oplus \alpha & \alpha^6 &= \alpha^2 + 1 \end{aligned}$$

and thus α generates the entire multiplicative group. An irreducible polynomial $p(\alpha)$ is *primitive* if α generates the entire multiplicative group.

A priori it is not clear that there always exist an invertible element that generates the group and that there exists a primitive polynomial of any degree. For instance, if we consider $\mathbb{F}_{16}$ specified by the polynomial $\alpha^4 \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1$. Then it is easy to see that

$$\begin{aligned} \alpha^5 = \alpha \times (\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1) = \alpha^4 \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha = 1 \\ (\alpha^3 \oplus \alpha^2)^3 = \alpha^6 \times (\alpha \oplus 1)^3 = \alpha \times (\alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1) = 1 \end{aligned}$$

and thus there are indeed elements that have order 5 and order 3. It is relatively easy to prove the following lemma.

Lemma 3. *If two group elements x and y have coprime orders p and q then their multiple xy has order pq. If two group elements x and y have orders p and q then there exists an element z which has order pq.*

This lemma is the key ingredient that allows us to prove that there exists always an element that generates the entire multiplicative group.

Theorem 4. *Let $\mathbb{F}_{p^k}$ be finite field. Then there exist an element $\xi \in \mathbb{F}_{p^k}$ that generates the entire multiplicative group.*

Sketch. First, note that we can find an element such that the orders of other invertible elements divide it. If this is not the case, then we can use Lemma 3 to resolve violations one by one. Let r be the final order of this element. Then obviously all elements from the field are roots of a polynomial $X \times (X^r - 1)$. By the Bezout theorem this polynomial can have at most $r+1$ distinct roots over any field. Consequently, $r = p^k - 1$ or otherwise the field must contain less than p^k elements. The claim follows as we have shown that $r = p^k - 1$. □

To illustrate the constructive element in the proof recall that $\text{ord}(\alpha) = 5$ and $\text{ord}(\alpha^3 \oplus \alpha^2) = 3$ in our example field. Thus the order of $\alpha^4 \oplus \alpha^3 = \alpha^2 \oplus \alpha \oplus 1$ must be 15 according to Lemma 3. To prove this, we must compute the third and fifth power of $\alpha^2 \oplus \alpha \oplus 1$:

$$\begin{aligned} (\alpha^2 \oplus \alpha \oplus 1)^3 = \alpha^3(\alpha^3 \oplus \alpha^2)^3 = \alpha^3 \\ (\alpha^2 \oplus \alpha \oplus 1)^5 = \alpha^5(\alpha^3 \oplus \alpha^2)^2 = \alpha^2 \oplus \alpha \oplus 1 \end{aligned}$$

indeed the order of $\alpha^2 \oplus \alpha \oplus 1$ is 15.

Exercise 5. *Find all elements of* $\mathbb{F}_{16}$ *that generate the entire multiplicative group if the field is specified by the polynomial* $\alpha^4 \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1$.

5 Subfields

A large field can contain a smaller field. For instance $\mathbb{F}_{2^k}$ always contains $\mathbb{F}_2 = \{0, 1\}$ as a sub-field, since addition and multiplication of scalars produces scalars. More generally, a set $\mathcal{F} \subseteq \mathbb{F}_{p^k}$ is a subfield if it is closed under multiplication and addition and it satisfies all field axioms. The simplest way to define a sub-field is to find an non-invertible element that generates the entire multiplicative group of a sub-field. Such an element must exists, since the multiplicative group is always generated by a single element. For example, $\mathbb{F}_{16}$ specified by the polynomial $\alpha^4 \oplus \alpha^3 \oplus \alpha^2 \oplus \alpha \oplus 1$ might contain subfields of size 2 and 4, as elements in the multiplicative group have orders 1, 3, 5 and field sizes must be a power of 2. Indeed, powers of $\alpha^3 \oplus \alpha^2$ together with zero generate $\mathbb{F}_4$, since

$$\begin{aligned}(\alpha^3 \oplus \alpha^2)(\alpha^3 \oplus \alpha^2) &= \alpha^3 \oplus \alpha \oplus 1\\ (\alpha^3 \oplus \alpha^2)(\alpha^3 \oplus \alpha \oplus 1) &= 1 \ .\end{aligned}$$

Obviously, we could use raw encoding of these polynomials 0000, 0001, 1100 and 1011 to represent $\mathbb{F}_4$. However, this representation is wasteful.

Exercise 6. *Prove that* $\mathbb{F}_{p^\ell}$ *is a sub-field of* $\mathbb{F}_{p^k}$ *if and only if* $p^\ell - 1$ *divides* $p^k - 1$. ***Hint:*** *Consider the orders of generators.*

Exercise 7. *Which fields* $\mathbb{F}_4, \ldots, \mathbb{F}_{256}$ *that contain subfields of size more than two?*

6 Isomorphisms and automorphisms

Only the field $\mathbb{F}_4$ is uniquely determined as there is a single quadratic irreducible polynomial with binary coefficients. A cubic polynomial is irreducible if neither zero nor one are roots. There are exactly two such polynomials[3] $\alpha^3 \oplus \alpha \oplus 1$ and $\beta^3 \oplus \beta^2 \oplus 1$. Both of them give a rise to 8 element fields. It is obvious to ask whether these two structures:

- $\mathcal{F}$ a set of cubic polynomials $f(\alpha)$ with reduction rules $2 = 0$ and $\alpha^3 \oplus \alpha \oplus 1 = 0$;
- $\mathcal{G}$ a set of cubic polynomials $g(\beta)$ with reduction rules $2 = 0$ and $\beta^3 \oplus \beta^2 \oplus 1 = 0$;

define different objects or not. Two objects are equivalent if computations done in one structure yield the same results as in the other structure. Of course, one cannot take this literally, as the first structure consists of polynomials of α and the second structure consists of polynomials of β. Therefore, we need a relabelling function $\kappa : \mathcal{F} \to \mathcal{G}$. It is easy to see that if a one-to-one function κ respects all field operations:

$$\begin{aligned}\kappa(0) &= 0 & \kappa(a+b) &= \kappa(a) + \kappa(b)\\ \kappa(1) &= 1 & \kappa(a \times b) &= \kappa(a) \times \kappa(b)\end{aligned} \tag{1}$$

[3]Here, we deliberately use different variables so that we would not later confuse ourselves.

then it provides the desired relabelling, as it does not matter whether we do computations before or after applying κ. To find such transformation from $\mathcal{F}$ to $\mathcal{G}$, it is sufficient to define the mapping $\alpha \mapsto s(\beta)$, since the equations (1) allow us to determine the behaviour of κ on all other elements. However, the mapping $\alpha \mapsto p(\beta)$ cannot be arbitrary, since α is a root of a polynomial $X^3 + X + 1$. Consequently, $s(\beta)^3 \oplus s(\beta) \oplus 1 = 0$ in the field $\mathcal{G}$. This is a general observation.

Theorem 5. *Let $p(\alpha)$ and $q(\beta)$ be irreducible polynomials that determine the fields of same size. Then a homomorphic mapping κ defined by $\alpha \mapsto s(\beta)$ is correctly defined if and only if $s(\beta)$ is a root of polynomial p.*

Proof. Necessity is evident, as $p(\alpha) = 0$ and thus $\kappa(p(\alpha)) = p(s(\beta)) = \kappa(0) = 0$. For sufficiency, note that the condition $\kappa(p(\alpha)) = 0$ also assures that the mapping is correctly defined, i.e., two equivalent representations $r(\alpha)$ and $r(\alpha) + t(\alpha) \times p(\alpha)$ are mapped to the same value, since

$$\kappa(r(\alpha) + t(\alpha) \times p(\alpha)) = \kappa(r(\alpha)) + \kappa(t(\alpha)) \times \kappa(p(\alpha)) = \kappa(r(\alpha)) \ .$$

□

Coming back to our example, note that $\beta \oplus 1$, $\beta^2 \oplus 1$ and $\beta^2 \oplus \beta$ are the roots of the polynomial in the field $\mathcal{G}$:

$$\begin{aligned}(\beta \oplus 1)^3 \oplus (\beta \oplus 1) \oplus 1 &= \beta^3 \oplus \beta^2 \oplus \beta \oplus 1 \oplus \beta \oplus 1 \oplus 1 = 0\\ (\beta^2 \oplus 1)^3 \oplus (\beta^2 \oplus 1) \oplus 1 &= \beta^2 \oplus \beta^2 \oplus 1 \oplus 1 = 0\\ (\beta^2 \oplus \beta)^3 \oplus (\beta^2 \oplus \beta) \oplus 1 &= \beta^2 \oplus \beta \oplus 1 \oplus \beta^2 \oplus \beta \oplus 1 = 0\end{aligned}$$

and thus polynomials $\beta \oplus 1$, $\beta^2 \oplus 1$ and $\beta^2 \oplus \beta$ are the only candidates for the mapping $\alpha \mapsto s(\beta)$. The first map is clearly invertible. For the others it takes a while to see that they are invertible, as well. They have to be invertible, since all mapped elements are closed under the addition and multiplication—either they cover all field elements or form a sub-field. As the multiplicative group has size 7 and none of the mappings map all non-zero elements to 1, the image cannot be a sub-field. Thus, all three maps are indeed relabellings which respect all field operations. These maps are called *isomorphisms*.

There is yet another curious thing to note. Note that $\alpha \oplus 1$ and $\alpha^2 \oplus \alpha$ are also the roots of the polynomial $X^3 + X + 1$ and thus correspondences $\alpha \mapsto \alpha \oplus 1$ and $\alpha \mapsto \alpha^2 \oplus 1$ define isomorphisms from $\mathcal{F}$ to $\mathcal{F}$. Such isomorphisms are *automorphisms*.

Exercise 8. *Our reasoning about the invertibility of mappings was indirect. Find correspondences $\beta \mapsto t(\alpha)$ that define direct inverses for all three mappings defined by correspondences $\alpha \mapsto \beta \oplus 1$, $\alpha \mapsto \beta^2 \oplus 1$ and $\alpha \mapsto \beta^2 \oplus \beta$.*

Exercise 9. *We know that $\mathbb{F}_4$ is a subfield of $\mathbb{F}_{16}$. Let $\mathbb{F}_4$ be specified by the polynomial $\alpha^2 \oplus \alpha \oplus 1$ and $\mathbb{F}_{16}$ by $\beta^4 \oplus \beta^3 \oplus \beta^2 \oplus \beta \oplus 1$. Define the corresponding isomorphism that maps elements from $\mathbb{F}_4$ to $\mathbb{F}_{16}$ and find its inverse.*