A BRIEF NOTE ON FINITE FIELDS

NAMAN KUMAR

ABSTRACT. We provide a brief primer on the introductory theory of finite fields from an abstract perspective, aimed primarily at computer scientists.

1. INTRODUCTION

We start by defining the notion of a field.

Definition 1.1. *A **field** is a set F together with two operations $(+, \cdot)$, which satisfies the following properties (termed as the field axioms)*[1]*:*

(1) *Associativity:* $a + (b + c) = (a + b) + c$, *and* $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
(2) *Commutativity:* $a + b = b + a$, $a \cdot b = b \cdot a$.
(3) *Additive Identity: There exists some* $0 \in F$ *such that* $a + 0 = a$ *for all* $a \in F$.
(4) *Multiplicative Identity: There exists some* $1 \in F$ *such that* $1 \cdot a = a$ *for all* $a \in F$, *where* $1 \neq 0$.
(5) *Additive Inverse: For each* $a \in F$, *there exists a unique* $-a$ *such that* $(-a) + a = 0$.
(6) *Multiplicative Inverse: For each nonzero* $a \in F$ *there exists a unique* a^{-1} *such that* $a \cdot a^{-1} = 1$.
(7) *Distributivity:* $a \cdot (b + c) = a \cdot b + a \cdot c$.

Definition 1.2. *A **finite field** is a field $(F, +, \cdot)$ such that the set F is finite.*

Note. For the remainder of this article, we will let 'field' refer primarily to finite fields; the theorems are not meant to be taken in their full generality, but rather restricted to the finite case.

1.3. **The field $\mathbb{F}_p$.** The stereotypical example of a finite field is the field $\mathbb{Z}/p\mathbb{Z}$, where p is prime. We denote this field by $\mathbb{F}_p$. Note that this field has exactly p elements, and every element vanishes under multiplication by p (or 0 in the context of the field). All the axioms are easy to verify, except the existence of a multiplicative inverse. To show this, we use Bézout's Identity.

Lemma 1.4 (Bézout's Identity)**.** *Let $a, b \in \mathbb{Z}$ and $a, b \neq 0$. Then there exist $x, y \in \mathbb{Z}$ such that $ax + by = \gcd(a, b)$.*

Proof. The proof follows from the correctness of the Euclidean Division Algorithm, which allows x and y to be computed explicitly. ☐

OREGON STATE UNIVERSITY

Date: October 18, 2024.

[1] A **commutative ring with unity** is a set $(R, +, \cdot)$ which satisfies all the axioms except the existence of multiplicative inverses. We will colloquially refer to R as a ring.

With this in hand, we take $a \in \mathbb{F}_p$ with $a \neq 0$ and p. By the previous identity it follows that there exists some x, y such that $ax + py = 1$, since $\gcd(a, p) = 1$ for $a \in \mathbb{F}_p$. Then it immediately follows that $\exists x$ such that $ax \equiv 1 \pmod{p}$, which is the multiplicative inverse.

In fact, we can go further: we can show that a^{p-2} is an inverse of a.

Theorem 1.5 (Fermat's Little Theorem)**.** *For prime p and all $a \in \mathbb{N}^+$, $a^{p-1} \equiv a \pmod{p}$.*

Proof. Consider the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$. This contains all a in $\mathbb{Z}/p\mathbb{Z}$ except 0. By Lagrange's theorem, the order of a must divide $|(\mathbb{Z}/p\mathbb{Z})^\times| = p - 1$. Hence, $a^{p-1} = a$ in $(\mathbb{Z}/p\mathbb{Z})^\times$, and thus $a^{p-1} = a \pmod{p}$. □

It immediately follows that a^{p-2} is the inverse of a in $\mathbb{F}_p$.

1.6. **Polynomial Rings.** Consider a ring R. Then $R[x]$ is the ring of polynomials with coefficients in R; this can be very easily verified to be a ring, and in particular, supports the euclidean division algorithm. We will briefly consider how to construct more rings from $R[x]$.

Definition 1.7 (Quotient Rings of Polynomials)**.** *Let $R[x]$ be a polynomial ring and $p(x) \in R[x]$ be some polynomial. Consider the set of equivalence classes $R[x]/p(x)$, where elements $a, b \in R[x]$ are said to be in the same equivalence class (designated as $a \sim b$) if $p(x) \mid a - b$; we denote the equivalence class as $[\![a]\!]$. Let $[\![a]\!] + [\![b]\!] = [\![a+b]\!]$ and $[\![a]\!] \cdot [\![b]\!] = [\![a \cdot b]\!]$. Then this set forms a ring.*

The above can be easily checked, we leave it as an exercise. It is easy to see that $F[x]$ is not a field even if F is a field. A natural question is whether $F[x]/p(x)$ can be a field: in general this is false. A counterexample is the ring $\mathbb{Z}/2\mathbb{Z}[x]/(x^2)$. The elements of this ring are $\{0, 1, x, x + 1\}$. However, x has no multiplicative inverse: $x \cdot 1 = x$, $x \cdot x = x^2 = 0$, and $x \cdot (x + 1) = x^2 + x = x$. However, $\mathbb{Z}/p\mathbb{Z}[x]/(x + 1)$ is a field (check)!

1.8. **Adjoining Elements.** We now see a way to construct fields. Consider the field $\mathbb{Q}$ of the rationals. Define the new field $\mathbb{Q}(\sqrt{2})$ as

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}.$$

Let addition and multiplication be defined in the obvious way. Seeing that this is a field is not too hard − the only nontriviality is seeing that a multiplicative inverse exists, but this can be easily found by rationalizing the denominator of the fractional inverse and obtained as

$$\frac{a - b\sqrt{2}}{a^2 - b^2}$$

in the closed form.

What exactly is this field? It is the field $\mathbb{Q}$ *adjoined* with an additional element, $\sqrt{2}$. We define $\sqrt{2}$ to be the element which is a root of the polynomial $x^2 - 2$. Consider now the ring $\mathbb{Q}[x]/(x^2 - 2)$. A way to interpret this ring is to consider all polynomials of the form $a(x) = (x^2 - 2)q(x) + r(x)$ and then send $x^2 - 2 \rightsquigarrow 0$, ie. $a(x) \sim r(x)$. Note that sending $x^2 - 2 \rightsquigarrow 0$ is equivalent to sending $x^2 \rightsquigarrow 2$, and thus $\mathbb{Q}[x]/(x^2 - 2)$ can be viewed as all polynomials in $\mathbb{Q}[x]$ with x^2 sent to 2. For example, $p(x) = x^3 + 2x^2 + x + 4 = (x^2)x + 2(x^2) + x + 4 \rightsquigarrow 2x + 4 + x + 4 = 3x + 8$. Thus, the equivalence class of $p(x)$ is $3x + 8$.

In particular, note that $x \cdot x = x^2 = 2$. Hence, $\mathbb{Q}[x]/(x^2 - 2)$ can be written as

$$\{a + bx : a, b \in \mathbb{Q}\}$$

where $x^2 = 2$. It follows that $x \mapsto \sqrt{2}$ is an isomorphism, and $\mathbb{Q}(\sqrt{2}) \cong \mathbb{Q}[x]/(x^2 - 2)$. We will see that whenever $p(x)$ is a monic irreducible polynomial in $F[x]$, then $F[x]/p(x)$ is a field which can also be interpreted as $F(\alpha)$ where α is an adjoint element that serves as a root of $p(x)$.

Theorem 1.9. *If $p(x)$ is a monic irreducible polynomial, then $F[x]/p(x)$ is a field*[2]*.*

Proof. The fact that $F[x]/p(x)$ is a ring is immediate; this follows from definition 1.7. The nontrivial part is showing the existence of a multiplicative inverse. That is, we have to show that for any nonzero element $a \in F[x]/p(x)$, there exists some b such that $ab = 1$. Note first that we can take the (polynomial) degree of g to be less than p: if not, we can use the euclidean algorithm in $F[x]$ to rewrite $g = pq + r$ for some polynomials q and r and take r to be the representative of g in $F[x]/p(x)$. It immediately follows that g and p share no common factors as p is irreducible (and no constant factors, either, since it is monic). Again using Bézout's Identity over $F[x]$, we can write $gx + py = 1$. However, recall that $p \rightsquigarrow 0$ in $F[x]/p(x)$, and so $gx = 1$, which gives a multiplicative inverse for g. □

The above proof uses Bézout's Identity over $F[x]$. In general, this identity holds over various domains of interest, and the proof over $\mathbb{Z}$ also applies to $F[x]$ without much change. We now consider the ring $F(\alpha)$ and set p to be the minimal polynomial of α, ie. p is the lowest degree polynomial in $F[x]$ such that α is a root. Note that this also means that p must be irreducible, else you could factor out the other factors to obtain a polynomial of lesser degree.

Corollary 1.10. *The field $F[x]/p$ is isomorphic to the ring $F(\alpha)$ obtained by adjoining the element α that is defined have minimal polynomial p.*

Proof. Consider the map from $F[x]$ to $F(\alpha)$ defined as $f : x \mapsto \alpha$. It is clear that this is a ring homomorphism. We consider the kernel of f. Clearly, $f(q(x)) = 0$ implies $q(\alpha) = 0$. Recall that α is *defined* as a root of $p(x)$, hence $p \mid q$ as p is the minimal polynomial of α. On the flipside, every multiple of $p(x)$ is sent to 0. It follows that the kernel of f is every multiple of $p(x)$. As rings, this is the ideal $(p(x))$, and the first ring isomorphism theorem states that the quotient ring $F[x]/(p(x))$ is isomorphic to $F(\alpha)$. Furthermore, $F(\alpha)$ is hence a field as well. □

Example 1.11. *Check that $\mathbb{Z}/2\mathbb{Z}[x]/(x^2 + x + 1)$ is a finite field of size 4.*

Proof. Let $p(x) = x^2 + x + 1$. Then $p(0) = 1$, and $p(1) = 1$. Hence this has no roots and is irreducible over $\mathbb{Z}/2\mathbb{Z}$. Thus, $\mathbb{Z}/2\mathbb{Z}[x]/p(x)$ is a field. Furthermore, each element must be a polynomial of degree at most 1. Hence, the field is $\{0, 1, x, x+1\}$. Note that unlike the counterexample $\mathbb{Z}/2\mathbb{Z}[x](x^2)$, here x does have an inverse: $x \cdot (x + 1) = x^2 + x = 1$. □

Corollary 1.12. *Let F be a field and p be a monic irreducible polynomial. Then $E = F[x]/p$ is a vector space over F. Furthermore, the dimension of E is $d = \deg p$.*

[2]We define a **monic irreducible** polynomial in F to be a polynomial which has leading coefficient 1 and does not have a nontrivial factorization into nonconstant polynomials. This definition is probably unsatisfying; note that there is no reason why monic irreducible polynomials need even exist. We will study more about this in a later version of this document.

Proof. We will show that $(E, +, \cdot)$ is a vector space over F. Let $a, b \in E$. The axioms of addition follow immediately from the fact that E is a field. Note that $F \subseteq E$; then the scalar multiplication axioms follow as well. Set $E = F(\alpha)$. Then we claim $1, \alpha, \alpha^2, \ldots, \alpha^{d-1}$ is a basis for E. If it is not, then $\alpha^{d-1} = \sum_{j=0}^{d-2} k_j \alpha^j$. Then α is a root of this polynomial of degree $d-1$, which contradicts the fact that p is irreducible. $\square$

2. Classification of Finite Fields

We will now turn our attention to the study of finite fields, and show several important results about the existence and uniqueness of finite fields.

2.1. **Characteristic.** We start by showing that the size of a finite field is associated with a prime called the characteristic.

Theorem 2.2. *If F is a finite field, then F contains a copy of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ for some prime p.*

Proof. We will show by a simple counting argument. F contains 1, and consider the elements $1, 1+1, 1+1+1, \ldots$. Since F is finite, this sequence must repeat; suppose that $n \cdot 1 = m \cdot 1$. It follows that $(n-m) \cdot 1 = 0$. Take the minimum such x for which $x \cdot 1 = 0$. We claim that x must be prime. Suppose it is not, and that $ab \cdot 1 = 0$ for some $a, b \leq x$. Then either $a \cdot 1 = 0$ or $b \cdot 1 = 0$, since $a, b \neq 0$. This contradicts the fact that x is the minimum such number, and thus x does not have a nontrivial factorization, and is prime. It is then easy to see that the elements $\{0, 1, 2 \cdot 1, \ldots, (p-1) \cdot 1\}$ are isomorphic to $\mathbb{F}_p$. $\square$

Definition 2.3. *The* ***characteristic*** *of a finite field F is the minimum x such that $x \cdot 1 = 0$.*

Corollary 2.4. *The characteristic of a finite field is always prime.*

2.5. **Size of finite fields.** We call the subfield $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ the **prime subfield** of F. It is easy to see that F is a vector space over $\mathbb{F}_p$; note that the proof actually follows from corollary 1.12, the first part of which only uses the fact that E is an extension field of F. Furthermore, since F is finite, the degree of F must be also be finite. Now take any basis $\mathcal{B} = \{b_1, \ldots, b_t\}$ of F over $\mathbb{F}_p$. It follows that

$$F = \left\{ \sum_{i=1}^{t} a_i b_i : a_i \in \mathbb{F}_p \right\},$$

and in particular that $|F| = p^t$.

Theorem 2.6. *The cardinality of any finite field F is p^t for some prime p.*

We first show a very useful lemma.

Lemma 2.7. *Let F be a finite field of size p^t. Then every element in F is a root of the polynomial $x^{p^t-1} - 1 = 0$.*

Proof. We will use Lagrange's theorem. Let $|F| = p^t$ and suppose that $\alpha \in F^\times$. Then it follows that $\alpha^{p^t-1} = 1$, since the order of any element in a group divides the size of the group. In particular, this implies that $\alpha^{p^t-1} - 1 = 0$. Hence, α is a root of the polynomial $x^{p^t-1} - 1 = 0$. $\square$

Theorem 2.8. *Let F be a finite field and denote by $F^\times$ its multiplicative subgroup $F \setminus \{0\}$. Then $F^\times$ is cyclic.*

Proof. First, note that any polynomial $x^d - 1$ has at most d solutions in $F^\times$, which follows from a simple quotienting argument.

Suppose now that $|F^\times| = n$. For each $d \mid n$, denote by F_d the elements of $F^\times$ of order d. It is possible that F_d may be empty. If it is not empty, then take $y \in F_d$. Let $\langle y \rangle$ be the subgroup generated by y; we claim that $\langle y \rangle \subseteq \{x \in F^\times : x^d = 1\}$. This follows immediately from the fact that $x \in \langle y \rangle \implies x = y^k$, and $x^d = y^{kd} = 1$. Furthermore, since y is of order d, indeed $|\langle y \rangle| = d$. Using the fact that the set is defined as all such elements which satisfy the polynomial $x^d - 1 = 0$, this set can have at most d elements, and hence $\langle y \rangle = \{x \in F^\times : x^d = 1\}$.

Now note that since every element of F_d satisfies $x^d - 1 = 0$, $F_d \subseteq \langle y \rangle$. In fact, $\langle y \rangle$ is nothing but a copy of $\mathbb{Z}/d\mathbb{Z}$. The number of elements of order *exactly* d in this group is $\phi(d)$ (known as the principal generators), where ϕ is Euler's totient function.

It follows that F_d either has size 0 or size $\phi(d)$. Every element of F is in *some* F_d. Since we now know the size of each F_d, we can add them up. Recall that

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

We then get that

$$n = |F^\times| = \sum_{d|n} |F_d| \geq \sum_{d|n} \phi(d)$$

with equality holding iff $|F_d| \neq 0$ for *any* F_d. However, this includes F_n. Hence $F_n \neq \emptyset$, and thus there are elements of order n in $F^\times$. It follows that $F^\times$ is cyclic. □

The above theorem has many interesting consequences. In particular, take E/F (which denotes that E is an extension field of F). If $E = F(\alpha)$, then in fact $\langle \alpha \rangle = E^\times$. If this were not true, the the order of α would be some $d < n$ implying that $\alpha^d - 1 = 0$, which contradicts the fact that $x^n - 1$ is the minimal polynomial of α – and hence $E \neq F(\alpha)$.

Stop and note here that the above implies that every finite field is an *algebraic* extension of $\mathbb{F}_p$. This means that every finite field can be written as $\mathbb{F}_p[x]/f(x) \cong \mathbb{F}_p(\alpha)$ for some irreducible polynomial $f(x)$, which is also the minimal polynomial of α. Furthermore, this minimal polynomial divides $x^{p^t-1} - 1 = 0$. Taking any element $\alpha \in F$, denote its minimal polynomial as $m_\alpha(x)$.

2.9. **Uniqueness of finite fields.** We have shown that finite fields can only be of size p^t for prime p. We will now show that every such field is unique.

Theorem 2.10. *Let $|E| = |F| = p^t$. Then $E \cong F$.*

Proof. Let $E = \mathbb{F}_p(\alpha)$. Consider $m_{\alpha_1}(x)$. This polynomial divides $x^{p^t} - 1$, because $\alpha \in E$. Furthermore, the degree of $m_{\alpha_1}(x)$ is t; this follows from corollary 1.12.

Now note that $m_{\alpha_1}(x)$ must have a root in F as well, since it divides $x^{p^t-1} - 1$. Call this root α_2. Consider $m_{\alpha_2}(x)$. Since $m_{\alpha_2}(x)$ is minimal, it must divide $m_{\alpha_1}(x)$, but by irreducibility of $m_{\alpha_1}(x)$, $m_{\alpha_2}(x) = cm_{\alpha_1}(x)$ for some constant factor $c \in \mathbb{F}_p$. The implication is that $\mathbb{F}_p(\alpha_2) \subseteq F$. But $\mathbb{F}_p(\alpha_2)$ was obtained by adjoining an element with minimal polynomial of degree t, and hence the size of this field must be p^t, which is equal to the size of F. Thus, $\mathbb{F}_p(\alpha_2) = F$.

Now map $\alpha_1 \mapsto \alpha_2$. It is easily seen that this is a field isomorphism. □

2.11. **Existence of Finite Fields.** Until now we have seen that if finite fields exist, they are of size p^t and there is a unique field of such size. We now show that this is in fact achieved: there is such a finite field for every p, t.

Theorem 2.12. *For any t, there is a finite field of size p^t.*

Before we move to the proof proper, we define the splitting field of $\mathbb{F}_p$.

Definition 2.13. *The **splitting field** of polynomial $f(x)$ over F is the minimal field extension such that the polynomial factors into linear factors over F.*

Such a field can be created by adjoining roots until every polynomial is factored.

Proof (of theorem 2.12). We claim that the splitting field F of $x^{p^t} - x$ over $\mathbb{F}_p$ has size p^t.

First we will show that $|F| \leq p^t$. Recall that F was made by adjoining all roots of $x^{p^t} - x$ to it. Let y, y' be two such roots. Then we show that all the additive/multiplicative combinations of y, y' are also roots of $x^{p^t} - 1$. Multiplication follows trivially. For addition, note that

$$(y + y')^{p^t} - 1 = \sum_{i=0}^{p^t} \binom{p^t}{i} y^i y'^{p^t - 1}.$$

Since $\binom{p^t}{i}$ is always divisible by p, in $\mathbb{F}_p$ this reduces to $y^{p^t} + y'^{p^t} = y + y'$. Note that every adjoined element is by definition a root, while the elements of $\mathbb{F}_p$ are trivially roots of $x^{p^t} - x$ (as they are raised to the power p). Thus, every element of F is a root of $x^{p^t} - x$. Since this has at most p^t roots, the size is bounded.

While we do know that there are p^t linear factors, this does not mean that they are not repeats. In particular, there could be multiple roots with the same value; this would mean that $|F| < p^t$. We will show that this is not possible, because $f(x) = x^{p^t} - x$ contains no repeat roots. To show this, recall that the repeat roots of a polynomial are roots of $\gcd(f, f')$. Now take $(x^{p^t} - x)' = -1$ in $\mathbb{F}_p$. However, this has no root. It follows that f has no repeat roots, and thus all the new elements we added were unique. We conclude that $|F| = p^t$. □

We have now proved the following theorem.

Theorem 2.14 (Classification of Finite Fields)**.** *For each prime p and $t \in \mathbb{N}$, there exists a unique finite field of size p^t, which is isomorphic to $\mathbb{F}_p[x]/f(x)$, where $f(x)$ is an irreducible polynomial over $\mathbb{F}_p$ of degree t.*

3. Galois Fields

We first make some common remarks on notation.

1. The finite field of order q is denoted $\mathbb{F}_q$ or $GF(q)$.
2. $\mathbb{F}_{p^k} \neq \mathbb{Z}/p^k\mathbb{Z}$ in general; it holds only for $k = 1$.
3. $\mathbb{F}_{p^k} \neq (\mathbb{Z}/p\mathbb{Z})^k$ in general; they *are* isomorphic as vector spaces, however.
4. $\mathbb{F}_{p^k} \leq \mathbb{F}_{p^l}$ iff $k \mid l$.

3.1. **Frobenius Map.** Let F be a finite field and recall the map $x \mapsto x^p$. Note that this map sends $\mathbb{F}_p$ to itself. Furthermore, it is additive as well as multiplicative. We call this map the **Frobenius Automorphism**, denoted by Frob.

We show that Frob is an automorphism. To do this, it is enough to see that it is an injection; we have already seen that it is a homomorphism. Let $x^p = y^p$. Then it follows that $x^p - y^p = (x-y)^p = 0 \implies x = y$, and we are done.

Definition 3.2. *Let* $\mathrm{Aut}(E/F)$ *be the group of automorphisms of* E *that preserve* F.

Note that $\mathsf{Frob} \in \mathrm{Aut}(E/F)$. However, it is not unique. Note that an automorphism that preserves F must send a root of the minimal polynomial to another root. It follows that

$$|\mathrm{Aut}(E/F)| \leq [E : F]$$

where $[E : F]$ is the degree of the field extension.

Definition 3.3. E/F *is* ***Galois*** *if* $|\mathrm{Aut}(E/F)| = [E : F]$.

We end this section by stating without proof that $\mathbb{F}_{p^t}/\mathbb{F}_p$ is always Galois, hence why it is sometimes represented as $GF(p^t)$.