FINITE FIELDS AND THE MÖBIUS FUNCTION

REBECCA GOLOVANOV

Abstract. In this paper, we discuss finite fields. We will prove their existence and uniqueness, followed by an application of the Möbius function to count the number of irreducible polynomials in a finite field.

Contents

1. Introduction

This paper is primarily self-contained, following the exposition of Galois Theory by David A. Cox [1]. It explores finite fields through the lens of abstract algebra and some group theory. Finite fields are especially interesting to work with due to their relation to prime numbers. We begin with Fermat's Little Theorem, followed by proving key results about finite fields. Finite fields make an especially nice area to study irreducible polynomials, but we will restrict that exploration to the Möbius function.

2. $\mathbb{Z}/p\mathbb{Z}$ is a finite field

The theory of finite fields is built upon the fact that $\mathbb{Z}/p\mathbb{Z}$ is a field. This is a consequence of Fermat's Little Theorem, which we will prove in this section.

The nontrivial aspect of proving that $\mathbb{Z}/p\mathbb{Z}$ is a field is in the existence of multiplicative inverses. Fermat's Little Theorem is what yields the existence of multiplicative inverses.

Theorem 2.1. *(Fermat's Little Theorem) If p is a prime number, then for any $a \in \mathbb{Z}$, we have that $p|a^p - a$. In modular arithmetic, we can write this as $a^p \equiv a \pmod{p}$.*

Date: October 11, 2021.

Proof. Let G denote $\mathbb{Z}/p\mathbb{Z}$.
Let k be the order of an element $a \in G$, i.e., the smallest $k \in \mathbb{N}$ such that $a^k \equiv 1 \pmod{p}$. Then the numbers $\{1, a, a^2, ..., a^{k-1}\}$ reduced modulo p form a subgroup of G whose order is k.

We wish to show that k will always divide $p-1$. We will prove by induction.

Base Case: Let A be the set with elements $1, a, ..., a^{k-1}$ reduced modulo p. If $A = G$ then $k = p-1$ so $k|p-1$. If $A \neq G$, then there exists a $b_1 \in G \setminus A$. Let $A_1 = \{b_1, ab_1, ..., a^{k-1}b_1\}$. If A_1 doesn't have k distinct elements, then there exist $m, n \in \{0, 1, ..., k-1\}$ such that $a^m b_1 \equiv a^n b_1 \pmod{p}$. It follows that $a^m \equiv a^n \pmod{p}$, which is impossible, since m, n are less than k. Thus, A_1 must have k distinct elements. If $A_1 \cap A \neq \emptyset$, then there would exist some $m, n \in \mathbb{N}$ such that $a^m \equiv a^n b_1 \pmod{p}$. However, this implies

$$b_1 \equiv a^n a^{k-m}$$
$$b_1 \equiv a^{n+k-m}$$

All powers of a are in A, so this would imply $b_1 \in A$, but it is defined to be in $G \setminus A$. Thus, A and A_1 are disjoint. Since both have k elements, $|A \cup A_1| = 2k$. If $A \cup A_1 = G$, then $2k = p-1$, so $k|p-1$.

We define some finite $A_2, ..., A_n$ inductively using the same strategy we used to define A_1.

Inductive Hypothesis: If $A \cup A_1 \neq G$, then there exists a b_2 such that $b_2 \in G \setminus (A \cup A_1)$. Let $A_2 = \{b_2, ab_2, ..., a^{k-1}b_2\}$. Using a similar argument to the base case, A_2 must have k distinct elements and be disjoint with $A \cup A_1$. Then $|A \cup A_1 \cup A_2| = 3k$. Following this logic, for some disjoint sets $A, A_1, ..., A_{n-1}$, we hypothesize $|A \cup A_1 \cup ... \cup A_{n-1}| = nk$.

Inductive Step: We wish to show that if $A \cup A_1 \cup ... \cup A_{n-1} \neq G$, then $A \cup A_1 \cup ... \cup A_{n-1} \cup A_n = (n+1)k$

If $A \cup A_1 \cup ... \cup A_{n-1} \neq G$, then there exists a b_n such that $b_n \in G \setminus (A \cup A_1 \cup ... \cup A_{n-1})$. Let $A_n = \{b_n, ab_n, ..., a^{k-1}b_n\}$. Using a similar argument to the base case, A_n must have k distinct elements and A_n must be disjoint with $A \cup A_1 \cup ... \cup A_{n-1}$. Then $A \cup A_1 \cup ... \cup A_{n-1} \cup A_n = nk + k = (n+1)k$.

If $A \cup A_1 \cup ... \cup A_{n-1} \cup A_n = G$, then $|G| = (n+1)k = p-1$, meaning $k|p-1$.

If $A \cup A_1 \cup ... \cup A_{n-1} \cup A_n \neq G$, consider that the cardinality of $|G|$ is bounded by p. This means the union of these sets A_i for $i \in \mathbb{N}$ will eventually equal G due to its upper bound, and the cardinality of G will be a multiple of k. Thus, $p-1$ is a multiple of k.

Let $p-1=km$ for some $m \in \mathbb{N}$. Then we have

$$\begin{aligned} a^{p-1} &\equiv a^{km} \pmod{p} \\ a^{p-1} &\equiv (a^k)^m \pmod{p} \\ a^{p-1} &\equiv 1^m \pmod{p} \\ a^{p-1} &\equiv 1 \pmod{p} \end{aligned}$$

From here, we can multiply both sides of the congruence to get

$$a^p \equiv a \pmod{p}.$$

□

Theorem 2.2. *$\mathbb{Z}/p\mathbb{Z}$ is a finite field*

Proof. We will refer to Definition 7.1 to demonstrate that $\mathbb{Z}/p\mathbb{Z}$ satisfies the conditions to be a field.

Associativity and commutativity of addition and multiplication are defined on the integers, as well as the distributivity of multiplication across addition. 0 is the additive identity and 1 is the multiplicative identity. Since the smallest prime is 2, the smallest $\mathbb{Z}/p\mathbb{Z}$ contains $\{0,1\}$ and thus contains the identities.

Additive Inverse: For any $a \in \mathbb{Z}/p\mathbb{Z}$, we know $a < p$. Subtraction is defined on $\mathbb{Z}$, so we know $p-a>0$ and that $0<p-a<p$. Then $p-a \in \mathbb{Z}/p\mathbb{Z}$. $a+p-a=p$ and $p \equiv 0 \pmod{p}$, so $a+b \equiv 0 \pmod{p}$, meaning a has an additive inverse.

Multiplicative Inverse: We will use Fermat's Little Theorem. For any $a \in \mathbb{Z} \setminus p\mathbb{Z}$, we know that $a^p \equiv a \pmod{p}$. Then we have

$$\begin{aligned} a^{p-1} &\equiv 1 \pmod{p} \\ a(a^{p-2}) &\equiv 1 \pmod{p} \\ a^{-1} &\equiv a^{p-2} \pmod{p} \end{aligned}$$

And so any $a \in \mathbb{Z}/p\mathbb{Z}$ has an inverse.

Thus, $\mathbb{Z}/p\mathbb{Z}$ is a field, and we will henceforth refer to it as $\mathbb{F}_p$. □

3. Field Extensions

The following definitions on field extensions are necessary to understand the theorems used to prove the existence and uniqueness of finite fields.

Definition 3.1. Let F, L be fields. Given a ring homomorphism of $\varphi : F \to L$, we say L is a *field extension* of F via φ. F will usually be identified with its image, $\varphi(F) = \{\varphi(a) | a \in F\} \subset L$, and write $F \subset L$. The map φ must be one-to-one, and L must be larger than F.

A specific form of a field extension is a *splitting field*, which is integral to proving the existence and uniqueness of finite fields.

Definition 3.2. $F[x]$ describes the ring of polynomials with variable x over a field. Let $f \in F[x]$ have degree $n>0$. Then an extension $F \subset L$ is a *splitting field* of f over F if

(1) $f = c(x - \alpha_1)...(x - \alpha_n)$ where $c \in F$ and $\alpha_i \in L$;
(2) $L = F(\alpha_1, ..., \alpha_n)$
(Lemma 4.1.9 of Galois Theory states: $F(\alpha_1, ..., \alpha_n)$ *is the smallest subfield of the field* L *containing* F *and* $\alpha_1, ..., \alpha_n$).

Definition 3.3. A polynomial is *separable* if it is nonconstant and its roots in a splitting field are distinct.

Example 3.4. In $\mathbb{R}[x]$, the polynomial $x^2 - x$ is separable because it has two roots, 0 and 1. Similarly, in $\mathbb{C}[x]$, the polynomial $x^3 - 2$ is separable because there are 3 different cube roots of 2 in the complex numbers. However, $x^3 - 2$ is not separable as a polynomial in $\mathbb{F}_3$ because it factors into $(x + 1)^3$, meaning it has a triple root.

Definition 3.5. If an $\alpha \in F$ is algebraic over F, the *minimal polynomial* of α is a unique nonconstant monic $p \in F[x]$ such that

(1) α is a root of p;
(2) if $f \in F[x]$ is any polynomial with α as a root, f is a multiple of p.

Example 3.6. 0 is algebraic over any field $F[x]$ because the polynomial x has a root of 0 and any multiple of x shares that root.

Definition 3.7. A polynomial is *irreducible* if it cannot be factored into nontrivial polynomials over the same field.

Example 3.8. The polynomial $x^2 - 2 \in \mathbb{Q}[x]$ is irreducible because it cannot be factored further over the rational numbers, but $x^2 - 2 \in \mathbb{R}[x]$ is reducible because it can be factored into $(x - \sqrt{2})(x + \sqrt{2})$. Similarly, $x^2 + 1 \in \mathbb{R}[x]$ is irreducible because it only has complex roots.

Definition 3.9. An extension $F \subset L$ is a *normal extension* if every irreducible polynomial in $F[x]$ that has a root in L splits completely over L.

Definition 3.10. An extension $F \subset L$ is a *separable extension* if every $\alpha \in L$ is separable over F. Every α is separable if its minimal polynomial over F is separable.

Example 3.11. The extension $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ is separable since the minimal polynomial of $a + b\sqrt{2}$ when $b \neq 0$, is $x^2 - 2ax + a^2 - 2b^2 = (x - a + b\sqrt{2})(x - a - b\sqrt{2})$. Every minimal polynomial has distinct roots, so every $\alpha \in \mathbb{Q}(\sqrt{2})$ is separable, making $\mathbb{Q}(\sqrt{2})$ a separable extension.

The next definition is of a *finite extension*, for which we need to understand basis and dimension.

Definition 3.12. Let V be a vector space over a field F. Then the set B is a *basis* of V if B is linearly independent and $V = \operatorname{span} S$. The *span* of a set of vectors is every linear combination of those vectors. Essentially, any $v \in V$ is a linear combination of the elements of B.

Example 3.13. The standard basis of $\mathbb{R}^2$ is $\left\{\binom{0}{1}, \binom{1}{0}\right\}$ because every vector in $\mathbb{R}^2$ can be written as a unique combination of the two listed.

Definition 3.14. The *dimension* of a vector space is the cardinality of its basis over its base field.
Note: The cardinality of a basis is constant, which is why the dimension of a vector space is well-defined.

Example 3.15. The dimension of $\mathbb{C}$ over $\mathbb{R}$ (notated as $\dim_{\mathbb{R}} \mathbb{C}$) is 2, while $\dim_{\mathbb{C}} \mathbb{C}$ is 1.

We can now define a finite extension as follows.

Definition 3.16. Let $F \subset L$ be a field extension.

(1) L is a *finite extension* if L is a finite-dimensional vector space over F.
(2) the degree of L over F is defined as follows

$$[L : F] = \begin{cases} \dim_F L & \text{if L is finite,} \\ \infty & \text{otherwise.} \end{cases}$$

Definition 3.17. A group G is *cyclic* if there is a $g \in G$ such that $G = \{g^l \mid l \in \mathbb{Z}\}$. When G is cyclic,

$$G \simeq \begin{cases} \mathbb{Z}, & \text{if } G \text{ is infinite,} \\ \mathbb{Z}/n\mathbb{Z} \text{ for some } n \in \mathbb{N}, & \text{if } |G| < \infty. \end{cases}$$

Example 3.18. The subgroup created in Theorem 2.1 is a cyclic group ($\{1, a, ..., a^{k-1}\}$) as it is composed exclusively of powers of a.

The following statement prepares us to define the *characteristic* of a field. Given a positive integer k, define $k \cdot 1 = 1 + 1 + ... + 1 \in F$, where 1 is the multiplicative identity of F.

Definition 3.19. F has *characteristic* 0 if $k \cdot 1 \neq 0$ for all positive integers k and has characteristic p if $p \cdot 1 = 0$ and p is prime.
Note that if a field has characteristic 0, it means there are infinite additive combinations of 1 without reaching 0, so any field of characteristic 0 must be infinite.

Example 3.20. The rings $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}$ all have characteristic 0, while $\mathbb{F}_p$ has characteristic p. Recall from the end of Section 2 that $\mathbb{F}_p$ is equivalent to $\mathbb{Z}/p\mathbb{Z}$ for some prime p.

Definition 3.21. Given a ring homomorphism $\varphi : R \to S$ where R, S are rings, its *kernel* is $\ker(\varphi) = \{r \in R \mid \varphi(r) = 0\}$ and its *image* is $\mathrm{Im}(\varphi) = \{\varphi(r) \mid r \in R\}$

Example 3.22. Consider $\mathbb{R}^2$ and the transformation $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then $\ker(A) = \left\{ v = ae + be \mid Av = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$, where e is the identity and v is any vector in $\mathbb{R}^2$. Then $\ker(A) = \begin{pmatrix} 0 \\ b \end{pmatrix}$ such that $b \in \mathbb{R}$

A theorem that follows from this is the Fundamental Theorem of Ring Homomorphisms (also known as The First Isomorphism Theorem)

Theorem 3.23. *Let $\varphi : R \to S$ be a ring homomorphism. Then there exists a unique ring homomorphism $\overline{\varphi} : R/\ker(\varphi) \simeq \mathrm{Im}(\varphi)$ such that $\overline{\varphi}(r + \ker(\varphi)) = \varphi(r)$ for all $r \in R$.*

4. Existence and Uniqueness of Finite Fields

The following proposition demonstrates how $\mathbb{F}_p$ relates to arbitrary finite fields.

Proposition 4.1. *Let F be a finite field. Then:*

(1) *There is a unique prime p such that F contains a subfield isomorphic to $\mathbb{F}_p$*
(2) *F is a finite extension of $\mathbb{F}_p$, and*

$$|F| = p^n, \text{ where } n = [F : \mathbb{F}_p].$$

Proof. Every field of characteristic 0 is infinite, so F must have characteristic p for a prime p. The set $p\mathbb{Z} \subset \mathbb{Z}$ is the kernel of the ring homomorphism sending $m \in \mathbb{Z}$ to $m \cdot 1 \in F$. Essentially, since F has characteristic p, any $m \in F$ multiplied by p sends it to 0, making $p\mathbb{Z}$ the kernel. By the Fundamental Theorem of Ring Homomorphisms (Theorem 3.23), F contains a subfield isomorphic to $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$.

The map $\mathbb{F}_p \to F$ makes F an extension field of $\mathbb{F}_p$. We identify $\mathbb{F}_p$ with its image (as shown in Definition 3.1) and write $\mathbb{F}_p \subset F$. We can consider F to be a vector space over $\mathbb{F}_p$. As such, the elements of F have finitely many vectors in F, whose span over $\mathbb{F}_p$ is F. Then F is a finite dimensional vector space over $\mathbb{F}_p$, so F is a finite extension of $\mathbb{F}_p$. If $n = [F : \mathbb{F}_p]$, then we can find a basis $\alpha_1, ..., \alpha_n$ of F over $\mathbb{F}_p$. By our definition of basis, any element $\beta \in F$ can be written uniquely as a linear combination of $\alpha_1, ..., \alpha_n$. Since the coefficients can be any of the p elements of $\mathbb{F}_p$, there are p^n possibilities for β. Then $|F| = p^n$. □

Given Proposition 4.1, we will assume any finite field F contains a subfield $\mathbb{F}_p$. The next major result is that F is a splitting field (Definition 3.2) over $\mathbb{F}_p$.

Theorem 4.2. *Let F be a finite field with $q = p^n$ elements. Then:*
(1) *$\alpha^q = \alpha$ for all $\alpha \in F$*
(2) *$x^q - x = \prod_{\alpha \in F}(x - \alpha)$*
(3) *F is a splitting field over $\mathbb{F}_p$ of $x^q - x \in \mathbb{F}_p[x]$*

Proof. We first prove part(1):

Since F has q elements, the multiplicative group $F^* = F \setminus \{0\}$ is a group with $q - 1$ elements. It follows that $\alpha^{q-1} = 1$ for all $\alpha \in F^*$, meaning $\alpha^q = \alpha$ for all $\alpha \in F$.

The above shows that q elements of F are roots of $x^q - x$. Then part (b) is proven since $x^q - x$ is monic of degree q, meaning $x^q - x$ splits completely over F. Since every element of F is a root, $x^q - x$ can't split completely over any strictly smaller field. Therefore, part (c) follows and F is a splitting field of $x^q - x \in \mathbb{F}_p[x]$. □

The conclusion of a splitting field of $x^q - x \in \mathbb{F}_p$ is critical to proving the uniqueness of finite fields.

Before we can prove uniqueness, we must first prove some useful results about splitting fields, especially their relation to homomorphisms.

For a polynomial f, $\langle f \rangle$ is the ideal (defined in Definition 7.4) generated by f in the polynomial ring $F[x]$.

Definition 4.3. Given a function h and $\alpha_1, ..., \alpha_n \in L$ where α_i such that $1 \leq i \leq n$ is algebraic over F, define $F[\alpha_1, ..., \alpha_n] = \{h(\alpha_1, ..., \alpha_n) \mid h \in F[x_1, ..., x_2]\}$. Essentially, $F[\alpha]$ consists of all polynomial expressions in L that can be formed using α with coefficients in F.

Lemma 4.4. *Assume $F \subset L$ is a field extension, and let $\alpha \in L$ be algebraic over F with minimal polynomial $q \in F[x]$. Then there is a unique ring isomorphism $F[\alpha] \simeq F[x]/\langle q \rangle$ that is the identity on F and maps α to the coset $x + \langle q \rangle$.*

Proof. Consider the ring homomorphism $\varphi : F[x] \to L$ that sends $h(x) \in F[x]$ to $h(\alpha) \in L$. By definition, the image of φ is $F[\alpha]$. We claim that $\ker(\varphi) = \langle q \rangle$. Note that $g \in F[x]$ implies that

$$\varphi(gq) = \varphi(g)\varphi(q) = g(\alpha)q(\alpha) = g(\alpha)0 = 0.$$

This shows that $\langle q \rangle \subset \ker(\varphi)$. Suppose there exists some $f \in \ker(\varphi)$. Then $f(\alpha) = 0$, which implies f is a multiple of q. Then $\ker(\varphi) \subset \langle q \rangle$, meaning $\ker(q) = \langle q \rangle$.

Since we know the image and kernel of φ, the Fundamental Theorem of Ring Homomorphisms (Theorem 3.23) gives a ring isomorphism

$$F[x]/\langle q \rangle \simeq F[\alpha].$$

This isomorphism is the identity on F and maps the coset $x + \langle q \rangle$ to α. Its inverse is the isomorphism described in the statement of the lemma.

Uniqueness follows since a ring homomorphism defined on $F[\alpha]$ is uniquely determined by its valued on F and α. □

Definition 4.5. $F(\alpha)$ describes the *quotient field* of $F[\alpha]$ (a field whose elements are quotients of polynomials).

Lemma 4.6. *Assume that $F \subset L$ is a field extension, and let $\alpha \in L$. Then α is algebraic over F if and only if $F[\alpha] = F(\alpha)$.*

Proof. Let $F(\alpha)$ denote the quotient field of the polynomial ring $F[\alpha]$, i.e., a ring whose elements contain F and α. It can be shown [1,Lemma 4,1,9] that $F(\alpha)$ is the smallest subfield of L containing F and α. It follows that $F(\alpha) \subset F[\alpha]$. The opposite inclusion always holds, meaning $F(\alpha) = F[\alpha]$ when α is algebraic over F.

For the second direction, suppose that $F[\alpha] = F(\alpha)$. We can assume $\alpha \neq 0$ since 0 is obviously algebraic over F. Then $1/\alpha \in F(\alpha) = F[\alpha]$ implies that

$$1/\alpha = a_0 + a_1\alpha + ... + a_m\alpha^m$$

for some $a_0, ..., a_m \in F$. Thus

$$0 = -1 + a_0\alpha + a_1\alpha^2 + ... + a_m\alpha^{m+1},$$

proving α is algebraic over F. □

Theorem 4.7. *Given a nonconstant polynomial $f_1 \in F_1[x]$, $\varphi : F_1 \simeq F_2$ and splitting fields L_1, L_2, there is an isomorphism $\overline{\varphi} : L_1 \simeq L_2$ such that $\varphi = \overline{\varphi}|_{F_1}$. Note that φ maps $f_1 \in F_1$ to a nonconstant polynomial $f_2 \in F_2$.*

Proof. This will be a proof by induction on $n = deg(f_1) = deg(f_2)$. When $n = 1$, $f_1 = ax + b$ has the root $-b/a \in F_1$, since $a \neq 0$. Then $L_1 = F_1$ and $L_2 = F_2$.

Now suppose $n > 1$. We know $L_1 = F_1(\alpha_1, ..., \alpha_n)$, where $\alpha_1, ..., \alpha_n$ are the roots of f_1. We will use the extensions

$$F_1 \subset F_1(\alpha_1) \subset L_1,$$

where $F_1(\alpha_1) \subset L_1$ is a splitting field of $g_1 = f_1/(x - \alpha_1)$. We proceed in the following five steps.

Step 1. Let $h_1 \in F_1[x]$ be the minimal polynomial of α_1. We know h_1 is an irreducible factor of $f_1 \in F_1[x]$, since α_1 is a root of f_1. Thus

$$F_1(\alpha_1) = F_1[\alpha_1] \simeq F_1[x]/\langle h_1 \rangle,$$

where we have used Lemma 4.4 for the isomorphism and Lemma 4.6 for the equality. The resulting isomorphism takes α_1 to $x + \langle h_1 \rangle$.

Step 2. The field isomorphism $\varphi : F_1 \simeq F_2$ induces a ring isomorphism $\tilde{\varphi} : F_1[x] \simeq F_2[x]$ that takes f_1 to f_2. This isomorphisms takes factors to factors and irreducibles to irreducibles. In particulaer, h_1 will map to an irreducible factor h_2 of f_2. Since f_2 splits completely over L_2. so does h_2. We can then label the roots of f_2 as $\beta_1, ..., \beta_n \in L_2$, where β_1 is a root of h_2.
Step 3. The root β_1 of f_2 gives the extensions

$$F_2 \subset F_2(\beta_1) \subset L_2,$$

where $F_2(\beta_1) \subset L_2$ is a splitting field of $g_2 = f_2/(x - \beta_1)$. As in Step 1, we have

$$F_2(\beta_1) = F_2[\beta_1] \simeq F_2[x]/\langle h_2 \rangle,$$

since h_2 is the minimal polynomial of β_1. This isomorphism takes β_1 to $x + \langle h_2 \rangle$.
Step 4. Since $\tilde{\varphi} : F_1[x] \simeq F_2[x]$ takes h_1 to h_2, it must take $\langle h_1 \rangle$ to $\langle h_2 \rangle$. This means we get an isomorphism of quotient rings,

$$F_1[x]/\langle h_1 \rangle \simeq F_2[x]/\langle h_2 \rangle,$$

that takes $x + \langle h_1 \rangle$ to $x + \langle h_2 \rangle$ and is φ on the coefficients. Combining this with steps 1 and 3, we get an isomorphism—

$$\varphi_1 : F_1(\alpha_1) \simeq F_1[x]/\langle h_1 \rangle \simeq F_2[x]/\langle h_2 \rangle \simeq F_2(\beta_1)|$$

that takes α_1 to β_1 and satisfies $\varphi_1|_{F_1} = \varphi$.
Step 5. Since $\varphi_1 : F_1(\alpha_1) \simeq F_2(\beta_1)$ takes α_1 to β_1 and f_1 to f_2, it also takes $g_1 = f_1/(x - \alpha_1)$ to $g_2 = g_2/(x - \beta_1)$. As noted above, L is a splitting field of g_1 over $F_1(\alpha_1)$, and similarly for L_2. Note that now we know

$$\begin{aligned} F_1 \subset F_1(\alpha_1) \subset L_1 \\ F_2 \subset F_2(\beta_1) \subset L_2 \\ \varphi_1 : F_1(\alpha_1) \simeq F_2(\beta_1) \\ \varphi : F_1 \simeq F_2. \end{aligned}$$

Since $g = f_1/(x - \alpha)$ has degree $n - 1$, we can apply the inductive hypothesis to $g_1 \in F_1(\alpha_1)[x]$ and $\varphi_1 : F_1(\alpha_1) \simeq F_2(\beta_1)$. This gives $\overline{\varphi_1} : L_1 \simeq L_2$, whose restriction to $F_1(\alpha_1)[x]$ is φ_1. But since $\varphi_1|_{F_1} = \varphi$, it follows that the restriction of φ_1 to F_1 is φ, meaing $\overline{\varphi_1}$ is the desired isomorphism. □

Corollary 4.8. *If L_1 and L_2 are splitting fields of $f \in F[x]$. then there is an isomorphism $L_1 \simeq L_2$ that is the identity on F.*

The above corollary allows us to speak of *the* splitting field of $f \in F[x]$, remembering that fields are unique up to isomorphism. It is essential to proving the uniqueness of finite fields, but we needed to do the previous inductive proof because otherwise φ_1 didn't need to be the identity.

Theorem 4.9. *[Uniqueness of Finite Fields] Two finite fields with the same number of elements are isomorphic.*

Proof. Uniqueness follows immediately from Corollary 4.8 and part(3) of Theorem 4.2, as any two splitting fields of $x^q - x \in \mathbb{F}_p[x]$ are isomorphic. □

Lemma 4.10. *[The Freshman's Dream] Let F be a field of characteristic p, and assume that $\alpha, \beta \in F$. Then $(\alpha + \beta)^p = \alpha^p + \beta^p$ and $(\alpha - \beta)^p = \alpha^p - \beta^p$.*

Proof. The binomial theorem implies that

$$(\alpha+\beta)^p - \alpha^p + \binom{p}{1}\alpha^{p-1}\beta + ... + \binom{p}{p-1}\alpha\beta^{p-1} + \beta^p.$$

We know that $\binom{p}{r}$ for $1 \leq r \leq p-1$ is equal to $\frac{p!}{r!(p-r)!} = \frac{p(p-1)...(p-r+1)}{r!}$. Then $\binom{p}{r}$ is divisible by p. Since F has characteristic p, the above binomial expansion reduces to $(\alpha+\beta)^p = \alpha^p + \beta^p$. The case for subtraction is argued similarly. □

Theorem 4.11. *[Existence of Finite Fields] Given any prime p and any positive integer n, there is a finite field with p^n elements .*

Proof. Let $q = p^n$, and let L be an extension of $\mathbb{F}_p$ such that $x^q - x$ splits completely over L. Since we are in characteristic p, the derivative of $x^q - x$ is -1, meaning $\gcd(x^q - x, (x^q - x)') = 1$. Thus $x^q - x$ is separable and has distinct roots in L. Then $F = \{\alpha \in L \mid \alpha^q = a\}$ is a subset of L consisting of q elements. Next, we must show F is a subfield of L, which requires use of Lemma 4.10.

First, we need to show sums and products of elements of F are in F.

Multiplication: We want to show that $\alpha\beta \equiv (\alpha\beta)^q \pmod{p}$ for any $\alpha, \beta \in F$.

We know that $\alpha\beta \equiv \alpha^q\beta^q \pmod{p}$. By the property of distribution of exponents across multiplication, we know $\alpha^q\beta^q = (\alpha\beta)^q$ for all integers α, β, so then we have

$$\alpha\beta \equiv (\alpha\beta)^q \pmod{p}.$$

Addition: We want to show that $\alpha + \beta \equiv (\alpha+\beta)^q \pmod{p}$ for any $\alpha, \beta \in F$.

We know that $\alpha + \beta \equiv \alpha^q + \beta^q \pmod{p}$. Since equivalences are bidirectional, I will proceed by proving $\alpha^q + \beta^q \pmod{p} \equiv (\alpha+\beta)^q$. For this, we will use Lemma 4.10.

Since we are in characteristic p, we can rewrite $(\alpha+\beta)^q$ to

$$(\alpha+\beta)^{p^n} = (\alpha+\beta)^{p(p^{n-1})}.$$

It follows from here that we have

$$(\alpha+\beta)^{p(p^{n-1})} = (\alpha^p + \beta^p)^{p^{n-1}}$$
$$(\alpha^p + \beta^p)^{p(p^{n-2})} = (\alpha^{p^2} + \beta^{p^2})^{p^{n-2}}$$

Since n is finite, we can repeat this process n times via induction, resulting in $(\alpha^{p^{n-1}} + \beta^{p^{n-1}})^p = \alpha^{p^n} + \beta^{p^n}$.

And so we get $(\alpha+\beta)^{p^n} = \alpha^{p^n} + \beta^{p^n}$, meaning

$$(\alpha+\beta)^q = \alpha^q + \beta^q.$$

Then we have $\alpha + \beta \equiv (\alpha+\beta)^q \pmod{p}$.

Next, we need to show that the additive and multiplicative identities are in F. Simply, we have that $0^n = 0$ for all $n \in \mathbb{Z}$ such that $n \neq 0$, and we know that $1^n = 1$ for all $n \in \mathbb{Z}$. Thus, we have

$$0 \equiv 0^q \pmod{p}$$
$$1 \equiv 1^q \pmod{p}$$

The multiplication and addition of integers is known to be commutative and associative, as well as the distribution of multiplication over addition.

Additive Inverse: For any $\alpha \in F$, we can have $\alpha - \alpha \equiv (\alpha - \alpha)^q \pmod{p}$, meaning $0 \equiv 0^q$, satisfying the conditions of F.

Multiplicative Inverse: Notice the condition of F: any $\alpha \in F$ satisfies $\alpha \equiv \alpha^q \pmod{p}$. This implies the following:

$$\begin{aligned} \alpha^{q-1} &\equiv 1 \pmod{p} \\ \alpha(\alpha^{q-2}) &\equiv 1 \pmod{p} \\ \alpha^{-1} &= \alpha^{q-2} \end{aligned}$$

Then any $\alpha \in F$ has a multiplicative inverse.

Thus, F is a subfield of L. It follows from here that F is a finite field with $q = p^n$ elements. □

5. Galois Correspondence and Application to Finite Fields

This section discusses abstract algebra and Galois theory necessary to analyze irreducible polynomials over finite fields.

Definition 5.1. Let $F \subset L$ be a finite extension. Then the *Galois group* notated $\mathrm{Gal}(L/F)$ is the set $\{\sigma : L \to L \mid \sigma \text{ is an automorphism, } \sigma(a) = a \text{ for all } a \in F\}$. Essentially, $\mathrm{Gal}(L/F)$ consists of all automorphisms of L that are the identity on F.

Definition 5.2. Suppose we have a finite extension $F \subset L$ with Galois group $\mathrm{Gal}(L/F)$. Given a subgroup $H \subset \mathrm{Gal}(L/F)$, we call $L^H = \{\alpha \in L \mid \sigma(\alpha) = \alpha \text{ for all } \sigma \in H\}$ the *fixed field* of H.

Definition 5.3. A finite extension $F \subset L$ is a *Galois extension* if any of the following are satisfied (the three statements are equivalent):

(1) L is the splitting field of separable polynomial in $F[x]$.
(2) F is a fixed field of $\mathrm{Gal}(L/F)$ acting on L.
(3) $F \subset L$ is a normal separable extension. (Note: (1) of the definition already requires L to be a normal extension)

Example 5.4. Consider $\mathbb{Q} \subset \mathbb{Q}(i, \sqrt[4]{2})$, which is the splitting field of $x^4 - 2$. Then it is a Galois extension. We have intermediate fields $\mathbb{Q}(i)$ and $\mathbb{Q}(\sqrt[4]{2})$. The extension $\mathbb{Q} \subset \mathbb{Q}(i)$ is Galois, as it is the splitting field of $x^2 + 1$. However, $\mathbb{Q} \subset \mathbb{Q}(\sqrt[4]{2})$ is not Galois, because $x^4 - 2$ is the minimal polynomial of $\sqrt[4]{2}$ but it doesn't split completely.

The following proposition discusses polynomials in relation to finite fields, which is essential for our exploration of irreducible polynomials in Section 6.

Proposition 5.5. *If $f \in \mathbb{F}_p[x]$ is nonconstant and $n \geq 1$, then the number of roots of f in $\mathbb{F}_{p^n}$ is the degree of the polynomial* $\gcd(f, x^{p^n} - x)$.

Proof. Let $g = \gcd(f, x^{p^n} - x)$, where the gcd is computed in $\mathbb{F}_p[x]$. Note that if the gcd is computed in $\mathbb{F}_p[x]$, the Euclidean Algorithm operates only on coefficients

contained in $\mathbb{F}_p[x]$. This means that $\gcd(f, x^{p^n} - x)$ computed over $\mathbb{F}_p[x]$ equals $\gcd(f, x^{p^n} - x)$ computed over $\mathbb{F}_{p^n}[x]$. Then we can compute the gcd in $\mathbb{F}_{p^n}[x]$. If we denote the elements of this field by α_i for $i = 1, ..., p^n$, then $x^{p^n} - x = (x-\alpha_1)...(x-\alpha_{p^n})$ by part(2) of Theorem 4.2. This is the irreducible factorization of $x^{p^n} - x$ in $\mathbb{F}_{p^n}[x]$. Hence g is the product of those $x - \alpha_i$ that divide f. Since $x - \alpha_i$ divides f if and only if $f(\alpha_i) = 0$, we obtain the product formula

$$g = \prod_{f(\alpha_i)=0} (x - \alpha_i).$$

The proposition follows immediately. □

Theorem 5.6. *If $q = p^n$, then:*

(1) $\mathbb{F}_p \subset \mathbb{F}_q$ is a Galois extension of degree n.
(2) The map $\mathrm{Frob}_p : \mathbb{F}_q \to \mathbb{F}_q$ defined by $\mathrm{Frob}_p(\alpha) = \alpha^p$ is an automorphism of $\mathbb{F}_q$ that is the identity on $\mathbb{F}_p$.
(3) Frob_p generates $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$. Thus there is a group isomorphism,

$$\mathrm{Gal}(\mathbb{F}_q/\,\mathrm{Gal}(\mathbb{F}_p) \simeq \mathbb{Z}/n\mathbb{Z},$$

that sends $\mathrm{Frob}_p \in \mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p$ to $[1] \in \mathbb{Z}/n\mathbb{Z}$.

Proof. Part(1): In the proof of Theorem 4.11, we observed $x^q - x$ is separable. Then Theorem 4.2 implies that $\mathbb{F}_q$ is the splitting field of a separable polynomial. Then $\mathbb{F}_p \subset \mathbb{F}_q$ is Galois. Proposition 4.1 implies that $[\mathbb{F}_q : \mathbb{F}_p] = n$ since $q = p^n$.

Part(2): By Lemma 4.10, notice

$$\mathrm{Frob}_p(\alpha + \beta) = (\alpha + \beta)^p = \alpha^p + \beta^p = \mathrm{Frob}_p(\alpha) + \mathrm{Frob}_p(\beta).$$

Since we also have $\mathrm{Frob}_p(1) = 1^p = 1$ and

$$\mathrm{Frob}_p(\alpha\beta) = \alpha^p\beta^p = \mathrm{Frob}_p(\alpha)\,\mathrm{Frob}_p(\beta),$$

we see that Frob_p is a ring homomorphism. A homomorphism is one-to-one if and only if its kernel is $\{0\}$. We know $a^p = 0$ implies $a = 0$ for $a \in \mathbb{F}_p$, and since we are in an integral domain, there is no other element β not equal to 0 such that $\alpha\beta = 0$, so 0 must be the kernel. Then Frob_p is a one to one map, and since it goes to itself it is also onto. Then Frob_p is an automorphism of $\mathbb{F}_q$. Since it is the identity on $\mathbb{F}_p$ by Theorem 2.1, we conclude that $\mathrm{Frob}_p \in \mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$.

Part(3): Since $\mathbb{F}_p \subset \mathbb{F}_q$ is Galois, we have

$$|\,\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)| = [\mathbb{F}_q : \mathbb{F}_p] = n,$$

where the second equality uses $q = p^n$ and Proposition 4.1. Then the order of Frob_p divides n. Suppose $(\mathrm{Frob}_p)^r$ is the identity, where $0 < r < n$. Then $\alpha^{p^r} = \alpha$ for all $\alpha \in \mathbb{F}_q$. Since $0 < r < n$, this implies that $x^{p^r} - x$ of degree $p^r < p^n = q$ has q roots, which is impossible. Thus, Frob_p has order n, yielding the desired isomorphism $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p) \simeq \mathbb{Z}/n\mathbb{Z}$. We call Frob_p the *Frobenius automorphism* of $\mathbb{F}_q$. □

In order to further analyze finite fields, we must introduce the Fundamental Theorem of Galois Theory, which has two parts.

Theorem 5.7. *Let $F \subset L$ be a Galois extension.*

(1) *For an intermediate field* $F \subset K \subset L$*, its Galois group* $\mathrm{Gal}(L/K) \subset \mathrm{Gal}(L/F)$ *has fixed field*

$$L^{Gal(L/K)} = K.$$

Furthermore, $|\mathrm{Gal}(L/K)| = [L : K]$ *and* $[\mathrm{Gal}(L/F) : \mathrm{Gal}(L/K)] = [K : F]$.

(2) *For a subgroup* $H \subset \mathrm{Gal}(L/F)$*, its fixed field* $F \subset L^H \subset L$ *has Galois group*

$$\mathrm{Gal}(L/L^H) = H.$$

Furthermore, $[L : L^H] = |H|$ *and* $[L^H : F] = [\mathrm{Gal}(L/F) : H]$.

Theorem 5.8. *Let* $F \subset L$ *be a Galois extension. Then the maps between intermediate fields* $F \subset K \subset L$ *and subgroups* $H \subset \mathrm{Gal}(L/F)$ *given by*

$$K \mapsto \mathrm{Gal}(L/K),$$
$$H \mapsto L^H$$

reverse inclusions and are inverses of each other. Furthermore, if a subfield K *corresponds to a subgroup* H *under these maps, then* K *is Galois over* F *if and only if* H *is normal in* $\mathrm{Gal}(L/F)$*, and when this happens, there is a natural isomorphism*

$$\mathrm{Gal}(L/F)/H \simeq \mathrm{Gal}(K/F).$$

Essentially, the subextensions of L/F *are isomorphic with subgroups of* $\mathrm{Gal}(L/F)$.

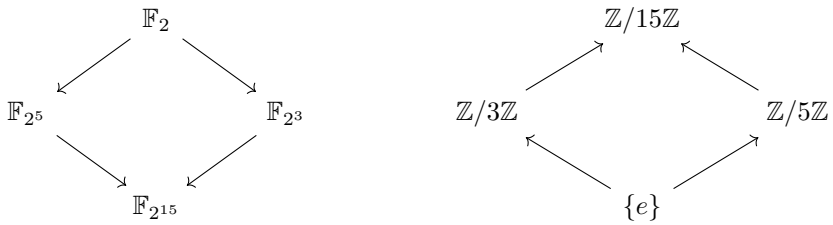
We will illustrate the second theorem with two examples. Recall that $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) \simeq \mathbb{Z}/n\mathbb{Z}$ (Theorem 5.6).

Example 5.9. For this example, we will use $\mathbb{F}_{2^3}$ as our L. The subextensions follow from prime factorization.

$$\begin{array}{ccc} \mathbb{F}_{2^3}/\mathbb{F}_2 & & \mathbb{Z}/3\mathbb{Z} \\ \downarrow & & \uparrow \\ \mathbb{F}_{2^3}/\mathbb{F}_{2^3} & & \{e\} \end{array}$$

The arrows are inclusion arrows, so $\{e\}$ is contained in $\mathbb{Z}/3\mathbb{Z}$. The isomorphisms are $\mathrm{Gal}(\mathbb{F}_{2^3}/\mathbb{F}_2) \simeq \mathbb{Z}/3\mathbb{Z}$ and $\mathrm{Gal}(\mathbb{F}_{2^3}/\mathbb{F}_{2^3}) \simeq \{e\}$, where the Galois groups are $\mathrm{Gal}(\mathbb{F}_{2^3}/\mathbb{F}_2)$ and $\mathrm{Gal}(\mathbb{F}_{2^3}/\mathbb{F}_{2^3})$, respectively.

Example 5.10. A slightly more complicated example is $\mathbb{F}_{2^{15}}/\mathbb{F}_2$, where $\mathrm{Gal}(L/F)$ corresponds to $\mathbb{Z}/15\mathbb{Z}$. We can illustrate the isomorphisms with the following diagram:



The isomorphisms of the subgroups reflected in the above diagram are as follows:

$$\begin{aligned}\mathrm{Gal}(\mathbb{F}_{2^{15}}/\mathbb{F}_2) &\simeq \mathbb{Z}/15\mathbb{Z}\\ \mathrm{Gal}(\mathbb{F}_{2^{15}}/\mathbb{F}_{2^5}) &\simeq \mathbb{Z}/3\mathbb{Z}\\ \mathrm{Gal}(\mathbb{F}_{2^{15}}/\mathbb{F}_{2^3}) &\simeq \mathbb{Z}/3\mathbb{Z}\\ \mathrm{Gal}(\mathbb{F}_{2^{15}}/\mathbb{F}_{2^15}) &\simeq \{e\}\end{aligned}$$

This theorem and the following lemma motivate a proof on discerning when one finite field is contained in another.

Lemma 5.11. *(Tower Theorem) Suppose that we have fields $F \subset K \subset L$.*

(1) If $[K : F] = \infty$ or $[L : K] = \infty$, then $[L : F] = \infty$.
(2) If $[K : F] < \infty$ and $[L : K] < \infty$, then $[L : F] = [L : K][K : F]$

Proof. We first prove the contrapositive of Part(1): if $[L : F] < \infty$, then $[K : F] < \infty$ and $[L : K] < \infty$. Then we may assume L has a finite dimension as a vector space over F. Let $\gamma_1, ..., \gamma_N$ be a basis.

Then we notice $K \subset L$ is a subspace of L over the field F. Since L has finite dimension over F, so does any subspace. Then $[K : F] = \dim_F K < \infty$.

Consider $\alpha \in L$. Since $\gamma_i, ..., \gamma_N$ span L over F, $\alpha = \sum_{i=1}^{N} a_i\gamma_i$, where $a_i \in F$. Since $F \subset K$, we can consider this a linear combination with coefficients in K. Thus L is spanned over K by a finite set, so that $[L : K] = \dim_K L < \infty$.
To prove part(2), let $m = [K : F]$ and $n = [L : K]$, and pick bases $\alpha_1, ..., \alpha_m$ of K over F and $\beta_1, ..., \beta_n$ of L over K. We will prove that the mn products

$$\alpha_i\beta_j,\ 1 \le i \le m,\ 1 \le j \le n,$$

form a basis of L over F.

Take $a \in L$. Since $\beta_1, ..., \beta_n$ span L over K, we can write $a = \sum_{j=1}^{n} b_j\beta_j$, where $b_1, ..., b_n \in K$. Then, since $\alpha_1, ..., \alpha_m$ span K over F, we have $b_j = \sum_{i=1}^{m} a_{ij}\alpha_i$, where $a_{ij} \in F$. Combining these equations, we get

$$\gamma = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}\alpha_i\right)\beta_j = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\alpha_i\beta_j.$$

Since $a_{ij} \in F$, this means $\alpha_i\beta_j$ span L over F.

To prove linear independence, suppose that we have a linear relation,

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

where $a_{ij} \in F$. As above, we can write this as

$$\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}\alpha_i\right)\beta_j = 0.$$

The expressions in the large parentheses all lie in K, and since the β_j are linearly independent over K, we conclude that

$$\sum_{i=1}^{m} a_{ij}\alpha_i = 0 \text{ for } 1 \le j \le n.$$

Since the α_i are linearly independent over F and $a_{ij} \in F$, we must have $a_{ij} = 0$ for all i and j, proving the desired linear independence.

Since the basis product forms a basis of L over F, we have proved the theorem. □

Now we prove a theorem on subfields of finite fields, which will be useful in the next section to count irreducible polynomials.

Theorem 5.12. *Let $\mathbb{F}_{p^m}$ and $\mathbb{F}_{p^n}$ be finite fields. Then $\mathbb{F}_{p^m}$ is isomorphic to a subfield of $\mathbb{F}_{p^n}$ if and only if $m|n$.*

Proof. As this is a bidirectional proof, we must prove both directions.

First suppose $\mathbb{F}_{p^m}$ is isomorphic to a subfield of $\mathbb{F}_{p^n}$. Then we know $\mathbb{F}_p \subset \mathbb{F}_{p^m} \subset \mathbb{F}_{p^n}$.

Proposition 4.1 and Lemma 5.11 imply that $n = [\mathbb{F}_{p^n} : \mathbb{F}_p] = [\mathbb{F}_{p^n} : \mathbb{F}_{p^m}][\mathbb{F}_{p^m} : \mathbb{F}_p] = [\mathbb{F}_{p^n} : \mathbb{F}_{p^m}]m$. Then we have that m divides n.

Now suppose that $m|n$. By Theorem 5.6, $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ is cyclic of order n. Then $\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ has a subgroup H of order $\frac{n}{m}$.

By The Fundamental Theorem of Galois Theory(Theorem 5.7,Theorem 5.8), the fixed field F of H is an extension $\mathbb{F}_p \subset F \subset \mathbb{F}_{p^n}$, satisfying $[F : \mathbb{F}_p] = [\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) : H] = \frac{n}{n/m} = m$. This equality follows from the end of Theorem 5.7.

Proposition 4.1 demonstrates F has order p^m. By Theorem 4.9, it follows that F is a subfield of $\mathbb{F}_{p^n}$ isomorphic to $\mathbb{F}_{p^m}$ □

When $m|n$, we have $\mathbb{F}_{p^m} \simeq F \subset \mathbb{F}_{p^n}$, yielding $\mathbb{F}_{p^m} \subset \mathbb{F}_{p^n}$.

6. Irreducible Polynomials and the Möbius Function

An interesting aspect of finite fields is that we can count how many irreducible polynomials the field contains using the Möbius function. Of course, in order to do this, we need to define the Möbius function.

Definition 6.1. The *Möbius function* is

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^s & \text{if } n = p_1 ... p_s \text{ for distinct primes } p_1, ..., p_s \\ 0 & \text{otherwise.} \end{cases}$$

Definition 6.2. An *arithmetic function* is a function whose domain is the positive integers and range is a subset of the complex numbers.

Example 6.3. The *divisor function* is a function whose value at the positive integer n is equal to the number of divisors of n.

Theorem 6.4. *If g and f are arithmetic functions satisfying $g(n) = \sum_{d|n} f(d)$ for all integers $n \geq 1$, then the Möbius inversion formula states*

$$f(n) = \sum_{d|n} \mu(d) g(\frac{n}{d}),$$

where μ is the Möbius function.

The proof to this formula can be found in An Introduction to the Theory of Numbers [3, Result 4.8].

Proposition 6.5. *Let $f \in \mathbb{F}_p[x]$ be irreducible of degree m. Then:*

(1) f divides $x^{p^m} - x$;
(2) f is separable;
(3) Given an integer $n \geq 1$, f divides $x^{p^n} - x \iff f$ has a root in $\mathbb{F}_{p^n} \iff m|n$.

Proof. We will first prove part(3). Let α be a root of f in a splitting field over $\mathbb{F}_p$. Since f is irreducible, $\mathbb{F}_p \subset F_p(\alpha)$ has degree m. Then $\mathbb{F}_p(\alpha) \simeq \mathbb{F}_{p^m}$ by Proposition 4.1 and Theorem 4.9. The second equivalence follows from Theorem 5.12. Since f is irreducible over $\mathbb{F}_p$, we also have $f | \gcd(f, x^{p^n} - x)$ if and only if the degree of $\gcd(f, x^{p^n} - x) > 0$. Then we get the second equivalence from Proposition 5.5.

part(1) follows by taking $n = m$ in part(3) and part(2) follows since $x^{p^m} - x$ is separable by the proof of Theorem 4.11. □

Our goal is to count the number of irreducible polynomials in $\mathbb{F}_p[x]$. This will be finite, as we are working in a finite field. Since any irreducible polynomial becomes monic after multiplying by a suitable constant, it is sufficient to compute

$$N_m = |\{f \in \mathbb{F}_p[x] | f \text{ is monic irreducible of degree } m\}|.$$

Theorem 6.6. *Let N_m be as defined above. Then, for any $n \geq 1$, we have*

$$\sum_{m|n} mN_m = p^n,$$

where the sum is over all positive divisors of n.

Proof. Since $x^p - x$ is separable, we know that it factors as a product of distinct irreducible polynomials in $\mathbb{F}_p[x]$. Since it is monic, we can assumed the factored polynomials are also monic. Additionally, part (c) of Proposition 6.5 shows that the polynomials in the factorization are all monic irreducible polynomials of $\mathbb{F}_p[x]$ whose degree m divides n.

We can write $x^{p^n} - x$ as follows. Let

$$\mathcal{N}_m = \{f \in \mathbb{F}_p[x] \mid f \text{ is monic irreducible of degree } m\},$$

so that $N_m = |\mathcal{N}_m|$. Then the above paragraph implies

$$x^p - x = \prod_{m|n} \prod_{f \in \mathcal{N}_m} f.$$

□

Example 6.7. Monic irreducible polynomials of degree 1 in $\mathbb{F}_p[x]$ are of the form $x - a$ for $a \in \mathbb{F}_p$, so $N_1 = p$. The theorem implies

$$p^2 = 2N_2 + N_1 = 2N_2 + p,$$

meaning $N_2 = \frac{1}{2}(p^2 - p)$.

Theorem 6.8. *The number of monic irreducible polynomials of degree n in $\mathbb{F}_p[x]$ is given by*

$$N_n = \frac{1}{n} \sum_{m|n} \mu(m) p^{\frac{n}{m}}.$$

Proof. Let F be a complex-valued function defined on the set of positive integers. Then a function G is defined by $G(n) = \sum_{m|n} F(m)$, where the sum is over all the positive divisors of n. The Möbius inversion formula (Theorem 6.4) demonstrates we can express F in terms of G as follows:

$$F(n) = \sum_{m|n} \mu(m)G(\frac{n}{m}.$$

If $F(n) = nN_n$, then Theorem 6.6 implies that

$$G(n) = \sum_{m|n} F(m) = \sum_{m|n} mN_m = p^n.$$

The inversion formula then yields

$$nN_n = F(n) = \sum_{m|n} \mu(m)G(\frac{n}{m} = \sum_{m|n} \mu(m)p^{\frac{n}{m}}.$$

Dividing both sides by n immediately gives the desired formula. □

Example 6.9. When n=4, Theorem 6.8 implies that

$$\begin{aligned} N_4 &= \frac{1}{4}(\mu(1)p^{\frac{4}{1}} + \mu(2)p^{\frac{4}{2}} + \mu(4)p^{\frac{4}{4}} \\ &= \frac{1}{4}(1 \cdot p^4 + (-1)^2 + 0 \cdot p) \\ &= \frac{1}{4}(p^4 - p^2). \end{aligned}$$

Thus, through our exploration of finite fields, we have derived the algorithm for counting the number of irreducible polynomials in a finite field.

7. Appendix

Definition 7.1. A *field* is a set F with two binary operations on F called addition, denoted $+$, and multiplication, denoted $\cdot$, satisfying the following *field axioms*.

(1) Commutativity of Addition;
(2) Associativity of Addition;
(3) Additive Identity;
(4) Additive Inverse;
(5) Commutativity of Multiplication;
(6) Associativity of Multiplication;
(7) Multiplicative Identity;
(8) Multiplicative Inverse (except for the additive identity);
(9) Distributivity of Multiplication over Addition;
(10) Distinct Additive and Multiplicative Identities $1 \neq 0$.

Definition 7.2. *Rings* are defined on fields. Multiplication is not necessarily commutative and the existence of multiplicative inverses is not guaranteed.

Example 7.3. $\mathbb{Z}$ is a commutative ring with respect to the usual addition and multiplication.
We specify that it is *commutative* because multiplication in $\mathbb{Z}$ is commutative, which is not a guarantee of a ring structure. We know $\mathbb{Z}$ contains 0 and 1, the additive and multiplicative identities, respectively. Since the usual addition and multiplication is defined on the integers, we know both operations are associative

and that addition is commutative. The additive inverses exist in $\mathbb{Z}$, as well as the distributivity of multiplication over addition. Note that even though multiplication is commmutative, multiplicative inverses don't exist in $\mathbb{Z}$ because they would require elements of $\mathbb{Q}$. As such, $\mathbb{Z}$ is not a field, but it is a ring, albeit a commutative one.

Definition 7.4. Let R be a ring. An *ideal* S of R is a subset $S \subset R$ such that:

(1) S is closed under addition.
(2) The 0 element of R is in S.
(3) S is closed under additive inverses. If $a \in S$, then $-a \in S$.
(4) If $r \in R$ and $x \in S$, then $rx \in S$ and $xr \in S$. In other words, S is closed under multiplication (on either side) by arbitrary ring elements.

Definition 7.5. A *quotient ring* is the quotient of a ring, A, and one of its ideals, $\mathfrak{a}$, denoted $A/\mathfrak{a}$. In general, a quotient ring is a set of equivalence classes where $[x] = [y]$ if and only if $x - y \in \mathfrak{a}$.

Example 7.6. Consider the ring $\mathbb{Z}$ and the ideal $6\mathbb{Z}$ (multiples of 6. The quotient ring would be $\mathbb{Z}_6 = \mathbb{Z}/6\mathbb{Z}$.
We are familiar with the notation $\mathbb{Z}/p\mathbb{Z}$ (used in Section 2) to describe the integers modulo a prime p, but note that the integers modulo any integer n is a quotient ring.

Definition 7.7. Given a subgroup H of a group G, the *left coset* determined by $g \in G$ is

$$gH = \{gh \mid h \in H\},$$

and the *right coset* determined by g is

$$Hg = \{hg \mid h \in H\}.$$

Example 7.8. Let G be the additive group of the integers and H be the subgroup $(3\mathbb{Z}, +)$. Then the right cosets of H in G are $3\mathbb{Z}, 3\mathbb{Z} + 1, 3\mathbb{Z} + 2$. Since addition on the integers is commutative, the left cosets equal the right cosets.

Definition 7.9. A map is *structure preserving* if it preserves the properties of the domain. For example, let G and F be groups. A map $\varphi : G \to F$ is structure preserving if $\varphi(xy) = \varphi(x)\varphi(y)$.

Definition 7.10. A polynomial is *monic* if the leading coefficient is 1.

Definition 7.11. Let F be a field. An $\alpha \in F$ is *algebraic* over F if there exists a polynomial $g(x)$ with coefficents in F such that $g(\alpha) = 0$.

Definition 7.12. A *homomorphism* is a structure preserving map between two algebraic structures of the same type.

Definition 7.13. An *isomorphism* is a structure preserving map between two structures of the same type that can be inverted. The inverse is also a homomorphism.

Definition 7.14. An *automorphism* is an isomorphism from an object to itself.

Definition 7.15. A *ring homomorphism* is a homomorphism between two rings. More explicitly, if R_1 and R_2 are rings, then a ring homomorphism is a map $f : R_1 \to R_2$ such that

(1) Addition is preserved: $f(a + b) = f(a) = f(b) \ \forall \ a, b \in R_1$;
(2) Multiplication is preserved: $f(ab) = f(a)f(b) \ \forall \ a, b \in R_1$;
(3) Multiplicative identity is preserved: $f(1_{R_1}) = 1_{R_2}$.

Acknowledgments. It is a pleasure to thank my mentor, Livia Xu, for her guidance and help writing this paper. I would also like to thank Peter May for organizing the University of Chicago REU and giving me this opportunity.

References

[1] David A. Cox. Galois Theory 2nd Edition. John Wiley and Sons, Inc. March 2012.
[2] André Weil. Number theory: An approach through history; from Hammurapi to Legendre Birkhäuser. 2007.
[3] Ivan Niven, Herbert S. Zuckerman. An Introduction to the Theory of Numbers, Fifth Edition. John Wiley and Sons, Inc. 1991