AN INTRODUCTION TO THE THEORY OF FIELD EXTENSIONS

SAMUEL MOY

ABSTRACT. Assuming some basic knowledge of groups, rings, and fields, the following investigation will introduce the reader to the theory of rings before proceeding to elaborate, in greater depth, on the theory of field extensions. Finally, a few consequences of the subject will be examined by solving classical straightedge and compass problems in a manner that effectively utilizes the material.

CONTENTS

1. THE BASICS

Definition 1.1. : A *ring* R is a set together with two binary operations $+$ and $\times$ (addition and multiplication, respectively) satisyfing the following axioms:

(i) $(R, +)$ is an *abelian* group,
(ii) $\times$ is associative: $(a \times b) \times c = a \times (b \times c) \quad$ for all $a, b, c \in R$,
(iii) the *distributive laws* hold in $R \quad$ for all $a, b, c \in R$:

$$(a + b) \times c = (a \times c) + (b \times c) \quad \text{and} \quad a \times (b + c) = (a \times b) + (a \times c).$$

Definition 1.2. The ring R is *commutative* if multiplication is commutative.

Definition 1.3. The ring R is said to have an *identity* (or contain a 1) if there is an element $1 \in R$ with

$$1 \times a = a \times 1 = a \quad \text{for all} \quad a \in R$$

Definition 1.4. A ring R with identity 1, where $1 \neq 0$, is called a *division ring* (or *skew field*) if $\forall$ nonzero element $a \in R$, $\exists\, b \in R$ such that $ab = ba = 1$.

Definition 1.5. A commutative division ring is called a *field*.

Example 1.6. $\mathbb{Z}$ is a commutative ring with 1(identity). $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}/p\mathbb{Z}$ (the integers modulo p, where p is prime) are all fields.

2. Ring Theory

Before beginning further study of fields, additional knowledge pertaining to rings is necessary. Most notably, the topics of ideals, ring homomorphisms and isomorphisms, and quotient rings must first be approached.

Definition 2.1. A *subring* of the ring R is a subgroup of R that is closed under multiplication (i.e. A subset S of a ring R is a subring if the operations of addition and multiplication in R when restricted to S give S the structure of a ring).

Definition 2.2. Let R and S be rings.

(1) A *ring homomorphism* is a map $\varphi : R \to S$ satisfying
   (a) $\varphi(a+b) = \varphi(a) + \varphi(b)$ for all $a, b \in R$, and
   (b) $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in R$.

(2) The *kernel* of the ring homomorphism φ, denoted kerφ, is the set of elements of R that map to 0 in S.

(3) A bijective ring homomorphism is called an *isomorphism.*

Definition 2.3. Let R be a ring, let I be a subset of R and let $r \in R$.

(1) $rI = \{ra \mid a \in I\}$ and $Ir = \{ar \mid a \in I\}$.

(2) A subset I of R is a *left ideal* (respectively *right ideal*) of R if
   (a) I is a subring of R, and
   (b) I is closed under left multiplication (respectively right multiplication) by elements from R.

(3) A subset I that is both a left ideal and a right ideal is called an *ideal* (or, a *two-sided ideal*) of R.

Definition 2.4. Let R be a ring and I an ideal of R. Then the *quotient ring* of R by I, denoted R/I is the ring defined by the following binary operations:

$$(r+I) + (s+I) = (r+s) + I \quad \text{and} \quad (r+I) \times (s+I) = (rs+I) \quad \forall r, s \in R.$$

Theorem 2.5. *(The First Isomorphism Theorem for Rings)*

(1) If $\varphi : R \to S$ is a homomorphism of rings, then the kernel of φ is an ideal of R, the image of φ is a subring of S and R/kerφ is isomorphic as a ring to $\varphi(R)$.

(2) If I is any ideal of R, then the map

$$R \to R/I \quad \textit{defined by} \quad r \mapsto r + I$$

is a surjective ring homomorphism with kernel I (this homomorphism is called the natural projection of R onto R/I). Thus, every ideal is the kernel of a ring homomorphism and vice versa.

Proof. Let I be the kernel of φ. Then the cosets under addition of I are exactly the fibers of φ (the sets of elements of R that map to a single element of S). In particular, the cosets $r+I$, $s+I$, and $rs+I$ are the fibers of φ over $\varphi(r), \varphi(s), \varphi(rs)$, respectively. Since φ is a ring homomorphism, $\varphi(r)\varphi(s) = \varphi(rs)$, hence $(r+I)(s+I) = rs+I$. Multiplication of cosets is well defined and so I is an ideal and R/I is a ring. The correspondance $r + I \mapsto \varphi(r)$ is a bijection between the rings R/I and $\varphi(R)$ which respects addition and multiplication. Hence, it is a ring isomorphism. If I is any ideal, then R/I is a ring (in particular is an abelian group) and the map $\pi : r \mapsto r + I$ is a group homomorphism with kernel I (natural projection for groups). It remains to check that π is a ring homomorphism. This is immediate from the definition of multiplication in R/I:

$$\pi : rs \mapsto rs + I = (r + I)(s + I) = \pi(r)\pi(s)$$

□

Theorem 2.6 (*The Lattice Isomorphism Theorem for Rings*)**.** *Let R be a ring and let I be an ideal of R. The correspondance $A \leftrightarrow A/I$ is an inclusion preserving bijection between the set of subrings A of R that contain I and the set of subrings of R/I. Furthermore, A (a subring containing I) is an ideal of R if and only if A/I is an ideal of R/I.*

The proof for this theorem will not be provided. However, it follows almost immediately from the Lattice Isomorphism Theorem for Groups. (In addition, rather than considering subgroups, we must consider ideals).

Definition 2.7. Let I and J be ideals of R.

(1) Define the *sum* of I and J by $I + J = \{a + b \mid a \in I, b \in J\}$.
(2) Define the *product* of I and J, denoted by IJ, to be the set of all *finite* sums of elements of the form ab with $a \in I$ and $b \in J$.

Definition 2.8. Let A be any subset of the ring R. Let (A) denote the smallest ideal of R containing A, called the *ideal generated by A*,

$$(A) = \bigcap_{I \supseteq A} I.$$

Definition 2.9. An ideal M in an arbitrary ring R is called a *maximal ideal* if $M \neq R$ and the only ideals containing M are M and R.

Definition 2.10. Assume R is a commutative ring. An ideal P is called a *prime ideal* if $P \neq R$ and whenever the product ab of two elements $a, b \in R$ is an element of P, then at least one of a and b is an element of P.

The following two propositions will be useful for later theorems regarding fields. Only Proposition 2.12 will be proved now (Proposition 2.11 will appear as a Lemma for a later theorem and will be proved then).

Proposition 2.11. *Assume R is a commutative ring. Then R is a field if and only if its only ideals are 0 and R.*

Proposition 2.12. *Assume R is a commutative ring. The ideal M is a maximal ideal if and only if the quotient ring R/M is a field.*

Proof. This follows from the Lattice Isomorphism Theorem for Rings along with Proposition 2.11. The ideal M is maximal if and only if there are no ideals I with $M \subset I \subset R$. By the Lattice Isomorphism Theorem the ideals of R containing M correspond bijectively with the ideals of R/M, so M is maximal if and only if the ideals of R/M are 0 and R/M. By Proposition 2.11 we see that M is a maximal ideal if and only if R/M is a field. □

Example 2.13. To assist in the understanding of what an ideal is, we will consider ideals in the commutative ring with 1, $\mathbb{Z}$. An example of an ideal in this ring is $(2) = \{2a \mid a \in \mathbb{Z}\}$ = multiples of 2 (the ideal generated by the element $2 \in \mathbb{Z}$).
Now, it should not be difficult to see that the only prime ideals in $\mathbb{Z}$ are those generated by prime numbers:
Suppose we have the ideal (n), n composite (and positive) (i.e. $n = ab, 0 < a, b \neq 1$). WLOG, assume $a < b < n$. Clearly $a \cdot n \in (n) \Rightarrow a \cdot a \cdot b = a^2 \cdot b \in (n)$. However, $0 < a^2 < n$ and $0 < b < n \Rightarrow$ neither a^2 nor $b \in (n)$. Therefore, (n) is not a prime ideal. Now take the ideal (p), p prime. Suppose $ab \in (p)$. Then $p \mid ab \Rightarrow p \mid a$ or $p \mid b$ (in other words, $a = n \cdot p$, $n \in \mathbb{N}$ or $b = n \cdot p$, $n \in \mathbb{N}$) $\Rightarrow a \in (p)$ or $b \in (p) \Rightarrow$ (p) must be a prime ideal.
In addition, we see that the only maximal ideals in $\mathbb{Z}$ generated by a single element are, again, those generated by prime numbers. To see why this is true, choose some ideal (p), p prime. The only ideals that contain the element $p \in \mathbb{Z}$ are (1) and (p). But $(1) = \mathbb{Z}$! Thus, (p) must then be a maximal ideal. And if we have some maximal ideal $(p^{'})$ generated by a single element, then it is contained only in $\mathbb{Z} = (1)$ and $(p^{'})$. But this implies that $p^{'} \in \mathbb{Z}$ is divisible by only $1, p^{'} \in \mathbb{Z} \Rightarrow p^{'}$ is prime.
Finally, if we take Proposition 2.12 into consideration, we obtain the result that, of the ideals for $\mathbb{Z}$ generated by a single element, those that are fields are exactly the quotients: $\mathbb{Z}/p\mathbb{Z}$, p a prime. That these quotients are fields is a result from Example 1.6.

With these theorems and propositions regarding rings at our disposal, we may now proceed to study the more specific case of fields and the field extensions that arise from them.

3. Fields and Field Extensions

Definition 3.1. The *characteristic* of a field, F, denoted $ch(F)$, is defined to be the smallest positive integer p such that $p \cdot 1_F = 0$ if such a p exists and is defined to be 0 otherwise.

Remark 3.2 (Definition 3.1). The characteristic of a field F, $ch(F)$, is either 0 or a prime p. If $ch(F) = p$, then for any $\alpha \in F$,

$$p \cdot \alpha = \underbrace{\alpha + \alpha + \ldots + \alpha}_{\text{p times}} = 0.$$

Proof. Suppose $ch(F) = n \in \mathbb{N}$, n not prime (i.e. n composite). Then $\exists\ a, b \in \mathbb{N}$ such that $n = ab \quad a, b \neq 1$.

$$0 = n \cdot 1 = (ab) \cdot 1 = (a \cdot 1) \cdot (b \cdot 1) \Rightarrow a \cdot 1 = 0 \quad \text{or} \quad b \cdot 1 = 0.$$

But $a, b < n \Rightarrow ch(F) \neq n$. Contradiction! □

Definition 3.3. The *prime subfield* of a field F is the subfield of F generated by the multiplicative identity 1_F of F. It is isomorphic to either $\mathbb{Q}$ (if $ch(F) = 0$) or $\mathbb{F}_p$ (if $ch(F) = p$).

Definition 3.4. If K is a field containing the subfield F, then K is said to be an *extension field* (or simply an *extension*) of F, denoted K/F. This notation is shorthand for "K over F."

Definition 3.5. The *degree* of a field extension K/F, denoted $[K:F]$, is the dimension of K as a vector space over F. The extension is said to be *finite* if $[K:F]$ is finite and is said to be *infinite* otherwise.

Example 3.6. The concept of field extensions can soon lead to very interesting and peculiar results. The following examples will illustrate this:

1. Take the field $\mathbb{Q}$. Now, clearly, we have the polynomial $p(x) = x^2 - 2 \in \mathbb{Q}[x]$; however, it should be evident that its roots, $\pm\sqrt{2} \notin \mathbb{Q}$. This polynomial is then said to be *irreducible* over $\mathbb{Q}$.
   Thus, by considering the quotient ring $\mathbb{Q}[x]/(x^2-2)$, we find that we obtain another field, denoted $\mathbb{Q}(\sqrt{2})$ (or $\mathbb{Q}(-\sqrt{2})$, which just so happens to be isomorphic to $\mathbb{Q}(\sqrt{2})$ | this, of course, is no coincidence).
2. Take the field $\mathbb{R}$. Again, we may easily find a polynomial, which is irreducible over our field. Choosing $p(x) = x^2 + 1 \in \mathbb{R}[x]$, it is obvious that the roots, $\pm i \notin \mathbb{R}$. Thus, if we consider the quotient ring, $\mathbb{R}[x]/(x^2+1)$, we obtain the field $\mathbb{R}(i)$ ($\cong \mathbb{C}$!).

Since both of the given examples are of polynomials that are irreducible over the particular fields, it will be of great benefit to examine the subject of irreducible polynomials (and the criteria to label them as irreducible) more closely.

Definition 3.7. A polynomial $p(x) \in R[x]$, a polynomial ring, is said to be *irreducible* if it cannot be factored as the product of two other polynomials of smaller degrees, both in the polynomial ring, $R[x]$.

Proposition 3.8. *Let F be a field and let $p(x) \in F[x]$. Then $p(x)$ has a factor of degree one if and only if $p(x)$ has a root in F, i.e. there is an $\alpha \in F$ with $p(\alpha) = 0$.*

Proposition 3.9. *A polynomial of degree two or three over a field F is reducible if and only if it has a root in F.*

Notation 3.10. If the polynomial is of degree $\geq$4, then it may still be reducible without necessarily having roots in the field (i.e. the polynomial may have factors of degree $\geq$2, yet still have no factors of degree 1). Fortunately, the explicit examples that will be shown will only require testing for the irreducibility of polynomials of degrees 2 and 3. In addition, the polynomials that will be tested for irreducibility will be elements of $\mathbb{Q}[x]$; therefore, the actual procedure of checking for irreducibility will be trivial.

Theorem 3.11. *Let $\varphi : F \to F^{'}$ be a homomorphism of fields. Then φ is either identically 0 or is injective, so that the image of φ is either 0 or isomorphic to F.*

Lemma 3.12. *Let F be a field. Then its only ideals are 0 and F.*

Proof. Let F be a field. Suppose $\exists$ a nonzero ideal I for F. Let $0 \neq a \in I$.
F a field $\Rightarrow \exists\ a^{-1}$ such that $a \cdot a^{-1} = a^{-1} \cdot a = 1$. Thus, $\forall\ r \in F$, we have $r = r \cdot 1 = r \cdot (a^{-1} \cdot a) = (r \cdot a^{-1}) \cdot a \in I$ (because $r \cdot a^{-1} \in F$). Hence, $I = F$.
This leaves $I = \{0\}$ as the only other possible ideal for F (it is very easy to check that this is ideal). □

Lemma 3.13. *If F is a field, then any nonzero ring homomorphism from F into another ring is injective.*

Proof. Let $\varphi : F \rightarrow F^{'}$ be a nonzero ring homomorphism. Since $ker\varphi$ is an ideal, and because φ is a nonzero ring homomorphism, then $ker\varphi$ is a proper ideal. As F is a field, then $ker\varphi = 0$ (because 0 is the only proper ideal of a field F).
Therefore, $\varphi : F \rightarrow F^{'}$ is injective. □

Proof of Theorem 3.11. Proving this theorem requires only combining the results of Lemmas 3.12 and 3.13. As F is a field, then its only ideals are 0 and F. Now, if $\varphi : F \rightarrow F^{'}$ is any nonzero ring homomorphism, by Lemma 3.13, it is injective, implying that the image of φ is isomorphic to F (i.e. $\varphi(F) \cong F$). And if we take φ to be a homomorphism that is *not* nonzero, then, it must be the zero homomorphism and therefore must be identically 0. □

Theorem 3.14. *Let F be a field and let $p(x)$ be an irreducible polynomial. Then there exists a field K containing an isomorphic copy of F in which $p(x)$ has a root. Thus, there exists an extension of F in which $p(x)$ has a root.*

Proof. Define $K = F[x]/(p(x))$. As $p(x)$ is irreducible over F, the ideal $(p(x))$ is a maximal ideal in $F[x] \Rightarrow K = F[x]/(p(x))$ is a field (by the *Lattice Isomorphism Theorem for Rings*). Let π be the natural projection (described in Theorem 2.5) of $F[x]$ to the quotient $F[x]/(p(x))$ restricted to the domain $F \subseteq F[x]$. We now have $\varphi = \pi\mid_F : F \rightarrow K$ a nonzero homomorphism, and thus, $\varphi(F)$ is an isomorphic copy of F contained in K. Last, it must be shown that $p(x)$ has a root in K. Let $\overline{x} = \pi(x)$ be the image of x in the quotient field K.

$$\begin{aligned} p(\overline{x}) &= \overline{p(x)} && \text{(as } \pi \text{ is a homomorphism)} \\ &= p(x)(modp(x)) && \text{(in} \quad K = F[x]/(p(x))) \\ &= 0 && \text{(in} \quad F[x]/(p(x))) \end{aligned}$$

Therefore, K is a field that contains an isomorphic copy of F in which $p(x)$ has a root. □

Theorem 3.15. *Let $p(x) \in F[x]$ be an irreducible polynomial of degree n over the field F and let K be the field $F[x]/(p(x))$. Let $\theta = x(modp(x)) \in K$. Then the elements*

$$1, \theta, \theta^2, \ldots, \theta^{n-1}$$

form a basis for K as a vector space over F so the degree of the extension is n. Hence,

$$K = \{a_0 + a_1\theta + a_2\theta^2 + \ldots + a_{n-1}\theta^{n-1} \mid a_0, a_1, a_2, \ldots, a_{n-1} \in F\}$$

consists of all polynomials of degree $< n$ in θ.

Proof. Let $a(x) \in F[x]$. Dividing $a(x)$ by $p(x)$, we obtain

$$a(x) = q(x) \cdot p(x) + r(x)$$

$r(x), q(x) \in F[x]$, and the remainder $r(x)$ with degree $< n$.
Now, since $q(x) \in F[x]$, then $q(x) \cdot p(x) \in (p(x))$, the ideal of $p(x)$, thus implying

$$a(x) \equiv r(x)(modp(x)).$$

Since $r(x)$ is of degree $< n$, we now have every residue class in $F[x]$ represented by a polynomial of degree $< n \Rightarrow 1, \theta, \theta^2, \ldots, \theta^{n-1}$, with $\theta \equiv x(modp(x))$, spans $K = F[x]/(p(x))$ as a vector space over F. It now remains to show that $1, \theta, \theta^2, \ldots, \theta^{n-1}$ are linearly independent. Suppose $1, \theta, \theta^2, \ldots, \theta^{n-1}$ are *not* linearly independent. Then $\exists$ $b_0, b_1, b_2, \ldots, b_{n-1} \in F$ such that

$$\begin{aligned} &b_0 + b_1\theta + b_2\theta^2 + \ldots + b_{n-1}\theta^{n-1} = 0 && \text{not all} \quad b_i = 0 \\ \Rightarrow\ &b_0 + b_1 x + b_2 x^2 + \ldots + b_{n-1}x^{n-1} \equiv 0(modp(x)) && \text{not all} \quad b_i = 0 \end{aligned}$$

$$\Rightarrow p(x) \quad \text{divides the polynomial} \quad b_0 + b_1 x + b_2 x^2 + \ldots + b_{n-1}x^{n-1}.$$

But $p(x)$ is of degree n whereas $b_0 + b_1 x + b_2 x^2 + \ldots + b_{n-1}x^{n-1}$ is of degree strictly less than n. Contradiction! Thus, $1, \theta, \theta^2, \ldots, \theta^{n-1}$ is a linearly independent spanning set (and thus a basis) for K as a vector space over $F \Rightarrow [K : F] = n$. □

Theorem 3.15 then leads to the following corollary.

Corollary 3.16. *Let K be as in Theorem 3.15, and let $a(\theta), b(\theta) \in K$ be two polynomials of degree $< n$ in θ. Then addition in K is defined simply by usual polynomial addition and multiplication in K is defined by*

$$a(\theta)b(\theta) = r(\theta)$$

where $r(x)$ is the remainder (of degree $< n$) obtained after dividing the polynomial $a(x)b(x)$ by $p(x)$ in $F[x]$.

Since the residue classes of $F[x]/(p(x))$ are all $r(\theta)$ of degree $< n$, we again see that K defined as $K = F[x]/(p(x))$ is a field over our chosen operations described above.

Definition 3.17. Let K be an extension of the field F and let $\alpha, \beta, \ldots \in K$ be a collection of elements of K. Then the smallest subfield of K containing both F and the elements $\alpha, \beta, \ldots$, denoted $F(\alpha, \beta, \ldots)$ is called the field *generated by $\alpha, \beta, \ldots$ over F*.

Definition 3.18. If the field K is generated by a single element α over F, $K = F(\alpha)$, then K is said to be a *simple extension* of F and the element α is called a *primitive element* for the extension.

Theorem 3.19. *Let F be a field and let $p(x) \in F[x]$ be an irreducible polynomial. Suppose K is an extension field of F containing the root α of $p(x)$, $p(\alpha) = 0$. Let $F(\alpha)$ denote the subfield of K generated by α over F. Then*

$$F(\alpha) \cong F[x]/(p(x)).$$

Proof. Consider the function

$$\begin{aligned} \varphi : F(x) &\to F(\alpha) \\ a(x) &\mapsto a(\alpha) \end{aligned}$$

This function can easily be shown to be a ring homomorphism by considering properties of polynomials:

$$\varphi(a(x)+b(x)) = \varphi((a+b)(x)) = (a+b)(\alpha) = a(\alpha)+b(\alpha) = \varphi(a(x))+\varphi(b(x)), \quad \text{and}$$

$$\varphi(a(x)\cdot b(x)) = \varphi(a(b(x))) = a(b(\alpha)) = a(\alpha)\cdot b(\alpha) = \varphi(a(x))\cdot\varphi(b(x)).$$

Now, α a root of p(x) $\Rightarrow$ p$(\alpha) = 0$. Thus, p$(x) \in \ker\varphi$, allowing us to obtain the induced homomorphism

$$\varphi^{'} : F[x]/(p(x)) \to F(\alpha)$$

As $p(x)$ is irreducible, then $F[x]/(p(x))$ is a field. Thus, since $\varphi^{'}$ is a nonzero ring homomorphism, then $\varphi^{'}$ is injective and $F[x]/(p(x)) \cong \varphi^{'}(F[x]/(p(x)))$. But $F[x]/(p(x))$ contains the root α in addition to an isomorphic copy of F (if we consider the natural projection from F to $F[x]/(p(x))$). Hence, $\varphi^{'} : F[x]/(p(x)) \to F(\alpha)$ is surjective in addition to being injective. Therefore, $\varphi^{'}$ is an isomorphism. □

Corollary 3.20. *Suppose in Theorem 3.19 that $p(x)$ is of degree n. Then*

$$F(\alpha) = \{a_0 + a_1\alpha + a_2\alpha^2 + \ldots + a_{n-1}\alpha^{n-1} \mid a_0, a_1, a_2, \ldots, a_{n-1} \in F\} \subseteq K$$

We obtain this corollary from Theorem 3.19 (where $1, \theta, \theta^2, \ldots, \theta^{n-1}$ is a basis for $F[x]/(p(x))$ with $\theta \equiv x(mod(p(x)))$). By applying the previously described map $\varphi^{'} : F[x]/(p(x)) \to F(\alpha)$, we see that

$$r(\theta) = a_0 + a_1\theta + \ldots + a_{n-1}\theta^{n-1} \mapsto a_0 + a_1\alpha + \ldots + a_{n-1}\alpha^{n-1} = r(\alpha).$$

Example 3.21. In order to make sense of all the previous theorems regarding field extensions, we will work with a simple example. Consider $p(x) = x^3 - 2 \in \mathbb{Q}[x]$. Having been exposed to the fields $\mathbb{R}$ and $\mathbb{C}$, we are well aware that $\nexists\ \alpha \in \mathbb{Q}$ such that $p(\alpha) = 0$.
Since $p(x)$ has no roots in $\mathbb{Q}$ and is of degree 3, then by our divisibility criteria, $p(x)$ must be irreducible over $\mathbb{Q}$.
Therefore, by Theorem 3.14, $\exists$ an extension of $\mathbb{Q}$ in which $p(x)$ has a root. Moreover, this extension may be written as the quotient $\mathbb{Q}[x]/(x^3-2)$. By Theorem 3.19, $\mathbb{Q}[x]/(x^3-2) \cong \mathbb{Q}(\alpha)$, α a root of $p(x) = x^3 - 2$.
If we now use our knowledge of $\mathbb{R}$, we may let this root, α, of $p(x)$ be $\alpha = \sqrt[3]{2}$. Thus, we obtain $\mathbb{Q}(\sqrt[3]{2})$, which is spanned by the basis $\{1, \sqrt[3]{2}, \sqrt[3]{2^2} = \sqrt[3]{4}\}$ over $\mathbb{Q}$. $\Rightarrow \mathbb{Q}(\sqrt[3]{2})$ is an extension of degree 3 over $\mathbb{Q}$.

Remark 3.22. The polynomial $p(x) = x^3 - 2$ has three roots over $\mathbb{C}$:

$$\sqrt[3]{2}, e^{2\pi i/3}\sqrt[3]{2}, e^{4\pi i/3}\sqrt[3]{2}.$$

A somewhat peculiar result is that an extension of $\mathbb{Q}$ by any one of these roots is isomorphic to the others. However, this result may be very easily justified if one considers that by Theorem 3.19, we have:

$$\mathbb{Q}[x]/(x^3-2) \cong \mathbb{Q}(\sqrt[3]{2}),$$
$$\mathbb{Q}[x]/(x^3-2) \cong \mathbb{Q}(e^{2\pi i/3}\sqrt[3]{2}),$$
$$\mathbb{Q}[x]/(x^3-2) \cong \mathbb{Q}(e^{4\pi i/3}\sqrt[3]{2}).$$

It then seems only natural that all of these extensions are isomorphic to one another. Thus, the three roots of x^3-2 are said to be *algebraically indestinguishable*, which makes sense, considering they are roots of the same irreducible polynomial!

4. Algebraic Field Extensions

Definition 4.1. The element $\alpha \in K$ is said to be *algebraic* over F if α is a root of some nonzero polynomial $f(x) \in F[x]$. If α is not algebraic over F, then α is said to be *transcendental* over F. The extension K/F is said to be *algebraic* if every element of K is algebraic over F.

Example 4.2. The number $\sqrt{2}$ is algebraic over the field $\mathbb{Q}$. The extension $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ is algebraic. Since every element $\alpha \in \mathbb{Q}(\sqrt{2})$ can be expressed as $\alpha = a+b\cdot\sqrt{2}$, $a,b \in \mathbb{Q}$, it can very easily be seen that α is the root of some polynomial in $\mathbb{Q}[x]$ (Simply take the polynomial $p(x) = x^2-2ax+a^2+b^2$. Its roots are $a+b\sqrt{2}$ and $a-b\sqrt{2}$). On the other hand, the extension $\mathbb{R}/\mathbb{Q}$ is not algebraic (because of the existence of *transcendental numbers* such as e and π which are not the roots of any polynomials in $\mathbb{Q}[x]$).

Proposition 4.3. *Let α be algebraic over F. Then there is a unique monic irreducible polynomial $m_{\alpha,F}(x) \in F[x]$, which has α as a root. Furthermore, a polynomial $f(x) \in F[x]$ has α as a root if and only if $m_{\alpha,F}(x)$ divides $f(x) \in F[x]$.*

Proof. Let $g(x)$ be a polynomial of minimal degree such that it has α as a root. We may multiply $g(x)$ by some constant to obtain a monic polynomial. Now, suppose this polynomial is reducible over F. Then $\exists\ a(x), b(x) \in F[x]$ such that $g(x) = a(x)b(x), deg(a(x)), deg(b(x)) < deg(g(x))$. Now $0 = g(\alpha) = a(\alpha)b(\alpha)$. Since F is a field (and thus has no zero divisors), then $a(\alpha) = 0$ or $b(\alpha) = 0$. But this contradicts the minimality of the degree of $g(x)$ having α as a root. Therefore, we have obtained $g(x)$, a monic irreducible polynomial with α as a root.

Now suppose $\exists\ f(x) \in F[x]$ with α as a root. We can express $f(x)$ as $f(x) = q(x)g(x) + r(x), \quad deg(r(x)) < deg(g(x)) \Rightarrow f(\alpha) = q(\alpha)g(\alpha) + r(\alpha) \Rightarrow r(\alpha) = 0$. But this contradicts the minimality of $g(x)$ unless $r(x) = 0$. Hence, $g(x)$ divides $f(x)$. To show the converse also holds, suppose $g(x)$ divides $f(x)$. Then $\exists\ a(x) \in F[x]$ such that $f(x) = a(x)g(x) \Rightarrow f(\alpha) = a(\alpha)g(\alpha) = a(\alpha)\cdot 0 = 0 \Rightarrow \alpha$ is a root of $f(x)$, thus proving the theorem. □

Definition 4.4. The polynomial $m_{\alpha,F}(x)$ in Proposition 4.3 is called the *minimal polynomial* for α over F. The *degree* of $m_{\alpha,F}(x)$ is called the *degree* of α.

Proposition 4.5. *Let α be algebraic over the field F and let $F(\alpha)$ be the field generated by α over F. Then*

$$F(\alpha) \cong F[x]/(m_\alpha(x))$$

or, in other words,

$$[F(\alpha) : F] = \deg m_\alpha(x) = \deg \alpha.$$

Proof. This proposition follows immediately from the result of Proposition 4.3. Since we saw $[K:F]=n$, where $K=F[x]/(p(x)), p(x)$ of degree n, then, as $m_\alpha(x)$ is the unique minimal monic polynomial, we easily obtain the result (after redefining K to be equal to $F[x]/(m_\alpha(x))$ that $[F(\alpha):F]=[K:F]=\deg m_\alpha(x)=\deg\alpha$. □

Proposition 4.6. *The element α is algebraic over F if and only if the simple extension $F(\alpha)/F$ is finite.*

Proof. Suppose α is algebraic. Then $\exists!\ m_\alpha(x)\in F[x]$ irreducible over F of degree $n\in\mathbb{N}$.

$$F(\alpha)\cong F[x]/(m_\alpha(x))\Rightarrow[F(\alpha):F]=n,\quad n\in\mathbb{N}\Rightarrow F(\alpha)/F\quad\text{is finite.}$$

Suppose $F(\alpha)/F$ is finite, $[F(\alpha):F]=n,\quad n\in\mathbb{N}$. Considering $\alpha\in F(\alpha)$, then $1,\alpha,\alpha^2,\ldots,\alpha^n$ form a linearly dependent set in the vector space $F(\alpha)$ over the field F (since the degree of $F(\alpha)$ as a vector space over F is n). Thus, $\exists\ b_0,b_1,\ldots,b_{n-1}$ not all 0 such that

$$b_0+b_1\alpha+b_2\alpha^2+\ldots+b_{n-1}\alpha^{n-1}=0$$

$\Rightarrow\exists$ some $b(x)\in F[x]$ of degree $\leq n$ for which α is a root $\Rightarrow\alpha$ is algebraic over F. □

Theorem 4.7. *Let $F\subseteq K\subseteq L$ be fields, $[L:K],[K:F]$ finite. Then,*

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

Proof. Suppose $[L:F]=m$ and $[K:F]=n$ n,m finite. L a vector space over K $\Rightarrow\exists$ a basis $\alpha_1,\alpha_2,\ldots,\alpha_m$ for L, and, similarly, K a vector space over $F\Rightarrow\exists$ a basis $\beta_1,\beta_2,\ldots,\beta_n$ for K. Now, every element of L may be expressed as a unique linear comination:

$$a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m\qquad a_1,a_2,\ldots,a_m\in K$$

And as $a_1,a_2,\ldots,a_m\in K$, they may each be expressed as by some linear combination:

$$a_i=b_{i1}\beta_1+b_{i2}\beta_2+\ldots+b_{in}\beta_n\qquad i=1,2,\ldots,m$$

Thus, every element of L may be expressed as a linear combination:

$$\sum_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}b_{ij}\alpha_i\beta_j$$

Thus, the mn elements, $\alpha_i\beta_j$, form a spanning set for L as a vector space over F (i.e. $[L:F]\leq mn$). It remains to be shown that the set of all $\alpha_i\beta_j$ is linearly independent. Suppose

$$\sum_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}b_{ij}\alpha_i\beta_j=0$$

Since $b_{ij}\beta_j\in K\ \forall i,j$, then we may view this summation as:

$$a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m=0\quad\Rightarrow\quad a_1,a_2,\ldots,a_m=0$$

since $\alpha_1,\alpha_2,\ldots,\alpha_m$ are linearly independent. This then implies

$$b_{i1}\beta_1 + b_{i2}\beta_2 + \ldots + b_{in}\beta_n = 0, \qquad i = 1, 2, \ldots, m.$$

But $\beta_1, \beta_2, \ldots, \beta_n$ forms a basis (and thus linearly independent set) for K as a vector space over $F \Rightarrow$

$$b_{i1} = b_{i2} = \ldots = b_{in} = 0, \qquad i = 1, 2, \ldots, m.$$

Therefore, the mn elements, $\alpha_i\beta_j$ form a linearly independent spanning set for L as a vector space over $F \Rightarrow [L : F] = mn = [L : K][K : F]$. □

Corollary 4.8. *Suppose L/F is a finite extension and let K be any subfield of L containing F, $F \subseteq K \subseteq L$. Then $[K : F]$ divides $[L : F]$.*

Proof. From the previous theorem, $[L : F] = [L : K][K : F]$ ($F \subset K \subset L$ fields with $[L : K], [K : F] \in \mathbb{N}$). Therefore, quite clearly, $[K : F] \mid [L : F]$. □

Definition 4.9. An extension K/F is *finitely generated* if there are elements $\alpha_1, \alpha_2, \ldots, \alpha_k$ in $K = F(\alpha_1, \alpha_2, \ldots, \alpha_k)$.

Lemma 4.10. *$F(\alpha, \beta) = (F(\alpha))(\beta)$. (i.e. The field generated over F by α and β is the field generated by β over the field $F(\alpha)$ generated by α over F).*

Proof. $F(\alpha, \beta)$ contains F and the element $\alpha \Rightarrow F(\alpha) \subseteq F(\alpha, \beta)$. Moreover, $F(\alpha, \beta)$ contains the element β as well $\Rightarrow (F(\alpha))(\beta) \subseteq F(\alpha, \beta)$ by the minimality of the field $(F(\alpha))(\beta)$ (over $F(\alpha)$ generated by β). Similarly, $(F(\alpha))(\beta)$ contains F and both the elements α and β, which implies $F(\alpha, \beta) \subseteq (F(\alpha))(\beta)$ by the minimality of $F(\alpha, \beta)$ (over F generated by α, β). Therefore, $F(\alpha, \beta) = (F(\alpha))(\beta)$. □

Theorem 4.11. *The extension K/F is finite if and only if K is generated by a finite number of algebraic elements over F. More precisely, a field generated over F by a finite number of algebraic elements of degrees $n_1, n_2, \ldots, n_k$ is algebraic of degree $\leq n_1 n_2 \ldots n_k$.*

Proof. Suppose the extension K/F is finite. Let $[K : F] = n$. Now take some basis for K over F, $\alpha_1, \alpha_2, \ldots, \alpha_n$. If we consider the extension $F(\alpha_i)$ over F, we see that

$$[K(\alpha_i) : F] \mid [K : F] \quad \forall i = 1, 2, \ldots, n$$

Thus, by Proposition 4.5, each α_i is algebraic. Now, since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are a basis for K over F, we obviously can generate K over F by the elements $\alpha_1, \alpha_2, \ldots, \alpha_n$. Thus, K is generated by a finite number of algebraic elements over F. To prove the converse, suppose that K is generated by a finite number of algebraic elements over F: $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then

$$K = F(\alpha_1, \ldots, \alpha_n) = (F(\alpha_1, \ldots, \alpha_{n-1})(\alpha_n) = \ldots = \underbrace{(\overbrace{(\overbrace{F(\alpha_1)}^{F_1})(\alpha_2))}^{F_2} \ldots (\alpha_n)}_{F_n}$$

Thus, we have the fields $F, F_1, F_2, \ldots, F_n$ satisyfing $F \subseteq F_1 \subseteq F_2 \subseteq \ldots \subseteq F_n = K$. Therefore, we have that

$$[K : F] = [F_n : F_{n-1}][F_{n-1} : F_{n-2}] \ldots [F_1 : F]$$

where $[F_i : F_{i-1}]$ is finite for $i = 1, 2, \ldots, n \Rightarrow$ the extension K/F is finite, concluding the proof. □

Proposition 4.12. *Let K_1 and K_2 be two finite extensions of a field F contained in K. Suppose that $[K_1 : F] = n$, $[K_2 : F] = m$, where n, m are relatively prime (i.e. $(n, m) = 1$). Then $[K_1K_2 : F] = [K_1 : F][K_2 : F] = nm$.*

Proof. Since K_1 and K_2 are both subfields of K_1K_2, then we know

$$[K_1 : F] \| [K_1K_2 : F] \quad \text{and} \quad [K_2 : F] \| [K_1K_2 : F]$$

But by assumption, $gcd(n, m) = 1 \Rightarrow [K_1K_2 : F]$ must be divisible by $lcm(n, m) = nm$ (i.e. $[K_1K_2 : F] \geq [K_1 : F][K_2 : F]$) To prove the equality of the expression, it must also be shown that $[K_1K_2 : F] \leq [K_1 : F][K_2 : F]$. Consider the fields K_1K_2, K_1, K_2, and F. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be a basis for K_1 over F and let $\beta_1, \beta_2, \ldots, \beta_m$ be a basis for K_2 over F. Since K_1K_2 is the *composite field* of K_1 and K_2, then it is the smallest field containing both K_1 and K_2. Hence, $K_1K_2 = F(\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_m) = K_1(\beta_1, \beta_2, \ldots, \beta_m)$. Thus, $[K_1K_2 : K_1] \leq m = [K_2 : F]$. Since $F \subseteq K_1 \subseteq K_1K_2$, then $[K_1K_2 : F] = [K_1K_2 : K_1][K_1 : F] \Rightarrow [K_1K_2 : F] \leq [K_1 : F][K_2 : F]$.
Combining both inequalities, we obtain $[K_1K_2 : F] = [K_1 : F][K_2 : F]$. □

5. Classical Straightedege and Compass Constructions

With all of these theorems now available to us, we may utilize the information on field extensions to prove the impossibility of certain classical straightedege and compass constructions. The problems we will consider are the following:

I. Is it possible using only straightedge and compass to construct a cube with precisely twice the volume of a given cube?
II. Is it possible using only straightedge and compass to trisect any given angle θ?
III. Is it possible using only straightedege and compass to construct a square whose area is precisely the area of a given circle?

Before we may answer these questions (all in the negative), we must first identify what it means for us to construct with only straightedege and compass. To begin, let us start with the distance 1 and the Cartesian plane $\mathbb{R}^2$. A point $(x, y) \in \mathbb{R}^2$ is constructible if and only if its coordinates, x and y, are both *constructible* elements of $\mathbb{R}$. Let us now consider which elements of $\mathbb{R}$ are constructible through a *finite* number of straightedege and compass operations.

By basic geometry, given two lengths a and b (and the assumed length of 1), we may construct $a \pm b$, ab, and a/b. Thus, quite clearly, we may construct all elements of $\mathbb{R}^2$ whose coordinates are rational.
Given lengths a and b, the construction of $a \pm b$ is quite obvious. The lengths ab and a/b may be obtained by constructing parallel lines.
For ab, the picture is a triangle with one side length of ab and another side length of a. The angle formed by these two sides is also the angle of another (perhaps smaller) similar triangle with side length 1 on the same line as the length a and side length b on the same line as the length ab (the construction of parallel lines is necessary because these triangles are similar to one another and also share an angle).
For a/b, the picture is a triangle with one side length of a and another side length of b. The angle formed by these two sides is also the angle of another (perhaps smaller) similar triangle with side length 1 on the same line as the length b and

side length a/b on the same line as the length a.

In addition, again by basic geometry, we may construct $\sqrt{a}$ given any length a. The diagram for this may be described as follows: Construct a circle of diameter $1+a$ and find the point on the diameter that separates into two segments of lengths 1 and a. Now, construct the perpendicular from this point to the edge of the circle. The length of this segment is $\sqrt{a}$.
We now have a field of *constructible* elements (a subset of $\mathbb{R}$) that is strictly larger than $\mathbb{Q}$. Let us call this field of constructible elements F.

Before proceeding to the problems, we must finally consider what sorts of extensions of our present field of constructible elements may be obtained by making our straightedge and compass operations, which are of the following four types:

(1) Connecting two given points by a straight line.
(2) Finding a point of intersection of two straight lines.
(3) Drawing a circle with given radius and center.
(4) Finding the point(s) of intersection of a straight line and a circle or the intersection of two circles.

If we consider operations of type (1), then we easily see that they will never extend our field because, by connecting two points, (a, b) and (c, d) (with a, b, c, d elements of our field), we must obviously end up with a length for the connecting segment that is already constructible.

If we consider operations of type (2), then we again see that they will also never extend our field (i.e. allow for the construction of any additional elements) because we are simply solving two linear equations simultaneously to find the point of intersection of the two lines. However, the coordinates of this point of intersection must already by elements of our field, meaning that the point is already constructible.

As we have shown, straightedge operations will not extend our field. On the other hand, compass will prove to be a bit more interesting.

Consider operations of type (3). In constructing a circle with center at (h, k) and radius r, we obtain the equation:

$$(x-h)^2 + (y-k)^2 = r^2$$

If we consider what effect this has on our field of constructible elements, we are essentially examining the case of the point(s) of intersection between a straight line and a circle. However, the the simultaneous solutions of a linear equation, $ax + by - c = 0$, and the circle, $(x-h)^2 + (y-k)^2 = r^2$ (with $a, b, c, h, k, r \in F$) are simply the simultaneous solutions of two quadratic equations. Thus, operations of type (3) result in at most an extension of degree 2.

Finally, consider operations of type (4). Again, we see that this results in at most a quadratic extension. Take two circles:

$$
\begin{aligned}
&(x-h)^2+(y-k)^2=r^2\\
\text{and}\quad &(x-h^{'})^2+(y-k^{'})^2={r^{'}}^2
\end{aligned}
$$

Subtracting the second equation from the first and then proceeding to manipulate the terms, we see that we obtain the same points of intersection with the following equations:

$$
\begin{aligned}
&(x-h)^2+(y-k)^2=r^2\\
\text{and}\quad &2(h^{'}-h)x+2(k^{'}-k)y=r^2-h^2-k^2-{r^{'}}^2+{h^{'}}^2+{k^{'}}^2
\end{aligned}
$$

This, however, is simply the intersection of a line and a circle, which, as previously shown, is *at most* a quadratic extension.

These results then lead us to the following proposition, which will allow us to show that none of the three classical straightedge and compass constructions is possible.

Proposition 5.1. *If the element $\alpha \in \mathbb{R}$ is obtained from a field $F \subset \mathbb{R}$ by a series of compass and straightedge operations, then $[F(\alpha) : F] = 2^k$ for some integer $k \geq 0$.*

Theorem 5.2 (Classical Straightedge and Compass Problems)**.** *None of the classical straightedge and compass problems (I. Doubling the Cube, II. Trisecting an Angle, III. Squaring the Circle) is possible.*

Proof. I. Doubling the cube in essence is just the construction of $\sqrt[3]{2} \in \mathbb{R}$ starting from the unit 1. However, the extension $\mathbb{Q}(\sqrt[3]{2})$ is an extension of degree 3 over $\mathbb{Q}$. Thus, by Proposition 5.1, this is impossible.

That certainly seemed like an excessive amount of exposition for such a short proof! Fortunately, the other two problems are a bit more interesting.

II. In order to prove that trisecting any arbitrary angle is impossible, we need only provide one counterexample. Now, if an angle θ can be constructed, then we can construct $\cos\theta$ by determining the point at distance 1 from the origin at angle θ from the positive x-axis. And, if $\cos\theta$ can be constructed, then, quite easily, $\sin\theta$ can be constructed. Additionally, the converse holds, i.e. if $\cos\theta$, $\sin\theta$ can be constructed, then the angle θ can be constructed. Now, let us consider $\theta = 60^\circ$. To show that this angle is not constructible, we need only show that, given $\cos 60^\circ$, it is impossible to construct $\cos 60/3^\circ = \cos 20^\circ$.
If $\theta = 60^\circ$, then $\cos\theta = 1/2$. Now, by the triple angle formula for cosines (this may be obtained by considering $\cos(2\theta+\theta)$:

$$\cos\theta = 4\cos^3\theta/3 - 3\cos\theta/3$$

Substituting $\theta = 60^\circ$, we see that $\beta = \cos 20^\circ$ satisfies the equation

$$4\beta^3 - 3\beta - 1/2 = 0$$

or $8(\beta)^3 - 6\beta - 1 = 0$. This can be written $(2\beta)^3 - 4(2\beta) - 1 = 0$. Let $\alpha = 2\beta$. Then we are left with:

$$\alpha^3 - 3\alpha - 1 = 0$$

However, by the rational root theorem, we can easily see that this has no roots in $\mathbb{Q}$ and, by Proposition 3.9, is therefore irreducible over the field $\mathbb{Q}$. Thus, an extension of the field $\mathbb{Q}$ by the root α will be of degree 3. Thus, by Proposition 5.1, α is not constructible, and thus, $\beta = \cos 20^\circ$ is not constructible. Therefore, an arbitrary angle cannot be trisected by straightedge and compass constructions.

III. Finally, we must show that it is impossible to square the circle with straightedge and compass operations. In order to do so, we must be able to construct some s such that $s^2 = \pi r^2$, $r \in F$. Hence, we must be able to construct $\sqrt{\pi}$, which would also mean that we could construct π. However, π is a transcendental number (a fact that we shall assume in this situation), and therefore the extension of $\mathbb{Q}$ by the element π is not finite (and certainly not a power of 2!). Thus, π is not constructible, which implies that $\sqrt{\pi}$ is not constructible. As a result, our desired element s is not constructible, proving that this third straightedge and compass problem is also impossible. □

6. Acknowledgements

I would like to extend my sincere thanks to my graduate mentor, Mike Miller for all the help he provided me throughout the REU. Just a few of his many assistances include helping me understand the above presented material, proofreading my paper, checking my proofs, and, finally, making me (somewhat) literate in the language of TeX.

7. References

[1] David S. Dummit and Richard M. Foote. Abstract Algebra. John Wiley and Sons, Inc. 2004.