Field Extensions and the Tower Law

Kitty Powell

April 2024

Abstract

In this essay I will introduce the concept of field extensions, which is a vital foundation for the study of Galois theory. I will begin by defining a field extension, classifying the various types of field extensions and deriving basic results involving these. I will then build up to proving the Tower Law and examine some of its consequences.

Contents

1 Introduction and motivating example

Field extensions and the Tower Law are key elements in the study of Galois theory, which has applications to geometry, algebraic number theory and, perhaps most famously, the insolubility of the quintic. Galois theory was developed by mathematician and French revolutionary Évariste Galois in the early 1800s. He famously died in a duel at the age of 20, but despite his young age, Galois had a tremendous impact on mathematics. [8, pp. 6–13]

Solving polynomials by radicals

Galois theory associates a group, called the *Galois group* to every polynomial, as detailed in [8, pp. 108–111]. We can study this group using field extensions and the Tower Law, which provides insight into the structure of the polynomial and how the roots relate to each other. If we do not know the roots of a polynomial, but we know its' Galois group, we can infer properties about the roots.

For polynomials of degree 4 and below, we have explicit formulas for the solutions to such polynomials in terms of their coefficients; the quadratic, cubic and 'quartic' formulas. Galois theory introduces a new way of solving polynomials, by studying the symmetry of permutations of their roots. The permutations of roots that preserve certain properties form the polynomial's *Galois group*. It is also through this method that we can show that polynomials of degree 5 and above are not solvable by radicals, in other words there are no formulas for solutions in terms of their coefficients. I will illustrate the core ideas behind Galois theory in the following example.

A simple example

This example is based on a similar example in Stewart's book [8, pp. 108–111]. I will present it following his method. The methods used here are not rigorous, as this example serves only to motivate the subject of this essay.

Let us consider the quartic polynomial

$$f(x) = x^4 - x^2 - 2 = (x^2 + 1)(x^2 - 2).$$

This is clearly a very 'nice' polynomial (the roots are obvious), but it serves as a good illustration of the basic principles of Galois theory. The roots of f are $\pm i, \pm\sqrt{2}$. Let us label these as follows: $\alpha = i, \beta = -i, \gamma = \sqrt{2}, \delta = -\sqrt{2}$. Clearly there are certain equations satisfied by these roots, such as:

$$\begin{aligned} \alpha^2 + 1 = 0, \qquad & \beta^2 + 1 = 0, \qquad & \alpha + \beta = 0, \\ \gamma^2 - 2 = 0, \qquad & \delta^2 - 2 = 0, \qquad & \gamma + \delta = 0, \\ \alpha\gamma - \beta\delta = 0, \qquad & \beta\gamma - \alpha\delta = 0 & \end{aligned}$$

as well as infinitely many others.

Let us now explore what happens to these 'valid equations' when we permute these four roots. Obviously, not all elements of $\mathbb{S}_4$ (the group of permutations on

4 elements) will preserve the validity of these equations. For example, swapping α and γ transforms the third equation into $\gamma + \beta = 0$, which is clearly not true for our values of β and γ. However, the permutation that swaps α and β, and the permutation that swaps γ and δ, both preserve the validity of these equations, as does their composition:

$$\begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \delta & \gamma \end{pmatrix}.$$

These permutations, along with the identity, form a group:

$$\left\{ \mathrm{id}, \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \gamma & \delta \end{pmatrix}, \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \alpha & \beta & \delta & \gamma \end{pmatrix}, \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \delta & \gamma \end{pmatrix} \right\}$$

In fact, these are the only permutations of $\mathbb{S}_4$ that preserve the validity of such equations, as stated in [8, p.109]. This group is called the *symmetry group* of the polynomial – in this particular context, we call it the *Galois group*. In this case, it is isomorphic to the Klein four-group, K_4. We can study the structure of the Galois group, and its subgroups, to determine how to solve polynomials such as the one above. In order to do this, one needs to have an understanding of field theory, which will be the main focus of this essay.

2 Key definitions and results

2.1 Field extensions

Consider a field L and a subfield K of L. For example, let's take $\mathbb{R}$ as a subfield of $\mathbb{C}$. We can define a very obvious map $\iota : \mathbb{R} \to \mathbb{C}$ with $\iota(a) = a$, an inclusion map. This is obviously an injective map and it can easily be shown to be a homomorphism.

Definition 2.1.1 (**Field extension**)**.** Let L be a subfield of $\mathbb{C}$. Let L have a subfield K. A *field extension* L/K is an injective homomorphism (also called a monomorphism) $\iota : K \to L$. We can then call K the *small field* and L the *large field*. [8, p. 64]

Note, it is perfectly well-defined to use field theory with finite fields, such as $\mathbb{F}_p$ where p is prime, as Swallow does in [9, pp. 186–187]. However, for the purposes of this essay, all major examples will work with subfields of $\mathbb{C}$.

Example. This example is taken from [8, pp. 65–66]. Let $K = \{a + b\sqrt{5} : a, b \in \mathbb{Q}\}$. We can show that K is a subfield of $\mathbb{C}$ as follows. Obviously, K is a subset of $\mathbb{C}$ and contains 0 and 1. The usual operations of $+$ and $\times$ follow the associative and distributive laws, inherited from $\mathbb{C}$. Clearly, K is closed under addition and any element $a + b\sqrt{5}$ has additive inverse $-a - b\sqrt{5} \in K$. Additionally,

$$(a + b\sqrt{5}) \times (c + d\sqrt{5}) = ac + 5bd + (ad + bc)\sqrt{5} \in K$$

and for $a + b\sqrt{5}$ non-zero,

$$(a + b\sqrt{5})^{-1} = \frac{1}{a + b\sqrt{5}} = \frac{a}{a^2 - 5b^2} - (\frac{b}{a^2 - 5b^2})\sqrt{5} \in K$$

so K is closed under multiplication and multiplicative inverses. The inclusion map $\iota : \mathbb{Q} \to K$ thus defines a field extension $K/\mathbb{Q}$.

We can see that in some fields, for example $\mathbb{F}_2$, if we add the multiplicative identity element to itself enough times, we get zero. For $\mathbb{F}_2$, this occurs when we add 1 twice. Similarly for $\mathbb{F}_p = \{0, 1, 2, ..., p - 1\}$ we get zero after adding the 1 p times. In other words, $p = 0$ when viewed as an element of $\mathbb{F}_p$.

Definition 2.1.2 (**Field characteristic**)**.** A field K has *characteristic* 0 if, for all $n \in \mathbb{Z}^+$, $n \neq 0$ when taken as an element of K. That is to say, adding 1 repeatedly within K will never yield zero.

Alternately, we say K has *characteristic* n, if n is the smallest positive integer such that $\underbrace{1 + 1 + \cdots + 1}_{\text{n}} = n = 0$ when considered as an element of K.

Examples

- $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ all have field characteristic 0. [6, p. 10]
- $\mathbb{F}_p$ has field characteristic p.

Lemma 2.1.3. If a field K has characteristic $n > 0$, then n is prime.

Proof. This proof is based on the proof used in [2, p. 26]. Assume, for a contradiction, that n is composite. Then $n = ab$, for some $a, b \in \mathbb{Z}^+$, where both a and b are strictly less than n. K has characteristic n, so $n = ab = 0$. Hence either $a = 0$ or $b = 0$, as elements in K. But n is defined as the smallest possible positive integer equal to zero, which gives a contradiction. Hence, n is prime. □

Definition 2.1.4 (**Prime subfields**)**.** The *prime subfield* of a field K is defined as the intersection of all subfields of K.

Theorem 2.1.5. Let K be a field. If K has characteristic 0, then its prime subfield is $\mathbb{Q}$. Otherwise K has characteristic p, where p is prime, and its prime subfield is $\mathbb{F}_p$.

Proof. Let P be the prime subfield of K. Since it is a field, it will contain the elements 0 and 1 by definition. Fields are closed under addition, so, by induction, for all positive integers $n \in \mathbb{Z}$,

$$n = \underbrace{1 + 1 + \cdots + 1}_{\text{n}} \in K$$

and hence the their additive inverses $-n \in K$.

Define a map $\phi : \mathbb{Z} \to P$ with

$$\phi(n) = \begin{cases} \overbrace{1+1+\cdots+1}^{\text{n}} = n & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ -\phi(-n) & \text{if } n < 0 \end{cases}$$

The map ϕ is based on the map $n \mapsto n^*$ from Stewart's book [8, p. 187] and I will examine the different cases arising from ϕ in the same way as his proof. Firstly, one can show that ϕ is a ring homomorphism, which Stewart leaves as an exercise. This is fairly trivial, as all is required is a simple check that the rules of a homomorphism are satisfied, however I outline the necessary calculations below.

- Clearly, $\phi(0) = 0$ and $\phi(1) = 1$.
- Take any $n, m \in \mathbb{Z}_{>0}$, then

  $$\phi(n+m) = n + m = \phi(n) + \phi(m)$$

  and $\phi(nm) = nm = n \times m = \phi(n) \times \phi(m)$.

  Similarly, if both $n, m \in \mathbb{Z}_{<0}$,

  $$\phi(n+m) = -\phi((-n)+(-m)) = -(\phi(-n) + \phi(-m)) = \phi(n) + \phi(m).$$

  $$\phi(nm) = \phi((-n)\cdot(-m)) = \phi(-n)\times\phi(-m) = -\phi(-n)\times -\phi(-m) = nm = \phi(n)\times\phi(m).$$

- Without loss of generality, let $n = 0$, $m \in \mathbb{Z}$. Then

  $$\phi(n+m) = \phi(m) = \phi(m) + \phi(0) = \phi(n) + \phi(m)$$

  $$\phi(nm) = \phi(0) = 0 = \phi(n) \times \phi(m)$$

- Without loss of generality, if $n \in \mathbb{Z}_{<0}$ and $m \in \mathbb{Z}_{>0}$, Then either $|m| > |n|$ so

  $$\phi(n+m) = n + m = -(-n) + m = -\phi(-n) + \phi(m) = \phi(n) + \phi(m)$$

  or $|m| < |n|$ so

  $$\phi(n+m) = -\phi(-(n+m)) = n + m = -(-n) + m = \phi(n) + \phi(m).$$

Now, we have 2 cases of ϕ to examine:

1. There exists an integer $a \neq 0$ such that $\phi(a) = 0$. Since $\phi(-a) = -\phi(a)$, we can assume that there is a smallest positive integer, p such that $\phi(p) = 0$. We can see that p is exactly what we defined to be the field characteristic of P, hence from Lemma 2.1.3, we know that p is prime. Therefore, the set $\{\phi(n) : n \in \mathbb{Z}\}$ forms a ring isomorphic to $\mathbb{F}_p$, which we know to also be a field. We know by the minimality of P that this must be the whole of P.

2. $\phi(a) \neq 0$ for all $a \neq 0$. Then since P is a field, it contains all elements of the form $\frac{\phi(a)}{\phi(b)}$ for $a, b \in \mathbb{Z}$, $b \neq 0$. This is equivalent equivalent to $\mathbb{Q}$, which must be the whole of P since $\mathbb{Q}$ is a field and P is the smallest subfield of K.

□

Primes subfields, as the name suggests, are essentially the 'field' equivalent of the prime numbers. Given any field, the smallest field you can reduce it to is its prime subfield. We will see an example of this later, in Proposition 2.2.2, which will give rise to a core piece of notation used for studying field extensions in the complex numbers.

2.2 Generating subfields

Definition 2.2.1 (**Subfield generated by a subset**)**.** Consider again a field L and a subset X of L. Note, X does not need any additional structure: we only need it to be a subset. We define the *subfield of L generated by X* as the intersection of all possible subfields of L that contain X.

Proposition 2.2.2. For every subset X of $\mathbb{C}$, the subfield of $\mathbb{C}$ generated by X contains $\mathbb{Q}$.

Proof. Let K be any subfield of $\mathbb{C}$. Note that $\mathbb{C}$ has characteristic 0, so by Theorem 2.1.5, its prime subfield is $\mathbb{Q}$. We can see this in more detail:

Clearly, all of $\mathbb{Z}$ is contained in K, since K is a field, so must contain 1 and be closed under addition and additive inverses. Additionally, K is closed under multiplicative inverses, so for every $n \neq 0 \in \mathbb{Z}$, $\frac{1}{n} \in K$. Hence by closure under multiplication, $\{\frac{p}{q} : p, q \in \mathbb{Z}, n \neq 0\} = \mathbb{Q}$ is a subfield of K. Since every subfield of $\mathbb{C}$ generated by X is a subfield of $\mathbb{C}$ by definition, it therefore must contain $\mathbb{Q}$ as a subfield. [8, p. 64] □

Because of this, we adopt the notation $\mathbb{Q}(X)$ to denote the subfield of $\mathbb{C}$ generated by X.

Example. The subfield of $\mathbb{C}$ generated by $\{i\}$ is $K = \{a + bi : a, b \in \mathbb{Q}\}$. This example is taken from [6, p. 10].

We know from Proposition 2.2.2 that every subfield of $\mathbb{C}$ contains $\mathbb{Q}$. Additionally, the subfield generated by $\{i\}$ must contain i by definition, and since it is a subfield, it will be closed under multiplication and addition of i with the rational numbers. Hence the subfield of $\mathbb{C}$ generated by $\{i\}$ must at least contain $K = \{a + bi : a, b \in \mathbb{Q}\}$. K is the Gaussian field [6, p. 10], i.e K is itself a field, hence it is the smallest subfield of $\mathbb{C}$ containing $\{i\}$. Therefore, as in Proposition 2.2.2, we denote K as $\mathbb{Q}(i)$.

Lemma 2.2.3. Let L be a field, with a non-empty subset X, where $X \neq \{0\}$. Let L have a subfield K. Then the following statements are equivalent:

(i) K is the subfield of L generated by X;

(ii) K is the smallest subfield of L containing X;

(iii) K is the smallest possible set of elements of L that can be obtained by elements of X by a finite sequence of field operations.

Proof. The above lemma is stated in [6, p. 11] without proof. The following proof is my original work.

- (ii) $\Rightarrow$ (iii): Let J be the smallest subfield of L containing X. J is a subfield so must contain the elements 0 and 1 and by definition, it contains all elements of X. J must be closed under any finite sequence of field operations, since it is itself a field. Hence, by the minimality of J, it must be the smallest possible set of elements of L that can be obtained by elements of X by a finite sequence of field operations.

- (ii) $\Leftarrow$ (iii): If J is the smallest possible set of elements of L that can be obtained by elements of X by a finite sequence of field operations, then it must be closed under those field operations, since all possible elements are contained in the set J. Additionally, by taking any element $x \in X$, we see that $0 = x - x \in J$ and $1 = \frac{x}{x} \in J$. Hence J is a subfield of L. By the minimality of J, it must be the smallest subfield of L containing X.

- (i) $\Leftrightarrow$ (iii): Let M be the smallest possible set of elements of L that can be obtained by elements of X by a finite sequence of field operations. Let K be the subfield of L generated by X. Since K is the intersection of all possible subfields of L that contain X (2.2.1) and M is a subfield of L containing X, we get that $K \subseteq M$. All subfields of L containing X must contain M, by the definition of a field. Hence their intersection will also contain M. So, $M \subseteq K$. Hence $M = K$.

□

Example. Consider again $\mathbb{Q}(i) = \{a + bi : a, b \in \mathbb{Q}\}$. $\mathbb{C}/\mathbb{Q}$ is a field extension with small field $\mathbb{Q}$ and large field $\mathbb{C}$. We can think of $\mathbb{Q}(i)$ as the subfield of $\mathbb{C}$ generated by $\mathbb{Q} \cup \{i\}$.

Definition 2.2.4. For a field extension L/K, and any subset $X \subset L$, we define the field obtained by *adjoining* X to K as the subfield of L generated by $K \cup X$. This is denoted $K(X)$.

2.3 Types of field extension

Theorem 2.3.1. If L/K is a field extension, then L is a vector space over K.

It can easily be shown that all properties of a vector space – additive identity, inverses, closure under addition and multiplication by scalars in K – are satisfied by L. I will omit the full calculations for brevity, as Garling does in [3, p. 40]

Example. $\mathbb{R}/\mathbb{Q}$ is a field extension. We know that $\mathbb{R}$ is an infinite-dimensional vector space over $\mathbb{Q}$.

- $\mathbb{R}$ is a commutative group under addition. We have $0 \in \mathbb{R}$, closure, inverses and commutativity and associativity of addition.
- The multiplicative identity $1 \in \mathbb{R}$.
- For any scalars $\lambda, \mu \in \mathbb{Q}$, $x, y \in \mathbb{R}$, we have that

$$\begin{aligned} &\lambda \cdot x \in \mathbb{R} \\ &(\lambda\mu) \cdot x = \lambda(\mu \cdot x) \\ &(\lambda + \mu) \cdot x = \lambda x + \mu x \\ &\lambda(x + y) = \lambda x + \lambda y \end{aligned}$$

- In her lecture notes for the University of Michigan [7, p. 3], Smith states that: "Because $\mathbb{R}$ itself is uncountable, no countable set can be a basis for $\mathbb{R}$ over $\mathbb{Q}$". So, as a vector space over $\mathbb{Q}$, $\mathbb{R}$ is infinite dimensional.

Definition 2.3.2 (**Degree of an extension**)**.** Considering a field extension L/K with this new vector space correspondence, we define the *degree* of the field extension L/K as the dimension of the vector space L, over the field K. We denote this as $[L : K]$.

Examples

- $\mathbb{C}/\mathbb{R}$ is a field extension. We can write any element of $\mathbb{C}$ in the form $a+bi$ for $a, b \in \mathbb{R}$ and if $a + bi = 0$ then both a and b must be zero. So $\{1, i\}$ is a basis of $\mathbb{C}$ over $\mathbb{R}$. Thus the degree of $\mathbb{C}/\mathbb{R}$ is 2.
- Similarly, $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ has degree 2. $\mathbb{Q}(\sqrt{2})$ is the subfield of $\mathbb{C}$ generated by $\mathbb{Q} \cup \{\sqrt{2}\}$. From Lemma 2.2.3, we can write this as $\{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$, with $\{1, \sqrt{2}\}$ as a basis.
- We have seen that $\mathbb{R}$ is an infinite dimensional vector space over $\mathbb{Q}$, so the degree of $\mathbb{R}/\mathbb{Q}$ is ∞.

Definition 2.3.3 (**Simple extension**)**.** Given a field extension L/K, we say the extension is *simple* (or *primitive*) if and only if $L = K(\alpha)$ for some $\alpha \in L$ [2, p. 40].

Example. Some simple field extensions do not appear to be simple until examined properly. For example, consider $\mathbb{Q}(i, \sqrt{2})$. This is a simple extension, since $\mathbb{Q}(i, \sqrt{2})$ can also be written as $\mathbb{Q}(i + \sqrt{2})$, which I will prove below.

Proof. This proof is adapted from a similar example in [8, pp. 67–68]. It is sufficient to show that $i \in \mathbb{Q}(i + \sqrt{2})$ and $\sqrt{2} \in \mathbb{Q}(i + \sqrt{2})$, since these imply both $\mathbb{Q}(i, \sqrt{2}) \subset \mathbb{Q}(i + \sqrt{2})$ and $\mathbb{Q}(i + \sqrt{2}) \subset \mathbb{Q}(i, \sqrt{2})$. Firstly,

$$(i + \sqrt{2})^2 = 1 + 2i\sqrt{2} \in \mathbb{Q}(i + \sqrt{2}), \text{ so}$$

$$(i + \sqrt{2})(1 + 2i\sqrt{2}) = -\sqrt{2} + 5i \in \mathbb{Q}(i + \sqrt{2})$$

$$\frac{(i+\sqrt{2})+(-\sqrt{2}+5i)}{6} = i \in \mathbb{Q}(i+\sqrt{2})$$

And $(i+\sqrt{2})-i = \sqrt{2} \in \mathbb{Q}(i+\sqrt{2})$. □

Definition 2.3.4 (**Algebraic extension**)**.** For a field extension L/K and an element $\alpha \in L$, we say α is *algebraic* over K if there exists a non-zero polynomial $f \in K[x]$ such that $f(\alpha) = 0$. If α is not algebraic over K, we say α is *transcendental* over K.
L/K is called an *algebraic extension* if every element of L is algebraic over K [6, p. 15].

Examples

- $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ is an algebraic extension. Take any $\alpha = a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$, where $a, b \in \mathbb{Q}$. Let $f \in \mathbb{Q}[x]$ be given by

  $$f(x) = x^2 - 2ax + a^2 - 2b^2.$$

  $f(\alpha) = 0$ for any choice of $\alpha \in \mathbb{Q}$, hence this extension is algebraic.

- $\mathbb{Q}(\pi)/\mathbb{Q}$ is not an algebraic extension, since $\pi \in \mathbb{Q}(\pi)$ is transcendental over $\mathbb{Q}$. The method of proving of π is transcendental over $\mathbb{Q}$ is not fully relevant to this essay, however a nice proof can be found in Baker's 1975 book [1, pp. 5-6]. Note that since π is transcendental over $\mathbb{Q}$, all powers of π will be linearly independent in $\mathbb{Q}(\pi)$, hence the degree of $\mathbb{Q}(\pi)/\mathbb{Q}$ is infinite.

Lemma 2.3.5. All finite field extensions are algebraic.

Proof. This proof is taken from [2, p. 45].

Suppose a field extension L/K is finite. So, the degree of L/K is $[L : K] = n$ for some $n \in \mathbb{N}$. For any $\alpha \in L$, consider $1, \alpha, \alpha^2, ..., \alpha^n$. These are all elements of L, which is an n-dimensional vector space over K, by definition. The set $\{1, \alpha, \alpha^2, ..., \alpha^n\}$ contains $n+1$ elements, hence they will be linearly dependent in the n-dimensional vector space K. So there exist constants $b_i \in K$ for $i \in \{0, 1, 2, ..., n\}$ such that

$$b_0 + b_1\alpha + b_2\alpha^2 + \cdots + b_n\alpha^n = 0$$

which is a non-trivial polynomial in α over K satisfying the requirements in Definition 2.3.4. Hence all elements of L are algebraic over K, which makes L/K an algebraic extension. □

3 The Tower Law

Now that I have introduced the types of field extensions and the *degree* of a field extension, I will introduce the main focus of this essay: the Tower Law. The Tower Law provides a way to understand the structure of more complicated field extensions by splitting them into a chain of smaller extensions.

Lemma 3.0.1. If we have two field extensions M/L and L/K, then we naturally have that M/K is also a field extension.

Proof. Since L/K is a field extension, we have a monomorphism $\iota_1 : K \to L$ and similarly for M/L we have a monomorphism $\iota_2 : L \to M$. The composition of two homomorphisms is also a homomorphism, and injectivity is preserved under composition of injective maps, hence $\iota_2 \circ \iota_1 : K \to M$ is also a monomorphism. Thus M/K is a field extension. □

Remark. If we have any two subfields K, L of $\mathbb{C}$ where $K \subseteq L$, we can define a monomorphism $\iota : K \to L$ by the inclusion map $\iota(a) = a$, as in Subsection 2.1. Hence if we have three subfields of $\mathbb{C}$, given as K, L, M, and $K \subseteq L \subseteq M$, we can automatically define field extensions L/K, M/L and M/K in this way.

Now we have a way of breaking down more complicated field extensions, we need a mechanism to relate their degrees to those of the simpler extensions. The following theorem has a widely-known standard proof, which I detail below based on [3, pp. 41–42] and [8, pp. 80–81].

Theorem 3.0.2 (A scaled-down Tower Law). For subfields K, L, M of $\mathbb{C}$ such that $K \subseteq L \subseteq M$, we have that

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

Proof. Since L/K is a field extension, by Theorem 2.3.1, we can consider L as a vector space over K. Let $(x_i)_{i\in I}$ be a basis for this vector space, where I is an arbitrary indexing set. Similarly, let $(y_j)_{j\in J}$ be a basis for M over L, for some indexing set J. The aim of this proof will be to show that $(x_i y_j)_{i\in I, j\in J}$ is a basis for the vector space M over K.

(i) The elements $(x_i y_j)_{i\in I, j\in J}$ are linearly independent over K:
Suppose that for $\alpha_{ij} \in K$,

$$\sum_{i\in I, j\in J} \alpha_{ij} x_i y_j = 0$$

This gives,

$$\sum_{j\in J} \left(\sum_{i\in I} \alpha_{ij} x_i \right) y_j = 0$$

Since $(x_i)_{i\in I}$ is a basis for L over K and each $\alpha_{ij} \in K$, we can see that $\sum_{i\in I} \alpha_{ij} x_i$ is an element of L, for every $j \in J$. Additionally, since $(y_j)_{j\in J}$ is a basis for M over L, we know the only way in which $\sum_{j\in J} a_j y_j = 0$ for $a_j \in L$ is to have $a_j = 0$ for all $j \in J$.

In this instance, we have

$$a_j = \sum_{i\in I} \alpha_{ij} x_i = 0 \ \ \forall j \in J.$$

In a similar fashion to the above argument, since each $\alpha_{ij} \in K$ and $(x_i)_{i\in I}$ is a basis for L over K, we conclude that each $\alpha_{ij} = 0$.

Hence, the set of $(x_i y_j)_{i\in I, j\in J}$ are linearly independent over K.

(ii) The elements $(x_i y_j)_{i\in I, j\in J}$ span M:
Any element, m, of M can be written as

$$m = \sum_{j\in J} \beta_j y_j,$$

where each $\beta_j \in L$. Additionally, every $\beta_j \in L$ can be written as

$$\beta_j = \sum_{i\in I} \beta_{ij} x_i,$$

where each $\beta_{ij} \in K$. Hence,

$$m = \sum_{i\in I, j\in J} \beta_{ij} x_i y_j$$

Since $(x_i)_{i\in I}$ is a basis for L over K and $(y_j)_{j\in J}$ is a basis for M over L, we know that $[L : K] = \#I$ and $[M : L] = \#J$. Similarly, I have shown that $(x_i y_j)_{i\in I, j\in J}$ is a basis for M over K, hence

$$[M : K] = \#J \cdot \#I = [M : L][L : K]$$

□

This theorem can easily be generalised by induction to account for any finite chain of field extensions.

Corollary 3.0.3. For a finite sequence of subfields K_i of $\mathbb{C}$, $i \in \{0, 1, ..., n\}$, where $K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n$, we have that

$$[K_n : K_0] = [K_n : K_{n-1}][K_{n-1} : K_{n-2}] \cdots [K_1 : K_0]$$

Here I use the notation adopted in [8, p. 82].

Definition 3.0.4. A sequence of subfields of $\mathbb{C}$, $(K_i)_{i=0}^n$, where $K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n$, is called a *tower* [3, p. 42]. $K_1, K_2, ..., K_{n-1}$ are called the *intermediate fields* between K_n and K_0.

3.1 Immediate applications

Let us now apply the Tower Law to some simple exercises and lemmas.

Lemma 3.1.1 (Intermediate fields of extensions with prime degree). Suppose that for a field extension L/K, we have that $[L : K] = p$, where p is a prime number. In [3, p. 42], Garling asks what intermediate fields there can be between L and K, leaving it as an exercise for the reader. It can easily be shown that the only intermediate fields are L and K themselves.

Proof. Suppose there are intermediate fields $K_1, K_2, ..., K_n$. By Corollary 3.0.3,

$$p = [L : K] = [K : K_1][K_1 : K_2] \cdots [K_n : L]$$

Since p is prime, its only factors are 1 and p. Hence all but one of $[K : K_1]$, $[K_n : L]$ and $[K_i : K_{i+1}]$, where $i \in \{1, ..., n-1\}$, must be equal to 1, and the other must be equal to p.
If a field extension has degree 1, that means the large field is a vector space of dimension 1, over the small field, hence the two fields will be isomorphic. Furthermore, since the small field is contained within the large field, they will be equivalent. Thus for every intermediate field extension

$$[K : K_1], [K_1 : K_2], ..., [K_{j-1} : K_j]$$

before the extension of degree p, each $K_i = K$ for $i \leq j$, the extension $[K_j : K_{j+1}]$ of degree p must be equivalent to the extension $[L : K]$. This is because K_j is equivalent to K and $K_j \subseteq L$, with $[L : K] = [K_j : K]$. All other $K_{j+1}, K_{j+2}, ..., K_n$ must also be equivalent to L, since they are all subfields of L containing $K_{j+1} = L$ as subfields. □

Example (**Explicit application of the Tower Law**)**.** This example is taken from both [4, p. 72] and [8, pp. 81–82].

Let us calculate the degree of $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}$. I will show that we can split $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}$ into a chain of two smaller field extensions: $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$. Then, we can calculate the degree of each of these extensions and apply the Tower Law.

Clearly, both of these are field extensions, defined by the obvious inclusion map. We have already seen, in Subsection 2.3, that $\mathbb{Q}(\sqrt{2})$ is a field extension, with $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$. Hence, it only remains to find a basis for $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ as a vector space over $\mathbb{Q}(\sqrt{2})$.

Note that Theorem 3.0.2 gives that $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ is a basis for $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$. This itself is sufficient to find the degree of this field extension, since the dimension of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$ will simply be the number of elements of the basis. Clearly, we expect the degree to be 4, however I will verify this more rigorously to further demonstrate the mechanics of the Tower Law.

- For $p, q, r, s \in \mathbb{Q}$, we can write any element of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ as

  $$p + q\sqrt{2} + r\sqrt{3} + s\sqrt{6} = p + q\sqrt{2} + (r + s\sqrt{2})\sqrt{3}$$

  Now, writing $p + q\sqrt{2} = k_1$ and $r + s\sqrt{2} = k_2$, such that $k_1, k_2 \in \mathbb{Q}(\sqrt{2})$, we can see that $\{1, \sqrt{3}\}$ spans $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ when considered as a vector space over $\mathbb{Q}(\sqrt{2})$.

- Now, we just need to check that $\{1, \sqrt{3}\}$ are linearly independent over $\mathbb{Q}(\sqrt{2})$. Let us consider the cases in which

  $$p + q\sqrt{2} + (r + s\sqrt{2})\sqrt{3} = 0 \ \ (*)$$

If $r + s\sqrt{2} = 0$, then $(*)$ reduces to $p + q\sqrt{2} = 0$. So in this instance, both the coefficients of 1 and $\sqrt{3}$ must be zero.

If $r + s\sqrt{2} \neq 0$, then we can divide $(*)$ by it and rearrange to obtain $\sqrt{3} = -(p + q\sqrt{2})/(r + s\sqrt{2})$. By rationalising the denominator of this expression, we determine that this implies $\sqrt{3} = a + b\sqrt{2}$, for some $a, b \in \mathbb{Q}$. After squaring and rearranging, this becomes $ab\sqrt{2} = 9 - a^2 - b^2 \in \mathbb{Q}$, which is only possible if either $a = 0$ or $b = 0$. This would imply that either $\sqrt{3} = a$ or $\sqrt{3} = b\sqrt{2}$. In his book, Stewart calls both of these cases "absurd" [8, p. 82]. Clearly, this gives us a contradiction, hence we need only consider the case when $r + s\sqrt{2} = 0$. Thus we arrive at the desired conclusion.

Now that we have found bases for $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$, all that remains is to apply the Tower Law:

$$[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \cdot 2 = 4$$

Lemma 3.1.2. The composition of algebraic extensions is also an algebraic extension.

Proof. Lemma 2.3.5 gives that all finite field extensions are algebraic, so it suffices to show that the composition of finite field extensions is also finite. Let M/L and L/K be finite field extensions, with $K \subseteq L \subseteq K$. Since M/L and L/K are finite, we can say they have degrees m and n respectively, where $n, m < \infty$. Then, by the Tower Law, we have that $[M : K] = [M : L][L : K] = nm < \infty$. This can easily be generalised by induction to any finite composition of algebraic extensions. □

4 Consequences of the Tower Law

4.1 An example from Galois theory in algebra (solving polynomials)

We are now sufficiently prepared to examine the example from the introduction in more detail.

We have the quartic polynomial $f(x)$, with roots α, β, γ and δ, and its Galois group:

$$G = \left\{ \mathrm{id}, \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \gamma & \delta \end{pmatrix}, \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \alpha & \beta & \delta & \gamma \end{pmatrix}, \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \delta & \gamma \end{pmatrix} \right\}$$

We can now use field extensions to understand how to solve $f(x) = 0$. Again, I will follow the explanation used in [8, pp. 110–112] to structure this discussion.

Abstracting from the idea of polynomials, we can think about G in terms of the field extensions

$$\mathbb{Q} \subseteq \mathbb{Q}(\gamma, \delta) \subseteq \mathbb{Q}(\alpha, \beta, \gamma, \delta)$$

Let us assume that we know nothing else about $f(x)$ other than the following properties of the structure of G:

1. For a subgroup $H \subseteq G$, with

$$H = \{\mathrm{id}, \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \gamma & \delta \end{pmatrix}\},$$

   the numbers fixed by H are exactly the numbers in $\mathbb{Q}(\gamma, \delta)$, as stated in [8, p. 111]. In other words, applying any element of H to any element of $\mathbb{Q}(\gamma, \delta)$ will yield another element of $\mathbb{Q}(\gamma, \delta)$.

2. The numbers fixed by G are exactly the numbers in $\mathbb{Q}$, which is also given in [8, p. 111].

We can now use these properties to relate the four roots to each other:

- Clearly, any element of H will preserve the validity of any equation that is symmetric in α and β. Hence, the expressions $\alpha + \beta$ and $\alpha\beta$ are fixed by H and therefore must be elements of $\mathbb{Q}(\gamma, \delta)$ by property 1.

- α and β solve the quadratic

$$g(x) = x^2 - (\alpha + \beta)x + \alpha\beta = (x - \alpha)(x - \beta).$$

  As we have seen, $g(x)$ is a quadratic with coefficients in $\mathbb{Q}(\gamma, \delta)$, hence we can use the quadratic formula to find radical expressions α and β in terms of elements of $\mathbb{Q}(\gamma, \delta)$.

- Similarly, the expressions $\gamma + \delta$ and $\gamma\delta$ are symmetric in γ and δ, and do not contain α or β, so are fixed by all of G. Hence, they belong to $\mathbb{Q}$ by property 2.

- γ and δ solve the quadratic

$$x^2 - (\gamma + \delta)x + \gamma\delta = (x - \gamma)(x - \delta),$$

  which has coefficients in $\mathbb{Q}$, as stated above. Hence we can use the quadratic formula to find radical expressions, in terms of rational numbers, for γ and δ.

We can now take the radical expressions over $\mathbb{Q}$ for γ and δ and substitute them into the radical expressions over $\mathbb{Q}(\gamma, \delta)$ for α and β. From this, we obtain radical expressions over $\mathbb{Q}$ for all four roots of $f(x)$. Hence, we have proved (non-rigourously) the existence of a formula for solution by radicals of any quartic polynomial with the Galois group G, as above. In a similar way, we can prove that polynomials of degree 5 and above have no such method of solution. As Stewart phrases it in [8, p. 111], quintic polynomials have "the wrong sort of Galois group".

4.2 Applications to Euclidean geometry

Galois theory, and field extensions in particular, have many applications outside of polynomials. In this section I will briefly detail how field theory can be applied to ruler and compass constructions. In particular, the Tower Law provides neat solutions to famous problems such as 'squaring the circle' and 'doubling the cube'.

Definition 4.2.1 (Ruler and compass constructions)**.** As stated in [9, p. 164], a *ruler and compass construction* is any result that can be obtained using a ruler and compass to draw on Euclidean space (e.g a flat sheet of paper), following any combination of only these two rules:

(i) The ruler can only be used to connect two points that already exist in the construction. Notably, one cannot use the ruler to extend a line indefinitely.

(ii) The compass can only be used to draw a circle centred at an existing point, which passes through another point in the construction. Therefore, the radius of the circle must be an existing length between the two points defining the circle – the centre and a point on its circumference – which are already contained in the construction.

We can view ruler and compass constructions through the lens of field theory in the following way.

Definition 4.2.2 (Constructible numbers)**.** Let P and Q be two points in a ruler and compass construction. We can first define the distance between these two points, in the Euclidean metric, to be the unit distance. Then, we say a number d is *constructible* if we can construct two points, A and B, that are a distance of d units apart.

Lemma 4.2.3. The constructible numbers form a field, containing $\mathbb{Q}$, which we will denote $\mathbb{K}$.
Proving this is fairly trivial. Obviously, $\mathbb{K}$ contains 1 and 0, by Definition 4.2.2. It then only remains to show that if $a, b \in \mathbb{R}$ are constructible then so are $a + b, a - b, ab$ and $\frac{a}{b}$. Swallow leaves this as an exercise in [9, p. 166].

To solve the problems mentioned above, we will use the following important lemma. It is stated here without proof, for brevity, however a full proof can be found in [9, pp. 167–168].

Lemma 4.2.4. If $c \in \mathbb{R}$ is constructible, then $[\mathbb{Q}(c) : \mathbb{Q}]$ is an integer power of 2.

We can use this to determine if certain constructions are possible based on whether the theoretical lengths you would need to construct agree with the result in Lemma 4.2.4. Notably, if we needed to construct a length $a \in \mathbb{R}$ and $[\mathbb{Q}(a) : \mathbb{Q}]$ was not a power of 2, then by contrapositive, the number a would not be constructible. Hence, the desired construction would be impossible.

Remark. These results also apply to constructions in the complex plane. As stated in [5, p. 50], a number $x + iy \in \mathbb{C}$ is constructible if $x, y \in \mathbb{R}$ are constructible. This means we can naturally apply the results explored in the earlier parts of this essay, since, as stated in Definition 2.1.1, this essay focuses primarily on subfields of $\mathbb{C}$.

Example. The following examples find their origins in Ancient Greek mathematics. I use [5, p. 54] as a basis for my explanations.

(i) **Squaring the circle**
Given a circle of radius r, the goal of this problem is to construct a square with an area equal to that of the circle. The area of the circle is πr^2, hence the side length of the desired square is $\sqrt{\pi}r$. We can choose r to be the unit length, hence this problem reduces to the question: is $\sqrt{\pi}$ a constructible number? Clearly $\mathbb{Q}(\pi) \subseteq \mathbb{Q}(\sqrt{\pi})$, since all elements of $\mathbb{Q}(\pi)$ are of the form

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

So we have the following tower of fields:

$$\mathbb{Q} \subseteq \mathbb{Q}(\pi) \subseteq \mathbb{Q}(\sqrt{\pi})$$

and hence can apply the Tower Law.

$$[\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}(\pi)] \cdot [\mathbb{Q}(\pi) : \mathbb{Q}]$$

We have already seen in Section 2.3 that $[\mathbb{Q}(\pi) : \mathbb{Q}] = \infty$. By the Tower Law, $[\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}] \geq [\mathbb{Q}(\pi) : \mathbb{Q}] = \infty$ so $\sqrt{\pi}$ is not constructible. Thus, squaring the circle is impossible.

(ii) **Doubling the cube**
Given a cube of side length l, the aim is to construct a cube of twice the volume. Note that this problem takes place in 3-dimensional Euclidean space, rather then the 2-dimensional space we have been working in previously, however all rules and definitions apply equivalently, as they had no dependence on dimension. Similarly to (i), we can take l to be the unit length. So from a cube of volume 1, we want to construct a cube of volume 2. This is only possible if $\sqrt[3]{2}$ is constructible. As stated in [5, p. 54], $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, which is not a power of 2. Hence, $\sqrt[3]{2}$ is not constructible and doubling the cube is impossible.

As we have seen, field theory has applications to an incredible variety of mathematics and can provide elegant solutions to numerous well-known problems. The Tower Law is a central part of some such solutions, as well as being vital to the study of more advanced concepts in Galois theory.

References

[1] Alan Baker. *Transcendental Number Theory*, pages 5–6. Cambridge Mathematical Library. Cambridge University Press, 1975.

[2] Gavin Brown. MA3D5 Galois Theory. University of Warwick.

[3] D.J.H. Garling. *A Course in Galois Theory*. Cambridge University Press, 1988.

[4] Tom Leinster. Galois theory. University of Edinburgh, March 2023.

[5] Matthew Macauley. Chapter 8: Field and Galois Theory, Math 4120, Modern Algebra. Clemson University, 2021.

[6] Samir Siksek. MA3D5 Galois Theory. University of Warwick.

[7] Karen E. Smith. Bases for Infinite Dimensional Vector Spaces, Math 513 Linear Algebra Supplement. University of Michigan, 2012.

[8] Ian Stewart. *Galois theory*. CRC press, 4th edition, 2015.

[9] John Swallow. *Exploratory Galois Theory*. Cambridge University Press, 2004.