Galois Theory and Polynomials

University of Texas at Dallas

Lendel Deguia

Contents

1 Introduction **3**
- 1.1 Prologue . . . 3
- 1.2 Summary . . . 5
- 1.3 Technical Matters . . . 6

I Relevant Ideas from Galois Theory **7**

2 Field Extensions **8**
- 2.1 General Properties . . . 8
- 2.2 Elements in Field Extensions . . . 10
- 2.3 Degree of a Field Extension . . . 15

3 Types of Extensions **17**
- 3.1 Algebraic Extensions . . . 17
- 3.2 Splitting Fields . . . 18
- 3.3 Normal Extensions . . . 20
- 3.4 Separable Extensions . . . 22

4 Galois Groups **27**
- 4.1 Properties . . . 27
- 4.2 Transitivity . . . 28

5 Galois Correspondence **32**
- 5.1 Galois Extensions . . . 32
- 5.2 Normality . . . 35
- 5.3 The Fundamental Theorem of Galois Theory . . . 38

6 Solvability **40**
- 6.1 Solvable Groups . . . 40
- 6.2 Solvable Extensions . . . 45
- 6.3 Galois's Theorem . . . 50

II Examining Solvability and Unsolvability of Polynomials 54

7 The Cubic and Quartic via Galois Theory 55
7.1 The Cubic 55
7.2 The Quartic 60

8 The Quintic 70
8.1 Unsolvability of the Universal Quintic 70
8.2 Solvable Quintics 71

A Concepts used from Abstract Algebra 80
A.1 Some Notable Theorems for Groups and Rings 80
A.2 Polynomials as Algebraic Objects 80
A.3 Further Properties of Polynomials 82
A.4 Symmetric Polynomials 86

References 90

1

Introduction

1.1 Prologue

Perhaps as far back as middle school we are taught to solve the equation $ax+b=0$, where a and b are typically rational numbers (though they can be elements of $\mathbb{R}$ or $\mathbb{C}$); the solution is, quite trivially, (of course assuming that $a \neq 0$),

$$x = -\frac{b}{a}$$

In high school algebra (or maybe middle school depending on the location), we are introduced to the quadratic formula

$$x = \frac{b \pm \sqrt{b^2 - 4ac}}{2a}$$

which is a pair of solutions to the equation $ax^2 + bx + c = 0$ where $a \neq 0$, b and c are, again, typically assumed to be elements of $\mathbb{Q}$, though this is valid even if they are elements of, say, $\mathbb{C}$. What we have so far are explicit general solutions to arbitrary polynomials of degree one and degree two with coefficients over some field ($\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all considered as mathematical entities called "fields"). On account of these expressions, can we find such solutions to an arbitrary polynomial of degree n with coefficients in some field F? That is, a general solution to the equation

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0$$

The answer is "yes" for $n = 1$, $n = 2$, $n = 3$, and $n = 4$, but "no*" for polynomials of degree $n \geq 5$. The asterisk on "no" is meant to indicate that the answer is "no" under the assumption we are restricted to the elementary operations of addition, subtraction, multiplication, division, and taking nth roots (i.e. $\sqrt[n]{\ }$, also called a "radical"). Furthermore, although the answer is "yes" for $n = 3$ and $n = 4$, the answers aren't very pretty (they are provided in Chapter 7).

However, why is this the case? Why is it that all of a sudden that solvability of polynomials stops as soon as $n = 5$ (polynomials of this degree are called "quintics")? First, this impossibility is inherent in our restriction to our allowed operations: $\{+, -, \times, \div, \sqrt[n]{}\}$. This means that allowing more operations will allow more solvability. In fact, for $n = 5$, it can be shown that the introduction of what is known as the "Bring Radical" permits a complete solution for any arbitrary quintic.

Even so, how do we know that a solution for the general quintic is impossible provided that we are restricted to the aforementioned operations? An attempt to acquire such a solution was carried out by the notable mathematician Lagrange. He had a strategy devised to solving polynomials that worked for polynomials of degree $n \leq 4$. The idea was to make substitutions that reduced the problem of solving for the roots of a polynomial of degree n to a problem of solving a polynomial of lower degree. For instance, he was able to apply this strategy to reducing the problem of solving a polynomial of $n = 3$ to a problem of solving for a polynomial of $n = 2$; likewise, he was able to use this strategy to reduce the problem of solving a for a polynomial of $n = 4$ to solving a polynomial of $n = 3$. Unfortunately, his strategy failed for arbitrary polynomials of degree $n = 5$; in fact, applying his strategy to quintic polynomials pushes the complexity of the problem into the opposite direction: having to solve a polynomial of degree $n = 6$. Despite this absurd observation, Lagrange believed that it was still possible to find a general solution to an arbitrary quintic.

The first official indications for the unsolvability of the quintic can be attributed to the works of Paolo Ruffini and Niels Abel. In summary, Ruffini laid down some ground work for Abel with an incomplete proof that the quintic is unsolvable by radicals. Abel built on top of the ideas in Ruffini's work (although he had a hard time understanding Ruffini's proof and suspected it may have not been complete) and provided a complete proof in 1824 referred today as the "Abel–Ruffini theorem". The idea of the proof is that if we assume any arbitrary quintic f is solvable by radicals and elementary arithmetic operations, then a solution will be of the form

$$x = p_0 + p_1 R^{\frac{1}{m}} + p_2 R^{\frac{2}{m}} + \ldots + p_{m-1} R^{\frac{m-1}{m}}$$

where m is a prime number and p_i for $i \in \{0, 1, \ldots, m-1\}$ and R are algebraic expressions that involve nested radicals, each with a nesting level lower than the degree of the polynomial in question; for instance, the quadratic formula has a nesting level of one; a nesting level of two would be something like

$$\sqrt[3]{\frac{b}{2} - \sqrt{\frac{b^2}{4} + \frac{a^2}{27}}}$$

The climax of the proof in Abel's work states that if we assume such a solution for the general quintic, then it necessarily follows that

$$R^{\frac{1}{5}} = \frac{1}{5}(x_1 + \zeta x_2 + \zeta^2 x_3 + \zeta^3 x_4 + \zeta^4 x_5) = \sqrt[5]{a + b\sqrt{\Delta_f}}$$

where ζ is a primitive fifth root of unity (I assume the reader knows what this is), x_i for $i \in \{1, 2, 3, 4, 5\}$ are assumed to be distinct roots of f, a and b are elements of the same field F as the coefficients of f and Δ_f is the discriminant of f (discussed in Appendix A). The issue that arises here is that since each x_i are assumed to be distinct, if we permute the indices of each of the roots, we find that there is a total of 120 unique possible values for $R^{\frac{1}{5}}$. However, the right-hand side indicates that there are 10 unique possible values (two because we can have $\pm\sqrt{\Delta_f}$ times the five possible values from taking a fifth root). Of course, this is a contradiction, so the Abel-Ruffini theorem establishes the impossibility of solving an arbitrary quintic [1] [2].

1.2 Summary

The work in this thesis is devoted to studying the solvability and unsolvability of polynomials through the lens of Galois theory originating from a brilliant and enigmatic character by the name of Évariste Galois. The Abel-Ruffini theorem establishes the "how" regarding impossibility of the solving the quintic, but does not address the "why"; Galois theory provides us with the ideas necessary to address the "why". Perhaps the most prominent underlying themes in Galois theory are permutations and symmetries/invariance (which are encapsulated in the mathematical branch known today as "Group Theory"). Essentially, the idea behind Galois theory is to study the hidden structure of polynomials and the relationships between their roots. We will find that the hidden structure entails the notion of field extensions and the relationships refer to groups of transformations on the roots (called Galois groups).

> *"Jump with both feet on calculations. put operations into groups, class them according to their difficulty and not according to their form; that is, according to me, the mission of future geometers, that is the path that I have entered in this work."*
>
> – Évariste Galois, Preface to *Two memoirs in pure analysis* [3]

The first part of the thesis is intended to survey relevant ideas of Galois theory to use in and serve as a comprehensive reference for Part II. If the reader is willing to accept the claims made in Part II, then this part may be skipped and only referenced on an as-needed basis. The first section starts off with an introduction to field extensions and builds up to Galois's insightful theorem: a polynomial is solvable by radicals if and only if its corresponding group (which we explain in Part I) is "solvable".

In Part II, we will analyze the Galois theory behind the cubic ($n = 3$), quartic ($n = 4$), and quintic ($n = 5$). Our objective there is to explicitly relate these hidden structures and relationships from Galois theory to the solvability of the cubic and quartic and the unsolvability of the quintic. Furthermore, regarding the cubic and quartic, we will reproduce exact solutions entirely based on Galois theory; that is, we will demonstrate that a solution approach independent of heuristic solutions such as those due to Cardan and Ferrari. We will also explore solvable quintics following an approach due to Dummit [4].

1.3 Technical Matters

Most of the content in the first part of the thesis is adapted from chapters 1 through 8 of the book "Galois Theory" (Cox, 2012) [5] with each idea reproduced in my paraphrasing; additionally, for the ideas presented here, parts of proof left as exercises and additional details were filled in with help from sources like "Github Repository: Cox Galois Theory Exercises" (Ganaye) [6], "A First Course in Abstract Algebra: 7th. Ed." (Fraleigh and Katz 2003) [7] and "Groupprops, the Group Properties Wiki" (Naik) [8].

Familiarity with the primitive roots of unity is assumed as mentioned earlier. Moreover, I will assume all fields discussed have characteristic zero unless said otherwise; even if this is the case, I will sometimes mention explicitly that fields have characteristic zero. Concepts from abstract algebra are used throughout this thesis. Rudimentary abstract algebra is assumed such as the Group, Ring, and Field axioms. Bigger notions such as the fundamental theorem of symmetric polynomials and fundamental theorem of group homomorphisms are also used; I will provide some notable ones in Appendix A.

In regards to group theory, I will use the conventions $H \triangleleft G$ to denote a normal subgroup and $\{G : H\}$ to denote the group index. Also, for a mapping $\phi\colon A \to B$, I will denote the kernel of ϕ as $\ker(\phi)$ and the image of A under ϕ as $\mathrm{im}(\phi)$. Unless said otherwise, we may assume that the capital letters F and L are fields throughout the rest of this paper; if we are working with two fields, F and L, assume specifically that the field corresponding to the letter L contains the field corresponding to the letter F. Moreover, the capital letters K and M will typically be used to denote intermediate fields (an exception is when we use K to denote the field $F(\sigma_1, \ldots, \sigma_n)$ as addressed in later sections).

Part I

Relevant Ideas from Galois Theory

2

Field Extensions

This chapter contains material addressed in chapters 3 and 4 in [5]. Section 2.1 introduces the notion of a field extension and addresses the material in Chapter 3.1 from [5]. Section 2.2 introduces the idea of a minimal polynomial (very important) and contains material from Chapter 4.1 in [5]. Section 2.3 addresses how we can consider field extensions as vector spaces, introduces the notion of a field extension degree and addresses the tower theorem (also very important); it contains material from chapter 4.3 in [5]. Parts of proof left as exercises and additional details were filled in with help from sources like "Github Repository: Cox Galois Theory Exercises" (Ganaye) [6], "A First Course in Abstract Algebra: 7th. Ed." (Fraleigh and Katz 2003) [7] and "Groupprops, the Group Properties Wiki" (Naik) [8].

2.1 General Properties

Proposition 2.1.1. *Let F and L be fields and $\phi\colon F \to L$ be a ring homomorphism. Then, ϕ is injective and $F \simeq \phi(F)$.*

Proof. Take $a \in F$ such that $a \neq 0_F$. Then, $a \cdot a^{-1} = 1_F \Rightarrow \phi(1_F) = \phi(a \cdot a^{-1}) = \phi(a) \cdot \phi(a^{-1}) = 1_L$. Since L is a field, it is an integral domain, so it follows that $\phi(a) \neq 0_L$. Thus, $\forall a \in F\colon a \neq 0_F$, we have that $\phi(a) \neq 0_L$ and hence that $\ker(\phi) = \{0_F\}$. Since a ring homomorphism is injective if and only if its kernel is zero, we conclude that ϕ is injective. Moreover, the corestriction $\phi\colon F \to \phi(F)$ is automatically surjective and hence bijective, so it constitutes an isomorphism; i.e. $F \simeq \phi(F)$. ■

By Proposition 2.1.1, we see that L either properly contains $\phi(F)$ or $L = \phi(F)$. This leads us to define the following:

Definition 2.1.2 (Extension Field). *If $\phi\colon F \to L$ is a ring homomorphism where F and L are fields such that $\phi(F) \subset L$, then L is a **field extension** of F via ϕ. Since $F \simeq \phi(F)$, we identify F with $\phi(F)$ and denote the extension as $F \subset L$.*

Theorem 2.1.3 (Kronecker's Theorem). *Let $f \in F[x]$ be irreducible; then, there is a field extension $F \subset L$ where $\exists \alpha \in L : f(\alpha) = 0$.*

Proof. Consider the ideal $I = \langle f \rangle$. Since, f is irreducible, I is maximal in $F[x]$ which implies that $L \coloneqq F[x]/I$ is a field where the additive identity is $\langle f \rangle$ and multiplicative identity is

$1 + \langle f \rangle$. Let $\phi\colon F \to L$ be defined such that for $a \in F$, $\phi(a) = a + I$. Here, we have a ring homomorphism since for $a, b \in F$:

$$\phi(a+b) = (a+b) + \langle f \rangle = (a + \langle f \rangle) + (b + \langle f \rangle) = \phi(a) + \phi(b)$$
$$\phi(ab) = (ab) + \langle f \rangle = (a + \langle f \rangle)(b + \langle f \rangle) = \phi(a)\phi(b)$$

Thus, we have a field extension $F \subset L$. Now, suppose $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ where $a_i \in F,\ i = 0, 1, \ldots, n$ and consider $\alpha = x + I$. Evaluation of f at α yields:

$$\begin{aligned} f(\alpha) &= (a_0 + I)\alpha^n + (a_1 + I)\alpha^{n-1} + \ldots + (a_{n-1} + I)\alpha + (a_n + I) \\ &= (a_0 + I)(x + I)^n + (a_1 + I)(x + I)^{n-1} + \ldots + (a_{n-1} + I)(x + I) + (a_n + I) \\ &= (a_0x^n + I) + (a_1x^{n-1} + I) + \ldots + (a_{n-1}x + I) + (a_n + I) \\ &= (a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n) + I \\ &= f(x) + \langle f \rangle = \langle f \rangle \end{aligned}$$

Here, we see that evaluation of f at α yields the additive identity of L; i.e. $f(\alpha) = 0$ ■

Example 2.1.4. *Define $\mathbb{C} = \mathbb{R}[x]/\langle x^2 + 1 \rangle$ with additive identity $\langle x^2 + 1 \rangle$ and multiplicative identity $1 + \langle x^2 + 1 \rangle$. By Theorem A.3.2, for any $f \in \mathbb{R}[x]$, there are $q, r \in \mathbb{R}[x]$ such that*

$$f(x) = (x^2 + 1)q(x) + r(x)$$

where $\deg(r) < 2$, so $\max(\deg(r)) = 1$. Therefore, any element of $\mathbb{C}$ is of the form

$$(ax + b) + \langle x^2 + 1 \rangle$$

where $a, b \in \mathbb{R}$. Also, define $\phi\colon \mathbb{R} \to \mathbb{C}$ by $\phi(a) = a + \langle x^2 + 1 \rangle,\ a \in \mathbb{R}$. Then, ϕ is a ring homomorphism since for $a, b \in \mathbb{R}$,

$$\phi(a+b) = (a+b) + \langle x^2 + 1 \rangle = (a + \langle x^2 + 1 \rangle) + (b + \langle x^2 + 1 \rangle) = \phi(a) + \phi(b)$$

$$\phi(ab) = (ab) + \langle x^2 + 1 \rangle = (a + \langle x^2 + 1 \rangle)(b + \langle x^2 + 1 \rangle) = \phi(a)\phi(b)$$

Here, both $\mathbb{R}$ and $\mathbb{C}$ are fields where ϕ is injective by Proposition 2.1.1 and hence, $\phi(\mathbb{R}) \subset \mathbb{C}$. By Definition 2.1.2, we say that $\mathbb{C}$ is an extension field of $\mathbb{R}$ via ϕ, so $\mathbb{R} \subset \mathbb{C}$.

Theorem 2.1.5. *Take $f \in F[x]$ and let $n = \deg(f)$ where $n > 0$. Then there is a field extension $F \subset L$ such that*

$$f = c(x - \alpha_1) \cdots (x - \alpha_n)$$

where $c \in F$ and $\alpha_1, \ldots, \alpha_n \in L$.

Proof. Let $P(n)$ be the statement of the theorem. Suppose $n = 1$. Then, $f = cx + a_1$ where $c, a_1 \in F$. Letting $L = F$, if $f(\alpha_1) = 0$ for some $\alpha_1 \in L$, it follows that $c\alpha_1 + a_1 = 0$ and hence, $\alpha_1 = -a_1/c$. Thus, $f = cx + a_1 = c(x + a_1/c) \Rightarrow f = c(x - \alpha_1)$, so $P(1)$ is true.

Now, suppose $n > 1$ where the statement is true for polynomials of degree $n - 1$; i.e. assume $P(n-1)$. Since $F[x]$ is a UFD, there is an irreducible factor f_1 of f. By Theorem 2.1.3, there is a field extension $F \subset F_1$ where $\exists \alpha_1 \in F_1 : f_1(\alpha_1) = 0$ in F_1. Since f_1 is a factor of f, it follows that $f(\alpha_1) = 0$ in F_1. By Corollary A.3.3, this necessarily means that

$$f = (x - \alpha_1)g$$

for some $g \in F_1[x]$ where $\deg(g) = n - 1$. Since we have assumed $P(n - 1)$, we have that there is a field extension $F_1 \subset L$ such that

$$g = c(x - \alpha_2) \cdots (x - \alpha_n)$$

where $c \in F$ and $\alpha_2, \ldots, \alpha_n \in L$. Thus, we have that $P(n - 1) \Rightarrow P(n)$ and hence, conclude $P(n)$ is true via the principle of mathematical induction. ■

Remark. *If there is a field extension $F \subset L$ for $f \in F[x]$ as described by Theorem 2.1.5, we say that f* ***splits completely over*** *L.*

2.2 Elements in Field Extensions

Definition 2.2.1. *Let $F \subset L$ be a field extension and $\alpha \in L$. Then α is said to be* ***algebraic*** *over F if there is a nonconstant polynomial $f \in F[x]$ such that $f(\alpha) = 0$; otherwise, α is said to be* ***transcendental*** *over F.*

Proposition 2.2.2. *Let $\alpha \in L$ be algebraic over F. Then there is a unique nonconstant $p \in F[x]$ where p is monic such that* (i) *$p(\alpha) = 0$ and* (ii) *if $f(\alpha) = 0$ for $f \in F[x]$, then $f \in \langle p \rangle$.*

Proof. Since $\alpha \in L$ is algebraic over F, we have that there is a nonconstant polynomial $p \in F[x]$ such that $p(\alpha) = 0$ which we may assume is monic on account of Eq. (A.2) and is of smallest degree. Suppose we have some other $f \in F[x]$ such that $f(\alpha) = 0$. Then, by Theorem A.3.2, there are $q, p \in F[x]$, such that

$$f(x) = q(x)p(x) + r(x)$$

where $\deg(r) < \deg(p)$ or $r(x) = 0$. Evaluating f at α yields:

$$f(\alpha) = r(\alpha)$$

since $p(\alpha) = 0$. If we assume $\deg(r) < \deg(p)$, this would contradict our assumption that p is a polynomial of smallest degree with α as a root, so it must be that $r(x) = 0$. In this case, $f = qp$, so $f \in \langle p \rangle$.

Now, suppose that p and $\tilde{p}$ satisify both (i) and (ii). Take $q\tilde{p} \in \langle \tilde{p} \rangle$ where $q \in F[x]$; since $\tilde{p}(\alpha) = 0$, we have that $q(\alpha)\tilde{p}(\alpha) = 0$, so $q\tilde{p} \in \langle p \rangle$ and hence, $\langle \tilde{p} \rangle \subset \langle p \rangle$. Similarly, take $qp \in \langle p \rangle$ where $q \in F[x]$; since $p(\alpha) = 0$, we have that $q(\alpha)p(\alpha) = 0$, so $qp \in \langle \tilde{p} \rangle$ and hence $\langle p \rangle \subset \langle \tilde{p} \rangle$. Thus, $\langle p \rangle = \langle \tilde{p} \rangle$, so $p = \tilde{p}$ and we see that p is unique. ■

Definition 2.2.3. *The polynomial $p \in F[x]$ described by Proposition 2.2.2, is called the* ***minimal polynomial*** *of algebraic element $\alpha \in L$ over F.*

Proposition 2.2.4. *The minimal polynomial $p \in F[x]$ of algebraic element $\alpha \in L$ is irreducible over F.*

Proof. Suppose p is not irreducible; then $\exists q, d \in F[x] : p = qd$. By the properties of the minimal polynomial according to Proposition 2.2.2, we have that p is of smallest degree such that $p(\alpha) = 0$. However, since $p(\alpha) = q(\alpha)d(\alpha)$, this means that $q(\alpha) = 0$ or $d(\alpha) = 0$; either way, we would have that $h \in \{q, d\}$ has the property that $\deg(h) < \deg(p)$ and $h(\alpha) = 0$. This yields a contradiction to our assumption about p, so p must be irreducible over F. ■

Definition 2.2.5. *Given the extension $F \subset L$, take $\alpha_1, \ldots, \alpha_n \in L$. Then, using evaluation define the following:*

$$F[\alpha_1, \ldots, \alpha_n] = \{f(\alpha_1, \ldots, \alpha_n) \mid f \in F[x_1, \ldots, x_n]\}$$

Here, $F[\alpha_1, \ldots, \alpha_n]$ is the set of all polynomials in variables $x_1, \ldots, x_n$ evaluated at $\alpha_1, \ldots, \alpha_n$. Furthermore,

$$F(\alpha_1, \ldots, \alpha_n) = \{f/g \mid f, g \in F[\alpha_1, \ldots, \alpha_n],\ g \neq 0\}$$

is the corresponding set of rational functions.

Using the field and ring axioms, one can check that $F[\alpha_1, \ldots, \alpha_n]$ is a ring and $F(\alpha_1, \ldots, \alpha_n)$ is a field.

Proposition 2.2.6. *Take $\alpha_1, \ldots, \alpha_n \in L$. Then, $F(\alpha_1, \ldots, \alpha_n)$ is the smallest subfield of L that contains F and $\alpha_1, \ldots, \alpha_n$.*

Proof. Take $f, g \in F[\alpha_1, \ldots, \alpha_n]$ so that $f/g \in F(\alpha_1, \ldots, \alpha_n)$. Note that f and g are polynomial expressions in $\alpha_1, \ldots, \alpha_n$ with coefficients in F; since $\alpha_1, \ldots, \alpha_n \in L$, it follows that $f, g \in L$ since both f and g are formed using addition and multiplication on elements of F with $\alpha_1, \ldots, \alpha_n$. Moreover, L is closed under division so $f/g \in L$ and it follows that $F(\alpha_1, \ldots, \alpha_n)$ is a subfield of L.

Now, suppose K is a subfield of L containing F and $\alpha_1, \ldots, \alpha_n$. Then, any $f \in F[\alpha_1, \ldots, \alpha_n]$ should also be in K since it is closed under multiplication and addition. Moreover, taking $f/g \in F(\alpha_1, \ldots, \alpha_n)$ where $f, g \in F[\alpha_1, \ldots, \alpha_n]$, we have that $f/g \in K$ since $f, g \in K$ and K is closed under division. Thus, $F(\alpha_1, \ldots, \alpha_n) \subset K$. ■

The subfield obtained in Proposition 2.2.6 is said to be the field obtained by adjoining $\alpha_1, \ldots, \alpha_n \in L$ to F. In this sense, we can adjoin elements of L to F in stages.

Corollary 2.2.7. *Take $\alpha_1, \ldots, \alpha_n \in L$ and let $K = F(\alpha_1, \ldots, \alpha_r)$ where $r \in \{1, \ldots, n-1\}$. Then,*

$$F(\alpha_1, \ldots, \alpha_n) = K(\alpha_{r+1}, \ldots, \alpha_n)$$

Proof. By Proposition 2.2.6, K is the smallest subfield of L containing F and $\alpha_1, \ldots, \alpha_r \in L$; likewise, $K(\alpha_{r+1}, \ldots, \alpha_n)$ is the smallest subfield of L containing K and $\alpha_{r+1}, \ldots, \alpha_n$. Also by Proposition 2.2.6, we have that $K(\alpha_{r+1}, \ldots, \alpha_n)$ is a subfield of L containing F and $\alpha_1, \ldots, \alpha_n$, so

$$F(\alpha_1, \ldots, \alpha_n) \subset K(\alpha_{r+1}, \ldots, \alpha_n)$$

Conversely, since F and $\alpha_1, \ldots, \alpha_r$ are in $F(\alpha_1, \ldots, \alpha_n)$, it follows that $K \subset F(\alpha_1, \ldots, \alpha_n)$. By Proposition 2.2.6, it follows that $K(\alpha_{r+1}, \ldots, \alpha_n)$ is the smallest subfield of $F(\alpha_1, \ldots, \alpha_n)$ containing K and $\alpha_{r+1}, \ldots, \alpha_n$ hence,

$$K(\alpha_{r+1}, \ldots, \alpha_n) \subset F(\alpha_1, \ldots, \alpha_n)$$

Thus, it follows that $F(\alpha_1, \ldots, \alpha_n) = K(\alpha_{r+1}, \ldots, \alpha_n)$. ■

Proposition 2.2.8. *Let $\alpha \in L$ be algebraic over F with minimum polynomial $p \in F[x]$. Then there is a unique ring isomorphism $\widetilde{\phi}$ such that*

$$F[\alpha] \simeq F[x]/\langle p \rangle$$

where $\widetilde{\phi}(F + \langle p \rangle) = F$ and $\widetilde{\phi}^{-1}(\alpha) = x + \langle p \rangle$.

Proof. Let $\phi\colon F[x] \to L$ be the evaluation homomorphism. Since $p \in F[x]$ is the minimum polynomial of ϕ, it follows that $\phi(p(x)) = p(\alpha) = 0$. Take $gp \in \langle p \rangle$; then, $\phi(gp) = \phi(g)\phi(p) = \phi(g(x))\phi(p(x)) = g(\alpha)p(\alpha) = g(\alpha) \cdot 0 = 0$, so it follows that $gp \in \ker(\phi)$ and hence, $\langle p \rangle \subset \ker(\phi)$. Conversely, take $f \in \ker(\phi)$; then, $\phi(f(x)) = f(\alpha) = 0$. However, by Proposition 2.2.2, this means that $f \in \langle p \rangle$, so $\ker(\phi) \subset \langle p \rangle$ and hence, $\ker(\phi) = \langle p \rangle$.

Let $R = F[x]$ and $S = L$; also, let $\gamma\colon R \to R/\ker(\phi)$ be given by $\gamma(f) = f + \langle p \rangle$. Since we have that $\phi\colon R \to S$ is a ring homomorphism, it follows by the fundamental theorem of ring homomorphisms that there is a unique ring isomorphism $\widetilde{\phi}\colon R/\ker(\phi) \to \mathrm{im}(\phi)$ where $\mathrm{im}(\phi) = \{\phi(f) \mid f \in R\} \subset S$. Note that by definition, $\mathrm{im}(\phi) = \phi(R) = \phi(F[x]) = F[\alpha] \subset L$. Thus, we have that $F[x]/\langle p \rangle \simeq F[\alpha]$, and since " $\simeq$ " is an equivalence relation,

$$F[\alpha] \simeq F[x]/\langle p \rangle$$

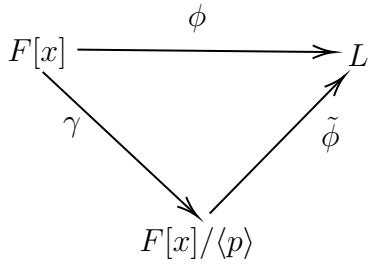

Figure 2.1: Ring Homomorphism Diagram

Now, note that $\phi(f) = (\widetilde{\phi} \circ \gamma)(f)$ as demonstrated by Fig. 2.1. First notice that for $x \in F[x]$, $\gamma(x) = x + \langle p \rangle$ and that $\widetilde{\phi}(x + \langle p \rangle) = \alpha$ since $\phi(x) = \alpha$. Since, $\widetilde{\phi}$ is an isomorphism, we have that $\widetilde{\phi}^{-1}(\alpha) = x + \langle p \rangle$. Moreover, let the fact that $\forall c \in F$, $\phi(c) = c$ be represented by $\phi(F) = F$. In this regard, $\gamma(F) = F + \langle p \rangle$, so it follows that $\widetilde{\phi}(F + \langle p \rangle) = F$ ■

Recall in the proof of Theorem 2.1.3 that if $f \in F[x]$ is irreducible, then $F[x]/\langle f \rangle$ is a field; also, the minimum polynomial $p \in F[x]$ over F of algebraic element $\alpha \in L$, is irreducible by Proposition 2.2.4. It follows that $F[x]/\langle p \rangle$ is a field and hence, $F[\alpha]$, by Proposition 2.2.8, is a field.

Proposition 2.2.9. *Let $\alpha \in L$. Then, α is algebraic over F if and only if $F[\alpha] = F(\alpha)$*

Proof. Suppose α is algebraic over F. First, note that it is always the case that $F[\alpha] \subset F(\alpha)$ according to Definition 2.2.5. Next, as mentioned, it is implicit by Proposition 2.2.8 that $F[\alpha]$ is a field; in fact, it is a field that contains F and α. By Proposition 2.2.6, $F(\alpha)$ is the smallest subfield of L that contains F and α, so $F(\alpha) \subset F[\alpha]$. Thus, it follows that $F[\alpha] = F(\alpha)$.

Now, suppose that $F[\alpha] = F(\alpha)$. If $\alpha = 0$, then for $f \in F[x]$ defined by $f(x) = x$, it follows that $f(\alpha) = 0$, so α is algebraic over F. Suppose $\alpha \neq 0$. Then, since $1/\alpha \in F[\alpha]$,

$$\begin{aligned} c_n\alpha^n + c_{n-1}\alpha^{n-1} + \ldots + c_1\alpha + c_0 &= \frac{1}{\alpha} \\ \Rightarrow c_n\alpha^{n+1} + c_{n-1}\alpha^n + \ldots + c_1\alpha^2 + c_0\alpha - 1 &= 0 \end{aligned}$$

for some $c_0, c_1, \ldots, c_n \in F$. Therefore, for $f \in F[x]$ defined by

$$f(x) = c_n x^{n+1} + c_{n-1}x^n + \ldots + c_1 x^2 + c_0 x - 1,$$

it follows that $f(\alpha) = 0$ and hence, that α is algebraic over F. ■

Proposition 2.2.10. *Let $\alpha_1, \ldots, \alpha_n \in L$ be algebraic over F. Then,*

$$F[\alpha_1, \ldots, \alpha_n] = F(\alpha_1, \ldots, \alpha_n)$$

Proof. Let $P(n)$ be the statement of the proposition. Note, by Proposition 2.2.9, $P(1)$ is true. Now, suppose $n > 1$ and $F[\alpha_1, \ldots, \alpha_{n-1}] = F(\alpha_1, \ldots, \alpha_{n-1})$. For some $f \in F[x_1, \ldots, x_n]$, we have that $f(\alpha_1, \ldots, \alpha_n) = 0$. Let $K = F[\alpha_1, \ldots, \alpha_{n-1}]$. Leaving f unevaluated at x_n yields $f(\alpha_1, \ldots, \alpha_{n-1}, x_n)$ which is an element of $K[x_n]$; i.e. f is a univariate polynomial in x_n with coefficients in K. It follows that α_n is algebraic over K, so by Proposition 2.2.9, $K[\alpha_n] = K(\alpha_n)$. Since $K = F(\alpha_1, \ldots, \alpha_{n-1})$ by assumption and by Corollary 2.2.7 we have

$$F[\alpha_1, \ldots, \alpha_{n-1}][\alpha_n] = F(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n)$$

Next, take $g \in F[\alpha_1, \ldots, \alpha_{n-1}, \alpha_n]$ and leave it unevaluated at x_n so that we have

$$g(\alpha_1, \ldots, \alpha_{n-1}, x_n) \in K[x_n] = F[\alpha_1, \ldots, \alpha_{n-1}][x_n]$$

Then, re-evaluating g at α_n, we get that $g \in K[\alpha_n]$ and hence we have shown that

$$F[\alpha_1, \ldots, \alpha_{n-1}, \alpha_n] \subset F[\alpha_1, \ldots, \alpha_{n-1}][\alpha_n]$$

Conversely, take $g \in K[\alpha_n]$. Then,

$$g = k_m {\alpha_n}^m + k_{m-1} {\alpha_n}^{m-1} + \ldots + k_0$$

where $k_i \in K$ for $i \in \{0, 1, \ldots, m\}$ for some $m \in \mathbb{N}$. Since the k_i are in K, they are polynomials in $\alpha_1, \ldots, \alpha_{n-1}$ with coefficients in F, so we may regard them as $k_i(\alpha_1, \ldots, \alpha_{n-1})$ and so we have

$$g = k_m(\alpha_1, \ldots, \alpha_{n-1}) {\alpha_n}^m + k_{m-1}(\alpha_1, \ldots, \alpha_{n-1}) {\alpha_n}^{m-1} + \ldots + k_0(\alpha_1, \ldots, \alpha_{n-1})$$

It is now apparent that g is a polynomial in $\alpha_1, \ldots, \alpha_{n-1}, \alpha_n$ with coefficients in F, so $g \in F[\alpha_1, \ldots, \alpha_{n-1}, \alpha_n]$ and hence,

$$F[\alpha_1, \ldots, \alpha_{n-1}][\alpha_n] \subset F[\alpha_1, \ldots, \alpha_{n-1}, \alpha_n]$$

It follows that it is the case that $P(n-1) \Rightarrow P(n)$ since we now have $F[\alpha_1, \ldots, \alpha_{n-1}][\alpha_n] = F[\alpha_1, \ldots, \alpha_{n-1}, \alpha_n]$. Thus,

$$F[\alpha_1, \ldots, \alpha_n] = F(\alpha_1, \ldots, \alpha_n)$$

■

2.3 Degree of a Field Extension

One can show that for the field extension $F \subset L$, that L forms a vector space over F; we denote this vector space as L/F and the dimension of this vector space $\dim_F(L)$.

Definition 2.3.1. *The **degree** of a field extension $F \subset L$ is $[L : F] = \dim_F(L)$. We say $F \subset L$ is a **finite extension** if $[L : F] < \infty$.*

Proposition 2.3.2. *If we have the field extension $F \subset L$, then $[L : F] = 1 \iff L = F$.*

Proof. Suppose $[L : F] = 1$; letting B denote the basis of L/F, we have that $B = \{l\}$ for some $l \in L$. Since $L = \text{span}(B) = \{a \cdot l \mid a \in F\}$, we have that $1 = a \cdot l$ for some $a \in F$. It follows that $l = 1/a$, and since $1/a \in F$, we have that $l \in F$. Thus, $L \subset F$ and since by assumption, $F \subset L$, we conclude that $F = L$.

Now, suppose that $F = L$. Let $l = 1$ where $l \in L$. Then, $\text{span}(\{l\}) = \{a \cdot 1 \mid a \in F\} \Rightarrow F = \text{span}(\{l\})$. Since $F = L$, we have that $L = \text{span}(\{l\})$, so $\dim_F(L) = 1$. ■

Recall from Proposition 2.2.6, that for $\alpha \in L$, $F(\alpha)$ is the smallest subfield of L that contains F and α.

Definition 2.3.3. *Take $\alpha \in L$ where $F \subset L$ is a field extension. Then, $F \subset F(\alpha)$ is called a **simple extension**.*

Theorem 2.3.4. *Take $\alpha \in L$ where α is algebraic over F. Let $p \in F[x]$ be the minimal polynomial of α where $n = \deg(p)$. Letting B denote a basis of the simple extension $F \subset F(\alpha)$, it follows that we can have $B = \{1, \alpha, \ldots, \alpha^{n-1}\}$ and hence, $[F(\alpha) : F] = n$.*

Proof. Since α is algebraic, it follows from Proposition 2.2.9 that $F(\alpha) = F[\alpha]$ and hence, that all elements of $F(\alpha)$ are of the form $f(\alpha)$ where $f \in F[x]$. Since $f, p \in F[x]$, by Theorem A.3.2, there must be some $q \in F[x]$ such that

$$f(x) = q(x)p(x) + r(x)$$

where, $r(x) = c_0 + c_1x + \ldots + c_{n-1}x^{n-1}$ and $c_0, c_1, \ldots, c_{n-1} \in F$; here, either $r(x) \neq 0$ where $\deg(r) \leq n - 1$ or $r(x) = c_0 = c_1 = \ldots = c_{n-1} = 0$. Suppose the former case for $r(x)$; then, evaluating f at α yields

$$f(\alpha) = c_0 + c_1\alpha + \ldots + c_{n-1}\alpha^{n-1}$$

since $p(\alpha) = 0$ by assumption. It follows that since, $f(\alpha)$ is an arbitrary element of $F(\alpha)$, we have that $F(\alpha) = \text{span}(\{1, \alpha, \ldots, \alpha^{n-1}\})$. Now, suppose $f(\alpha) = 0$; then, $r(\alpha) = 0$. However, if we maintain the assumption that $r(x) \neq 0$, then we have a contradiction since p is the polynomial of smallest degree such that $p(\alpha) = 0$, yet $\deg(r) \leq n - 1 < n = \deg(p)$. Thus, we can only have

$$c_0 + c_1\alpha + \ldots + c_{n-1}\alpha^{n-1} = 0$$

when $r(x) = 0$ and hence, only when $c_0 = c_1 = \ldots = c_{n-1} = 0$. Thus, we have that $\{1, \alpha, \ldots, \alpha^{n-1}\}$ consists of linearly independent vectors that span $F(\alpha)$. Furthermore, we see that $\dim_F(F(\alpha)) = n$, so $[F(\alpha) : F] = n$. ■

Corollary 2.3.5. *Let $\alpha \in L$ where $F \subset L$ is a field extension. Then, α is algebraic over F if and only if $[F(\alpha) : F] < \infty$.*

Proof. Suppose α is algebraic over F. By Theorem 2.3.4, if $p \in F[x]$ is the minimal polynomial of α where $\deg(p) = n$, then, $[F(\alpha) : F] = n < \infty$.

Now, suppose $[F(\alpha) : F] = n < \infty$. Then, the vectors in $\{1, \alpha, \ldots, \alpha^n\}$ are linearly dependent since this is a set of $n+1$ vectors, so there are $c_0, c_1, \ldots, c_n$ where at least one $c_i \neq 0$ for $i \in \{1, \ldots, n\}$ such that

$$c_0 + c_1\alpha + \ldots + c_{n-1}\alpha^{n-1} + c_n\alpha^n = 0$$

Here, we have that α is a root of some nonzero $f \in F[x]$ given by

$$f(x) = c_0 + c_1 x + \ldots + c_{n-1}x^{n-1} + c_n x^n$$

Thus, it follows that α is algebraic over F. ■

Theorem 2.3.6 (Tower Theorem). *Suppose that we have successive field extensions given by the fields $F \subset K \subset L$. If $[K : F] < \infty$ and $[L : K] < \infty$, then $[L : F] < \infty$ where $[L : F] = [L : K][K : F]$. Furthermore, if $[K : F] = \infty$ or $[L : K] = \infty$, then $[L : F] = \infty$.*

Proof. Suppose $[K : F] < \infty$ and $[L : K] < \infty$. Then, $[K : F] = m$ and $[L : K] = n$ for some $m, n \in \mathbb{N}$. Let $B_{K/F} = \{\alpha_i\}_{i=1}^m$ and $B_{L/K} = \{\beta_j\}_{j=1}^n$ be the bases of the vector spaces K over F and L over K respectively. Take $l \in L$; then, $l = \sum_{j=1}^n b_j\beta_j$ where each $b_j \in K$. Since for each $j \in \{1, \ldots, n\}$, $b_j \in K$, we have that $b_j = \sum_{i=1}^m a_{ij}\alpha_i$. It follows that

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

Thus, we have that $L = \text{span}_F(\{\alpha_i\beta_j\}_{i=1,j=1}^{n,m})$ where the subscript indicates scalars are in F. Moreover, suppose

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

Then,

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

Since $\{\beta_j\}_{j=1}^n$ is a basis of L/K, it is a set of linearly independent vectors, so we conclude for each j that $\sum_{i=1}^m a_{ij}\alpha_i = 0$. Furthermore, since $\{\alpha_i\}_{i=1}^m$ is a basis of K/F, it is also a set of linearly independent vectors and it follows that $a_{ij} = 0$ for each i and j. Therefore, $\{\alpha_i\beta_j\}_{i=1,j=1}^{n,m}$ is a set of linearly independent vectors that forms a basis for L/F where $[L : F] = mn$ and hence, $[L : F] = [L : K][K : F]$. Furthermore, since it is the case that $[L : F] = [L : K][K : F]$ while both $[L : K]$ and $[K : F]$ are finite, we see that if either $[L : K] = \infty$ or $[K : F] = \infty$, then $[L : F] = \infty$. ■

3

Types of Extensions

This chapter introduces the notion of algebraic extensions (section 3.1; corresponds to chapter 4.4 in [5]), splitting fields (section 3.2; corresponds to chapter 5.1 in [5]), normal extensions (section 3.3; corresponds to chapter 5.2 in [5]), and separable extensions (section 3.4; corresponds to both chapter 5.3 and 5.4 in [5]). Normal and separable extensions will be especially important when we discuss the solvability of polynomials. Parts of proof left as exercises and additional details were filled in with help from sources like "Github Repository: Cox Galois Theory Exercises" (Ganaye) [6], "A First Course in Abstract Algebra: 7th. Ed." (Fraleigh and Katz 2003) [7] and "Groupprops, the Group Properties Wiki" (Naik) [8].

3.1 Algebraic Extensions

Definition 3.1.1. *A field extension $F \subset L$ is called an **algebraic extension** if $\forall \alpha \in L$, α is algebraic over F.*

Example 3.1.2. *Note that the field extension given by $\mathbb{Q} \subset \mathbb{R}$ is not an algebraic extension since, for instance, $\pi \in \mathbb{R}$, and there is no $f \in \mathbb{Q}[x]$ such that $f(\pi) = 0$. We will see that this is related to the fact that the degree of the field extension $[\mathbb{R} : \mathbb{Q}]$ is infinite; on top of that, the extension is also uncountably infinite since $\mathbb{Q}$ is countable while $\mathbb{R}$ is not.*

Proposition 3.1.3. *Let $F \subset L$ be a finite extension. Then:* (i) *$F \subset L$ is an algebraic extension and* (ii) *if $\alpha \in L$ where α is algebraic over F and $p \in F[x]$ is its minimal polynomial where $n = deg(p)$, then $[L : F]/n \in \mathbb{N}$.*

Proof. (i) By Proposition 2.2.6, we have that $F \subset F(\alpha) \subset L$ for any $\alpha \in L$. Since $F \subset L$ is finite, $F \subset F(\alpha)$ is also finite by the Tower Theorem (Theorem 2.3.6; specifically, the contrapositive of the infinite statement), and it follows from Corollary 2.3.5 that any $\alpha \in L$ will be algebraic over F, so $F \subset L$ is an algebraic extension.

(ii) Since α is algebraic over F, it follows from Theorem 2.3.4 that $[F(\alpha) : F] = n$. Using the Tower Theorem again, we have that

$$[L : F] = n[L : F(\alpha)]$$

and since $[L : F(\alpha)] \in \mathbb{N}$, it follows that $[L : F]/n \in \mathbb{N}$. ■

By Proposition 3.1.3, we can be sure that if a field extension $F \subset L$ is finite, then it is algebraic. Contrapositively, this means that if the extension is not algebraic, then it is not finite. Thus, since $\mathbb{Q} \subset \mathbb{R}$ is not algebraic, it is not finite.

Theorem 3.1.4. *A field extension $F \subset L$ is finite if and only if there are $\alpha_1, \ldots, \alpha_n \in L$ each algebraic over F such that $L = F(\alpha_1, \ldots, \alpha_n)$.*

Proof. Suppose $F \subset L$ is finite so $[L : F] < \infty$ and that $\dim_F(L) = n$ where we suppose $\alpha_1, \ldots, \alpha_n \in L$ form a basis of L over F. Then, $L = \{a_1\alpha_1 + \ldots + a_n\alpha_n \mid a_1, \ldots, a_n \in F\}$ where by Definition 2.2.5, we see that $L \subset F(\alpha_1, \ldots, \alpha_n)$. However, by Proposition 2.2.6, $F(\alpha_1, \ldots, \alpha_n) \subset L$, so $L = F(\alpha_1, \ldots, \alpha_n)$ where by Proposition 3.1.3, each α_i is algebraic where $i \in \{1, \ldots, n\}$.

Conversely, suppose there are $\alpha_1, \ldots, \alpha_n \in L$ each algebraic over F such that $L = F(\alpha_1, \ldots, \alpha_n)$. Let, $L_0 = F$ and $L_i = F(\alpha_1, \ldots, \alpha_i)$ where $i \in \{1, \ldots, n\}$. Then,

$$F \subset L_1 \subset \ldots \subset L_{n-1} \subset L$$

By Corollary 2.2.7, $L_i = L_{i-1}(\alpha_i)$ where α_i is algebraic over L_{i-1} since it is algebraic over F; also, by Corollary 2.3.5, since α_i is algebraic over L_{i-1}, it follows $L_{i-1} \subset L_{i-1}(\alpha_i) = L_i$ is finite. Thus, by consecutive applications of the tower theorem, we find that $[L : F]$ $= [L : L_{n-1}] \cdots [L_2 : L_1][L_1 : F]$ is also finite. ■

3.2 Splitting Fields

Definition 3.2.1. *Take $f \in F[x]$ where $n = \deg(f) \in \mathbb{N}$. Then, the field extension $F \subset L$ is called a **splitting field** of f over F if f splits completely over L with roots $\alpha_1, \ldots, \alpha_n \in L$ and $L = F(\alpha_1, \ldots, \alpha_n)$.*

Remark. *It is important to note that a splitting field is the smallest field extension such that a polynomial splits completely; i.e. just because a polynomial splits completely in a field extension does not necessarily mean that field extension is a splitting field.*

Theorem 3.2.2. *Take $f \in F[x]$ where $\deg(f) = n > 0$ and let L be a splitting field of f over F. Then, $[L : F] \leq n!$*

Proof. Let $P(n)$ be the statement of the theorem. Suppose $n = 1$; then, $f = ax+b = a(x+\frac{b}{a})$ where $a \neq 0$. Since $f(-b/a) = 0$ and $-b/a \in F$, it follows that $L = F$ and hence, by Proposition 2.3.2, $[L : F] = 1 \leq 1!$, so $P(1)$ is true.

Now suppose $n > 1$ and $P(n-1)$ is true. Let $\alpha_1, \ldots, \alpha_n$ be the roots of f; by assumption, $L = F(\alpha_1, \ldots, \alpha_n)$ is the splitting field of f over F. Expressing f as $f = (x - \alpha_1)g$, it is implicit by Theorem A.3.2 that $g \in F(\alpha_1)[x]$ where $\deg(g) = n - 1$; this means that the splitting field of g over $F(\alpha_1)$ is $F(\alpha_1)(\alpha_2, \ldots, \alpha_n)$. By Corollary 2.2.7, it follows that L is the splitting field of g over $F(\alpha_1)$. Since $P(n-1)$ is true and $\deg(g) = n - 1$, we have $[L : F(\alpha_1)] \leq (n-1)!$. Also, by Theorem 2.3.4, since α_1 is algebraic over F, $[F(\alpha_1) : n] \leq n$. By the tower theorem, $[L : F] = [L : F(\alpha_1)][F(\alpha_1) : F] \leq n[L : F(\alpha_1)] \leq n(n-1)!$. Thus, $P(n-1) \Rightarrow P(n)$ is true and $[L : F] \leq n!$. ■

Remark. *If $\varphi\colon F_1 \to F_2$ is a field isomorphism we will denote it as $\varphi\colon F_1 \simeq F_2$. Also, provided with some $\phi\colon S \to T$, denote the restriction of ϕ to $A \subset S$ mapped onto $\phi(A) \subset T$ as $\phi|_A$*

Theorem 3.2.3. *Suppose $\varphi\colon F_1 \simeq F_2$ and $f_1 \in F_1[x]$ where $n = \deg(f_1)$ and φ acting on the coefficients of f_1 yields some $f_2 \in F_2[x]$ such that $\deg(f_2) = \deg(f_1)$. Let L_1 and L_2 be the splitting fields of f_1 over F_1 and f_2 over F_2 respectively; then, there is some $\overline{\varphi}\colon L_1 \simeq L_2$ where $\varphi = \overline{\varphi}\big|_{F_1}$.*

Proof. Let $P(n)$ be the statement of the theorem pertaining to when $n = \deg(f_1) = \deg(f_2)$. Suppose $n = 1$; then, as in the proof of Theorem 3.2.2, we have that $L_1 = F_1$ and $L_2 = F_2$, so by setting $\overline{\varphi} = \varphi$, it follows that $P(1)$ is true.

Now, suppose $n > 1$ and that $P(n-1)$ is true. Letting $\alpha_1, \ldots, \alpha_n \in L_1$ be the roots of f_1, we have $L_1 = F_1(\alpha_1, \ldots, \alpha_n)$. As in the proof of Theorem 3.2.2, if we express f_1 as $f_1 = (x - \alpha_1)g_1$ then, $\deg(g_1) = n - 1$ and L_1 is the splitting field of g_1 over $F_1(\alpha_1)$.

Since α_1 is algebraic over F_1, it follows from Proposition 2.2.8 and Proposition 2.2.9 that

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

where $h_1 \in F_1[x]$ is the minimal polynomial of α_1 and hence, an irreducible factor of f_1 according to Proposition 2.2.2.

Since $\varphi\colon F_1 \simeq F_2$ is a field isomorphism, if we define $\widetilde{\varphi}\colon F_1[x] \to F_2[x]$ by having φ acting on the coefficients of a polynomial of $F_1[x]$ so we get a polynomial in $F_2[x]$ of the same degree, we find that $\widetilde{\varphi}$ is a ring isomorphism induced by the field isomorphism φ. It follows that $\widetilde{\varphi}(f_1) = f_2 \in F_2[x]$ where $\widetilde{\varphi}(h_1) \coloneqq h_2$ is an irreducible factor of f_2. Moreover, since f_2 splits completely over L_2 by assumption, it follows that h_2 also splits completely over L_2.

Let $\beta_1, \ldots, \beta_n \in L_2$ be the roots of f_2 and let β_1 also be a root of h_2. Since h_2 is irreducible, it is the minimal polynomial of β_1 according to Proposition 2.2.4. Similar to α_1, it follows from Proposition 2.2.8 and Proposition 2.2.9 that

$$F_2(\beta_2) \simeq F_2[x]/\langle h_2 \rangle$$

Since $\widetilde{\varphi}$ is an isomorphism we have that $\forall p_2 \in F_2[x]$, there is a $p_1 \in F_1[x]$ such that $\widetilde{\varphi}(p_1) = p_2$, so that $\widetilde{\varphi}(p_1 h_1) = p_2 h_2$ and hence, $\widetilde{\varphi}(\langle h_1 \rangle) = \langle h_2 \rangle$; furthermore, we have that $\widetilde{\varphi}(p_1 + \langle h_1 \rangle) = p_2 + \langle h_2 \rangle$. Keeping in mind the first isomorphism theorem as in the proof of Proposition 2.2.8, we can conclude that

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

Using the above results and the transitive property of "$\simeq$", it follows that there is a field isomorphism φ_1 such that

$$\varphi_1\colon F_1(\alpha_1) \simeq F_2(\beta_1)$$

where $\varphi_1(\alpha_1) = \beta_1$ and $\varphi_1\big|_{F_1} = \varphi$.

In a similar manner to how the field isomorphism φ induced the ring isomorphism $\widetilde{\varphi}$, the field isomorphism φ_1 induces a ring isomorphism $\widetilde{\varphi}_1 \colon F_1(\alpha_1)[x] \simeq F_2(\beta_1)[x]$ defined by having φ_1 acting on the coefficients of a polynomial of $F_1(\alpha_1)[x]$ so we get a polynomial in $F_2(\beta_1)[x]$ of the same degree. Applying $\widetilde{\varphi}_1$ to $f_1 = (x - \alpha_1)g_1$ as expressed earlier, we get

$$\widetilde{\varphi}_1(f_1) = \widetilde{\varphi}_1((x - \alpha_1)g_1) = (x - \varphi_1(\alpha_1))\widetilde{\varphi}_1(g_1) \coloneqq (x - \beta_1)g_2$$

where $\deg(g_2) = n - 1$. Additionally, note that $\widetilde{\varphi}_1\big|_{F_1[x]} = \widetilde{\varphi}$, so we may conclude that $\widetilde{\varphi}_1(f_1) = f_2$ and hence, $f_2 = (x - \beta_1)g_2$; also, as in the proof of Theorem 3.2.2, we have that L_2 is the splitting field of g_2 over $F_2(\beta_1)$.

Since we assumed $P(n-1)$ is true and $\deg(g_1) = \deg(g_2) = n - 1$, it follows that there is some $\overline{\varphi}_1 \colon L_1 \simeq L_2$ where $\varphi_1 = \overline{\varphi}_1\big|_{F_1(\alpha_1)}$. However, note that $\varphi_1\big|_{F_1} = \varphi$, so it follows that $\overline{\varphi}_1\big|_{F_1} = \varphi$. Therefore, $\overline{\varphi}_1$ is also the described isomorphism for when $\deg(f_1) = \deg(f_2) = n$ and we conclude that $P(n-1) \Rightarrow P(n)$ is true. ■

Corollary 3.2.4. *Let L be the splitting field of $f \in F[x]$ and let $h \in F[x]$ be irreducible with roots $\alpha, \beta \in L$. Then, there is some $\sigma \colon L \simeq L$ such that $\sigma(\alpha) = \beta$ and $\forall a \in F,\ \sigma(a) = a$.*

Proof. Since h is irreducible, it is the minimal polynomial of $\alpha, \beta \in L$; also, since α and β are roots of a polynomial in $F[x]$, they are algebraic over F. Therefore, it follows by Proposition 2.2.8 and Proposition 2.2.9 that

$$F(\alpha) \simeq F[x]/\langle h \rangle \quad \textit{and} \quad F(\beta) \simeq F[x]/\langle h \rangle$$

which implies by transitivity of "$\simeq$" that $F(\alpha) \simeq F(\beta)$, so there is an isomorphism ϕ that maps $F(\alpha)$ one-to-one and onto $F(\beta)$ where $\phi(\alpha) = \beta$ and $\forall a \in F,\ \phi(a) = a$.

Since $f \in F[x]$, it follows that $f \in F(\alpha)[x]$ and $f \in F(\beta)[x]$. Also, note that we have $F \subset F(\alpha) \subset L$ and $F \subset F(\beta) \subset L$, so we can conclude that L is a splitting field of f over both $F(\alpha)$ and $F(\beta)$ and that ϕ induces an isomorphism that maps f over $F(\alpha)$ to itself over $F(\beta)$. By Theorem 3.2.3, it follows that there is some $\overline{\phi} \colon L \simeq L$ such that $\overline{\phi}\big|_{F(\alpha)} = \phi$; thus, by setting $\sigma = \overline{\phi}$, we have some $\sigma \colon L \simeq L$ such that $\sigma(\alpha) = \beta$ and $\forall a \in F,\ \sigma(a) = a$. ■

3.3 Normal Extensions

Definition 3.3.1. *An algebraic extension $F \subset L$ is said to be a **normal extension** if every irreducible polynomial over F with a root in L splits completely over L.*

Proposition 3.3.2. *Suppose L is the splitting field of $f \in F[x]$ and $g \in F[x]$ is irreducible. Then, if there is $\beta \in L$ such that $g(\beta) = 0$, then g splits completely over L.*

Proof. WLOG, we may assume f and g are monic. Suppose $\alpha_1, \ldots, \alpha_n$ are the roots of f so that $L = F(\alpha_1, \ldots, \alpha_n)$ and $f = (x - \alpha_1) \cdots (x - \alpha_n)$ over L. Since g is irreducible, it follows that it is the minimal polynomial of β by Proposition 2.2.4. Also, by Proposition 2.2.10,

$L = F[\alpha_1, \ldots, \alpha_n]$, so $\beta = h(\alpha_1, \ldots, \alpha_n)$ for some $h \in F[x_1, \ldots, x_n]$. Now, consider $S(x) \in F[x_1, \ldots, x_n][x]$ where

$$S(x) = \prod_{\sigma \in S_n} (x - h(x_{\sigma(1)}, \ldots, x_{\sigma(n)}))$$

Here, the action of S_n on $h(x_1, \ldots, x_n)$ yields a total of $n!$ values (where each one may or may not be unique); denote these values as $h_1, \ldots, h_{n!}$ where $h_1 \coloneqq h(x_1, \ldots, x_n)$. Thus, it follows that

$$S(x) = \prod_{i=1}^{n!} (x - h_i) = \sum_{i=0}^{n!} p_i(x_1, \ldots, x_n) x^i$$

where the last equality follows from multiplying out the linear factors as in Appendix A.4 for f; furthermore, from Appendix A.4, we see that multiplying out linear factors yields symmetric coefficients, so we can conclude that each p_i is symmetric. By Corollary A.4.5, since $\alpha_1, \ldots, \alpha_n \in L$ are roots of f, it follows that for each $i \in \{1, \ldots, n\}$, $p_i(\alpha_1, \ldots, \alpha_n) \in F$. Thus, evaluation of $S(x)$ at the roots of f yields $s(x) = (x - \beta) \prod_{i=2}^{n!} (x - h_i(\alpha_1, \ldots, \alpha_n)) \in L[x]$ and hence,

$$s(x) = \sum_{i=0}^{n!} p_i(\alpha_1, \ldots, \alpha_n) x^i$$

where we see that $s(x) \in F[x]$. Since g is the minimal polynomial of β it divides $s(x)$ according to Proposition 2.2.2 and hence, can be considered as a factor of $s(x)$. Furthermore, we see that $s(x)$ splits completely over L, so it follows that g splits completely over L. ■

Theorem 3.3.3. *Suppose $F \subset L$ is a field extension. Then, L is the splitting field of some $f \in F[x]$ if and only if $F \subset L$ is normal and finite.*

Proof. Suppose L is the splitting field of some $f \in F[x]$. Then, by Theorem 3.1.4, we see that $F \subset L$ is finite and by Proposition 3.1.3, $F \subset L$ is also algebraic. Moreover, by Proposition 3.3.2, every irreducible polynomial that has a root in L splits completely over L so $F \subset L$ is also normal.

Conversely, suppose that $F \subset L$ is normal and finite. Using Theorem 3.1.4 again, since the extension is finite, it follows that there are $\alpha_1, \ldots, \alpha_n$ each algebraic over F such that $L = F(\alpha_1, \ldots, \alpha_n)$. For each $i \in \{1, \ldots, n\}$, let $p_i \in F[x]$ be the minimal polynomial of α_i and define $f = p_1 p_2 \cdots p_n$. Since $F \subset L$ is normal and each p_i is irreducible by Proposition 2.2.4 with $\alpha_i \in L$ as a root, it follows that each p_i splits completely over L and hence, f splits completely over L. Now, let L' be a subfield of L containing the roots of f and F. Then, $F(\alpha_1, \ldots, \alpha_n) \subset L' \subset L$, but $L = F(\alpha_1, \ldots, \alpha_n)$ so it follows that $L = L'$ and hence, L is the splitting field of f over F. ■

3.4 Separable Extensions

Definition 3.4.1. *Let $F \subset L$ be an algebraic extension.*
(i) *A nonconstant polynomial $f \in F[x]$ is said to be* ***separable*** *if its roots in a splitting field are distinct.*
(ii) *$\alpha \in L$ is said to be* ***separable*** *over F if the minimal polynomial of α over F is separable.*
(iii) *$F \subset L$ is a* ***separable extension****, if for all $\alpha \in L$, α is separable over F.*

Definition 3.4.2. *Let $g = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \in F[x]$; then, the* ***formal derivative*** *of g is defined to be*

$$g' = na_0x^{n-1} + (n-1)a_1x^{n-2} + \ldots + a_{n-1} \tag{3.1}$$

Remark. *Eq.* (3.1) *is simply the derivative of g if we consider the principles of calculus. For simplicity, let's assume corresponding theorems from calculus such as the product and chain rule are true here, although they can be proved from an algebraic point of view when considering Eq.* (3.1)*.*

Theorem 3.4.3. *Let $f \in F[x]$ be monic and nonconstant. Then, f is separable if and only if $\Delta(f) \neq 0$. Also, f is separable if and only if* $\gcd(f, f') = 1$.

Proof. The first statement is quite trivial when considering Eq. (A.8); indeed, if for some $i < j$, $\alpha_i = \alpha_j$, then $\Delta(f) = 0$. Conversely, it cannot be the case that $\Delta(f) = 0$ if for all $i < j$, $\alpha_i \neq \alpha_j$, so we must conclude that for some $i < j$, $\alpha_i = \alpha_j$ when $\Delta(f) = 0$.

For the second statement, suppose L is a splitting field of f over F so that $f = (x - \alpha_1)\ldots(x - \alpha_n)$ in $L[x]$. For some $i \in \{1, \ldots, n\}$, let

$$f = (x - \alpha_i)h_i(x)$$

where $h_i(x) = \prod_{j \neq i}(x - \alpha_j)$. Taking the derivative of f via the product rule yields $f' = (x - \alpha_i)h_i'(x) + h_i(x)$ where evaluation at α_i yields $f'(\alpha_i) = h_i(\alpha_i) = \prod_{j \neq i}(\alpha_i - \alpha_j)$. Now, suppose that $\gcd(f, f') \neq 1$. Then, there is some nonconstant $g \in F[x]$ where $\deg(g) > 0$ that divides both f and f'. In this regard, g is a factor of f so for some $i \in \{1, \ldots, n\}$, $g(\alpha_i) = 0$; since g is also a factor of f', it follows that $f'(\alpha_i) = 0$. Thus, $f'(\alpha_i) = \prod_{j \neq i}(\alpha_i - \alpha_j) = 0$ which means that $\alpha_i = \alpha_j$ for some $i \neq j$ which contradicts our assumption of the separability of f.

Conversely, suppose $\gcd(f, f') = 1$. Then, for some $a, b \in F[x]$, $af + bf' = 1$ where upon evaluation at α_i for some $i \in \{1, \ldots, n\}$, $1 = bf'(\alpha_i)$, so it cannot be the case that $f'(\alpha_i) = 0$. Since $f'(\alpha_i) = \prod_{j \neq i}(\alpha_i - \alpha_j)$, it follows that $\alpha_i \neq \alpha_j$ for $i \neq j$, so all the roots of f are distinct. ■

Proposition 3.4.4. *Let F be a field of characteristic zero. Then, if $f \in F[x]$ is irreducible where* $\deg(f) = n > 0$*, then f is separable.*

Proof. WLOG, suppose the leading coefficient of f is nonzero. By the definition of the formal derivative, $f' \neq 0$ and since $\deg(f) = n$, it follows that $\deg(f') = n - 1$. Since f is irreducible, its only divisors up to a constant factor are itself and one. Let $g = \gcd(f, f')$; then $g \in \{1, f\}$. Since g divides f' and $f' \neq 0$, it follows that $\deg(g) \leq \deg(f') = n - 1$, so $g \neq f$. Thus, it must be that $g = \gcd(f, f') = 1$. ■

Theorem 3.4.5 (Theorem of the Primitive Element). *Let $F \subset L$ be a finite extension where $L = F(\alpha_1, \ldots, \alpha_n)$ and each α_i, $i \in \{1, \ldots, n\}$ is separable over F. Then, there is an $\alpha \in L$, called the **primitive element** of the extension, separable over F such that $L = F(\alpha)$. Also, there are $t_1, \ldots, t_n \in F$ such that α can be expressed as*

$$\alpha = t_1\alpha_1 + \ldots + t_n\alpha_n$$

Proof. Let $P(n)$ be the statement of the theorem. Then, P(1) is fairly trivial since $L = F(\alpha_1)$ and setting $t_1 = 1$ yields $\alpha = \alpha_1$.

Suppose $n = 2$ (which is useful to consider for the induction step); then $L = F(\beta, \gamma)$ where β and γ are separable over F. Let $f, g \in F[x]$ be the minimal polynomials of β and γ respectively where $\deg(f) = l$ and $\deg(g) = m$. Separability of β and γ means that f has distinct roots $\beta_1, \beta_2, \ldots, \beta_l$ and g has distinct roots $\gamma_1, \gamma_2, \ldots, \gamma_m$ where we set $\beta_1 = \beta$ and $\gamma_1 = \gamma$. Since F is presumed to have characteristic zero, it is infinite in cardinality, and so we can find some $\lambda \in F$ such that

$$\lambda \neq \frac{\beta_i - \beta_r}{\gamma_s - \gamma_j},\ (i,j),(r,s) \in \{1, \ldots, l\} \times \{1, \ldots, m\} \text{ where } s \neq j$$

which implies by setting $r = 1$ and $s = 1$ that

$$\beta + \lambda\gamma \neq \beta_i + \lambda\gamma_j \tag{3.2}$$

where $(i, j) \in \{1, \ldots, l\} \times \{2, \ldots, m\}$. By Definition 2.2.5, it's clear that $F(\beta + \lambda\gamma) \subset F(\beta, \gamma)$. It remains to show the reverse inclusion to conclude that $F(\beta + \lambda\gamma) = F(\beta, \gamma)$ and that $\beta + \lambda\gamma$ is the primitive element of the extension when $n = 2$.

To start, take $\beta, \gamma \in F(\beta, \gamma)$. Note that $g \in F[x] \subset F(\beta + \lambda\gamma)[x]$ and $g(\gamma) = 0$; also, since $f(\beta) = 0$, it follows $f(\beta + \lambda\gamma - \lambda\gamma) = 0$, so in this sense, γ is a root of $f(\beta + \lambda\gamma - \lambda x) \in F(\beta + \lambda\gamma)[x]$. Now, consider the greatest common divisor (gcd) of $g(x)$ and $f(\beta + \lambda\gamma - \lambda x)$; we claim that the gcd is a monic linear polynomial $h(x) \in F(\beta + \lambda\gamma)[x]$. To show this, suppose the gcd is 1. Then, there are $a, b \in F(\beta + \lambda\gamma)[x]$ such that:

$$a(x)g(x) + b(x)f(\beta + \lambda\gamma - \lambda x) = 1$$

However, upon evaluation at γ that

$$a(\gamma)g(\gamma) + b(\gamma)f(\beta + \lambda\gamma - \lambda\gamma) = 1 \ \Rightarrow\ 0 = 1$$

since γ is a root of both $g(x)$ and $f(\beta + \lambda\gamma - \lambda x)$; clearly, this is a contradiction. Thus, we must have that

$$h(x) = \gcd(g(x), f(\beta + \lambda\gamma - \lambda x))$$

where $\deg(h) \geq 1$. Note that by default, γ is a root of h. We further claim that $\deg(h) = 1$. Suppose $\deg(h) > 1$; then, since h divides g, it is a factor of g. Since g is separable, h is separable so it must have at least one other root γ_j where $j \in \{2, \ldots, m\}$. However, since h also divides $f(\beta + \lambda\gamma - \lambda x)$, it is a factor of $f(\beta + \lambda\gamma - \lambda x)$, so $f(\beta + \lambda\gamma - \lambda\gamma_j) = 0$ since $h(\gamma_j) = 0$; since $\beta_1, \ldots, \beta_l$ are the only roots of f, there is some $i \in \{1, \ldots, l\}$ such that $\beta_i = \beta + \lambda\gamma - \lambda\gamma_j$ which implies

$$\beta + \lambda\gamma = \beta_i + \lambda\gamma_j$$

for some $(i, j) \in \{1, \ldots, l\} \times \{1, \ldots, m\}$. However, this contradicts Eq. (3.2), so we conclude that $\deg(h) = 1$ where $h(x) = x - \gamma$ since γ is the only root of h, and since $h(x) \in F(\beta + \lambda\gamma)[x]$, it follows that $\gamma \in F(\beta + \lambda\gamma)$; moreover, $F(\beta + \lambda\gamma)$ is closed under subtraction so

$$(\beta + \lambda\gamma) - \lambda\gamma \in F(\beta + \lambda\gamma) \Rightarrow \beta \in F(\beta + \lambda\gamma)$$

We conclude that $\beta, \gamma \in F(\beta + \lambda\gamma)$, so $F(\beta, \gamma) \subset F(\beta + \lambda\gamma)$ and hence $F(\beta, \gamma) = F(\beta + \lambda\gamma)$.

Now, we will show that $\beta + \lambda\gamma$ is separable over F. Suppose p is the minimal polynomial of $\beta + \lambda\gamma$ over F. Consider

$$s(x) = \prod_{j=1}^{m} f(x - \lambda\gamma_j)$$

where we can see that $s(\beta + \lambda\gamma) = 0$. Moreover, consider $s(x)$ as an evaluation of $q(x_1, \ldots, x_m, x) \in F[x_1, \ldots, x_m][x]$ at $(x_1, \ldots, x_m) = (\gamma_1, \ldots, \gamma_m)$. In this sense,

$$q = \sum_{k=0}^{d} w_k(x_1, \ldots, x_m)x^k = \prod_{j=1}^{m} f(x - \lambda x_j)$$

where each $w_k \in F[x_1, \ldots, x_m]$ and $d \in \mathbb{N}$. Also, define the action of some $\sigma \in S_m$ on some $u(x_1, \ldots, x_m, x) \in F[x_1, \ldots, x_m][x]$ by $\sigma(u) = u(x_{\sigma(1)}, \ldots, x_{\sigma(m)}, x)$. Now observe

$$\begin{aligned}\sigma(q) &= \sigma\bigg(\prod_{j=1}^{m} f(x-\lambda x_j)\bigg) \\ &= \prod_{j=1}^{m} \sigma(f(x-\lambda x_j))) \\ &= \prod_{j=1}^{m} f(x-\lambda x_{\sigma(j)}) \\ &= \prod_{j=1}^{m} f(x-\lambda x_j) = q\end{aligned}$$

Thus, $\sigma(q) = q$ which means that

$$q = \sum_{k=0}^{d} w_k x^k = \sigma(q) = \sum_{k=0}^{d} \sigma(w_k) x^k$$

Therefore, each $w_k \in F[x_1, \ldots, x_m]$ is a symmetric polynomial with coefficients in F and since g is a monic polynomial with roots $\gamma_1, \ldots, \gamma_m$ in a larger field L, it follows by Corollary A.4.5 that evaluation of each w_k at the roots of g implies that $w_k(\gamma_1, \ldots, \gamma_m) \in F$. Thus, evaluation of $q(x)$ at $\gamma_1, \ldots, \gamma_m$ yields

$$q(\gamma_1, \ldots, \gamma_m, x) = s(x) = \prod_{j=1}^{m} f(x-\lambda\gamma_j) = \sum_{k=0}^{d} w_k(\gamma_1, \ldots, \gamma_m) x^k$$

where the last equality shows that $s(x) \in F[x]$. Furthermore, since p is the minimal polynomial of $\beta + \lambda\gamma$ where $s(\beta + \lambda\gamma) = 0$, it follows by Proposition 2.2.2 that $p(x)$ divides $s(x)$. Since $f(x) = (x-\beta_1)\cdots(x-\beta_l)$, we get that $f(x-\lambda\gamma_j) = ((x-\lambda\gamma_j)-\beta_1)\cdots((x-\lambda\gamma_j)-\beta_l)$, so

$$s(x) = \prod_{i=1}^{l}\prod_{j=1}^{m}(x-(\beta_i+\lambda\gamma_j))$$

and since each $\beta_i + \lambda\gamma_j$ are distinct as indicated above, it follows that s is separable which must mean that p is separable since it divides s. Therefore, we see that $\beta + \lambda\gamma$ is separable over F and setting $t_1 = 1$ and $t_2 = \lambda$, we can conclude that $P(2)$ is true.

Now, suppose $n > 2$ and $P(n-1)$ is true. Then, there are $t_1, \ldots, t_{n-1} \in F$ where $F(\alpha_1, \ldots, \alpha_{n-1}) = F(\alpha_0)$ such that

$$\alpha_0 = t_1\alpha_1 + \ldots + t_{n-1}\alpha_{n-1}$$

is separable over F. Also, by Corollary 2.2.7,

$$L = F(\alpha_1, \ldots, \alpha_n) = F(\alpha_1, \ldots, \alpha_{n-1})(\alpha_n) = F(\alpha_0)(\alpha_n) = F(\alpha_0, \alpha_n)$$

and according to the proof for $P(2)$, $F(\alpha_0, \alpha_n) = F(\alpha_0 + \lambda\alpha_n)$ for some $\lambda \in F$ where $\alpha_0 + \lambda\alpha_n$ is separable. Furthermore, setting $t_n = \lambda$ and $\alpha = \alpha_0 + t_n\alpha_n$, we have that

$$\alpha = t_1\alpha_1 + \ldots + t_{n-1}\alpha_{n-1} + t_n\alpha_n$$

is a separable element over F where $F(\alpha_1, \ldots, \alpha_n) = F(\alpha)$. Thus, $P(n-1) \Rightarrow P(n)$ is true. ■

4

Galois Groups

The material in this chapter corresponds to chapter 6 in [5]. Section 4.1 introduces our main tool in analyzing the relationship between polynomial roots mentioned in Chapter 1; this section corresponds to chapter 6.1 and chapter 6.2 in [5]. Section 4.2 addresses the idea of transitive groups and establishes the isomorphism between the Galois group of the universal extension and the symmetric group; this section corresponds to chapter 6.3 and 6.4 in [5]. Parts of proof left as exercises and additional details were filled in with help from sources like "Github Repository: Cox Galois Theory Exercises" (Ganaye) [6], "A First Course in Abstract Algebra: 7th. Ed." (Fraleigh and Katz 2003) [7] and "Groupprops, the Group Properties Wiki" (Naik) [8].

4.1 Properties

Definition 4.1.1. *Let $\sigma\colon L \to L$ be a field isomorphism. Then, we say σ is an **automorphism** of L where we denote it as $\sigma\colon L \simeq L$.*

Definition 4.1.2 (Galois Group). *Suppose $F \subset L$ is a finite extension. Then, the **Galois Group** of $F \subset L$ is given by*

$$\mathrm{Gal}(L/F) = \{\sigma\colon L \simeq L \mid \forall a \in F,\ \sigma(a) = a\}$$

*If L is the splitting field of $f \in F[x]$, then $\mathrm{Gal}(L/F)$ is the **Galois Group of f over F**.*

Remark. *One can show that the Galois Group is indeed a group via composition.*

Lemma 4.1.3. *Let $F \subset L$ be a finite extension and take $\sigma \in \mathrm{Gal}(L/F)$. Then, taking $p \in F[x_1, \ldots, x_n]$ and $\beta_1, \ldots, \beta_n \in L$,*

$$\sigma(p(\beta_1, \ldots, \beta_n)) = p(\sigma(\beta_1), \ldots, \sigma(\beta_n))$$

Proof. Since $\mathrm{Gal}(L/F)$ is a group, σ acting on p preserves the multiplication and addition involved among the monomials that p consists of; also, by definition, σ maps the coefficients of the monomials of p to themselves. ■

Proposition 4.1.4. *Let $F \subset L$ be a finite extension and $\sigma \in \mathrm{Gal}(L/F)$.* (i) *Suppose that $p \in F[x]$ is nonconstant with $\alpha \in L$ as a root. Then, $\sigma(\alpha) \in L$ is also a root of p.* (ii) *If $L = F(\alpha_1, \ldots, \alpha_n)$, then the action of σ on any element of L is uniquely determined by $\sigma(\alpha_1), \ldots, \sigma(\alpha_n)$.*

Proof. (i) Since $p(\alpha) = 0$, it follows by Lemma 4.1.3 that $0 = \sigma(0) = \sigma(p(\alpha)) = p(\sigma(\alpha))$
(ii) By Proposition 3.1.3, since $F \subset L$ is finite, it follows that it is also algebraic which implies by Proposition 2.2.10 that $L = F[\alpha_1, \ldots, \alpha_n]$. It follows that any $\beta \in L$ can be expressed as $\beta = h(\alpha_1, \ldots, \alpha_n)$ where $h \in F[x_1, \ldots, x_n]$. By Lemma 4.1.3,

$$\sigma(\beta) = \sigma(h(\alpha_1, \ldots, \alpha_n)) = h(\sigma(\alpha_1), \ldots, \sigma(\alpha_n))$$

Thus, we see that σ is uniquely determined by $\sigma(\alpha_1), \ldots, \sigma(\alpha_n)$. ■

Theorem 4.1.5. *Let L be the splitting field of a separable polynomial $f \in F[x]$. Then, $|\mathrm{Gal}(L/F)| = [L : F]$.*

Proof. Since f is separable, it has distinct roots $\alpha_1, \ldots, \alpha_n$ where $L = F(\alpha_1, \ldots, \alpha_n)$. For each $i \in \{1, \ldots, n\}$, the minimum polynomial p_i of each α_i is irreducible so by Proposition 3.4.4, p_i is also separable; therefore, each α_i is separable. By Theorem 3.4.5, it follows that $\exists \beta \in L$ such that $L = F(\beta)$. Let $h \in F[x]$ be the minimum polynomial of β. Then, by Theorem 2.3.4, if $\deg(h) = m$ then, $[L : F] = m$. Again, by Proposition 3.4.4, since h is irreducible, it has distinct roots $\beta_1, \ldots, \beta_m \in L$ where we set $\beta_1 = \beta$. By Corollary 3.2.4, for each $i \in \{1, \ldots, m\}$, there is some $\sigma_i \colon L \simeq L$ such that $\sigma_i(\beta) = \beta_i$ and $\forall a \in F,\ \sigma(a) = a$. It follows $\sigma_1, \ldots, \sigma_m \in \mathrm{Gal}(L/F)$ and due to separability of h, for $i \neq j$, $\sigma_i \neq \sigma_j$ since $\sigma_i(\beta) = \beta_i \neq \beta_j = \sigma_j(\beta)$, so $m \leq |\mathrm{Gal}(L/F)|$.

Now, for any $\sigma \in \mathrm{Gal}(L/F)$, since $L = F(\beta)$, Proposition 4.1.4 implies that σ is uniquely determined by $\sigma(\beta) \in \{\beta_1, \ldots, \beta_m\}$, so it follows that $\sigma = \sigma_i$ for some $i \in \{1, \ldots, m\}$ and hence, $|\mathrm{Gal}(L/F)| = m$. ■

4.2 Transitivity

Suppose $F \subset L$ is the splitting field of a separable polynomial $f \in F[x]$ where $\deg(f) = n$; let $\alpha_1, \ldots, \alpha_n \in L$ be the distinct roots of f so that over L,

$$f = a_o(x - \alpha_1) \cdots (x - \alpha_n)$$

where $a_0 \in F \backslash \{0\}$. Now, take $\sigma \in \mathrm{Gal}(L/F)$; by Proposition 4.1.4, for $i \in \{1, \ldots, n\}$, we have that $\sigma(\alpha_i)$ is also a root of f. Since $\sigma(\alpha_i)$ is another one of the n roots of f, $\sigma(\alpha_i) = \alpha_j$ for some $j \in \{1, \ldots, n\}$. In this sense, $\sigma \colon \alpha_i \mapsto \alpha_j$ corresponds to $\tau \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ where $\tau(i) = j$, so $\sigma(\alpha_i) = \alpha_{\tau(i)}$; furthermore, by Proposition 4.1.4, $\tau(1), \ldots, \tau(n)$ are uniquely determined. Also, since the roots are distinct, $\alpha_i \neq \alpha_j \iff i \neq j$ and since σ is an automorphism, $\alpha_i \neq \alpha_j \Rightarrow \sigma(\alpha_i) \neq \sigma(\alpha_j)$ which implies that $\alpha_{\tau(i)} \neq \alpha_{\tau(j)}$ and hence, $\tau(i) \neq \tau(j)$; thus, we have that $i \neq j \Rightarrow \tau(i) \neq \tau(j)$, so τ is injective. Since

$\tau\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ is a one to one mapping where for each $i \in \{1,\ldots,n\}$, $\tau(i)$ is uniquely determined, we conclude that τ is a permutation on $\{1,\ldots,n\}$, so $\tau \in S_n$. Here, we have a mapping $\rho\colon \mathrm{Gal}(L/F) \to S_n$ where $\rho(\sigma) = \tau$ when $\sigma(\alpha_i) = \alpha_{\tau(i)}$.

Proposition 4.2.1. *Let $f \in F[x]$ be the separable polynomial mentioned in the above paragraph and α_i be one of its distinct roots. Then, the mapping given by $\rho\colon \mathrm{Gal}(L/F) \to S_n$ is a group homomorphism.*

Proof. Take $\sigma_1, \sigma_2 \in \mathrm{Gal}(L/F)$, where $\sigma_1(\alpha_i) = \alpha_{\tau_1(i)}$ and $\sigma_1(\alpha_i) = \alpha_{\tau_2(i)}$ and $\tau_1, \tau_2 \in S_n$; this means that $\rho(\sigma_1) = \tau_1$ and $\rho(\sigma_2) = \tau_2$. Then,

$$\sigma_1 \circ \sigma_2(\alpha_i) = \sigma_1(\sigma_2(\alpha_i)) = \sigma_1(\alpha_{\tau_2(i)}) = \alpha_{\tau_1(\tau_2(i))} = \alpha_{\tau_1 \circ \tau_2(i)}$$

Therefore $\sigma_1 \circ \sigma_2(\alpha_i) = \alpha_{\tau_1 \circ \tau_2(i)}$, so $\rho(\sigma_1 \circ \sigma_2) = \tau_1 \circ \tau_2$ and hence, $\rho(\sigma_1 \circ \sigma_2) = \rho(\sigma_1) \circ \rho(\sigma_2)$. ■

Definition 4.2.2. *Let H be a subgroup of S_n. Then H is **transitive** if for all $i, j \in \{1,\ldots,n\}$, there is some $\tau \in H$ such that $\tau(i) = j$.*

Proposition 4.2.3. *Let $f \in F[x]$ be a separable polynomial such that $\deg(f) = n$ and let L be the splitting field of f. Then, letting H denote the subgroup of S_n that corresponds to $\mathrm{Gal}(L/F)$, it follows that H is transitive if and only if f is irreducible.*

Proof. Suppose f is irreducible with distinct roots $\alpha_1, \ldots, \alpha_n \in L$. Then, by Corollary 3.2.4, $\forall i, j \in \{1,\ldots,n\}$, there is an automorphism $\sigma \in \mathrm{Gal}(L/F)$ such that, $\sigma(\alpha_i) = \alpha_j$; this corresponds to $\tau \in H$ where $\tau(i) = j$. Thus, we have that $\forall i, j \in \{1,\ldots,n\}$, $\exists\ \tau \in H\colon \tau(i) = j$, so H is transitive.

Now, suppose that H is transitive and let h be an irreducible nonconstant factor of f where the distinct roots of f are $\alpha_1, \ldots \alpha_n \in L$. Since h is a factor of f, there is some α_i such that $h(\alpha_i) = 0$ and since H is transitive, we have that $\forall i, j \in \{1,\ldots,n\}$, there is some $\sigma \in \mathrm{Gal}(L/F)$ such that $\sigma(\alpha_i) = \alpha_j$. Since $h \in F[x]$, by Proposition 4.1.4, $\sigma(\alpha_i) = \alpha_j$ is also a root of h which means that h has at least n roots, so $\deg(h) \geq n$. It follows that f is irreducible since it must be that $\deg(h) \leq \deg(f)$. ■

In Appendix A.4, we introduced the rings

$$k = F[\sigma_1, \ldots, \sigma_n]$$
$$l = F[x_1, \ldots, x_n]$$

Also, in Appendix A.4, we defined the universal polynomial $\widetilde{f} \in l[x]$ Eq. (A.3)

$$\widetilde{f} = (x - x_1) \cdots (x - x_n)$$

where its expansion Eq. (A.7) is in $k[x]$ and is given by

$$\widetilde{f} = x^n - \sigma_1 x^{n-1} + \ldots + (-1)^r \sigma_r x^{n-r} + \ldots + (-1)^n \sigma_n$$

Let $K = F(\sigma_1, \ldots, \sigma_n)$ and $L = F(x_1, \ldots, x_n)$, so by Eq. (A.5), $K \subset L$. We see that $\widetilde{f} \in K[x]$ splits completely over the field L according to Eq. (A.3) where the distinct roots of $\widetilde{f}$ are $x_1, \ldots, x_n$, so $\widetilde{f}$ is separable. Since this is the case, we can conclude that L is the splitting field of $\widetilde{f}$ over K and hence,

$$L = K(x_1, \ldots, x_n)$$

Definition 4.2.4. *From the above paragraph, $K \subset L$ is deemed the **universal extension** in degree n.*

Lemma 4.2.5. *Suppose R is an integral domain and K is its corresponding field of rational functions. Then, if $\phi\colon R \to R$ is a ring isomorphism, it extends uniquely to an automorphism $\widetilde{\phi}\colon K \simeq K$.*

Proof. Take $f, g, p, q \in R$ where $g \neq 0$ and $q \neq 0$ so that $fq = pg \Rightarrow \phi(f)\phi(q) = \phi(p)\phi(g)$; this implies that

$$\frac{f}{g} = \frac{p}{q} \Rightarrow \frac{\phi(f)}{\phi(g)} = \frac{\phi(p)}{\phi(q)}$$

Thus, for any $f/g \in K$, we can define the mapping $\widetilde{\phi}\colon K \to K$ given by $\widetilde{\phi}(f/g) = \phi(f)/\phi(g)$. Now, consider $a/b, c/d \in K$; then,

$$\widetilde{\phi}\left(\frac{a}{b}\frac{c}{d}\right) = \frac{\phi(ac)}{\phi(bd)} = \frac{\phi(a)}{\phi(b)}\frac{\phi(c)}{\phi(c)} = \widetilde{\phi}\left(\frac{a}{b}\right)\widetilde{\phi}\left(\frac{c}{d}\right)$$

$$\widetilde{\phi}\left(\frac{a}{b} + \frac{c}{d}\right) = \widetilde{\phi}\left(\frac{ad+bc}{bd}\right) = \frac{\phi(ad+bc)}{\phi(bd)} = \frac{\phi(a)}{\phi(b)} + \frac{\phi(c)}{\phi(d)} = \widetilde{\phi}\left(\frac{a}{b}\right) + \widetilde{\phi}\left(\frac{c}{d}\right)$$

Additionally, $\widetilde{\phi}(1) = \phi(1) = 1$ and $\widetilde{\phi}(0) = \phi(0) = 0$, so $\widetilde{\phi}$ is a ring homomorphism. Also, using the fact that ϕ is injective, note that for all $a/b, c/d \in K$

$$\widetilde{\phi}\left(\frac{a}{b}\right) = \widetilde{\phi}\left(\frac{c}{d}\right) \Rightarrow \frac{\phi(a)}{\phi(b)} = \frac{\phi(c)}{\phi(d)} \Rightarrow \phi(ad) = \phi(bc) \Rightarrow ad = bc \Rightarrow \frac{a}{b} = \frac{c}{d}$$

Furthermore, using the fact that ϕ is surjective, we have that for all $p, q \in R$, there are $a, b \in R$ such that $\phi(a) = p$ and $\phi(b) = q$, so

$$\frac{\phi(a)}{\phi(b)} = \frac{p}{q} \Rightarrow \widetilde{\phi}\left(\frac{a}{b}\right) = \frac{p}{q}$$

Thus, for all $p/q \in K$, there is $a/b \in K$ such that

$$\widetilde{\phi}\left(\frac{a}{b}\right) = \frac{p}{q}$$

so it follows that $\widetilde{\phi}$ is bijective and is hence, a field automorphism $\widetilde{\phi}\colon K \simeq K$.

Now, suppose $\psi\colon K \simeq K$ is a field automorphism that extends ϕ. Then, for any $a/b \in K$

$$\psi\left(\frac{a}{b}\right) = \frac{\psi(a)}{\psi(b)} = \frac{\phi(a)}{\phi(b)} = \widetilde{\phi}\left(\frac{a}{b}\right)$$

so it follows that $\psi = \widetilde{\phi}$ and hence, $\widetilde{\phi}$ is unique. ■

Theorem 4.2.6. *Let $K \subset L$ be the universal extension in degree n. The action of the Galois Group $\mathrm{Gal}(L/K)$ of separable polynomial $\widetilde{f} \in K[x]$ on the roots of $\widetilde{f}$ yields the isomorphism*

$$\mathrm{Gal}(L/K) \simeq S_n$$

Proof. Recall that the action of some $\tau \in S_n$ on some $f \in F[x_1, \ldots, x_n]$ is given by

$$\tau f(x_1, \ldots, x_n) = f(x_{\tau(1)}, \ldots, x_{\tau(n)})$$

In this sense, τ can be regarded as the map $\tau\colon F[x_1, \ldots, x_n] \to F[x_1, \ldots, x_n]$ that evaluates $f(x_1, \ldots, x_n)$ at $(x_{\tau(1)}, \ldots, x_{\tau(n)})$. By Theorem A.3.1, we see that τ is an evaluation homomorphism and hence, it is a ring homomorphism. Thus, it follows that for $f, g \in F[x_1, \ldots, x_n]$,

$$\begin{aligned}\tau(f+g) &= \tau(f) + \tau(g)\\ \tau(fg) &= \tau(f)\tau(g)\end{aligned}$$

Also, for $\tau, \gamma \in S_n$, since

$$\tau(\gamma(f)) = \tau(f(x_{\gamma(1)}, \ldots, x_{\gamma(n)})) = f(x_{\tau(\gamma(1)}, \ldots, x_{\tau(\gamma(n)}) = f(x_{\tau\circ\gamma(1)}, \ldots, x_{\tau\circ\gamma(n)})$$

it follows that $\tau(\gamma(f)) = \tau \circ \gamma(f)$. Since S_n is a group, there is a $\tau^{-1} \in S_n$ such that $\tau \circ \tau^{-1} = i_d$ where i_d is the identity mapping, so for any $f \in F[x_1, \ldots, x_n]$,

$$\tau \circ \tau^{-1}(f) = \tau^{-1} \circ \tau(f) = i_d(f) = f$$

and since it is the case that a mapping can have an inverse mapping if and only if the mapping is bijective, it follows that $\tau\colon F[x_1, \ldots, x_n] \to F[x_1, \ldots, x_n]$ is bijective and hence, is a ring isomorphism that is the identity on F since it only permutes the variables. Now, consider $L = F(x_1, \ldots, x_n)$, the corresponding field of rational functions of $l = F[x_1, \ldots, x_n]$. By Lemma 4.2.5, it follows that τ extends to a unique automorphism on L that is the identity on F. Moreover, the elementary symmetric polynomials are fixed under the action of τ, so it follows that τ is the identity on $K = F(\sigma_1, \ldots, \sigma_n)$, so $\tau \in \mathrm{Gal}(L/K)$. Therefore, when considering ρ as in Proposition 4.2.1, we have that $\rho(\tau) = \tau$, so ρ is an identity mapping on τ and since $\tau \in S_n$ was arbitrary, we conclude that ρ is bijective, so $\mathrm{Gal}(L/K) \simeq S_n$ ■

5

Galois Correspondence

This chapter addresses the relationship between Galois groups and field extensions. Section 5.1 establishes properties of what is known as a Galois extension; this section corresponds to chapter 7.1 in [5]. Section 5.2 addresses the relationship between normal subgroups and normal field extensions; this section corresponds to chapter 7.2 in [5]. Finally, section 5.3 collects all of the material from the previous sections designating the entire package as the Fundamental Theorem of Galois Theory which solidifies the relationship between field extensions and Galois Groups and addresses the reverse inclusions that entail from the correspondence; this section corresponds to chapter 7.3 from [5]. Parts of proof left as exercises and additional details were filled in with help from sources like "Github Repository: Cox Galois Theory Exercises" (Ganaye) [6], "A First Course in Abstract Algebra: 7th. Ed." (Fraleigh and Katz 2003) [7] and "Groupprops, the Group Properties Wiki" (Naik) [8].

5.1 Galois Extensions

Definition 5.1.1. *Let $F \subset L$ be a finite extension with Galois group* $\mathrm{Gal}(L/F)$ *where H is a subgroup. Then,*

$$L_H = \{\alpha \in L \mid \forall \sigma \in H,\ \sigma(\alpha) = \alpha\}$$

is called the ***fixed field*** *of H.*

Lemma 5.1.2. *Let L_H be the fixed field as in Definition 5.1.1. Then, $F \subset L_H \subset L$.*

Proof. By definition, $\alpha \in L_H \Rightarrow \alpha \in L$, so $L_H \subset L$. Also, for all $\sigma \in H$ and for all $\alpha \in F$, $\sigma(\alpha) = \alpha$, so $\alpha \in F \Rightarrow \alpha \in L_H$. Thus, $F \subset L_H$. ■

Definition 5.1.3. *If the finite extension $F \subset L$ is normal and separable, then it is called a* ***Galois extension****.*

Theorem 5.1.4. *Let $F \subset L$ be a finite extension. Then, the following are equivalent:* ***(i)*** *L is the splitting field of some separable polynomial $f \in F[x]$.* ***(ii)*** *F is the fixed field of* $\mathrm{Gal}(L/F)$. ***(iii)*** *$F \subset L$ is a Galois extension.*

Proof. Suppose (i) is true and that K is the fixed field of $\mathrm{Gal}(L/F)$. Then, by Lemma 5.1.2, $F \subset K \subset L$. Also, note that for any $\sigma \in \mathrm{Gal}(L/K)$, we have that for all $a \in F$, $\sigma(a) = a$ since $F \subset K$, so $\sigma \in \mathrm{Gal}(L/F)$ and hence, $\mathrm{Gal}(L/K) \subset \mathrm{Gal}(L/F)$. Additionally, by assumption, K is the fixed field of $\mathrm{Gal}(L/F)$, so $\sigma \in \mathrm{Gal}(L/F) \Rightarrow \sigma \in \mathrm{Gal}(L/K)$ by definition; thus, $\mathrm{Gal}(L/F) \subset \mathrm{Gal}(L/K)$ and hence, $\mathrm{Gal}(L/F) = \mathrm{Gal}(L/K)$. Since f is a separable polynomial in $F[x]$, it is also a separable polynomial in $K[x]$, so by Theorem 4.1.5, $|\mathrm{Gal}(L/F)| = [L : F]$ and $|\mathrm{Gal}(L/K)| = [L : K]$. Therefore, $[L : F] = [L : K]$ and by the tower theorem, $[L : F] = [L : K][K : F]$, so $[K : F] = 1$ and it follows by Proposition 2.3.2, that $F = K$, so (ii) follows from (i).

Next, suppose (ii) is true. Take $\alpha \in L$ and suppose that the action of $\mathrm{Gal}(L/F)$ on α yields the distinct elements $\alpha_1, \alpha_2, \ldots, \alpha_m$ in L where $\alpha_1 \coloneqq \alpha$. Define $h \in L[x]$ given by

$$h(x) = \prod_{i=1}^{m} (x - \alpha_i) \tag{5.1}$$

Take $\sigma \in \mathrm{Gal}(L/F)$. Note that $\alpha_i = \tau(\alpha)$ for some $\tau \in \mathrm{Gal}(L/F)$ where $i \in \{1, \ldots, m\}$; then, $\sigma(\alpha_i) = \sigma(\tau(\alpha_i)) = \sigma \circ \tau(\alpha_i) = \alpha_j$ for some $j \in \{1, \ldots, m\}$. Since σ is an automorphism on L, it is a self mapping on $\{\alpha_1, \ldots, \alpha_m\}$ and is hence a permutation of the roots of h. By Eq. (5.1), we can see that permutation of the roots of h amounts to permuting the linear factors $(x - \alpha_i)$ for $i \in \{1, \ldots, n\}$. However, multiplication is commutative on $L[x]$, so permutation of the linear factors of h yields the same polynomial. This means that the coefficients of h are fixed by elements of $\mathrm{Gal}(L/F)$ which allows us to conclude that they lie in F, so $h \in F[x]$.

Now, suppose $g \in F[x]$ is an irreducible factor of h such that $g(\alpha) = 0$. Then, by Proposition 4.1.4, we have that for all $\sigma \in \mathrm{Gal}(L/F)$, $\sigma(\alpha)$ is also a root of g, but by assumption, this means that α_i is a root of g for all $i \in \{1, \ldots, m\}$ which means by Eq. (5.1) that h divides g which means h is a factor of g. However, g is an irreducible factor of h yet h divides g, so it follows that $h = g$ and that h is irreducible. Thus, we may conclude that h is the minimal polynomial of α. Here, we have shown that any irreducible $f \in F[x]$ such that $f(\alpha) = 0$ can be expressed as $f = ch$ for some $c \in F$. Also, for any $\alpha \in L$, we can construct its minimal polynomial with distinct roots as we did for Eq. (5.1). Therefore, we may conclude that every irreducible polynomial in $F[x]$ with a root in L splits completely over L and that for all $\alpha \in L$, the roots of its minimal polynomial are distinct; i.e., $F \subset L$ is normal and separable and hence, is a Galois extension, so (iii) follows from (ii).

Finally, suppose (iii) is true. Since $F \subset L$ is finite, by Theorem 3.1.4, we can set $L = F(\alpha_1, \ldots, \alpha_n)$ where the minimal polynomial p_i for each α_i over F are separable since $F \subset L$ is Galois. Then, let $q_1, \ldots, q_m$, where $m \leq n$, be the distinct minimal polynomials in $\{p_1, \ldots, p_n\}$ and let $f = q_1 \cdots q_m$; here, since each q_i are pairwise distinct and monic (since they are minimal polynomials), they cannot be constant multiples of each other. Furthermore, each q_i has distinct roots in L which is a splitting field of f. Suppose that there is some $\alpha \in L$ such that $q_i(\alpha) = q_j(\alpha)$ for $i \neq j$. Since q_i and q_j are irreducible, it would follow from Proposition 2.2.2 that each would be a constant multiple of each other which yields a contradiction since that is not the case. Therefore the pairwise distinct q_i cannot share a root, so f has distinct roots and is, hence, separable.

Note that f splits completely over L since $F \subset L$ is normal so each q_i splits completely over L. Denote the roots of f as $\beta_1, \ldots, \beta_r$ where $n \leq r$; also, let $L' = F(\beta_1, \ldots, \beta_r) \subset L$ be the splitting field of f. Note that for each $i \in \{1, \ldots, n\}$, $\alpha_i = \beta_k$ for some $k \in \{1, \ldots, r\}$, so $\{\alpha_1, \ldots, \alpha_n\} \subset \{\beta_1, \ldots, \beta_r\}$, and hence, $L = F(\alpha_1, \ldots, \alpha_n) \subset L' = F(\beta_1, \ldots, \beta_r) \subset L$. It follows that $L' = L$ and that L is the splitting field of separable polynomial $f \in F[x]$. Thus, (i) follows from (iii). ■

Corollary 5.1.5. *If $F \subset L$ is a Galois extension and there is an intermediate field K such that $F \subset K \subset L$, then $K \subset L$ is a Galois extension.*

Proof. By Theorem 5.1.4, since $F \subset L$ is a Galois extension, then L is the splitting field of some separable polynomial $f \in F[x]$. Since $F \subset K$, it is valid to say that $f \in K[x]$, so L is the splitting field of some separable polynomial $f \in K[x]$. Thus, again, by Theorem 5.1.4, $K \subset L$ is a Galois extension. ■

Corollary 5.1.6. *$F \subset L$ is a Galois extension if and only if $|\mathrm{Gal}(L/F)| = [L : F]$.*

Proof. Suppose $F \subset L$ is Galois; then by Theorem 5.1.4, L is the splitting field of some separable polynomial $f \in F[x]$. By Theorem 4.1.5, it follows that $|\mathrm{Gal}(L/F)| = [L : F]$.

Conversely, suppose that $|\mathrm{Gal}(L/F)| = [L : F]$; also, suppose that K is the fixed field of $|\mathrm{Gal}(L/F)|$; then, by Lemma 5.1.2, $F \subset K \subset L$. As in the proof of Theorem 5.1.4 in the first paragraph, it follows that $\mathrm{Gal}(L/F) = \mathrm{Gal}(L/K)$, so $|\mathrm{Gal}(L/F)| = |\mathrm{Gal}(L/K)|$. Therefore, we have that K is fixed by $\mathrm{Gal}(L/K)$, so by Theorem 5.1.4, $K \subset L$ is a Galois extension; additionally, since $K \subset L$ is Galois, by Theorem 4.1.5, we have $|\mathrm{Gal}(L/K)| = [L : K]$. It follows that $[L : F] = [L : K]$ but by the tower theorem, $[L : F] = [L : K][K : F]$, so we must have $[K : F] = 1$, and by Proposition 2.3.2, $K = F$. Thus, we conclude that F is fixed by $\mathrm{Gal}(L/F)$, so by Theorem 5.1.4, $F \subset L$ is Galois. ■

Proposition 5.1.7. *If $F \subset L$ is finite and separable, then there is an extension $L \subset M$ such that $F \subset M$ is a Galois extension that is unique up to isomorphism.*

Proof. Since $F \subset L$ is finite and separable, by Theorem 3.1.4, we can set $L = F(\alpha_1, \ldots, \alpha_n)$ where each α_i is separable for $i \in \{1, \ldots, n\}$. For each α_i, let p_i be its minimum polynomial over F and let $q_1, \ldots, q_m$ be the distinct minimum polynomials in $\{p_1, \ldots, p_n\}$. Then, if we define $f = q_1 \cdots q_m$, just as in the proof of Theorem 5.1.4 in the second to last paragraph, it follows that f is separable.

Consider f as a polynomial in $L[x]$, and suppose M is the splitting field of f so that, denoting the roots of f as $\beta_1, \ldots, \beta_r$, $M = L(\beta_1, \ldots, \beta_r)$. Note that

$$F(\beta_1, \ldots, \beta_r) \subset L(\beta_1, \ldots, \beta_r) = M$$

and that since for all $i \in \{1, \ldots, n\}$, $\beta_k = \alpha_i$ for some $k \in \{1, \ldots, r\}$, it follows that

$$L = F(\alpha_1, \ldots, \alpha_n) \subset F(\beta_1, \ldots, \beta_r)$$

Here, we have that $F(\beta_1, \ldots, \beta_r)$ contains both L and $\beta_1, \ldots, \beta_r$, so it follows that $M \subset F(\beta_1, \ldots, \beta_r)$ and hence, by the first inclusion, that $M = F(\beta_1, \ldots, \beta_r)$. Thus, we have that

M is the splitting field over F of separable polynomial f, and by Theorem 5.1.4, it follows that we have an extension $L \subset M$ such that $F \subset M$ is a Galois extension.

Now, suppose we have another extension $L \subset U$ such that $F \subset U$ is Galois. Then, each p_i splits completely over U, so f splits completely over U. Let M' be the subfield of U obtained by appending the roots of f to F. Then, M' is a splitting field of f over L; also, since for all $i \in \{1, \ldots, n\}$, α_i is also a root of f, it follows that $L \subset M'$. According to Theorem 3.2.3 (letting $F_1 = F_2 = F$), it follows there is an isomorphism $\varphi\colon M \to M'$ such that for all $a \in L$, $\varphi(a) = a$. However, recall that $f \in F[x]$, so by Theorem 5.1.4, since M' is the splitting field of some separable polynomial $f \in F[x]$, it follows that $M' \subset F$ is Galois, so both M and M' are Galois and isomorphic. ■

Definition 5.1.8. *The Galois extension $F \subset M$ described in Proposition 5.1.7 is called the* ***Galois Closure*** *of the finite and separable extension $F \subset L$.*

The Galois closure $F \subset M$ of a finite separable extension $F \subset L$ can be regarded as the smallest field extension of L that is Galois over F.

5.2 Normality

Definition 5.2.1. *Let $F \subset K \subset L$ be a sequence of finite extensions. Then, taking some $\sigma \in \mathrm{Gal}(L/F)$,*

$$\sigma K = \{\sigma(\alpha) \mid \alpha \in K\}$$

is called the ***conjugate field of*** *K.*

Lemma 5.2.2. *For the sequence of finite extensions $F \subset K \subset L$, take $\sigma \in \mathrm{Gal}(L/F)$. Then, $F \subset \sigma K \subset L$ and $[K : F] = [\sigma K : F]$.*

Proof. Since $F \subset K$ and σ is the identity on F, we may conclude that $F \subset \sigma K$; also, since σ is a field automorphism on L, it is a field isomorphism when restricted to K, so we can also conclude that $\sigma K \subset L$, so $F \subset \sigma K \subset L$.

Furthermore, by definition, we may regard both K and σK as vector spaces over F. In this sense, the restriction of σ to K can be thought of as an isomorphism of vector spaces; i.e. $\sigma|_K\colon K \to \sigma K$. It follows that $\dim_F(K) = \dim_F(\sigma K)$ and hence that, $[K : F] = [\sigma K : F]$. ■

Lemma 5.2.3. *If $F \subset K \subset L$ is a sequence of finite extensions, then $\mathrm{Gal}(L/K)$ is a subgroup of $\mathrm{Gal}(L/F)$; also, if we take $\sigma \in \mathrm{Gal}(L/F)$, then $\mathrm{Gal}(L/\sigma K) = \sigma\mathrm{Gal}(L/K)\sigma^{-1}$.*

Proof. Suppose $\sigma \in \mathrm{Gal}(L/K)$; then for all $a \in K$, $\sigma(a) = a$, but since $F \subset K$, we also have that for all $a \in F$, $\sigma(a) = a$ so $\sigma \in \mathrm{Gal}(L/F)$ and thus, $\mathrm{Gal}(L/K) \subset \mathrm{Gal}(L/F)$.

Now, take $\beta \in \sigma K$; then for some $\alpha \in K$, $\beta = \sigma(\alpha)$. Consider $\sigma\tau\sigma^{-1} \in \sigma\mathrm{Gal}(L/K)\sigma^{-1}$ where $\tau \in \mathrm{Gal}(L/K)$. Then, $\tau(\alpha) = \alpha$, so

$$(\sigma\tau\sigma^{-1})(\beta) = \sigma\tau\sigma^{-1}(\sigma(\alpha)) = \sigma\tau(\alpha) = \sigma(\alpha) = \beta$$

Thus, $\sigma\tau\sigma^{-1}(\beta) = \beta$, so $\sigma\tau\sigma^{-1} \in \text{Gal}(L/\sigma K)$ and hence, $\sigma\text{Gal}(L/K)\sigma^{-1} \subset \text{Gal}(L/\sigma K)$. Conversely, take $\rho \in \text{Gal}(L/\sigma K)$ and consider $\sigma^{-1}\rho\sigma \in \text{Gal}(L/F)$; note that for $\beta \in \sigma K$, $\beta = \sigma(\alpha)$ for some $\alpha \in K$. Thus, since $\rho(\beta) = \beta$,

$$(\sigma^{-1}\rho\sigma)(\alpha) = \sigma^{-1}(\rho(\sigma(\alpha))) = \sigma^{-1}(\sigma(\alpha)) = \alpha$$

so $(\sigma^{-1}\rho\sigma)(\alpha) = \alpha$ and $\sigma^{-1}\rho\sigma \in \text{Gal}(L/K)$ hence, $\rho \in \sigma\text{Gal}(L/K)\sigma^{-1}$. Therfore, it follows that $\sigma\text{Gal}(L/K)\sigma^{-1} = \text{Gal}(L/\sigma K)$. ■

Theorem 5.2.4. *Suppose we have a sequence of fields $F \subset K \subset L$ where $F \subset L$ is a Galois extension. Then, the following statements are equivalent:*
(a) For all $\sigma \in \text{Gal}(L/F)$, $\sigma K = K$.
(b) $\text{Gal}(L/K) \triangleleft \text{Gal}(L/F)$.
(c) $F \subset K$ is a Galois extension.
(d) $F \subset K$ is a normal extension.

Proof. To show that (a) and (b) are equivalent, first, suppose that for all $\sigma \in \text{Gal}(L/F)$, $K = \sigma K$. Then, by Lemma 5.2.3, $\text{Gal}(L/K) = \text{Gal}(L/\sigma K) = \sigma\text{Gal}(L/K)\sigma^{-1}$, so $\text{Gal}(L/K)$ is closed under conjugation by $\text{Gal}(L/F)$, so $\text{Gal}(L/K) \triangleleft \text{Gal}(L/F)$. Conversely, suppose $\text{Gal}(L/K) \triangleleft \text{Gal}(L/F)$; then, according to Lemma 5.2.3, for any $\sigma \in \text{Gal}(L/F)$, $\text{Gal}(L/K) = \sigma\text{Gal}(L/K)\sigma^{-1} = \text{Gal}(L/\sigma K)$. Since $\text{Gal}(L/\sigma K)$ is the identity on σK and $\text{Gal}(L/K)$ is the identity on K, by Corollary 5.1.5, it follows that $K \subset L$ and $\sigma K \subset L$ are Galois extensions. Now, we have by Theorem 5.1.4 that K is the fixed field of $\text{Gal}(L/K)$ and σK is the fixed field of $\text{Gal}(L/\sigma K)$. However, $\text{Gal}(L/\sigma K) = \text{Gal}(L/K)$, so it follows that $K = \sigma K$.

Demonstrating the equivalence of (c) and (d), first suppose that $F \subset K$ is Galois. Then, by definition, $F \subset K$ is normal. Now, conversely, suppose that $F \subset K$ is normal. Since $F \subset L$ is Galois, it follows that it is also separable over F; then, since any $\alpha \in K$ implies that $\alpha \in L$ where all $\alpha \in L$ is separable over F, it follows that K is separable over F and thus, $F \subset K$ is Galois.

Finally, we assert the equivalence of (d) and (a). Suppose that for all $\sigma \in \text{Gal}(L/F)$, $K = \sigma K$ and let $f \in F[x]$ be irreducible over F with a root $\alpha \in K$. Recall from the proof of Theorem 5.1.4, that any irreducible $f \in F[x]$ can be expressed as $f = ch$ for some $c \in F$ where h is given by

$$h(x) = \prod_{i=1}^{m}(x - \alpha_i)$$

from Eq. (5.1) whenever F is the fixed field of $\text{Gal}(L/F)$; here, for $i \in \{1, \ldots, m\}$ the α_i are the distinct roots acquired from applying the autmorphisms in $\text{Gal}(L/F)$ to α. Since $\alpha \in K$, it follows that each α_i will lie in a conjugate field of K, but $K = \sigma K$ for all $\sigma \in \text{Gal}(L/F)$, so it follows that all α_i are in K. Thus, h and hence, f splits completely over K, so we have that any irreducible polynomial over F splits completely over K. Therefore, $F \subset K$ is normal. Conversely, suppose $F \subset K$ is normal. Take $\alpha \in K$ and $\sigma \in \text{Gal}(L/F)$. Let p be the minimal polynomial of α over F. Then, by Proposition 4.1.4, it follows that $\sigma(\alpha)$ is also a root of p and since $F \subset K$ is normal, $\sigma(\alpha) \in K$ because p splits completely over K. Thus, $\sigma K \subset K$. Furthermore, by Lemma 5.2.2, $[K : F] = [\sigma K : F]$, so $\sigma K = K$. ■

Theorem 5.2.5. *Suppose we have a sequence of finite extensions where $F \subset K$ and $F \subset L$ are Galois. Then,*

$$\text{(i) } \mathrm{Gal}(L/K) \triangleleft \mathrm{Gal}(L/F)$$
$$\text{(ii) } \mathrm{Gal}(L/F)/\mathrm{Gal}(L/K) \simeq \mathrm{Gal}(K/F)$$

Proof. For (i), since $F \subset K$ is Galois, it follows by Theorem 5.2.4 that $\mathrm{Gal}(L/K) \triangleleft \mathrm{Gal}(L/F)$.

Next, for (ii), take $\sigma \in \mathrm{Gal}(L/F)$; then, when we restrict σ to K, we get an isomorphism $\sigma|_K \colon K \simeq \sigma K$, but by Theorem 5.2.4, $\sigma K = K$, so $\sigma|_K$ is an automorphism on K. Also, since for all $a \in F$, $\sigma(a) = a$ and $F \subset K$, this also holds true for $\sigma|_K$ which means we are able to conclude $\sigma|_K \in \mathrm{Gal}(K/F)$.Define the map

$$\Phi \colon \mathrm{Gal}(L/F) \to \mathrm{Gal}(L/K)$$

given by $\Phi(\sigma) = \sigma|_K$. Then, for $\sigma_1, \sigma_2 \in \mathrm{Gal}(L/F)$, $\Phi(\sigma_1) = \sigma_1|_K$ and $\Phi(\sigma_2) = \sigma_2|_K$. Now, consider $\sigma_1 \circ \sigma_2 \in \mathrm{Gal}(L/F)$, so that $\Phi(\sigma_1 \circ \sigma_2) = (\sigma_1 \circ \sigma_2)|_K$. Note that for all $a \in K$, $\sigma_1(a) \in K$ and $\sigma_2(a) \in K$; then, $\sigma_1 \circ \sigma_2(a) = \sigma_1(\sigma_2(a))$ and since $\sigma_2(a) \in K$, we have $\sigma_1(\sigma_2(a)) \in K$. Thus, for all $a \in K$

$$(\sigma_1 \circ \sigma_2)|_K(a) = (\sigma_1 \circ \sigma_2)(a) = \sigma_1(\sigma_2(a)) = \sigma_1|_K(\sigma_2|_K(a)) = (\sigma_1|_K \circ \sigma_2|_K)(a)$$

and therefore, $\sigma_1 \circ \sigma_2 = \sigma_1|_K \circ \sigma_2|_K$, so $\Phi(\sigma_1 \circ \sigma_2) = \Phi(\sigma_1) \circ \Phi(\sigma_2)$ and hence, Φ is a group homomorphism. Furthermore, denote the identity mapping in $\mathrm{Gal}(L/F)$ by i_d; then,

$$\ker(\Phi) = \{\sigma \in \mathrm{Gal}(L/F) \mid \Phi(\sigma) = i_d|_K\}$$

Notice that this set consists of the set of automorphisms in $\mathrm{Gal}(L/F)$ that is the identity on K which is precisely $\mathrm{Gal}(L/K)$. By the fundamental theorem of group homomorphisms, it follows that $\mathrm{im}(\Phi) \subset \mathrm{Gal}(K/F)$ where $\mathrm{im}(\Phi) \simeq \mathrm{Gal}(L/F)/\mathrm{Gal}(L/K)$. However, using Corollary 5.1.6 and the tower theorem, we find

$$|\mathrm{im}(\Phi)| = \frac{|\mathrm{Gal}(L/F)|}{|\mathrm{Gal}(L/K)|} = \frac{[L:F]}{[L:K]} = [K:F] = |\mathrm{Gal}(K/F)|$$

so it follows that $\mathrm{im}(\Phi) = \mathrm{Gal}(K/F)$. Thus,

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

■

5.3 The Fundamental Theorem of Galois Theory

So far, we have most of the framework for the Fundamental Theorem of Galois Theory. Here, we will collect the results we acquired and state the theorem explicitly. For the following, let $F \subset L$ be a Galois extension.

Theorem 5.3.1. *The Galois group* $\mathrm{Gal}(L/K)$ *of an intermediate field extension* $F \subset K \subset L$ *has* K *as its fixed field. Also,* $\{\mathrm{Gal}(L/F) : \mathrm{Gal}(L/K)\} = [K : F]$ *and* $|\mathrm{Gal}(L/K)| = [L : K]$.

Proof. Since $F \subset L$ is Galois, it follows by Corollary 5.1.5 that $K \subset L$ is Galois, so by Theorem 5.1.4, it follows that K is the fixed field of $\mathrm{Gal}(L/K)$.

Now, we have that both $F \subset L$ and $K \subset L$ are Galois, so by Corollary 5.1.6 that $[L : F] = |\mathrm{Gal}(L/F)|$ and $[L : K] = |\mathrm{Gal}(L/K)|$. Since

$$\{\mathrm{Gal}(L/F) : \mathrm{Gal}(L/K)\} = |\mathrm{Gal}(L/F)|/|\mathrm{Gal}(L/K)|$$

it follows by the tower theorem that $\{\mathrm{Gal}(L/F) : \mathrm{Gal}(L/K)\} = \dfrac{[L : F]}{[L : K]} = [K : F]$. ■

Theorem 5.3.2. *Suppose that the fixed field of* $H \subset \mathrm{Gal}(L/F)$ *is* L_H *where* $F \subset L_H \subset L$. *Then the Galois group of* $L_H \subset L$ *is* $\mathrm{Gal}(L/L_H) = H$. *Also,* $[L : L_H] = |H|$ *and* $[L_H : F] = \{\mathrm{Gal}(L/F) : H\}$.

Proof. By Corollary 5.1.5, $L_H \subset L$ is Galois since $F \subset L$ is Galois which implies by Corollary 5.1.6 that $|\mathrm{Gal}(L/L_H)| = [L : L_H]$. For every $\sigma \in H$, by definition of fixed fields, we have that for all $a \in L_H$, $\sigma(a) = a$. This implies that $H \subset \mathrm{Gal}(L/L_H)$. By the theorem of the primitive element, there is some $\alpha \in L$ such that $L = L_H(\alpha)$. Now, the action of $\sigma \in H$ on α will yield $\sigma(\alpha) \in L$. Define

$$h(x) = \prod_{\sigma \in H} (x - \sigma(\alpha))$$

just as in the proof of Theorem 5.1.4 except every $\sigma(\alpha)$ is not necessarily pairwise distinct. Recall that the action of any permutation on such a polynomial leaves it invariant as the resulting polynomial consists of the same linear factors, so $h \in L_H[x]$ where $h(\alpha) = 0$. Let p be the the minimal polynomial of α. Then, by Proposition 2.2.2 and Theorem 2.3.4, it follows that p divides h and $\deg(p) = [L : L_H]$. Since p divides h, we have $\deg(h) \geq \deg(p)$. Notice that $\deg(h) = |H|$ and that $|H| \leq |\mathrm{Gal}(L/L_H)|$ since $H \subset \mathrm{Gal}(L/L_H)$, so we have

$$[L : L_H] \leq |H| \leq |\mathrm{Gal}(L/L_H)|$$

but this means that $|H| = |\mathrm{Gal}(L/L_H)|$ since $[L : L_H] = |\mathrm{Gal}(L/L_H)|$, so $H = \mathrm{Gal}(L/L_H)$. Furthermore, by Theorem 5.3.1, it follows that $[L_H : F] = \{\mathrm{Gal}(L/F) : H\}$ since L_H is an intermediate field of the extensions $F \subset L_H \subset L$. ■

Theorem 5.3.3. *For the sequence of extensions, $F \subset K \subset L$, define the map*

$$\phi\colon K \mapsto \mathrm{Gal}(L/K)$$

and for the subgroup $H \subset \mathrm{Gal}(L/F)$, define the map

$$\rho\colon H \mapsto L_H$$

Then, $\rho = \phi^{-1}$ and $\phi = \rho^{-1}$. Additionally, if $K_1 \subset K_2 \subset L$, then $\mathrm{Gal}(L/K_2) \subset \mathrm{Gal}(L/K_1)$ and if $\mathrm{Gal}(L/L) \subset H_1 \subset H_2$, then $L_{H_2} \subset L_{H_1} \subset L$. Furthermore, if for subfield K where $F \subset K \subset L$, $\phi(K) = H$ and $\rho(H) = K$, it follows that $H \triangleleft \mathrm{Gal}(L/F)$ if and only if $F \subset K$ is Galois and

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

Proof. Suppose $F \subset K \subset L$. Since $F \subset L$ is Galois, it follows by Theorem 5.1.4 that $\mathrm{Gal}(L/K) \subset \mathrm{Gal}(L/F)$. Then, $\phi(K) = \mathrm{Gal}(L/K)$ and by Theorem 5.3.1 $\rho(\mathrm{Gal}(L/K)) = L_{\mathrm{Gal}(L/K)} = K$, so $\rho(\phi(K)) = K$. Also, suppose $H \subset \mathrm{Gal}(L/F)$, so by Lemma 5.1.2, $F \subset L_H \subset L$. Then, $\rho(H) = L_H$ and by Theorem 5.3.2, $\phi(L_H) = \mathrm{Gal}(L/L_H) = H$, so $\phi(\rho(H)) = H$. Thus, we conclude that ϕ and ρ are inverse mappings of each other; i.e. $\rho = \phi^{-1}$ and $\phi = \rho^{-1}$.

Now, suppose that $K_1 \subset K_2 \subset L$; by Theorem 5.2.5, it follows that $\mathrm{Gal}(L/K_2) \subset \mathrm{Gal}(L/K_1)$ which means that $\phi(K_2) \subset \phi(K_1)$. Also, suppose $\mathrm{Gal}(L/L) \subset H_1 \subset H_2$; then, also by Theorem 5.2.5, it follows $L_{H_2} \subset L_{H_1} \subset L$ which means that $\rho(H_2) \subset \rho(H_1)$. Therefore, both ϕ and ρ reverse inclusions. Furthermore, the rest of the theorem follows directly from Theorem 5.3.1, Theorem 5.3.2, Theorem 5.2.4, and Theorem 5.2.5. ■

We will refer to the above theorems collectively as the **Galois correspondence**.

6

Solvability

Here, we establish the idea of a solvable group (section 6.1; corresponds to chapter 8.1 in [5]) and a solvable extension (section 6.2; corresponds to chapter 8.2 in [5]). Section 6.3 contains Galois's solvability theorem which is the climax of Part I (also corresponds to chapter 8.2 of [5]). It is precisely this solvability theorem that gives us insight into the solvability of the cubic and quartic and the unsolvability of the quintic. Parts of proof left as exercises and additional details were filled in with help from sources like "Github Repository: Cox Galois Theory Exercises" (Ganaye) [6], "A First Course in Abstract Algebra: 7th. Ed." (Fraleigh and Katz 2003) [7] and "Groupprops, the Group Properties Wiki" (Naik) [8].

6.1 Solvable Groups

Definition 6.1.1. *A finite group G is said to be a* ***solvable group*** *if there are subgroups*

$$G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0$$

where $G_n = \{e\}$ ("e" denotes the group identity element), $G_0 = G$, and G_i for $i \in \{1, \ldots, n\}$ satisfies the following properties: (i) $G_i \triangleleft G_{i-1}$ *and* (ii) G_{i-1}/G_i *is a cyclic group of prime order (or equivalently $\{G_{i-1} : G_i\} = p$ where p is prime since finite groups of prime order are cyclic).*

Proposition 6.1.2. *If H is a subgroup of a finite solvable group G, then H is solvable.*

Proof. Let G have the sequence of subgroups as described in Definition 6.1.1 and define $H_i = G_i \cap H$ for $i \in \{0, 1, \ldots, n\}$. Also, define the mapping

$$\pi \colon H_{i-1} \to G_{i-1}/G_i$$

given by $\pi(h) = hG_i$ for $h \in H_{i-1}$; here, π is a group homomorphism by the definition of quotient groups where for $h_1, h_2 \in H_{i-1}$, $\pi(h_1h_2) = (h_1h_2)G_i = (h_1G_i)(h_2G_i) = \pi(h_1)\pi(h_2)$. Also, notice that

$$\ker(\pi) = \{h \in H_{i-1} \mid \pi(h) = G_i\}$$

so $h \in \ker(\pi)$ if and only if $hG_i = G_i$ which is the case if and only if $h \in G_i$ and hence if and only if $h \in G_i \cap H_{i-1}$; since $G_i \subset G_{i-1}$, it follows

$$G_i \cap H_{i-1} = G_i \cap (G_{i-1} \cap H) = (G_i \cap G_{i-1}) \cap (G_i \cap H) = G_i \cap (G_i \cap H) = G_i \cap H = H_i$$

Therefore, $h \in G_i \cap H_{i-1}$ if and only if $h \in H_i$, so we can conclude that $\ker(\pi) = H_i$. Additionally, $\ker(\pi) = H_i \triangleleft H_{i-1}$, so by the fundamental theorem of group homomorphisms,

$$H_{i-1}/H_i \simeq \text{im}(\pi) \subset G_{i-1}/G_i$$

Since G_{i-1}/G_i is cyclic, H_{i-1}/H_i must also be cyclic, but G_{i-1}/G_i is also of prime order so H_{i-1}/H_i must be trivial or must be of the same order as G_{i-1}/G_i and hence directly isomorphic to G_{i-1}/G_i; i.e. $H_i = H_{i-1}$ or H_{i-1}/H_i is a cyclic group of prime order. WLOG, assuming it is not the case for $i \in \{1, \ldots, n\}$ that $H_i = H_{i-1}$, we have that

$$H_n \subset \ldots \subset H_i \subset H_{i-1} \subset \ldots \subset H_0$$

where $H_n = H \cap G_n = H \cap \{e\} = \{e\}$, $H_0 = H \cap G_0 = H \cap G = H$, and the sequence of subgroups satisfies the conditions of Definition 6.1.1, so H is solvable. ■

Remark. *If it were the case that $H_i = H_{i-1}$ for some $i \in \{1, \ldots, n\}$, then we would just have a smaller sequence of subgroups still satisfying the same properties hence, solvability of H would still follow.*

Lemma 6.1.3. *Suppose $f \colon A \to B$ is a well-defined surjective group homomorphism such that $|A| = p$ where p is a prime number. Then $|B| = p$ or $|B| = 1$.*

Proof. Since f is well-defined, $\text{im}(f) \subset B$. Also, suppose $b \in B$; since f is surjective, there is an $a \in A$ such that $f(a) = b$, but by definition, $f(a) \in \text{im}(f)$, so $b \in \text{im}(f)$ and hence, $B = \text{im}(f)$. It follows by the fundamental theorem of group homomorphisms that since f is a group homomorphism,

$$A/\ker(f) \simeq B$$

However, since $\ker(f)$ is a subgroup of A, we must have that $|\ker(f)|$ divides $|A|$ according to Lagrange's theorem, but since $|A| = p$ where p is prime, it is either the case that $|\ker(f)| = p$ or $|\ker(f)| = 1$. Thus, $\{A : \ker(f)\} = p$ or $\{A : \ker(f)\} = 1$ and hence, we conclude that $|B| = p$ or $|B| = 1$ ■

Theorem 6.1.4. *Suppose G is a finite group where $H \triangleleft G$. Then, G is solvable if and only if both H and G/H are solvable.*

Proof. Suppose G is solvable with a sequence of subgroups as in Definition 6.1.1; then, by Proposition 6.1.2, H is solvable. Furthermore, to show that G/H is solvable, define the group homomorphism $\pi\colon G \to G/H$ given by $\pi(g) = gH$ which is clearly surjective since for all $gH \in G/H$, there is a $g \in G$ such that $\pi(g) = gH$; also, it is well-defined since for $g_1, g_2 \in G$, if $g_1 = g_2$, then $g_1H = g_2H$ hence, $\pi(g_1) = \pi(g_2)$. Let $\widetilde{G}_i \coloneqq \pi(G_i) = \mathrm{im}(\pi|_{G_i})$ for $i \in \{0, 1, \ldots, n\}$ where $G_n = \{e\}$ and $G_0 = G$. By the fundamental theorem of group homomorphisms, $G/\ker(\pi) \simeq \mathrm{im}(\pi) \subset G/H$ where

$$\ker(\pi) = \{g \in G \mid gH = H\}$$

so $g \in \ker(\pi) \iff gH = H \iff g \in H$; thus, $\ker(\pi) = H$, so $G/H \simeq \mathrm{im}(\pi) \subset G/H$. Since $\widetilde{G}_0 = \pi(G) = \mathrm{im}(\pi)$, it follows $\widetilde{G}_0 = G/H$; moreover, since $G_n = \{e\}$, then $\widetilde{G}_n = \pi(\{e\}) = \{eH\} = \{H\}$, so $\widetilde{G}_n = \{H\}$ (keep in mind that H is the identity of G/H).

Now, we must show that $\widetilde{G}_i \triangleleft \widetilde{G}_{i-1}$ for $i \in \{1, \ldots, n\}$. By definition, we have that $\widetilde{G}_i = \{\pi(g_i) \mid g_i \in G_i\}$ and $\widetilde{G}_{i-1} = \{\pi(g_{i-1}) \mid g_{i-1} \in G_{i-1}\}$ where also by definition, $\pi(g_i) = g_iH$ and $\pi(g_{i-1}) = g_{i-1}H$. Note that since $G_i \subset G_{i-1}$, $g_i \in G_i \Rightarrow g_i \in G_{i-1}$, so $g_iH \in \widetilde{G}_i \Rightarrow g_iH \in \widetilde{G}_{i-1}$ hence, $\widetilde{G}_i \subset \widetilde{G}_{i-1}$. Furthermore, take $\pi(g_i) \in \widetilde{G}_i$ and $\pi(g_{i-1}) \in \widetilde{G}_{i-1}$ and consider $\pi(g_{i-1})\pi(g_i)\pi^{-1}(g_{i-1}) \in \widetilde{G}_{i-1}$. It follows that

$$\pi(g_{i-1})\pi(g_i)\pi^{-1}(g_{i-1}) = (g_{i-1}H)(g_iH)(g_{i-1}H)^{-1} = (g_{i-1}g_iH)(g_{i-1}^{-1}H) = (g_{i-1}g_ig_{i-1}^{-1})H$$

By assumption, $G_i \triangleleft G_{i-1}$, so $(g_{i-1}g_ig_{i-1}^{-1}) \in G_i$ which means that $(g_{i-1}g_ig_{i-1}^{-1})H \in \widetilde{G_i}$, so $\pi(g_{i-1})\pi(g_i)\pi^{-1}(g_{i-1}) \in \widetilde{G}_i$. Therefore, for any $\pi(g_i) \in \widetilde{G}_i$ we have $\pi(g_{i-1})\pi(g_i)\pi^{-1}(g_{i-1}) \in \widetilde{G}_i$ for any $\pi(g_{i-1}) \in \widetilde{G}_{i-1}$ hence, $\widetilde{G}_i \triangleleft \widetilde{G}_{i-1}$.

Next, consider the map

$$\Phi\colon G_{i-1}/G_i \to \widetilde{G}_{i-1}/\widetilde{G}_i$$

given by $\Phi(gG_i) = \pi(g)\widetilde{G}_i$ where $g \in G_{i-1}$. Observe that for $(g_1G_i), (g_2G_i) \in G_{i-1}/G_i$ where $g_1, g_2 \in G_{i-1}$, the mapping Φ is well-defined since, keeping in mind that π is well-defined where $g_1 = g_2 \Rightarrow \pi(g_1) = \pi(g_2)$, we have

$$g_1G_i = g_2G_i \Rightarrow \pi(g_1)\widetilde{G}_i = \pi(g_2)\widetilde{G}_i \Rightarrow \Phi(g_1G_i) = \Phi(g_2G_i)$$

Moreover, Φ is a group homomorphism since

$$\Phi((g_1G_i)(g_2G_i)) = \Phi((g_1g_2G_i)) = \pi(g_1g_2)\widetilde{G}_i = (\pi(g_1)\pi(g_2))\widetilde{G}_i = (\pi(g_1)\widetilde{G}_i)(\pi(g_2)\widetilde{G}_i)$$

so $\Phi((g_1G_i)(g_2G_i)) = \Phi(g_1G_i)\Phi(g_2G_i)$. Additionally, Φ is surjective since for any $\pi(g)\widetilde{G}_i \in \widetilde{G}_{i-1}/\widetilde{G}_i$ where $g \in G_{i-1}$, there is $gG_i \in G_{i-1}/G_i$ such that $\Phi(gG_i) = \pi(g)\widetilde{G}_i$. Thus, by Lemma 6.1.3, since by assumption, for some prime number p, $\{G_{i-1} : G_i\} = p$, it follows

that either $\{\widetilde{G}_{i-1} : \widetilde{G}_i\} = p$ or $\{\widetilde{G}_{i-1} : \widetilde{G}_i\} = 1$. As in the proof of Proposition 6.1.2, WLOG, we will assume it is not the case that $\{\widetilde{G}_{i-1} : \widetilde{G}_i\} = 1$, so that for $i \in \{0, 1, \ldots, n\}$,

$$\{H\} = \widetilde{G}_n \triangleleft \ldots \triangleleft \widetilde{G}_i \triangleleft \widetilde{G}_{i-1} \ldots \triangleleft \widetilde{G}_0 = G/H$$

where $\{\widetilde{G}_{i-1} : \widetilde{G}_i\}$ is prime. Thus, G/H is solvable.

Conversely, suppose H and G/H are solvable where we have the sequence of subgroups

$$\{e\} = H_l \triangleleft \ldots \triangleleft H_i \triangleleft H_{i-1} \triangleleft \ldots H_0 = H$$
$$\{H\} = \widetilde{G}_m \triangleleft \ldots \triangleleft \widetilde{G}_j \triangleleft \widetilde{G}_{j-1} \triangleleft \ldots \widetilde{G}_0 = G/H$$

with each sequence defined as in Definition 6.1.1. Let $\pi \colon G \to G/H$ be defined same as in the first half of the proof and let its pre-image for some $K \subset G/H$, denoted by $\pi^{-1}(K)$, be the set $\pi^{-1}(K) = \{g \in G \mid \pi(g) \in K\}$. Note that by definition, $g \in \pi^{-1}(K) \Rightarrow g \in G$, so $\pi^{-1}(K) \subset G$; also, if $K = \{H\}$ then $\pi^{-1}(K) = \pi^{-1}(\{H\}) = \{g \in G \mid \pi(g) \in \{H\}\}$, so $g \in \pi^{-1}(\{H\}) \iff \pi(g) \in \{H\} \iff gH = H \iff g \in H$ hence, $\pi^{-1}(\{H\}) = H$.

Now, consider $\pi^{-1}(G/H) = \{g \in G \mid \pi(g) \in G/H\}$; in this case, $g \in \pi^{-1}(G/K) \iff gH \in G/H \iff g \in G$, so it follows that $\pi^{-1}(G/H) = G$. Also, if $\{H\} \subset K$, then $\pi(g) \in \{H\} \Rightarrow \pi(g) \in K$, so $g \in \pi^{-1}(\{H\}) \Rightarrow g \in \pi^{-1}(K)$. It follows that we have $\{g \in G \mid \pi(g) \in \{H\}\} \subset \{g \in G \mid \pi(g) \in K\}$, so

$$H \subset \pi^{-1}(K) \subset G$$

From this, we can conclude that the direction of containment in the image is the same as the direction of containment of the pre-image, so according to our sequence of subgroups for G/H, it follows

$$H = \pi^{-1}(\widetilde{G}_m) \subset \ldots \pi^{-1}(\widetilde{G}_j) \subset \pi^{-1}(\widetilde{G}_{j-1}) \subset \ldots \pi^{-1}(\widetilde{G}_0) = G$$

Furthermore, H has the sequence of normal subgroups from our assumption of its solvability, so if we define

$$G_i = \begin{cases} \pi^{-1}(\widetilde{G}_i), & 0 \leq i \leq m \\ H_{i-m}, & m \leq i \leq l+m \end{cases}$$

then, we have

$$\{e\} = G_{l+m} \triangleleft \ldots \triangleleft G_m \subset \ldots \subset G_0 = G$$

Notice that for $0 \leq i \leq m$, since $\widetilde{G}_i \triangleleft \widetilde{G}_{i-1}$, taking $g_i \in G_i$ and $g_{i-1} \in G_{i-1}$ so that $\pi(g_i) \in \widetilde{G}_i$ and $\pi(g_{i-1}) \in \widetilde{G}_{i-1}$, it follows

$$\pi(g_{i-1})\pi(g_i)\pi^{-1}(g_{i-1}) \in \widetilde{G}_i \Rightarrow \pi(g_{i-1}g_ig_{i-1}^{-1}) \in \widetilde{G}_i \Rightarrow g_{i-1}g_ig_{i-1}^{-1} \in \pi^{-1}(\widetilde{G}_i) = G_i$$

Therefore, for any $g_{i-1} \in G_{i-1}$, it follows that $g_{i-1}g_ig_{i-1}^{-1} \in G_i$ for any $g_i \in G_i$, so $G_i \triangleleft G_{i-1}$. It follows the G_i form a sequence of normal subgroups of G for $0 \leq i \leq l+m$. Now, consider Φ from the first half of the proof; we have shown that it is a well-defined surjective homomorphism. However, observe that

$$\ker(\Phi) = \{gG_i \in G_i/G_{i-1} \mid \Phi(gG_i) = \widetilde{G}_i\}$$

Thus, it follows $gG_i \in \ker(\Phi) \iff \Phi(gG_i) = \widetilde{G}_i \iff \pi(g)\widetilde{G}_i = \widetilde{G}_i \iff \pi(g) \in \widetilde{G}_i \iff g \in \pi^{-1}(\widetilde{G}_i) \iff gG_i = G_i$ hence, $\ker(\Phi) = \{G_i\}$ which is the identity on G_{i-1}/G_i, so we conclude that Φ is injective. Therefore, $G_{i-1}/G_i \simeq \widetilde{G}_{i-1}/\widetilde{G}_i$ and since $\{\widetilde{G}_{i-1} : \widetilde{G}_i\}$ is prime, it follows that $\{G_{i-1} : G_i\}$ is also prime, so according to Definition 6.1.1, we conclude that G is solvable. ■

Proposition 6.1.5. *If finite group G is Abelian, then G is solvable.*

Proof. Here, we will use the principle of strong induction where $P(n)$ for $n \in \mathbb{N}$ is the statement "G is solvable where $|G| = n$". Suppose $n = 1$; then $|G| = 1$ so $G = \{e\}$ which is trivially solvable, so $P(1)$ is true.

Now, suppose that $n > 1$ and that $P(k)$ is true for all $k < n$. Let p be a prime divisor of $|G|$; note that if $|G| = p$, then by Lagrange's theorem, G has no nontrivial subgroups, so

$$\{e\} \triangleleft G$$

where $\{G : \{e\}\} = p$ hence, G is solvable. Therefore, suppose that $p < n$ where $|G| = n$. By Cauchy's theorem, it follows that there is a $g \in G$ such that $|\langle g \rangle| = p$; let $H = \langle g \rangle$, so that $H \subset G$. Take $h \in H$ and consider $ghg^{-1} \in G$. Since G is Abelian,

$$ghg^{-1} = gg^{-1}h = eh = h \in H$$

so $ghg^{-1} \in H$. Thus, $H \triangleleft G$. Also, note that since p is prime, $1 < p \Rightarrow 1/p < 1 \Rightarrow n/p < n$ where $|G/H| = n/p$. Therefore we have that $|H| < n$ and $|G/H| < n$, so by assumption, it follows that H and G/H are solvable. Furthermore, by Theorem 6.1.4, this implies that G is solvable, so the fact that $P(n)$ is true for some $n \in N$ follows from the assumption that $P(k)$ is true for all $k < n$. We conclude that $P(n)$ is true for all $n \in \mathbb{N}$. ■

Proposition 6.1.6. *Suppose that G is a finite group with a sequence of normal subgroups where*

$$\{e\} = G_n \triangleleft G_{n-1} \triangleleft \ldots \triangleleft G_1 \triangleleft G_0 = G$$

If for $i \in \{1, \ldots, n\}$, G_{i-1}/G_i is Abelian, then G is solvable.

Proof. By Proposition 6.1.5, since G_{i-1}/G_i is finite and Abelian, it follows that it is solvable, so we have the sequence of normal subgroups

$$\{G_i\} = \widetilde{G}_{i,m} \triangleleft \widetilde{G}_{i,m-1} \triangleleft \ldots \triangleleft \widetilde{G}_{i,1} \triangleleft \widetilde{G}_{i,0} = G_{i-1}/G_i$$

where $\{\widetilde{G}_{i,j-1} : \widetilde{G}_{i,j}\}$ for $1 \leq j \leq m$. Let π be the same mapping from the proof of Theorem 6.1.4 except that $\pi \colon G_{i-1} \to G_{i-1}/G_i$ (in terms of the proof of Theorem 6.1.4, $G = G_{i-1}$, $H = G_i$, and $G/H = G_{i-1}/G_i$) which is defined by $\pi(g) = gG_i$ where $g \in G_{i-1}$. Recall from the proof that the preimages of subgroups of G_{i-1}/G_i form a sequence of normal subgroups with pairwise prime subgroup indexes; i.e.

$$G_i = \pi^{-1}(\widetilde{G}_{i,m}) \triangleleft \pi^{-1}(\widetilde{G}_{i,m-1}) \triangleleft \ldots \triangleleft \pi^{-1}(\widetilde{G}_{i,1}) \triangleleft \pi^{-1}(\widetilde{G}_{i,0}) = G_{i-1}$$

where $\{\pi^{-1}(\widetilde{G}_{i,j-1}) \colon \pi^{-1}(\widetilde{G}_{i,j})\}$ is prime. It follows that when we combine each sequence of normal subgroups formed by G_i and G_{i-1} for $i \in \{1, \ldots, n\}$, we get a sequence of normal subgroups for G (that may or may not be longer than the sequence of normal subgroups formed by the G_i themselves) that satisfies the conditions of Definition 6.1.1, so G is solvable. ■

6.2 Solvable Extensions

Definition 6.2.1. *A field extension $F \subset L$ is called a **radical extension** when there are intermediate field extensions*

$$F = F_0 \subset F_1 \subset \ldots F_{n-1} \subset F_n = L$$

such that for $i \in \{1, \ldots, n\}$, there is $\alpha_i \in F_i$ where $F_i = F_{i-1}(\alpha_i)$ and ${\alpha_i}^{m_i} \in F_{i-1}$, $m_i \in \mathbb{N}$.

Definition 6.2.2. *A field extension $F \subset L$ is considered a **solvable extension** if there is a field extension $L \subset M$ such that $F \subset M$ is a radical extension. In this case, any $f \in F[x]$ is deemed to be **solvable by radicals**.*

According to the above definitions, it is easy to infer that every radical extension is a solvable extension. However, it isn't necessarily the case that every solvable extension is a radical extension.

Definition 6.2.3. *Suppose that the field L has two subfields $K_1 \subset L$ and $K_2 \subset L$. Then, the **compositum** of K_1 and K_2, which we will denote by K_1K_2, is the smallest subfield of L that contains K_1 and K_2.*

Proposition 6.2.4. *Suppose we have a sequence of finite extensions $F \subset L \subset M$ where $F \subset M$ is Galois. Then, the compositum of all conjugate fields of L in M is the Galois closure of $F \subset L$.*

Proof. By the theorem of the primitive element, there is some $\alpha \in L$ such that $L = F(\alpha)$. Since, $F \subset M$ is Galois, the minimum polynomial of α $h \in F[x]$ splits completely over M and is separable over M. Denote the roots of h to be $\alpha_1, \ldots, \alpha_n$ where $\alpha_1 \coloneqq \alpha$ so that in M,

$$h = (x - \alpha_1) \cdots (x - \alpha_n)$$

Let $K = F(\alpha_1, \ldots, \alpha_n)$ which contains L since $\alpha \in K$. By Proposition 2.2.6, K is the smallest subfield of M that contains $\alpha_1, \ldots, \alpha_n$ and since h splits completely over K, we conclude that K is the splitting field of separable polynomial h. By Theorem 5.1.4, it follows that $F \subset K$ is Galois and since $L \subset K$, we can also see that $F \subset K$ is the Galois closure of the extension $F \subset L$. Furthermore, we have that h is separable and is irreducible over F where K is its splitting field. By Proposition 4.2.3, $\mathrm{Gal}(K/F)$ acts transitively on the roots of h, so the conjugate fields of L in K are $F(\alpha_1), \ldots, F(\alpha_n)$. Also, by Corollary 2.2.7, $F(\alpha_1, \ldots, \alpha_n) = F(\alpha_1) \cdots F(\alpha_n)$, so K is precisely the compositum of the conjugate fields of L in K.

Now, for any $\tau \in \mathrm{Gal}(M/F)$, since any root α_i for $i \in \{1, \ldots, n\}$ of $h \in F[x]$ is in M, then $\tau(\alpha_i)$ is also a root of h, but the only roots of h are $\alpha_1, \ldots, \alpha_n$, so the conjugate fields of L in K are the same as the conjugate fields of L in M. Thus, the compositum of all conjugate fields of L in M is the Galois closure of $F \subset L$. ■

Proposition 6.2.5. *If we have field extensions $F \subset K_1 \subset L$ and $F \subset K_2 \subset L$ where $F \subset K_1$ is radical, then $K_1 \subset K_1K_2$ is radical.*

Proof. Since $F \subset K_1$ is radical, there are subfields

$$F = F_0 \subset F_1 \subset \ldots \subset F_n = K_1$$

such that for $i \in \{1, \ldots, n\}$, $F_i = F_{i-1}(\gamma_i)$ where ${\gamma_i}^{m_i} \in F_{i-1}$ for $m_i \in \mathbb{N}$.

Now, define

$$\begin{aligned}
U_0 &= K_2, \\
U_1 &= U_0(\gamma_1) = K_2(\gamma_1), \\
U_2 &= U_1(\gamma_2) = K_2(\gamma_2, \gamma_2), \\
&\vdots \\
U_n &= U_{n-1}(\gamma_n) = K_2(\gamma_1, \gamma_2, \ldots, \gamma_n)
\end{aligned}$$

Also note that since $F \subset K_2$,

$$\begin{aligned}
F_0 &= F \subset K_2 = U_0 \\
F_1 &= F(\gamma_1) \subset K_2(\gamma_1) = U_1 \\
F_2 &= F(\gamma_1, \gamma_2) \subset K_2(\gamma_1, \gamma_2) = U_2 \\
&\vdots \\
F_n &= F(\gamma_1, \gamma_2, \ldots, \gamma_n) \subset K_2(\gamma_1, \gamma_2, \ldots, \gamma_n) = U_n
\end{aligned}$$

so for $i \in \{0, 1, \ldots, n\}$, it follows that $F_i \subset U_i$ which means that ${\gamma_i}^{m_i} \in F_{i-1} \Rightarrow {\gamma_i}^{m_i} \in U_{i-1}$ for $i \in \{1, \ldots, n\}$. Therefore, we have

$$K_2 = U_0 \subset U_1 \subset \ldots \subset U_n$$

such that for $i \in \{1, \ldots, n\}$, $U_i = U_{i-1}(\gamma_i)$ where ${\gamma_i}^{m_i} \in U_{i-1}$ for $m_i \in \mathbb{N}$, so $K_2 \subset U_n$ is a radical extension.

Note that since $K_1 \subset L$ and $K_1 = F(\gamma_1, \ldots, \gamma_n)$, we have that $\gamma_1, \ldots, \gamma_n \in L$. Moreover, we have that $F_n = K_1 \subset K_2(\gamma_1, \ldots, \gamma_n) \subset L$. By definition of the compositum, it follows that

$$K_1 \subset K_1K_2 \subset K_2(\gamma_1, \ldots, \gamma_n) \subset L$$

Also by the definition of the compositum, we have $K_2 \subset K_1K_2 \subset L$, but by Proposition 2.2.6, $K_2(\gamma_1, \ldots, \gamma_n)$ is the smallest subfield of L that contains K_2 and $\gamma_1, \ldots, \gamma_n$, so

$$K_2 \subset K_2(\gamma_1, \ldots, \gamma_n) \subset K_1K_2 \subset L$$

Here, we can conclude that $U_n = K_1K_2$, so $K_2 \subset K_1K_2$ is a radical extension. ■

Theorem 6.2.6. *Suppose the finite extension $F \subset L$ is separable and radical. Then, the Galois closure of $F \subset L$ is radical.*

Proof. By Proposition 5.1.7, there is $L \subset M$ such that $F \subset M$ is Galois. Now, take any $\sigma \in \mathrm{Gal}(M/F)$, so that we have conjugate field $F \subset \sigma L \subset M$. Since $F \subset L$ is radical, we have subfields

$$F = F_0 \subset F_1 \subset \ldots \subset F_n = L$$

such that for $i \in \{1, \ldots, n\}$, $F_i = F_{i-1}(\gamma_i)$ where ${\gamma_i}^{m_i} \in F_{i-1}$ for $m_i \in \mathbb{N}$. Let $U_i = \sigma F_i$ and $\sigma(\gamma_i) = \rho_i$; here, $U_0 = \sigma F_0 = F_0$ and $U_n = \sigma F_n = \sigma L$. Then

$$U_i = \sigma F_i = \sigma(F_{i-1}(\gamma_i)) = U_{i-1}(\sigma(\gamma_i)) = U_{i-1}(\rho_i)$$

so $U_{i-1} \subset U_i$. Also,

$$\sigma({\gamma_i}^{m_i}) = (\sigma(\gamma_i))^{m_i} = {\rho_i}^{m_i} \in U_{i-1}$$

Thus, we have

$$F = U_0 \subset U_1 \subset \ldots \subset U_n = \sigma L$$

such that for $i \in \{1, \ldots, n\}$, $U_i = U_{i-1}(\rho_i)$ where ${\rho_i}^{m_i} \in U_{i-1}$ for $m_i \in \mathbb{N}$, so $F \subset \sigma L$ is a radical extension. It follows that all conjugate fields are radical and by Proposition 6.2.5, the compositum of all of these conjugate fields is radical over F. Additionally, by Proposition 6.2.4, this compositum is the Galois closure of $F \subset L$. Therefore, the Galois closure of $F \subset L$ is radical over F. ■

Lemma 6.2.7. *Suppose F has characteristic zero where $F \subset L$ is Galois and ζ is a primitive mth root of unity. Then,*
(i) *$L \subset L(\zeta)$ and $F \subset F(\zeta)$ are Galois.*
(ii) *$\mathrm{Gal}(L(\zeta)/L)$ and $\mathrm{Gal}(F(\zeta)/F)$ are Abelian.*
(iii) *$F \subset L(\zeta)$ and $F(\zeta) \subset L(\zeta)$ are Galois.*

Proof. Suppose $M \in \{F, L\}$ and consider $M \subset M(\zeta)$. For any M, note that $x^m - 1 \in M[x]$ and is also separable. By Theorem 5.1.4, since $M(\zeta)$ is the splitting field of $x^m - 1$, it follows that $M \subset M(\zeta)$ is Galois, so (i) is proven. Additionally, note that for any $\sigma \in \mathrm{Gal}(M(\zeta)/M)$, since $M(1, \zeta, \ldots, \zeta^{m-1}) = M(\zeta)$, by Proposition 4.1.4, σ is uniquely determined by its action on ζ. Thus, since the roots of $x^m - 1$ are $1, \zeta, \ldots, \zeta^{m-1}$, when considering any $\sigma, \tau \in \mathrm{Gal}(M(\zeta)/M)$ it must be that $\sigma(\zeta) = \zeta^i$ and $\tau(\zeta) = \zeta^j$ for some $i, j \in \{0, 1, \ldots, m-1\}$. Now, observe that

$$\sigma\tau(\zeta) = \sigma(\zeta^j) = (\sigma(\zeta))^j = (\zeta^i)^j = \zeta^{ij}$$
$$\tau\sigma(\zeta) = \tau(\zeta^i) = (\tau(\zeta))^i = (\zeta^j)^i = \zeta^{ij}$$

Therefore, $\sigma\tau = \tau\sigma$, so $\mathrm{Gal}(M(\zeta)/M)$ is Abelian and hence, (ii) is proven.

Finally, consider the extensions $F \subset L \subset L(\zeta)$. Since $F \subset L$ is Galois, it is normal so every irreducible polynomial over F that has a root in L splits completely over L; however, if every irreducible polynomial over F that has a root in L splits completely over L, then surely they split completely over $L(\zeta)$, so $F \subset L(\zeta)$ is normal. Also, by Proposition 3.4.4, since F has characteristic zero, $F \subset L(\zeta)$ is separable since every irreducible polynomial over F is separable and hence it is also Galois. By Corollary 5.1.5, since we have $F \subset F(\zeta) \subset L(\zeta)$ where $F \subset L(\zeta)$ is Galois, it follows that $F(\zeta) \subset L(\zeta)$ is Galois, thus (iii) is proven. ■

Proposition 6.2.8. *Suppose F has characteristic zero where $F \subset L$ is Galois and ζ is a primitive mth root of unity. Then, the following are equivalent:*
(a) *$\mathrm{Gal}(L/F)$ is solvable*
(b) *$\mathrm{Gal}(L(\zeta)/F)$ is solvable*
(c) *$\mathrm{Gal}(L(\zeta)/F(\zeta))$ is solvable*

Proof. First, consider $F \subset L \subset L(\zeta)$. Since $F \subset L(\zeta)$ is Galois by Lemma 6.2.7 and $F \subset L$ is Galois, it follows by Theorem 5.2.5 that

$$\mathrm{Gal}(L/F) \simeq \mathrm{Gal}(L(\zeta)/F)/\mathrm{Gal}(L(\zeta)/L)$$

and $\mathrm{Gal}(L(\zeta)/L) \triangleleft \mathrm{Gal}(L(\zeta)/F)$. Recall that $\mathrm{Gal}(L(\zeta)/L)$ is Abelian by Lemma 6.2.7, so by Proposition 6.1.5, it is solvable. By Theorem 6.1.4, since we can assume solvability of $\mathrm{Gal}(L(\zeta)/L)$, it follows that $\mathrm{Gal}(L/F)$ is solvable if and only if $\mathrm{Gal}(L(\zeta)/F)$ is solvable hence, (a) and (b) are equivalent.

Next, consider $F \subset F(\zeta) \subset L(\zeta)$. Since, again, by Lemma 6.2.7, $F \subset F(\zeta)$ and $F \subset L(\zeta)$ are Galois, it follows by Theorem 5.2.5 that

$$\mathrm{Gal}(F(\zeta)/F) \simeq \mathrm{Gal}(L(\zeta)/F)/\mathrm{Gal}(L(\zeta)/F(\zeta))$$

and $\mathrm{Gal}(L(\zeta)/F(\zeta)) \triangleleft \mathrm{Gal}(L(\zeta)/F)$. Also by Lemma 6.2.7, $\mathrm{Gal}(F(\zeta)/F)$ is Abelian, so it is also solvable according to Proposition 6.1.5. By Theorem 6.1.4, since we can assume solvability of $\mathrm{Gal}(F(\zeta)/F)$, it follows that $\mathrm{Gal}(L(\zeta)/F)$ is solvable if and only if $\mathrm{Gal}(L(\zeta)/F(\zeta))$ is solvable hence, (b) and (c) are equivalent. ■

Lemma 6.2.9. *Suppose $K \subset M$ is a Galois extension and $\mathrm{Gal}(M/K) \simeq \mathbb{Z}/p\mathbb{Z}$ where p is prime. Then, if $\zeta \in K$ where ζ is a pth primitive root of unity, there is some $\alpha \in M$ such that $M = K(\alpha)$ and $\alpha^p \in K$.*

Proof. By assumption, it follows that $\mathrm{Gal}(M/K)$ is cyclic so there is some $\sigma \in \mathrm{Gal}(M/K)$ that generates the entire group. Now, take $\beta \in M \backslash K$; for each $i \in \{0, 1, \ldots, p-1\}$, consider

$$\alpha_i = \beta + \zeta^{-i}\sigma(\beta) + \zeta^{-2i}\sigma^2(\beta) + \ldots + \zeta^{-(p-1)i}\sigma^{p-1}(\beta) \tag{6.1}$$

Here, Eq. (6.1) is called the **Lagrange Resolvent**. Then,

$$\begin{aligned} \sigma(\alpha_i) &= \sigma(\beta) + \zeta^{-i}\sigma^2(\beta) + \zeta^{-2i}\sigma^3(\beta) + \ldots + \zeta^{-(p-1)i}\sigma^p(\beta) \\ \Rightarrow \zeta^{-i}\sigma(\alpha_i) &= \zeta^{-i}\sigma(\beta) + \zeta^{-2i}\sigma^2(\beta) + \zeta^{-3i}\sigma^3(\beta) + \ldots + \zeta^{-(p-1)i}\sigma^p(\beta) + \zeta^{-ip}\sigma^p(\beta) \end{aligned}$$

However, the last term on right hand side of the last line is just β since $\sigma^p = i_d$ where i_d is the identity mapping since σ is of order p; also, $\zeta^{-ip} = (\zeta^p)^{-i} = 1^{-i} = 1$, so

$$\zeta^{-i}\sigma(\alpha_i) = \zeta^{-i}\sigma(\beta) + \zeta^{-2i}\sigma^2(\beta) + \zeta^{-3i}\sigma^3(\beta) + \ldots + \zeta^{-(p-1)i}\sigma^p(\beta) + \beta \tag{6.2}$$

but we see that Eq. (6.2) is exactly Eq. (6.1). Thus, we have $\zeta^{-i}\sigma(\alpha_i) = \alpha_i$ and hence,

$$\sigma(\alpha_i) = \zeta^i\alpha_i \tag{6.3}$$

Also, since $\sigma({\alpha_i}^p) = (\sigma(\alpha_i))^p = (\zeta^i\alpha_i)^p = {\alpha_i}^p$, we see that ${\alpha_i}^p$ remains invariant under the generator of $\mathrm{Gal}(M/K)$, so it follows that ${\alpha_i}^p \in K$; additionally, by Eq. (6.2), $\sigma(\alpha_0) = \alpha_0$, so it is also the case that $\alpha_0 \in K$.
Now, it has to be that for at least one $i \in \{1, \ldots, p-1\}$ that $\alpha_i \neq 0$. Suppose it is the case that for all of the α_i is zero. Then, adding up all of the α_i for $i \in \{0, 1, \ldots, p-1\}$ and taking into account Eq. (6.1), we get

$$\begin{aligned} \alpha_0 = \alpha_0 + \alpha_1 + \ldots + \alpha_{p-1} &= (\beta + \sigma(\beta) + \sigma^2(\beta) + \ldots + \sigma^{p-1}(\beta)) \\ &\quad + (\beta + \zeta^{-1}\sigma(\beta) + \zeta^{-2}\sigma^2(\beta) + \ldots + \zeta^{-(p-1)}\sigma^{p-1}(\beta)) \\ &\quad \vdots \\ &\quad + (\beta + \zeta^{-(p-1)}\sigma(\beta) + \zeta^{-2(p-1)}\sigma^2(\beta) + \ldots + \zeta^{-(p-1)(p-1)}\sigma^{p-1}(\beta)) \\ &= p\beta + (1 + \zeta^{-1} + \zeta^{-2} + \ldots + \zeta^{-(p-1)})\sigma(\beta) \\ &\quad + (1 + \zeta^{-2} + \zeta^{-4} + \ldots + \zeta^{-2(p-1)})\sigma^2(\beta) \\ &\quad \vdots \\ &\quad + (1 + \zeta^{-(p-1)} + \zeta^{-2(p-1)} + \ldots + \zeta^{-(p-1)(p-1)})\sigma^{p-1}(\beta) = p\beta \end{aligned}$$

where the last equality is due to the fact that for all $i \in \{1, \ldots, p-1\}$,

$$1 + \zeta^{-i} + \zeta^{-2i} + \ldots + \zeta^{-(p-1)i} = \frac{1 - \zeta^{-ip}}{1 - \zeta^{-i}} = \frac{1-1}{1-\zeta^{-i}} = 0$$

Thus, we have that $\alpha_0 = p\beta$ which is in K as established earlier in the proof. This means that $\beta = \alpha_0/p \in K$, but this is a contradiction since we assumed that $\beta \in M \backslash K$. Therefore, we must have at least one $i \in \{1, \ldots, p-1\}$ such that $\alpha_i \neq 0$. We also have that $\zeta^i \neq 1$, so $\zeta^i \alpha_i \neq \alpha_i$ for $i \in \{1, \ldots, p-1\}$, and by Eq. (6.3), it follows that $\sigma(\alpha_i) \neq \alpha_i$ and hence $\alpha_i \notin K$. Since $|\mathrm{Gal}(M/K)| = p$, it follows by Corollary 5.1.6 that $[M : K] = p$ since $K \subset M$ is Galois, and since ${\alpha_i}^p \in K$, α_i is algebraic over K with minimal polynomial $f_i = x^p - {\alpha_i}^p \in K[x]$. Therefore, by Theorem 2.3.4, we are able to conclude that $M = K(\alpha_i)$. ■

6.3 Galois's Theorem

Theorem 6.3.1. *Suppose F has characteristic zero where $F \subset L$ is a Galois extension. Then*

$$F \subset L \text{ is solvable} \iff \mathrm{Gal}(L/F) \text{ is solvable}$$

Proof. Suppose $F \subset L$ is solvable. Then, there is a radical extension $F \subset L'$ such that $L \subset L'$; also, since $F \subset L$ is normal then surely $F \subset L'$ is also normal since every irreducible polynomial over F that splits completely over L also splits completely over L' since. Furthermore, by Proposition 3.4.4, every irreducible polynomial over F is separable, so $F \subset L'$ is separable and hence, a Galois extension. By Theorem 5.2.5, since we have that $F \subset L \subset L'$, it follows that

$$\mathrm{Gal}(L'/F)/\mathrm{Gal}(L'/L) \simeq \mathrm{Gal}(L/F)$$

so if we find that $\mathrm{Gal}(L'/F)$ is solvable, it will follow by Theorem 6.1.4 that $\mathrm{Gal}(L/F)$ is solvable. Thus, for convenience, we may relabel L' as L and consider $F \subset L$ as the radical extension. By Proposition 6.2.5, since $F \subset L$ is radical, it follows that $F(\zeta) \subset L(\zeta)$ is radical since $L(\zeta)$ is the compositum of L and $F(\zeta)$. Also, by Proposition 6.2.8, solvability of $\mathrm{Gal}(L(\zeta)/F(\zeta))$ necessarily implies the solvability of $\mathrm{Gal}(L/F)$, so we can assume WLOG that F has any primitive mth root of unity as needed. Now, since $F \subset L$ is radical, there are subfields

$$F_0 \subset F_1 \subset \ldots \subset F_{n-1} \subset F_n$$

where $F_0 \coloneqq F$, $F_n \coloneqq L$, and for $i \in \{1, \ldots, n\}$, there is $\gamma_i \in F_i$ such that $F_i = F_{i-1}(\gamma_i)$ and ${\gamma_i}^{m_i} \in F_{i-1}$ where $m_i \in \mathbb{N}$. Consider the intermediate extension $F_{i-1} \subset F_i$ where we assume $\zeta_i \in F$ and hence, $\zeta_i \in F_{i-1}$ where ζ_i is a primitive m_ith root of unity. The minimum

polynomial of γ_i in $F_{i-1}[x]$ is $f_{i-1} \coloneqq x^{m_i} - {\gamma_i}^{m_i}$ which has roots $\gamma_i, \zeta_i\gamma_i, \ldots, {\zeta_i}^{m_i-1}\gamma_i$ in F_i. Therefore, F_i is the splitting field of $f_{i-1} \in F_{i-1}[x]$ since

$$F_i = F_{i-1}(\gamma_i, \zeta_i\gamma_i, \ldots, {\zeta_i}^{m_i-1}\gamma_i) = F_{i-1}(\gamma_i)$$

It follows that every irreducible polynomial over F_{i-1} splits completely over F_i, so $F_{i-1} \subset F_i$ is normal and hence, is a Galois extension.

Furthermore, consider $\sigma \in \mathrm{Gal}(F_i/F_{i-1})$; by Proposition 4.1.4, $\sigma(\gamma_i)$ is also a root of f_{i-1}, so there is some $l \in \{1, \ldots, m_i - 1\}$ such that $\sigma(\gamma_i) = {\zeta_i}^l\gamma_i$. Now, define the mapping

$$\phi\colon \mathrm{Gal}(F_i/F_{i-1}) \to \mathbb{Z}/m_i\mathbb{Z}$$

where $\phi(\sigma) = [l]$ when $\sigma(\gamma_i) = {\zeta_i}^l\gamma_i$. Take $\sigma_1, \sigma_2 \in \mathrm{Gal}(F_i/F_{i-1})$ and also take $l_1, l_2 \in \{1, \ldots, m_i - 1\}$, so that $\phi(\sigma_1) = [l_1]$ and $\phi(\sigma_2) = [l_2]$; then,

$$\sigma_1 \circ \sigma_2(\gamma_i) = \sigma_1({\zeta_i}^{l_2}\gamma_i) = {\zeta_i}^{l_1}({\zeta_i}^{l_2}\gamma_i) = {\zeta_i}^{l_1+l_2}\gamma_i$$

so $\phi(\sigma_1 \circ \sigma_2) = [l_1 + l_2] = [l_1] + [l_2]$. Therefore, $\phi(\sigma_1 \circ \sigma_2) = \phi(\sigma_1) + \phi(\sigma_2)$. Moreover, suppose that $\phi(\sigma_1) = \phi(\sigma_2)$; then, $[l_1] = [l_2]$ and

$${\zeta_i}^{l_1}\gamma_i = {\zeta_i}^{l_2}\gamma_i \Rightarrow \sigma_1(\gamma_i) = \sigma_2(\gamma_i) \Rightarrow \sigma_1 = \sigma_2$$

Therefore, $\phi(\sigma_1) = \phi(\sigma_2) \Rightarrow \sigma_1 = \sigma_2$, so we are able to conclude that ϕ is an injective homomorphism. By Cayley's theorem, it follows that $\mathrm{Gal}(F_i/F_{i-1})$ is isomorphic to a subgroup of $\mathbb{Z}/m_i\mathbb{Z}$. Since $\mathbb{Z}/m_i\mathbb{Z}$ is cyclic, it follows that every subgroup of $\mathbb{Z}/m_i\mathbb{Z}$ is cyclic which means that $\mathrm{Gal}(F_i/F_{i-1})$ is cyclic.

For $i \in \{0, 1, \ldots, n\}$, let $\mathrm{Gal}(L/F_i) = G_i$; then, by the Galois correspondence,

$$G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0$$

where by Theorem 5.2.5, $G_i \triangleleft G_{i-1}$ and

$$G_{i-1}/G_i \simeq \mathrm{Gal}(F_i/F_{i-1})$$

so G_{i-1}/G_i is cyclic. Additionally, G_{i-1}/G_i is abelian since all cyclic groups are abelian. Thus, by Proposition 6.1.6, it follows that $\mathrm{Gal}(L/F)$ is solvable.

We will divide the converse into two parts. Suppose $\mathrm{Gal}(L/F)$ is solvable. **Part (i)**: Additionally, suppose that F contains a primitive pth root of unity for every prime p that divides $|\mathrm{Gal}(L/F)|$. Let $G_0 = \mathrm{Gal}(L/F)$ and $G_n = \{e\}$; since G_0 is solvable, we have subgroups

$$G_n \triangleleft G_{n-1} \triangleleft \ldots G_1 \triangleleft G_0$$

where $\{G_{i-1} : G_i\}$ is prime for $i \in \{1, \ldots, n\}$; also, let F_i be the fixed field under the action of G_i. Then, by the Galois correspondence,

$$F_0 \subset F_1 \subset \ldots \subset F_{n-1} \subset F_n$$

where $F_0 = F$ and $F_n = L$. Moreover, by Theorem 5.2.5 together with the Galois correspondence, we have that

$$\mathrm{Gal}(L/F_{i-1})/\mathrm{Gal}(L/F_i) \simeq \mathrm{Gal}(F_i/F_{i-1})$$

and since $\{G_{i-1} : G_i\}$ is prime, it follows that $\mathrm{Gal}(F_i/F_{i-1}) \simeq \mathbb{Z}/p\mathbb{Z}$ where p is prime. Consider the extensions $F \subset F_{i-1} \subset F_i \subset L$. Note that since $F \subset L$ is Galois, by Corollary 5.1.5 it follows that both $F_{i-1} \subset L$ and $F_i \subset L$ are Galois. By Corollary 5.1.6, it follows that $|\mathrm{Gal}(L/F_{i-1})| = [L : F_{i-1}]$ and $|\mathrm{Gal}(L/F_i)| = [L : F_i]$, so according to the isomorphisms,

$$[L : F_{i-1}] = p[L : F_i]$$

Furthermore, since $[L : F] = [L : F_{i-1}][F_{i-1} : F]$, it follows

$$[L : F] = p[L : F_i][F_{i-1} : F]$$

so $|\mathrm{Gal}(F_i/F_{i-1})| = p$ divides $|\mathrm{Gal}(L/F)|$, and by assumption, F contains a primitive pth root of unity ζ_p, so $\zeta_p \in F_{i-1}$. Also, since $\mathrm{Gal}(L/F_i) \triangleleft \mathrm{Gal}(L/F_{i-1})$ by Theorem 5.2.5, it follows by Theorem 5.2.4 that $F_{i-1} \subset F_i$ is a Galois extension. Thus, by Lemma 6.2.9, since $F_{i-1} \subset F_i$ is Galois, $\zeta_p \in F_{i-1}$, and $\mathrm{Gal}(F_i/F_{i-1}) \simeq \mathbb{Z}/p\mathbb{Z}$ where p is prime, it follows that there is a $\gamma_i \in F_i$ such that $F_i = F_{i-1}(\gamma_i)$ and ${\gamma_i}^p \in F_{i-1}$. It follows that $F \subset L$ is a radical extension.

Part (ii): Now, suppose the assumption made in the first sentence of part (i) is not given. By Proposition 6.2.8 if we let ζ be a primitive mth root of unity where $m \coloneqq |\mathrm{Gal}(L/F)|$, then $\mathrm{Gal}(L(\zeta)/F(\zeta))$ is solvable since $\mathrm{Gal}(L/F)$ is solvable. As in the proof of Proposition 6.2.8, we have

$$\mathrm{Gal}(L(\zeta)/F)/\mathrm{Gal}(L(\zeta)/L) \simeq \mathrm{Gal}(L/F)$$

which was induced via the group homomorphism $\mathrm{Gal}(L(\zeta)/F) \to \mathrm{Gal}(L/F)$ by imposing a restriction to L on any $\sigma \in \mathrm{Gal}(L(\zeta)/F)$ in the same manner as in the proof of Theorem 5.2.5. Since $\mathrm{Gal}(L(\zeta)/F(\zeta)) \subset \mathrm{Gal}(L(\zeta)/F)$, this yields a similar homomorphism

$$\Phi\colon \mathrm{Gal}(L(\zeta)/F(\zeta)) \to \mathrm{Gal}(L/F)$$

defined by $\Phi(\sigma) = \sigma|_L$. Note that

$$\ker(\Phi) = \{\sigma \in \text{Gal}(L(\zeta)/F(\zeta)) \mid \Phi(\sigma) = i_d|_L\} \subset \text{Gal}(L(\zeta)/L)$$

(where $i_d|_L$ is the identity mapping on $L(\zeta)$ restricted to L) so $\ker(\Phi)$ is the set of all automorphisms in $\text{Gal}(L(\zeta)/F(\zeta))$ that fixes both $F(\zeta)$ and L. If $\sigma \in \text{Gal}(L(\zeta)/L)$, then σ permutes the mth roots of unity or is the identity mapping. However, if σ permutes the mth roots of unity, then $F(\zeta)$ cannot be fixed by the action of σ. Thus, the only automorphism in $\ker(\Phi)$ is the identity mapping so the kernel of Φ is trivial. It follows that Φ is an injective homomorphism since a mapping is injective if and only if the kernel is trival; by Cayley's theorem, the image of Φ is a subgroup of $\text{Gal}(L/F)$. By the fundamental theorem of group homomorphisms, $\text{Gal}(L(\zeta)/F(\zeta)) \simeq \text{im}(\Phi)$, so we essentially have that $\text{Gal}(L(\zeta)/F(\zeta))$ is a subgroup of $\text{Gal}(L/F)$ and by Lagrange's theorem, $|\text{Gal}(L(\zeta)/F(\zeta))|$ divides $m = |\text{Gal(L/F)}|$.

Now, let p be a prime that divides $|\text{Gal}(L(\zeta)/F(\zeta))|$; then p should also divide m. Consider $\zeta^{\frac{m}{p}}$; note that, taking $k \in \mathbb{N}$,

$$(\zeta^{\frac{m}{p}})^k = 1 \Rightarrow e^{2\pi i \frac{k}{p}} = 1 \Rightarrow (k/p)(2\pi i) = n(2\pi i), n \in \mathbb{N} \Rightarrow k/p = n \Rightarrow k = np$$

so k would be a positive integer multiple of p which would mean that $\zeta^{m/p}$ is a primitive pth root of unity. Since p divides m, $\zeta^{m/p} \in F(\zeta)$ and by part (i), it follows that $F(\zeta) \subset L(\zeta)$ is a radical extension. Also, $F \subset F(\zeta)$ is radical because $\zeta^m = 1 \in F$, so it follows that $F \subset L(\zeta)$ is radical since we can combine the radical sequences of $F \subset F(\zeta)$ and $F(\zeta) \subset L(\zeta)$. Therefore, by definition, since $F \subset L \subset L(\zeta)$, we conclude that $F \subset L$ is solvable. ■

We will refer to sequence of fields formed by a radical extension as a solvable chain of field extensions and the sequence of normal subgroups formed by a solvable group as a solvable chain of normal subgroups.

Part II

Examining Solvability and Unsolvability of Polynomials

7

The Cubic and Quartic via Galois Theory

This chapter is written with assistance from chapter 8.3 and chapter 12.1 from [5]. Suppose we are unaware of the solution process for both the cubic and quartic; can we reproduce these solutions with only knowledge of Galois theory?

Here, we will focus on applying Galois theory to the process of solving polynomials. Consider the universal polynomial $\widetilde{f} \in F(\sigma_1, \ldots, \sigma_n)[x]$ from Appendix A.4. The roots $x_1, \ldots, x_n$ of $\widetilde{f}$ are distinct by assumption which lie in the field $F(x_1, \ldots, x_n)$. This is clearly the splitting field of $\widetilde{f}$. Thus, $F(x_1, \ldots, x_n)$ is the splitting field of separable polynomial $\widetilde{f} \in F(\sigma_1, \ldots, \sigma_n)$, so $F(\sigma_1, \ldots, \sigma_n) \subset F(x_1, \ldots, x_n)$ is a Galois extension by Theorem 5.1.4. Computational matters are handled with the assistance of computer algebra program Mathematica.

7.1 The Cubic

Let $\omega = e^{2\pi i/3} = -\frac{1}{2} + \frac{i\sqrt{3}}{2}$ and $F = \mathbb{Q}(\omega)$. Also, set $L = F(x_1, x_2, x_3)$ where x_1, x_2, x_3 are the roots of the universal cubic polynomial over the field $K = F(\sigma_1, \sigma_2, \sigma_3) \subset L$ given by

$$\widetilde{f}(x) = x^3 - \sigma_1 x^2 + \sigma_2 x - \sigma_3$$

where $\sigma_1, \sigma_2, \sigma_3$ are the elementary symmetric polynomials. Here, L is the splitting field of $\widetilde{f}$, so $\mathrm{Gal}(L/K) \simeq S_3$ according to Theorem 4.2.6 and hence, $[L : K] = 6$ by Corollary 5.1.6; thus, $K \subset L$ is a sixth degree Galois extension.

It is known that S_3 forms a solvable chain given by:

$$\{e\} \triangleleft A_3 \triangleleft S_3$$

since $\{S_3 : A_3\} = 2$ which is prime and $\{A_3 : \{e\}\} = 3$ which is also prime.

By the Galois Correspondence, this indicates there is an intermediate field M such that $K \subset M \subset L$ fixed by the action of A_3, so $\mathrm{Gal}(L/M) \simeq A_3$ and hence, $[L : M] = 3$, so $M \subset L$ is a third degree extension. By the Tower Theorem,

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

so we must have $6 = 3[M : K]$ which implies that $[M : K] = 2$ and hence, that $M \subset K$ is a second degree extension. In fact, this is apparent by considering the quotient group $S_3/A_3 \simeq \mathbb{Z}/2\mathbb{Z}$ and noting that this implies that $\text{Gal}(M/K) \simeq \mathbb{Z}/2\mathbb{Z}$.

Now, we may begin the solution process of finding the roots of $\widetilde{f}$ expressed in terms of elements in K (where every element is some rational expression of the elementary symmetric polynomials) by first making use of the fact that since $M \subset L$ is a Galois extension, $\text{Gal}(L/M) \simeq A_3 \simeq \mathbb{Z}/3\mathbb{Z}$ where 3 is prime, and M contains a primitive third root of unity, ω, it follows by Lemma 6.2.9 that $\exists \phi \in L : \phi^3 \in M$ and $L = M(\phi)$. Moreover, this means that ϕ is a primitive element of $M \subset L$, so since $[L : M] = 3$, according to what we established in Section 2.3, we may say that the vector space that L forms over M has basis vectors $\{1, \phi, \phi^2\}$, and so explicitly:

$$L = \{a + b\phi + c\phi^2 \mid a, b, c \in M\}$$

Similarly, since $K \subset M$ is a Galois extension, $\text{Gal}(M/K) \simeq \mathbb{Z}/2\mathbb{Z}$ where 2 is prime, and K contains a primitive second root of unity (which is $e^{2\pi i/2} = e^{\pi i} = -1$), it follows by Lemma 6.2.9 that $\exists \psi \in M : \psi^2 \in K$ and $M = K(\psi)$. Again, we have a primitive element ψ of a field extension $K \subset M$, so since $[M : K] = 2$, we may say that the vector space that M forms over K has basis vectors $\{1, \psi\}$, and explicitly, we have:

$$M = \{a + b\psi \mid a, b \in K\}$$

Using the explicit expressions of L over M and M over K, we can also see that the vector space that L forms over K has basis vectors $\{1, \phi, \phi^2, \psi, \psi\phi, \psi\phi^2\}$, and thus, explicitly:

$$L = \{a + b\phi + c\phi^2 + p\psi + q\psi\phi + r\psi\phi^2 \mid a, b, c, p, q, r \in K\}$$

So far, for $K \subset M \subset L$, we have that for L/M, $\phi^3 \in M$ where $\phi \in L$ and $L = M(\phi)$; also, for M/K, $\psi^2 \in K$ where $\psi \in M$ and $M = K(\psi)$. This observation makes it apparent that $K \subset L$ is a solvable extension and hence, $\widetilde{f}$ is indeed solvable by radicals over K. We can now express our sequence of field extensions as

$$K \subset K(\psi) \subset K(\psi, \phi)$$

By Proposition A.4.7, we know that for any $\sigma \in S_3$,

$$\sigma\sqrt{\Delta} = \text{sgn}(\sigma)\sqrt{\Delta}$$

Notice that for every $\sigma \in A_3$, $\text{sgn}(\sigma) = +1$, so $\sqrt{\Delta}$ is invariant under the action of A_3 which means that $\sqrt{\Delta} \in M$. Furthermore, since $\Delta \in F(\sigma_1, \sigma_2, \sigma_3)$, we know that

$$x^2 - \Delta \in K[x]$$

which is irreducible over K where $\sqrt{\Delta}$ is a root. This means that $x^2 - \Delta \in K[x]$ is the minimal polynomial of $\sqrt{\Delta}$ over K and we may conclude that, by Section 2.3, $K \subset K(\sqrt{\Delta})$ is a degree two extension; i.e. $[K(\sqrt{\Delta}) : K] = 2$.

Thus, $\sqrt{\Delta}$, satisfies the properties of $\psi \in M$, so we can conclude that $M = K(\sqrt{\Delta})$. Also, take $\phi \in L$ where $\phi = x_1 + \omega x_2 + \omega^2 x_3$. Then, denoting the orbit of the action of S_3 on ϕ by $\text{Orb}(S_3, \phi)$, we have that $|\text{Orb}(S_3, \phi)| = 6$, so ϕ does not lie in a proper subfield of L. However, taking $\tau \in A_3$ to be $\tau = (123)$, we find that $\tau(\phi) = x_2 + \omega x_3 + \omega^2 x_1 = \omega^2(x_1 + \omega x_2 + \omega^2 x_3) = \omega^2\phi$, so that $\tau(\phi^3) = (x_2 + \omega x_3 + \omega^2 x_1)^3 = (\omega^2(x_1 + \omega x_2 + \omega^2 x_3))^3 = (\omega^2\phi)^3 = \omega^6\phi^3 = \phi^3$. Therefore, we have $\tau(\phi^3) = \phi^3$; it can also be shown for the other nontrivial element $\rho \in A_3$ given by $\rho = (132)$ that $\rho(\phi^3) = \phi^3$. Thus, ϕ^3 is invariant under the action of A_3 and we find that $\phi^3 \in M$. This means that according to the structure of the extension $K \subset M$ that for some $a, b \in K$,

$$\phi^3 = a + b\sqrt{\Delta}$$

Since $[M : K] = 2$, it follows that the minimum polynomial of ϕ^3 over K is of degree two, so ϕ^3 is a root of a quadratic polynomial with coefficients in K. Here, a, b, and Δ are all expressible in terms of the elementary symmetric functions since they all lie in K. Since $\sqrt{\Delta}$ is a root of $x^2 - \Delta$, note that $-\sqrt{\Delta}$ is its conjugate, so the conjugate of ϕ^3 is given by

$$\bar{\phi^3} = a - b\sqrt{\Delta}$$

so we can conclude that ϕ^3 is a root of the quadratic polynomial p over M given by $p = (x - \phi^3)(x - \bar{\phi^3})$. Also, notice that for $\gamma = (23)$, $\gamma(a + b\sqrt{\Delta}) = a - b\sqrt{\Delta}$ since $\text{sgn}(\gamma) = -1$. Therefore, $\gamma\phi^3 = \bar{\phi^3}$ where

$$\bar{\phi^3} = (x_1 + \omega x_3 + \omega^2 x_2)^3$$

Expanding p we get

$$p = x^2 - (\phi^3 + \bar{\phi^3})x + \phi^3\bar{\phi^3} = x^2 - 2ax + (a^2 - b^2\Delta)$$

so we can see that $p \in K[x]$ and hence that $(\phi^3 + \bar{\phi^3}) \in K$ and $\phi^3\bar{\phi^3} \in K$. Indeed, we can find via a symmetric reduction algorithm implicit by Theorem A.4.4 that

$$\phi^3 + \bar{\phi^3} = 2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3 = 2a$$
$$\phi^3\bar{\phi^3} = ({\sigma_1}^2 - 3\sigma_2)^3 = a^2 - b^2\Delta$$

where, using the same algorithm and Definition A.4.6

$$\Delta = \sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2$$

Observe that

$$\begin{aligned}\phi^3\bar{\phi^3} &= a^2 - b^2\Delta \\ \Rightarrow b &= \sqrt{\frac{a^2 - \phi^3\bar{\phi^3}}{\Delta}}\end{aligned}$$

Then,

$$\begin{aligned}\phi^3 &= a + b\sqrt{\Delta} \\ &= a + \sqrt{a^2 - \phi^3\bar{\phi^3}}\end{aligned}$$

From the above equations, it easily follows that $a = (2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3)/2$, so

$$\begin{aligned}a^2 - \phi^3\bar{\phi^3} &= -\frac{27\sigma_1^2\sigma_2^2}{4} + 27\sigma_1^3\sigma_3 - \frac{243\sigma_1\sigma_2\sigma_3}{2} + 27\sigma_2^3 + \frac{729\sigma_3^2}{4} \\ &= -\frac{27}{4}(\sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2) = -\frac{27}{4}\Delta\end{aligned}$$

Thus,

$$\begin{aligned}\phi^3 &= a + \sqrt{-\frac{27}{4}\Delta} \\ &= a + \frac{3\sqrt{3}i}{2}\sqrt{\Delta}\end{aligned}$$

which clearly implies

$$\bar{\phi^3} = a - \frac{3\sqrt{3}i}{2}\sqrt{\Delta}$$

Now, we can set up the following systems of equations:

$$\begin{aligned}x_1 + x_2 + x_3 &= \sigma_1 \\ x_1 + \omega x_2 + \omega^2 x_3 &= \lambda_1 \\ x_1 + \omega^2 x_2 + \omega x_3 &= \lambda_2\end{aligned}$$

where

$$\lambda_1 = \sqrt[3]{a + \frac{3\sqrt{3}i}{2}\sqrt{\Delta}}\,, \quad \lambda_2 = \sqrt[3]{a - \frac{3\sqrt{3}i}{2}\sqrt{\Delta}}$$

Casting the linear system of equations as a matrix equation, we get:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \sigma_1 \\ \lambda_1 \\ \lambda_2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \lambda_1 \\ \lambda_2 \end{bmatrix}$$

which implies that

$$\begin{aligned} x_1 &= \frac{1}{3}(\sigma_1 + \lambda_1 + \lambda_2) \\ x_2 &= \frac{1}{3}(\sigma_1 + \omega^2\lambda_1 + \omega\lambda_2) \\ x_2 &= \frac{1}{3}(\sigma_1 + \omega\lambda_1 + \omega^2\lambda_2) \end{aligned}$$

Therefore

$$\begin{aligned} x_1 = \frac{1}{3}\sigma_1 &+ \frac{1}{3}\sqrt[3]{\frac{1}{2}(2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3) + \frac{3\sqrt{3}i}{2}\sqrt{\sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2}} \\ &+ \frac{1}{3}\sqrt[3]{\frac{1}{2}(2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3) - \frac{3\sqrt{3}i}{2}\sqrt{\sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2}} \end{aligned}$$

$$\begin{aligned} x_2 = \frac{1}{3}\sigma_1 &+ \frac{\omega^2}{3}\sqrt[3]{\frac{1}{2}(2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3) + \frac{3\sqrt{3}i}{2}\sqrt{\sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2}} \\ &+ \frac{\omega}{3}\sqrt[3]{\frac{1}{2}(2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3) - \frac{3\sqrt{3}i}{2}\sqrt{\sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2}} \end{aligned}$$

$$\begin{aligned} x_3 = \frac{1}{3}\sigma_1 &+ \frac{\omega}{3}\sqrt[3]{\frac{1}{2}(2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3) + \frac{3\sqrt{3}i}{2}\sqrt{\sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2}} \\ &+ \frac{\omega^2}{3}\sqrt[3]{\frac{1}{2}(2\sigma_1^3 - 9\sigma_1\sigma_2 + 27\sigma_3^3) - \frac{3\sqrt{3}i}{2}\sqrt{\sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 - 4\sigma_2^3 - 27\sigma_3^2}} \end{aligned}$$

7.2 The Quartic

We will follow a similar process as for the cubic, but will focus primarily on acquiring an explicit solution and explicitly expressing the chain of solvable extensions here. Similar to the cubic, let $K = F(\sigma_1, \sigma_2, \sigma_3, \sigma_4)$ and $L = F(x_1, x_2, x_3, x_4)$. Here, we will let F be a field that contains $\mathbb{Q}$ in addition to any primitive roots of unity that is necessary (recall from Proposition 6.2.8 that appending primitive roots of unity does not affect solvability of a solvable extension) to find the roots of the universal quartic $\widetilde{f} \in K[x]$ given by

$$\widetilde{f} = x^4 - \sigma_1 x^3 + \sigma_2 x^2 - \sigma_3 x + \sigma_4$$

By Theorem 4.2.6, $\mathrm{Gal}(L/K) \simeq S_4$. One can show that a solvable chain for S_4 is

$$\{e\} \triangleleft \langle (12)(34) \rangle \triangleleft V_4 \triangleleft A_4 \triangleleft S_4$$

which have orders $1, 2, 4, 12$ and 24 respectively. Also, by the Galois Correspondence, the solvable chain yields the fixed fields

$$K_0 \subset K_1 \subset K_2 \subset K_3 \subset K_4$$

where $K_0 = K$ and $K_4 = L$. Just as in our work for the cubic, we can easily find that the fixed field of $A_4 \simeq \mathrm{Gal}(L/K_1)$ is $K(\sqrt{\Delta})$, so $K_1 = K(\sqrt{\Delta})$. In order to express the other extension fields in a similar manner so as to explicitly express the radical solvability of $K \subset L$, let $\tau = (12)(34)$. Here,

$$\langle \tau \rangle = \{(), (12)(34)\}$$

where "()" denotes the identity element. Using Theorem 6.3.1 we can conclude the sequence of field extensions is solvable; furthermore, we can infer that every subsequent field extension in the solvable chain of extension fields is Galois of prime order. In the case of L/K_3, by Corollary 5.1.6, since $\mathrm{Gal}(L/K_3) \simeq \langle \tau \rangle$, it follows that $[L : K_3] = 2$; also, notice that $\mathrm{Gal}(L/K_3) \simeq \langle \tau \rangle \simeq \mathbb{Z}/2\mathbb{Z}$, so by Lemma 6.2.9, there is some $\alpha_1 \in L$ such that ${\alpha_1}^2 \in K_3$ and $L = K_3(\alpha_1)$. We can find an explicit form of α_1 via the Lagrange resolvent as in the proof of Lemma 6.2.9. Let $\beta_1 \in L$ be given by $\beta_1 = (x_1 - x_2) - (x_3 + x_4)$; notice that $\tau(\beta_1) = (x_2 - x_1) - (x_3 + x_4) \neq \beta_1$, so $\beta_1 \notin K_3$. However, following along the construction of the Lagrange resolvent, where we set $\zeta_2 = e^{\frac{2\pi i}{2}} = -1$,

$$\begin{aligned} \alpha_1 &= \beta_1 + {\zeta_2}^{-1}\tau(\beta_1) = \beta_1 - \tau(\beta_1) \\ &= (x_1 - x_2) - (x_3 + x_4) - ((x_2 - x_1) - (x_3 + x_4)) \\ &= 2(x_1 - x_2) \end{aligned}$$

Indeed, we see that $\tau({\alpha_1}^2) = {\alpha_1}^2$, so ${\alpha_1}^2 \in K_3$ and by Lemma 6.2.9, $L = K_3(\alpha_1)$. The factor of 2 in α_1 is not necessary since $K_3(\alpha_1) = K_3(\alpha_1/2)$, so let $\phi_4 \coloneqq \alpha_1/2 = (x_1 - x_2)$. Thus, we have that

$$L = K_3(\phi_4)$$

We can follow a similar approach for lower field extensions as they are all of prime degree. Recall from Theorem 5.2.5 that

$$\mathrm{Gal}(L/K_2)/\mathrm{Gal}(L/K_3) \simeq \mathrm{Gal}(K_3/K_2)$$

which means that

$$\mathrm{Gal}(K_3/K_2) \simeq V_4/\langle\tau\rangle \simeq \mathbb{Z}/2\mathbb{Z} \simeq \langle\tau\rangle$$

This time, let $\beta_2 = (x_1 + x_2) - (x_3 + x_4)$; here, it's clear that $\beta_2 \in K_3$, but $\beta_2 \notin K_2$ since, for example, the action of $(13)(24) \in V_4$ on β_2 yields $(x_3 + x_4) - (x_1 + x_2) = -\beta_2 \neq \beta_2$. Using the Lagrange resolvent approach again, it follows

$$\begin{aligned}
\alpha_2 &= \beta_2 + {\zeta_2}^{-1}\tau(\beta_2) \\
&= (x_1 + x_2) - (x_3 + x_4) + {\zeta_2}^{-1}((x_3 + x_4) - (x_1 + x_2)) \\
&= 2(x_1 + x_2) - 2(x_3 + x_4) = 2\beta_2
\end{aligned}$$

so by Lemma 6.2.9, $K_3 = K_2(\alpha_2)$. One can easily check that ${\alpha_2}^2 \in K_2$ since for any $\sigma \in V_4$, $\sigma({\alpha_2}^2) = {\alpha_2}^2$. Furthermore, α_2 only differs from β_2 by a constant factor of 2, so letting $\phi_3 = \beta_2$, we have that ${\phi_3}^2 \in K_2$ and

$$K_3 = K_2(\phi_3)$$

Now, continuing with the same approach, let $\beta_3 = {\phi_3}^2$ where $\beta_3 \in K_2$, but $\beta_3 \notin K_1$ since, for example, the action of $(123) \in A_4$ on β_3 yields

$$(123)\beta_3 = (123)[(x_1 + x_2) - (x_3 + x_4)]^2 = [(x_2 + x_3) - (x_1 + x_4)]^2 \neq \beta_3$$

Also,

$$\mathrm{Gal}(L/K_1)/\mathrm{Gal}(L/K_2) \simeq \mathrm{Gal}(K_2/K_1) \simeq \mathbb{Z}/3\mathbb{Z}$$

and since (as one could check) $\langle(123)\rangle \simeq \mathbb{Z}/3\mathbb{Z}$, we may conclude that

$$\mathrm{Gal}(K_2/K_1) \simeq \langle(123)\rangle$$

Let $\rho = (123)$. Using the Lagrange resolvent where now, we use $\zeta_3 = e^{\frac{2\pi i}{3}} = -\frac{1}{2} + \frac{i\sqrt{3}}{2}$, we get (with the help of Mathematica),

$$\begin{aligned}
\alpha_3 &= \beta_3 + {\zeta_3}^{-1}\rho(\beta_3) + {\zeta_3}^{-2}\rho^2(\beta_3) \\
&= \beta_3 + {\zeta_3}^{2}\rho(\beta_3) + \zeta_3\rho^2(\beta_3) \\
&= 4x_1x_2 - 2i\sqrt{3}x_3x_2 - 2x_3x_2 + 2i\sqrt{3}x_4x_2 - 2x_4x_2 \\
&+ 2i\sqrt{3}x_1x_3 - 2x_1x_3 - 2i\sqrt{3}x_1x_4 - 2x_1x_4 + 4x_3x_4
\end{aligned}$$

so, again, by Lemma 6.2.9, letting $\phi_2 = \alpha_3$, $K_2 = K_1(\phi_2)$ where ${\phi_2}^3 \in K_1$ (the invariance of ${\phi_2}^3$ under the action of all elements of A_4 can be checked explicitly using Mathematica or any other algebraic software). Finally, let $\phi_1 = \sqrt{\Delta}$; now, we have $K_1 = K_0(\phi_1)$, $K_2 = K_1(\phi_2)$, $K_3 = K_2(\phi_3)$, and $K_4 = K_3(\phi_4)$ where for $i \in \{1, 2, 3, 4\}$, ${\phi_i}^p \in K_{i-1}$ where $p \in \{2, 3\}$. We can now observe that $K \subset L$ is indeed a solvable extension and that $\widetilde{f}$ is indeed solvable by radicals over K. Collecting our results, we can now say that

$$K \subset K(\phi_1) \subset K(\phi_1, \phi_2) \subset K(\phi_1, \phi_2, \phi_3) \subset K(\phi_1, \phi_2, \phi_3, \phi_4)$$

where, expressing each ϕ_i explicitly in terms of x_1, x_2, x_3 and x_4,

$$\phi_1 = \sqrt{(x_1 - x_2)^2 (x_1 - x_3)^2 (x_2 - x_3)^2 (x_1 - x_4)^2 (x_2 - x_4)^2 (x_3 - x_4)^2}$$

$$\begin{aligned}
\phi_2 &= 4x_1x_2 - 2i\sqrt{3}x_3x_2 - 2x_3x_2 + 2i\sqrt{3}x_4x_2 - 2x_4x_2 \\
&+ 2i\sqrt{3}x_1x_3 - 2x_1x_3 - 2i\sqrt{3}x_1x_4 - 2x_1x_4 + 4x_3x_4
\end{aligned}$$

$$\phi_3 = x_1 + x_2 - x_3 - x_4$$

$$\phi_4 = x_1 - x_2$$

Now, let's focus on the fields extensions

$$K \subset K_3 \subset L$$

where we have established that $\mathrm{Gal}(L/K_3) \simeq \langle\tau\rangle$ and that $\phi_3 \in K_3$ since it is invariant under the action of of $\langle\tau\rangle$. Also, since $\langle\tau\rangle \triangleleft S_4$, it is also true that $\mathrm{Gal}(L/K_3) \triangleleft \mathrm{Gal}(L/K)$, so by Theorem 5.2.4, it follows that for all $\sigma \in \mathrm{Gal}(L/K)$, $\sigma K_3 = K_3$ which means that $\sigma(\phi_3) \in K_3$. Consider $\mathrm{Orb}(S_4, \phi_3)$ which consists of the distinct elements $\{\psi_k\}_{k=1}^6$ where setting $\psi_1 = \phi_3$, we have

$$\begin{aligned}
\psi_1 &= x_1 + x_2 - x_3 - x_4 & \psi_4 &= -x_1 - x_2 + x_3 + x_4 \\
\psi_2 &= x_1 - x_2 + x_3 - x_4 & \psi_5 &= -x_1 + x_2 - x_3 + x_4 \\
\psi_3 &= x_1 - x_2 - x_3 + x_4 & \psi_6 &= -x_1 + x_2 + x_3 - x_4
\end{aligned}$$

where each $\psi_k \in K_3$. Notice that $\psi_4 = -\psi_1$, $\psi_5 = -\psi_2$, and $\psi_6 = -\psi_3$, so we can just focus on ψ_1, ψ_2, and ψ_3. Next, let $\tau_1 = \tau = (12)(34)$, $\tau_2 = (13)(24)$, and $\tau_3 = (14)(23)$. Note that for each $i \in \{1, 2, 3\}$, $\langle \tau_i \rangle \simeq \mathbb{Z}/2\mathbb{Z}$, so $\langle \tau_i \rangle \simeq \mathrm{Gal}(L/K_3)$. This means that we can conclude that each ${\psi_i}^2$ is in K_2; indeed, the invariance of each ${\psi_i}^2$ under the action of V_4 can be checked quite easily by noting that each ψ_i is of the form $(x_a + x_b) - (x_c + x_d)$ with stabalizer $\langle (ab)(cd) \rangle$ and the action of any element in V_4 not in the stabalizer yields $-\psi_i = (x_c + x_d) - (x_a + x_b)$.

By assumption, $K_1 \subset K_2$ is an algebraic extension (implicit by the fact that it is Galois), so considering ${\psi_1}^2 \in K_2$, let $h \in K_1[x]$ be its minimal polynomial so that $h({\psi_1}^2) = 0$. Recall that $\mathrm{Gal}(K_2/K_1) \simeq \langle (123) \rangle = \{(), (123), (132)\}$. By Proposition 4.1.4, for any $\sigma \in \mathrm{Gal}(K_2/K_1)$, it follows $h(\sigma({\psi_1}^2)) = 0$. Then, the action of elements in $\langle (123) \rangle$ yields

$$\begin{aligned} ()\,{\psi_1}^2 &= {\psi_1}^2 \\ (123)\,{\psi_1}^2 &= {\psi_3}^2 \\ (132)\,{\psi_1}^2 &= {\psi_2}^2 \end{aligned}$$

Since ${\psi_1}^2, {\psi_2}^2, {\psi_3}^2 \in K_2$, h splits completely over K_2 where

$$h = (x - {\psi_1}^2)(x - {\psi_2}^2)(x - {\psi_3}^2)$$

and multiplying out the linear factors, $h \in K_1[x]$ where

$$h = x^3 - \left(\psi_1^2 + \psi_2^2 + \psi_3^2\right) x^2 + \left(\psi_1^2\psi_2^2 + \psi_3^2\psi_2^2 + \psi_1^2\psi_3^2\right) x - \psi_1^2\psi_2^2\psi_3^2$$

It turns out that the coefficients of h are symmetric functions, so we have

$$h = x^3 + c_1 x^2 + c_2 x + c_3 \tag{7.1}$$

where

$$\begin{aligned} c_1 &= 8\sigma_2 - 3\sigma_1^2 \\ c_2 &= 3\sigma_1^4 - 16\sigma_2\sigma_1^2 + 16\sigma_3\sigma_1 + 16\sigma_2^2 - 64\sigma_4 \\ c_3 &= -\left(\sigma_1^3 - 4\sigma_2\sigma_1 + 8\sigma_3\right)^2 \end{aligned} \tag{7.2}$$

Thus, we can see that $h \in K[x]$. Now, considering that we can construct the system of equations

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= \sigma_1 \\ x_1 + x_2 - x_3 - x_4 &= \psi_1 \\ x_1 - x_2 + x_3 - x_4 &= \psi_2 \\ x_1 - x_2 - x_3 + x_4 &= \psi_3 \end{aligned}$$

which, when casted as a matrix equation yields

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \sigma_1 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{bmatrix}$$

so that

$$\begin{aligned} x_1 &= \frac{1}{4}(\sigma_1 + \psi_1 + \psi_2 + \psi_3) \\ x_2 &= \frac{1}{4}(\sigma_1 + \psi_1 - \psi_2 - \psi_3) \\ x_3 &= \frac{1}{4}(\sigma_1 - \psi_1 + \psi_2 - \psi_3) \\ x_4 &= \frac{1}{4}(\sigma_1 - \psi_1 - \psi_2 + \psi_3) \end{aligned} \tag{7.3}$$

we have essentially reduced the problem of solving a quartic to solving a cubic (Eq. (7.1)) where in our solutions for the cubic, we make the following substitutions: $\sigma_1 \mapsto -c_1$, $\sigma_2 \mapsto c_2$ and $\sigma_3 \mapsto -c_3$. This yields

$$\begin{aligned} {\psi_a}^2 = \frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) &+ \frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\ &+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\ &+ \frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\ &- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \end{aligned}$$

$$\begin{aligned} {\psi_b}^2 = \frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) &+ \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\ &+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\ &+ \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\ &- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \end{aligned}$$

$$\psi_c{}^2 = \frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1-i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4$$
$$+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}$$
$$+ \frac{\left(-1+i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4$$
$$- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}$$

Whichever of ${\psi_a}^2$, ${\psi_b}^2$, or ${\psi_c}^2$ corresponds to ${\psi_1}^2$, ${\psi_2}^2$, or ${\psi_3}^2$ is arbitrary up to the requirement that, according to Eq. (7.2),

$${\psi_1}^2{\psi_2}^2{\psi_3}^2 = \left(\sigma_1^3 - 4\sigma_2\sigma_1 + 8\sigma_3\right)^2$$

Thus, it follows that we may require that:

$$\psi_1\psi_2\psi_3 = \sigma_1^3 - 4\sigma_2\sigma_1 + 8\sigma_3 \tag{7.4}$$

(where we choose positive $\sqrt{-c_3}$ in Eq. (7.4); we can also set the requirement to be negative $\sqrt{-c_3}$). Suppose that ψ_a, ψ_b, and ψ_c are the positive square roots of the above expressions. Letting $\psi_1^2 = \psi_a^2$, $\psi_2^2 = \psi_b^2$, and $\psi_3^2 = \psi_c^2$, this means that we can have $\psi_1 = \psi_a$, $\psi_2 = \psi_b$, and $\psi_3 = \psi_c$; alternatively, we may have two of ψ_a, ψ_b, and ψ_c be negative where, for example, $\psi_1 = -\psi_a$, $\psi_2 = \psi_b$, and $\psi_3 = -\psi_c$. We choose the former case where all of ψ_a, ψ_b, and ψ_c are positive. Thus, by Eq. (7.3):

$$
\begin{aligned}
x_1 = \frac{\sigma_1}{4} + \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
&+ \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
&+ \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
x_2 = \frac{\sigma_1}{4} + \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
- \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
- \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
x_3 = \frac{\sigma_1}{4} - \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
&+ \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
&- \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
x_4 = \frac{\sigma_1}{4} - \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
- \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2} \\
+ \frac{1}{4}\Bigg(\frac{1}{3}\left(3\sigma_1^2 - 8\sigma_2\right) + \frac{\left(-1 - i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(&128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&+ \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3} \\
&+ \frac{\left(-1 + i\sqrt{3}\right)}{2}\frac{1}{3\sqrt[3]{2}}\Bigg(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4 \\
&- \sqrt{4\left(-16\sigma_2^2 + 48\sigma_1\sigma_3 - 192\sigma_4\right)^3 + \left(128\sigma_2^3 - 576\sigma_1\sigma_3\sigma_2 - 4608\sigma_4\sigma_2 + 1728\sigma_3^2 + 1728\sigma_1^2\sigma_4\right)^2}\Bigg)^{1/3}\Bigg)^{1/2}
\end{aligned}
$$

8

The Quintic

This chapter primarily addresses material from the paper "Solving Solvable Quintics" by Dummit [4]. Just like in Chapter 7, computational assistance is provided by Mathematica.

8.1 Unsolvability of the Universal Quintic

Suppose F is a field that contains $\mathbb{Q}$ in addition to any primitive roots of unity that is necessary. For the Universal Cubic and Quartic, solvability of the Galois groups eventually led us to construct a linear system of the form

$$M\vec{x} = \vec{b}$$

that led to an explicit solution in terms of the elementary symmetric polynomials. Here, M is a $n \times n$ matrix with $n = \deg \widetilde{f}$, $\vec{x} = (x_1, \ldots, x_n)^T$, and $\vec{b} = (b_1, \ldots, b_n)^T$ with $b_1 = \sigma_1$. For $i \in \{2, \ldots, n\}$, the b_i consisted of nested radical expressions of elements in $F(x_1, \ldots, x_n)$ expressible as nested radicals of elements in $F(\sigma_1, \ldots, \sigma_n)$; each b_i was some kth root ($k = 3$ in the case of solving the cubic and $k = 2$ in the case of solving the quartic) of elements in the orbit of some element in $F(x_1, \ldots, x_n)$ under the action of S_n.

For instance, recall that for the cubic, we derived the element ${\lambda_1}^3 = \phi^3$ which had an orbit order of 2 under the action of elements in S_3; i.e. for all $\sigma \in S_3$,

$$\sigma {\lambda_1}^3 \in \{{\lambda_1}^3, {\lambda_2}^3\}$$

where ${\lambda_2}^3 = \bar{\phi}^3$. Here, we would have $b_1 = \sigma_1$, $b_2 = \lambda_1$, and $b_3 = \lambda_3$ where b_2 and b_3 are cube roots of elements in the orbit of ${\lambda_1}^3 \in F(x_1, x_2, x_3)$ under the action of S_3.

Likewise, recall that for the quartic, we derived the element ${\psi_1}^2$ which had an orbit order of 3; i.e. for all $\sigma \in S_4$,

$$\sigma {\psi_1}^2 \in \{{\psi_1}^2, {\psi_2}^2, {\psi_3}^2\}$$

Here, we would have $b_1 = \sigma_1$, $b_2 = \psi_1$, $b_3 = \psi_2$, $b_4 = \psi_3$ where b_2, b_3, and b_4 are square roots of elements in the orbit of ${\psi_1}^2 \in F(x_1, x_2, x_3, x_4)$ under the action of S_4.

Now, let us suppose we can find an element $\rho_1 \in F(x_1, x_2, x_3, x_4, x_5)$ that would lead us to solving the universal quintic like ${\lambda_1}^3$ for the cubic and ${\psi_1}^2$ for the quartic. Provided that $b_1 = \sigma_1$, it should be the case that ρ_1 has an orbit order of 4; i.e. for all $\sigma \in S_5$,

$$\sigma\rho_1 \in \{\rho_1, \rho_2, \rho_3, \rho_4\}$$

However, this poses an issue. In the case of the cubic, an orbit order of 2 for ${\lambda_1}^3$ implies that its stabilizer has a group order of 3; for the quartic, an orbit order of 3 for ${\psi_1}^2$ implies that its stabilizer has a group order of 8. This means that up to isomorphism, the stabilizer of ${\lambda_1}^3$ is $\mathbb{Z}/3\mathbb{Z}$ and the stabilizer of ${\psi_1}^2$ is D_8 (the dihedral group) which are subgroups of S_3 and S_4 respectively.

Since ρ_1 has an orbit order of 4, this means that its stabilizer should have a group order of 30. In other words, finding an element $\rho_1 \in F(x_1, x_2, x_3, x_4, x_5)$ with an orbit order of 4 under the action of S_5 implies the existence of a subgroup $H \subset S_5$ such that $|H| = 30$. This is a contradiction because **no such subgroup exists**. This means we cannot construct a linear system in the same manner as we did for the cubic and quartic.

Furthermore, it can be shown that A_5 is a simple group which means that A_5 has no nontrivial normal subgroups. This leads us to conclude something unfortunate. Let $K = F(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5)$ and $L = F(x_1, x_2, x_3, x_4, x_5)$. It can be shown that S_5 is not solvable since the only normal chain of S_5 is

$$\{e\} \triangleleft A_5 \triangleleft S_5$$

yet $\{A_5 : \{e\}\} = 60$ which is obviously not prime. Thus, by Theorem 6.3.1, the universal extension $K \subset L$ is not a solvable extension, so we conclude that the universal quintic

$$\widetilde{f} = x^5 - \sigma_1 x^4 + \sigma_2 x^3 - \sigma_3 x^2 + \sigma_4 x - \sigma_5$$

is not solvable by radicals.

8.2 Solvable Quintics

It is important to keep in mind that it is only the universal quintic with coefficients over $F(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5)$ that is unsolvable by radicals. The Galois group of $F(x_1, x_2, x_3, x_4, x_5)$ over $F(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5)$ is S_5 which is set in stone by definition. On the other hand, if we consider an arbitrary irreducible quintic polynomial $f \in F[x]$ such that

$$f = x^5 + px^4 + qx^3 + rx^2 + sx + t \tag{8.1}$$

where $p, q, r, s, t \in F$ with roots $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$, then the Galois group of $F(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ over F can be any transitive subgroup of S_5 according to Proposition 4.2.3; here, let's relabel L as $L = F(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$.

This distinction is important because for instance, we can easily see that the polynomial $x^5 - 1 \in \mathbb{Q}[x]$ is solvable by radicals, so its Galois group is not S_5 itself (which would make it unsolvable). Indeed, the splitting field of $x^5 - 1$ is $\mathbb{Q}(1, \zeta_5, {\zeta_5}^2, {\zeta_5}^3, {\zeta_5}^4)$ where ζ_5 is a primitive 5th root of unity, but it is easy to see that this splitting field is just $\mathbb{Q}(\zeta_5)$. Clearly, the minimal polynomial of ζ_5 is just $x^5 - 1$, so by Theorem 2.3.4, $[\mathbb{Q}(\zeta_5) : \mathbb{Q}] = 5$ which means by Corollary 5.1.6, that $|\mathrm{Gal}(\mathbb{Q}(\zeta_5)/\mathbb{Q})| = 5$; from this, we can see that $\mathrm{Gal}(\mathbb{Q}(\zeta_5)/\mathbb{Q}) \simeq \mathbb{Z}/5\mathbb{Z}$.

Considering that a solvable subgroup of S_5 must be transitive for an irreducible polynomial over F, according to Cox [5], our options are limited. The following proposition is adapted from "Galois Theory" on page 368 as **Lemma 13.2.1**:

Proposition 8.2.1. *Let $G \subset S_5$ be a subgroup. Then the following are equivalent:*
(a) *G is transitive*
(b) *$|G|$ is divisible by 5*

This helps rule out a lot of subgroups of S_5 as a solvable transitive subgroup. First, we consider the allowed orders of subgroups of S_5 which, by Lagrange's theorem, should have group orders that are divisors of 120: 120, 60, 24, 20, 12, 10, 8, 6, 5, 4, 3, 2, 1 (it is well known and can be shown by the Sylow theorems that not every divisor of 120 corresponds to a subgroup of S_5, so we did not consider those here; e.g. recall that there is no subgroup of order 30 as we pointed out in Section 8.1). Obviously, orders less than 5 can be ruled out along with 120 and 60 since these orders correspond to S_5 and A_5 respectively (which, again, are unsolvable). Now, by Proposition 8.2.1, we can rule out orders of 24, 12, 8, and 6; this leaves us with allowed group orders of 20, 10, and 5.

This means that if a transitive subgroup of S_5 is solvable, up to isomorphism, it is one of $\mathrm{AGL}(1, \mathbb{F}_5)$, $\mathrm{AGL}(1, \mathbb{F}_5) \cap A_5$, or $\mathbb{Z}/5\mathbb{Z}$ (AGL stands for "Affine Linear Group" according to notation of [5]; it is notated as F_{20} and called the Frobenius Group of order 20 according to [4]); the first subgroup is the group $\langle (12345), (1243) \rangle$ and the second subgroup can be identified up to isomorphism with D_{10}, the dihedral group of order 10. In fact, it can be shown all of these subgroups together with the identity $\{e\}$ form a solvable chain:

$$\{e\} \triangleleft \mathbb{Z}/5\mathbb{Z} \triangleleft D_{10} \triangleleft \mathrm{AGL}(1, \mathbb{F}_5) \tag{8.2}$$

We will assume that $\mathrm{Gal}(L/F) \simeq \mathrm{AGL}(1, \mathbb{F}_5)$, so by the Galois correspondence along with Theorem 6.3.1, we must have the solvable extension chain

$$F \subset K \subset M \subset L \tag{8.3}$$

where K and M are intermediate fields that we have yet to identify. Considering the group indices in Eq. (8.2) (5, 2, 2 from left to right), we will only need a primitive fifth root of unity appended to $\mathbb{Q}$ since the nontrivial primitive second root of unity is just -1 which is already in $\mathbb{Q}$; thus, we let $F = \mathbb{Q}(\zeta_5)$. Moreover, we can identify K fairly easily based on our work in Chapter 7 since

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

which, by taking note of Eq. (8.2), implies that $|\mathrm{Gal}(K/F)| = 2$ and hence, that $[K : F] = 2$. Thus, $F \subset K$ is a quadratic extension which should be the discriminant appended to F, except we are no longer working with the universal extension in this case so instead of $\sqrt{\Delta}$, we consider $\sqrt{\Delta(f)}$ (from Eq. (A.8)). Here, we will use the notation $\Delta_f \coloneqq \Delta(f)$ to make the notation cleaner. Of course, $\sqrt{\Delta_f} \neq 0$ since we are assuming that f is separable (F is of characteristic zero), so by Theorem 3.4.3, $\Delta(f) \neq 0$. Therefore, it follows that

$$K = F\big(\sqrt{\Delta_f}\big)$$

Moving forward, let $\zeta = \zeta_5$ from the beginning of this section; we note that in [4], Dummit introduces $r_1, r_2, r_3, r_4 \in L$ given by

$$\begin{aligned} r_1 &= \alpha_1 + \zeta\alpha_2 + \zeta^2\alpha_3 + \zeta^3\alpha_4 + \zeta^4\alpha_5 \\ r_2 &= \alpha_1 + \zeta^2\alpha_2 + \zeta^4\alpha_3 + \zeta^1\alpha_4 + \zeta^3\alpha_5 \\ r_3 &= \alpha_1 + \zeta^3\alpha_2 + \zeta\alpha_3 + \zeta^4\alpha_4 + \zeta^2\alpha_5 \\ r_4 &= \alpha_1 + \zeta^4\alpha_2 + \zeta^3\alpha_3 + \zeta^2\alpha_4 + \zeta\alpha_5 \end{aligned} \tag{8.4}$$

which are the Lagrange resolvents from Eq. (6.1) where r_1 corresponds to $i = 4$, r_2 corresponds to $i = 3$, r_3 corresponds to $i = 2$, and r_4 corresponds to $i = 1$. By the same approach in Chapter 7, we can conclude that for any $j \in \{1, 2, 3, 4\}$, $L = M(r_j)$ where ${r_j}^5 \in M$ since (as one can check easily using Mathematica), for all $\sigma \in \mathrm{Gal}(L/M) \simeq \langle(12345)\rangle$, $\sigma {r_j}^5 = {r_j}^5$. Following the notation of [4], let $R_1 = {r_1}^5$, so that

$$R_1 = l_0 + \zeta l_1 + \zeta^2 l_2 + \zeta^3 l_3 + \zeta^4 l_4$$

where

$$\begin{aligned} l_0 = {} & \alpha_1^5 + 20\alpha_3\alpha_4\alpha_1^3 + 20\alpha_2\alpha_5\alpha_1^3 + 30\alpha_2\alpha_3^2\alpha_1^2 + 30\alpha_3\alpha_5^2\alpha_1^2 + 30\alpha_2^2\alpha_4\alpha_1^2 + 30\alpha_4^2\alpha_5\alpha_1^2 + 20\alpha_2\alpha_4^3\alpha_1 \\ & + 20\alpha_4\alpha_5^3\alpha_1 + 30\alpha_3^2\alpha_4^2\alpha_1 + 30\alpha_2^2\alpha_5^2\alpha_1 + 20\alpha_2^3\alpha_3\alpha_1 + 20\alpha_3^3\alpha_5\alpha_1 + 120\alpha_2\alpha_3\alpha_4\alpha_5\alpha_1 \\ & + \alpha_2^5 + \alpha_3^5 + \alpha_4^5 + \alpha_5^5 + 20\alpha_2\alpha_3\alpha_5^3 + 30\alpha_2^2\alpha_3\alpha_4^2 + 30\alpha_2\alpha_4^2\alpha_5^2 + 30\alpha_3^2\alpha_4\alpha_5^2 \\ & + 20\alpha_2\alpha_3^3\alpha_4 + 20\alpha_3\alpha_4^3\alpha_5 + 30\alpha_2^2\alpha_3^2\alpha_5 + 20\alpha_2^3\alpha_4\alpha_5 \end{aligned}$$

$$\begin{aligned} l_1 = {} & 5\alpha_2\alpha_1^4 + 10\alpha_4^2\alpha_1^3 + 20\alpha_3\alpha_5\alpha_1^3 + 10\alpha_3^3\alpha_1^2 + 30\alpha_4\alpha_5^2\alpha_1^2 + 60\alpha_2\alpha_3\alpha_4\alpha_1^2 + 30\alpha_2^2\alpha_5\alpha_1^2 + 5\alpha_5^4\alpha_1 \\ & + 20\alpha_3\alpha_4^3\alpha_1 + 30\alpha_2^2\alpha_3^2\alpha_1 + 60\alpha_2\alpha_3\alpha_5^2\alpha_1 + 20\alpha_2^3\alpha_4\alpha_1 + 60\alpha_2\alpha_4^2\alpha_5\alpha_1 + 60\alpha_3^2\alpha_4\alpha_5\alpha_1 \\ & + 10\alpha_2^2\alpha_4^3 + 10\alpha_3^2\alpha_5^3 + 20\alpha_2\alpha_4\alpha_5^3 + 30\alpha_2\alpha_3^2\alpha_4^2 + 10\alpha_2^3\alpha_5^2 + 30\alpha_3\alpha_4^2\alpha_5^2 \\ & + 5\alpha_2^4\alpha_3 + 5\alpha_3^4\alpha_4 + 5\alpha_4^4\alpha_5 + 20\alpha_2\alpha_3^3\alpha_5 + 60\alpha_2^2\alpha_3\alpha_4\alpha_5 \end{aligned}$$

$$\begin{aligned} l_2 = {} & 5\alpha_3\alpha_1^4 + 10\alpha_2^2\alpha_1^3 + 20\alpha_4\alpha_5\alpha_1^3 + 10\alpha_5^3\alpha_1^2 + 30\alpha_2\alpha_4^2\alpha_1^2 + 30\alpha_3^2\alpha_4\alpha_1^2 + 60\alpha_2\alpha_3\alpha_5\alpha_1^2 + 5\alpha_4^4\alpha_1 \\ & + 20\alpha_2\alpha_3^3\alpha_1 + 30\alpha_3^2\alpha_5^2\alpha_1 + 60\alpha_2\alpha_4\alpha_5^2\alpha_1 + 60\alpha_2^2\alpha_3\alpha_4\alpha_1 + 20\alpha_2^3\alpha_5\alpha_1 + 60\alpha_3\alpha_4^2\alpha_5\alpha_1 \\ & + 5\alpha_2\alpha_5^4 + 20\alpha_2\alpha_3\alpha_4^3 + 20\alpha_3\alpha_4\alpha_5^3 + 10\alpha_2^3\alpha_3^2 + 10\alpha_3^3\alpha_4^2 + 10\alpha_4^3\alpha_5^2 + 30\alpha_2^2\alpha_3\alpha_5^2 \\ & + 5\alpha_2^4\alpha_4 + 5\alpha_3^4\alpha_5 + 30\alpha_2^2\alpha_4^2\alpha_5 + 60\alpha_2\alpha_3^2\alpha_4\alpha_5 \end{aligned}$$

$$
\begin{aligned}
l_3 = {} & 5\alpha_4\alpha_1^4 + 10\alpha_5^2\alpha_1^3 + 20\alpha_2\alpha_3\alpha_1^3 + 10\alpha_2^3\alpha_1^2 + 30\alpha_3\alpha_4^2\alpha_1^2 + 30\alpha_3^2\alpha_5\alpha_1^2 + 60\alpha_2\alpha_4\alpha_5\alpha_1^2 + 5\alpha_3^4\alpha_1 \\
& + 20\alpha_2\alpha_5^3\alpha_1 + 30\alpha_2^2\alpha_4^2\alpha_1 + 60\alpha_3\alpha_4\alpha_5^2\alpha_1 + 60\alpha_2\alpha_3^2\alpha_4\alpha_1 + 20\alpha_4^3\alpha_5\alpha_1 + 60\alpha_2^2\alpha_3\alpha_5\alpha_1 \\
& + 5\alpha_2\alpha_4^4 + 5\alpha_3\alpha_5^4 + 10\alpha_2^2\alpha_3^3 + 10\alpha_3^2\alpha_4^3 + 10\alpha_4^2\alpha_5^3 + 30\alpha_2\alpha_3^2\alpha_5^2 + 30\alpha_2^2\alpha_4\alpha_5^2 + 20\alpha_2^3\alpha_3\alpha_4 \\
& + 5\alpha_2^4\alpha_5 + 60\alpha_2\alpha_3\alpha_4^2\alpha_5 + 20\alpha_3^3\alpha_4\alpha_5
\end{aligned}
$$

$$
\begin{aligned}
l_4 = {} & 5\alpha_5\alpha_1^4 + 10\alpha_3^2\alpha_1^3 + 20\alpha_2\alpha_4\alpha_1^3 + 10\alpha_4^3\alpha_1^2 + 30\alpha_2\alpha_5^2\alpha_1^2 + 30\alpha_2^2\alpha_3\alpha_1^2 + 60\alpha_3\alpha_4\alpha_5\alpha_1^2 + 5\alpha_2^4\alpha_1 \\
& + 20\alpha_3\alpha_5^3\alpha_1 + 60\alpha_2\alpha_3\alpha_4^2\alpha_1 + 30\alpha_4^2\alpha_5^2\alpha_1 + 20\alpha_3^3\alpha_4\alpha_1 + 60\alpha_2\alpha_3^2\alpha_5\alpha_1 + 60\alpha_2^2\alpha_4\alpha_5\alpha_1 \\
& + 5\alpha_2\alpha_3^4 + 5\alpha_3\alpha_4^4 + 5\alpha_4\alpha_5^4 + 10\alpha_2^2\alpha_5^3 + 10\alpha_2^3\alpha_4^2 + 10\alpha_3^3\alpha_5^2 + 60\alpha_2\alpha_3\alpha_4\alpha_5^2 + 30\alpha_2^2\alpha_3^2\alpha_4 \\
& + 20\alpha_2\alpha_4^3\alpha_5 + 30\alpha_3^2\alpha_4^2\alpha_5 + 20\alpha_2^3\alpha_3\alpha_5
\end{aligned}
$$

It can be checked that for all $\sigma \in \mathrm{AGL}(1, \mathbb{F}_5)$, that $\sigma l_0 = l_0$, so it follows that $l_0 \in F$. Furthermore, it can be checked that

$$
l_0 + l_1 + l_2 + l_3 + l_4 = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5)^5 \tag{8.5}
$$

and for $i \in \{1, 2, 3, 4\}$, it can also be checked that for all $\sigma \in \mathrm{Gal}(L/M) \simeq \langle (12345) \rangle$, $\sigma l_i = l_i$, so $l_i \in M$. Furthermore, under the action of $\mathrm{AGL}(1, \mathbb{F}_5)$, R_1 yields an orbit of order 4 which consists of

$$
\begin{aligned}
R_1 &= l_0 + \zeta l_1 + \zeta^2 l_2 + \zeta^3 l_3 + \zeta^4 l_4 \\
R_2 &= l_0 + \zeta l_3 + \zeta^2 l_1 + \zeta^3 l_4 + \zeta^4 l_2 \\
R_3 &= l_0 + \zeta l_2 + \zeta^2 l_4 + \zeta^3 l_1 + \zeta^4 l_3 \\
R_4 &= l_0 + \zeta l_4 + \zeta^2 l_3 + \zeta^3 l_2 + \zeta^4 l_1
\end{aligned} \tag{8.6}
$$

Here, the R_i are indexed such that $R_i = {r_i}^5$ where the r_i are from Eq. (8.4). Dummit indicates that $\mathrm{Gal}(M/F) \simeq \langle (2354) \rangle \simeq \mathbb{Z}/4\mathbb{Z}$; since $[K : F] = 2$, it should be the case that $[M : K] = 2$. Since $\mathrm{Gal}(M/K) \triangleleft \mathrm{Gal}(M/F) \simeq \langle (2354) \rangle$, we may conclude that

$$
\mathrm{Gal}(M/K) \simeq \langle (25)(34) \rangle = \{(), (25)(34)\}
$$

In order to determine the element $\phi \in M$ such that $M = K(\phi)$ and $\phi^2 \in K$, we will again consider the Lagrange resolvent on l_1; it is the case that $l_1 \notin K$ since it can be checked that the action of $(25)(34)$ on l_1 is l_4, so by Eq. (6.1) (where the second primitive root of unity is just -1):

$$
\phi = l_1 - l_4
$$

Indeed, since $(25)(34)$ takes l_1 to l_4 and vice-versa, it is apparent that ϕ^2 is invariant under the action of $\langle (25)(34) \rangle$ so $\phi^2 \in K$. Now, we have $K = F(\sqrt{\Delta_f})$ where $\Delta_f \in F$, $M = K(\phi)$ where $\phi^2 \in K$, and (choosing $j = 1$ from Eq. (8.4)) $L = M(r_1)$ where ${r_1}^5 \in M$; we can now

observe that $F \subset L$ is a solvable extension and that f is indeed solvable by radicals over F. Thus, our extension chain can be expressed as

$$F \subset F(\epsilon_1) \subset F(\epsilon_1, \epsilon_2) \subset F(\epsilon_1, \epsilon_2, \epsilon_3)$$

where

$$\begin{aligned}
\epsilon_1 &= \sqrt{\Delta_f} \\
\epsilon_2 &= l_1 - l_4 \\
\epsilon_3 &= \alpha_1 + \zeta\alpha_2 + \zeta^2\alpha_3 + \zeta^3\alpha_4 + \zeta^4\alpha_5
\end{aligned}$$

Just like the action of (25)(34) on l_1, it can also be shown that its action on l_2 is l_3. Let $h_1 \in K[x]$ and $h_2 \in K[x]$ be the minimal polynomials of l_1 and l_2 respectively. By Proposition 4.1.4, it should be the case that for all $\sigma \in \mathrm{Gal}(M/K)$, $h_1(\sigma(l_1)) = 0$ and $h_2(\sigma(l_2)) = 0$ which means that over M,

$$\begin{aligned}
h_1(x) &= (x - l_1)(x - l_4) \\
h_2(x) &= (x - l_2)(x - l_3)
\end{aligned}$$

and hence, over K

$$\begin{aligned}
h_1(x) &= x^2 - (l_1 + l_4)x + l_1l_4 \\
h_2(x) &= x^2 - (l_2 + l_3)x + l_2l_3
\end{aligned} \tag{8.7}$$

so $l_1 + l_4$ and l_1l_4 along with $l_2 + l_3$, and l_2l_3 are elements of K. Also, it can be checked that the action of $\mathrm{Gal}(M/F)$ yields that $l_1 + l_4$ and $l_2 + l_3$ are conjugate in M and hence, in K; similarly, the action also yields that l_1l_4 and l_2l_3 are conjugate in K. Since for all $k \in K$, $k = a + b\sqrt{\Delta_f}$ where $a, b \in F$, using Dummit's notation, we can express these conjugate pairs as

$$\begin{aligned}
l_1 + l_4 &= -\big(T_1 + T_2\sqrt{\Delta_f}\big) \\
l_2 + l_3 &= -\big(T_1 - T_2\sqrt{\Delta_f}\big) \\
l_1l_4 &= T_3 + T_4\sqrt{\Delta_f} \\
l_2l_3 &= T_3 - T_4\sqrt{\Delta_f}
\end{aligned} \tag{8.8}$$

where $T_1, T_2, T_3, T_4 \in F$. It follows that

$$\begin{aligned}
T_1 &= -\frac{1}{2}(l_1 + l_4 + l_2 + l_3) \\
T_2 &= -\frac{1}{2\sqrt{\Delta_f}}(l_1 + l_4 - l_2 - l_3) \\
T_3 &= \frac{1}{2}(l_1l_4 + l_2l_3) \\
T_4 &= \frac{1}{2\sqrt{\Delta_f}}(l_1l_4 - l_2l_3)
\end{aligned} \tag{8.9}$$

Next, we consider $\theta \in L$ as defined by Dummit in [4]:

$$\begin{aligned}\theta = \alpha_3\alpha_4\alpha_1^2 + \alpha_2\alpha_5\alpha_1^2 + \alpha_2\alpha_4^2\alpha_1 + \alpha_4\alpha_5^2\alpha_1 + \alpha_2^2\alpha_3\alpha_1 \\ + \alpha_3^2\alpha_5\alpha_1 + \alpha_2\alpha_3\alpha_5^2 + \alpha_2\alpha_3^2\alpha_4 + \alpha_3\alpha_4^2\alpha_5 + \alpha_2^2\alpha_4\alpha_5\end{aligned} \tag{8.10}$$

It can be shown that for all $\sigma \in \mathrm{AGL}(1, \mathbb{F}_5)$, $\sigma\theta = \theta$, so $\theta \in F$. The way to proceed from here in order to acquire an explicit solution for a solvable quintic is to note that for each $i \in 1, 2, 3, 4$, each T_i has a unique orbit of order 6 under the action of S_5. It is also the case for θ that under the action of S_5, θ forms an orbit of order 6. We will use Dummit's definition for the elements in the orbit of θ via S_5; that is,

$$\begin{aligned}\theta_1 &= \theta, & \theta_2 &= (123)\theta, & \theta_3 &= (132)\theta \\ \theta_4 &= (12)\theta, & \theta_5 &= (23)\theta, & \theta_6 &= (13)\theta\end{aligned}$$

The key here is Eq. (8.10); note how our assumption of solvability of Eq. (8.1) by radicals over F necessarily entailed that $\theta \in F$. Dummit [4] (in his paper "Solving Solvable Quintics" on page 389 as **Theorem 1**) and Cox [5] (in his book "Galois Theory" on page 377 as **Corollary 13.2.11**) demonstrate proofs that the converse also holds true; although Cox uses a different definition of θ from Eq. (8.10), his definition of θ also has a stabilizer of $\mathrm{AGL}(1, \mathbb{F}_5)$ which is the important idea here.

We will also demonstrate explicitly why the converse should also be true. From Section 8.1, we noted that it is impossible to set up a linear system based on an orbit of order 4 when solving the quintic. However, it is possible to set up a linear system based on an orbit of order 6 since this entails the existence of a subgroup of order 20, which we should know exists by now. Of course, this might not be very helpful considering the linear systems we have had for the cubic and quartic involved orbits of strictly lower orders than the degree of the target polynomial.

However, Dummit cleverly exploits the fact that θ, T_1, T_2, T_3, and T_4 are in F. Since F is closed under addition and multiplication, we should be able to express each T_i or any element of F as a linear combination of powers of θ; e.g.

$$T_i = c_{i0} + c_{i1}\theta + c_{i2}\theta^2 + c_{i3}\theta^3 + c_{i4}\theta^4 + c_{i5}\theta^5 \tag{8.11}$$

Here, we could express each T_i as a power series up to any power of θ, but we have chosen to express it up to a power of 5 because we also want to take advantage of the fact that both T_i and θ have an orbit order of 6 which will allow us to construct the following linear system:

$$\begin{aligned}T_i &= c_{i0} + c_{i1}\theta_1 + c_{i2}{\theta_1}^2 + c_{i3}{\theta_1}^3 + c_{i4}{\theta_1}^4 + c_{i5}{\theta_1}^5 \\ (123)T_i &= c_{i0} + c_{i1}\theta_2 + c_{i2}{\theta_2}^2 + c_{i3}{\theta_2}^3 + c_{i4}{\theta_2}^4 + c_{i5}{\theta_2}^5 \\ (132)T_i &= c_{i0} + c_{i1}\theta_3 + c_{i2}{\theta_3}^2 + c_{i3}{\theta_3}^3 + c_{i4}{\theta_3}^4 + c_{i5}{\theta_3}^5 \\ (12)T_i &= c_{i0} + c_{i1}\theta_4 + c_{i2}{\theta_4}^2 + c_{i3}{\theta_4}^3 + c_{i4}{\theta_4}^4 + c_{i5}{\theta_4}^5 \\ (23)T_i &= c_{i0} + c_{i1}\theta_5 + c_{i2}{\theta_5}^2 + c_{i3}{\theta_5}^3 + c_{i4}{\theta_5}^4 + c_{i5}{\theta_5}^5 \\ (13)T_i &= c_{i0} + c_{i1}\theta_6 + c_{i2}{\theta_6}^2 + c_{i3}{\theta_6}^3 + c_{i4}{\theta_6}^4 + c_{i5}{\theta_6}^5\end{aligned} \tag{8.12}$$

The goal here is to solve for each c_{ij} for each $(i,j) \in \{1,2,3,4\} \times \{0,1,2,3,4,5\}$. Considering Eq. (8.11), each T_i would be expressed purely as an element of F under the assumption that θ is in F. Also, when we solve both quadratic equations from Eq. (8.7) in terms of Eq. (8.8), we will get

$$l_{\pm} = \frac{1}{2}\Bigg(-(T_1 + T_2\sqrt{\Delta_f}) \pm \sqrt{(T_1 + T_2\sqrt{\Delta_f})^2 - 4(T_3 + T_4\sqrt{\Delta_f})}\Bigg) \tag{8.13}$$

$$l^*_{\pm} = \frac{1}{2}\Bigg(-(T_1 - T_2\sqrt{\Delta_f}) \pm \sqrt{(T_1 - T_2\sqrt{\Delta_f})^2 - 4(T_3 - T_4\sqrt{\Delta_f})}\Bigg) \tag{8.14}$$

where $l_1, l_4 \in \{l_+, l_-\}$ and $l_2, l_3 \in \{l^*_+, l^*_-\}$; the choice in each case here is decided by an ordering condition which, according to Dummit, must satisfy

$$(l_1 - l_4)(l_2 - l_3) = \mathcal{O}\sqrt{\Delta_f} \tag{8.15}$$

for some fixed $\mathcal{O} \in F$. One can check that the left-hand side of Eq. (8.15) yields an orbit order of 12 under the action of S_5; of course, by Proposition A.4.7, $\sqrt{\Delta_f}$ yields an orbit order of 2 under the action of S_5. This must mean that $\mathcal{O}$ yields an orbit order of 6. Thus, just like Eq. (8.12), we can construct

$$\begin{aligned}
\mathcal{O} &= d_0 + d_1\theta_1 + d_2{\theta_1}^2 + d_3{\theta_1}^3 + d_4{\theta_1}^4 + d_5{\theta_1}^5 \\
(123)\mathcal{O} &= d_0 + d_1\theta_2 + d_2{\theta_2}^2 + d_3{\theta_2}^3 + d_4{\theta_2}^4 + d_5{\theta_2}^5 \\
(132)\mathcal{O} &= d_0 + d_1\theta_3 + d_2{\theta_3}^2 + d_3{\theta_3}^3 + d_4{\theta_3}^4 + d_5{\theta_3}^5 \\
(12)\mathcal{O} &= d_0 + d_1\theta_4 + d_2{\theta_4}^2 + d_3{\theta_4}^3 + d_4{\theta_4}^4 + d_5{\theta_4}^5 \\
(23)\mathcal{O} &= d_0 + d_1\theta_5 + d_2{\theta_5}^2 + d_3{\theta_5}^3 + d_4{\theta_5}^4 + d_5{\theta_5}^5 \\
(13)\mathcal{O} &= d_0 + d_1\theta_6 + d_2{\theta_6}^2 + d_3{\theta_6}^3 + d_4{\theta_6}^4 + d_5{\theta_6}^5
\end{aligned} \tag{8.16}$$

After finding the c_{ij} and d_j for $j \in \{0,1,2,3,4,5\}$ in Eq. (8.12) and *Eq.* (8.16), since $\theta \in F$, each T_i for $i \in \{1,2,3,4\}$ and $\mathcal{O}$ will be expressible in terms of the elementary symmetric polynomials evaluated at the roots of f hence, they will be expressible in terms of the coefficients of Eq. (8.1). It is the fact that $\theta \in F$ that allows us to move forward from here towards a radical solution. If not, then we cannot use the fundamental theorem of symmetric polynomials to express $\mathcal{O}$ and the T_i in terms of the coefficients of Eq. (8.1).

Before moving on, we should note first that both Eq. (8.12) and Eq. (8.16) are computationally heavy for an arbitrary solvable quintic. Thankfully, Dummit has these values explicitly listed in the appendix of "Solving Solvable Quintics". Even so, the expressions for $\mathcal{O}$ and the T_i are monstrous and just formatting these expressions to display here seems to be a formidable task on its own. Instead, we will take advantage of the fact that any quintic (so long as we are working with a base field of characteristic zero) can be transformed via a a Tschirnhaus transformation into what is known as the Bring—Jerrard normal form:

$$f = x^5 + ax + b \tag{8.17}$$

We will assume that for $i \in \{1,2,3,4,5\}$, the α_i are roots for Eq. (8.17). It also follows that $\Delta_f = 256a^5 + 3125b^4$. Fortunately, Dummit has these expressions for Eq. (8.17) also listed out in his paper; we provide them here:

$$\begin{aligned}\mathcal{O} = \frac{1}{256a^5+3125b^4}(&-1291500a^3\theta^2 - 399500a^2\theta^3 - 2280000a^4\theta - 1036800a^5 \\ &- 76625a\theta^4 + 488281 25b^4 - 16100\theta^5)\end{aligned} \tag{8.18}$$

$$\begin{aligned}T_1 &= \frac{1}{50b^3}(416a^3\theta^2 + 112a^2\theta^3 + 768a^4\theta + 512a^5 + 24a\theta^4 - 15625b^4 + 4\theta^5) \\ T_2 &= \frac{2480a^3\theta^2 + 760a^2\theta^3 + 4480a^4\theta + 3840a^5 + 140a\theta^4 - 78125b^4 + 30\theta^5}{512a^5b + 6250b^5} \\ T_3 &= \frac{1}{2b^2}(-21260a^3\theta^2 - 5980a^2\theta^3 - 34240a^4\theta - 18880a^5 - 1255a\theta^4 \\ &\qquad + 781250b^4 - 240\theta^5) \\ T_4 &= \frac{11500a^3\theta^2 + 3250a^2\theta^3 + 25000a^4\theta + 68800a^5 + 375a\theta^4 + 100\theta^5}{512a^5 + 6250b^4}\end{aligned} \tag{8.19}$$

Since we are dealing with the simplified quintic from Eq. (8.17), this implies via the elementary symmetric polynomial σ_1 evaluated at the roots of f that

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = 0 \tag{8.20}$$

which implies along with Eq. (8.5), Eq. (8.13) and Eq. (8.14),

$$l_0 = -(l_1 + l_2 + l_3 + l_4) = 2T_1$$

Now, for $i \in \{1,2,3,4\}$, let r_i be the real 5th root of the polynomial

$$p_i = x^5 - R_i$$

where the R_i are given in Eq. (8.6). Using the r_i and Eq. (8.20) together with Eq. (8.4) implies that

$$\begin{aligned}
\alpha_1 &= \frac{1}{5}(r_1 + r_2 + r_3 + r_4) \\
\alpha_2 &= \frac{1}{5}(\zeta r_1 + \zeta^2 r_2 + \zeta^3 r_3 + \zeta^4 r_4) \\
\alpha_3 &= \frac{1}{5}(\zeta^2 r_1 + \zeta^4 r_2 + \zeta r_3 + \zeta^3 r_4) \\
\alpha_4 &= \frac{1}{5}(\zeta^3 r_1 + \zeta r_2 + \zeta^4 r_3 + \zeta^2 r_4) \\
\alpha_5 &= \frac{1}{5}(\zeta^4 r_1 + \zeta^3 r_2 + \zeta^2 r_3 + \zeta r_4)
\end{aligned} \tag{8.21}$$

Assuming that $\theta \in F$, Eq. (8.21) provides us with all five roots of Eq. (8.17) expressed as radical solutions over F; that is why solvability of Eq. (8.1) follows (since we can undo the transformation that led to Eq. (8.17)). The important thing to note here is that for $i \in \{1, 2, 3, 4, 5\}$, each α_i has a radical term of the form:

$$\sqrt[5]{\ldots + \sqrt{\ldots + c\theta\sqrt{\Delta_f}} + \ldots}$$

Here, we have a series of nested radicals. Notice that each nested radical corresponds to each subsequent field extension in Eq. (8.3): $\sqrt{\Delta_f}$ corresponds to the field extension $F \subset K$ where $[K : F] = 2$, $\sqrt{\ldots + c\theta\sqrt{\Delta_f}}$ corresponds to the field extension $K \subset M$ where $[M : K] = 2$, and $\sqrt[5]{\ldots + \sqrt{\ldots + c\theta\sqrt{\Delta_f}} + \ldots}$ corresponds to the field extension $M \subset L$ where $[L : M] = 5$. In other words, the field extensions from Eq. (8.3) correspond to the most amount of nested radicals we can have, so we can nest radicals at most three times (since there are three field extensions). Since $\theta \in F$, this poses no problems and we can consider ourselves done; otherwise, we might have a series of further nested radicals, perhaps an infinite amount, from the expression of θ which would require more field extensions and contradict our assumption of the solvability of $F \subset L$ since the highest order solvable subgroup of S_5 only permits three field extensions by virtue of Theorem 6.3.1.

Appendix A

Concepts used from Abstract Algebra

A.1 Some Notable Theorems for Groups and Rings

(**Lagrange's Theorem**)

Theorem A.1.1. *Suppose $H \subset G$ is a subgroup of G. Then, $|H|$ divides $|G|$.*

(**Fundamental Theorem of Group Homomorphisms**)

Theorem A.1.2. *Suppose $\phi\colon G_1 \to G_2$ is a group homomorphism. Then, there is a unique group isomorphism $\tilde{\phi}\colon G_1/\ker(\phi) \to \mathrm{im}(\phi)$ where $\tilde{\phi}(g\ker(\phi)) = \phi(g)$ for all $g \in G_1$.*

(**Cayley's Theorem**)

Theorem A.1.3. *Every group G is isomorphic to a subgroup of a symmetric group.*

(**Fundamental Theorem of Ring Homomorphisms**)

Theorem A.1.4. *Suppose $\phi\colon R_1 \to R_2$ is a ring homomorphism. Then, there is a unique ring isomorphism $\tilde{\phi}\colon R_1/\ker(\phi) \to \mathrm{im}(\phi)$ where $\tilde{\phi}(r + \ker(\phi)) = \phi(r)$ for all $r \in R_1$.*

A.2 Polynomials as Algebraic Objects

Definition A.2.1 (Monomials and Polynomials). *Let F be a field and $x_1, \ldots, x_n$ where $n \in \mathbb{N}$ be distinct indeterminates/variables (we will refer to them as variables). Take $c \in F$ and $a_1, \ldots, a_n \in \mathbb{N}_0$ where $\mathbb{N}_0$ is the set $\mathbb{N} \cup \{0\}$. Then, a **monomial** in n variables is an expression of the form:*

$$c{x_1}^{a_1} \cdots {x_n}^{a_n}$$

*Here, $c \in F$ is called a **coefficient**. A **polynomial** is a sum of monomials.*

Definition A.2.2 (Set of Polynomials over a Field). *Let F be a field and let $x_1, \ldots, x_n$ be distinct indeterminates/variables. Then, $F[x_1, \ldots, x_n]$ is the ring of all polynomials in $x_1, \ldots, x_n$ over F.*

Definition A.2.3. *For a monomial $m \in F[x_1, \ldots, x_n]$ where $m(x_1, \ldots, x_n) = c{x_1}^{a_1} \cdots {x_n}^{a_n}$, the **total degree** is $\deg(m) = a_1 + \ldots + a_n$. The **degree of a polynomial** is the monomial term in a polynomial of highest total degree; the coefficient of this monomial is called the **leading coefficient**. If the degree of the polynomial is zero, it is a **constant polynomial**; i.e. it is purely an element of $F \backslash \{0\}$ where the degree of $0 \in F$ is left undefined.*

Definition A.2.4. *A polynomial $f \in F[x_1, \ldots, x_n]$ is **irreducible** if it cannot be expressed as a product of polynomials of strictly lower degree; i.e. f is irreducible if it cannot be expressed in the form $f = h_1 \cdots h_m$ such that $\forall k \in \{1, \ldots, m\},\ \deg(h_k) < \deg(f)$*

In one variable x (univariate), $F[x]$ is the ring of polynomials over a field F where $f \in F[x]$ such that $deg(f) = n$ (here, the notations "f" and "$f(x)$" are synonymous) and is of the form:

$$f(x) = a_n x^n + \ldots + a_1 x + a_0 \tag{A.1}$$

One can show that $F[x_1, \ldots, x_n]$ is an integral domain by making use of the fact that it is also a ring under addition and multiplication as well as the fact that F is an integral domain (which is implicit since F is a field). Furthermore, it can also be shown that $F[x_1, \ldots, x_n]$ is a Unique Factorization Domain (UFD) by using a generalized version of Gauss's Lemma (reference later). By assuming $F[x_1, \ldots, x_n]$ is a UFD, we necessarily assume that any $f \in F[x_1, \ldots, x_n]$ can be expressed as the product of irreducible polynomials $g_1, \ldots, g_r \in F[x_1, \ldots, x_n]$ such that $f = g_1 \cdots g_r$ where $r \in \mathbb{N}$.

Definition A.2.5 (Field of Rational Functions). *Since $F[x_1, \ldots, x_n]$ is an integral domain, we may define its field of rational functions as*

$$F(x_1, \ldots, x_n) = \{f/g \mid f, g \in F[x_1, \ldots, x_n], g \neq 0\}$$

Example A.2.6. *Consider the field of rational numbers $\mathbb{Q}$ and two variables $x_1 \coloneqq x,\ x_2 \coloneqq y$. Then, if we have*

$$p(x, y) = 69x^4y^{20} + 6x^6y^6 + 80085,$$

it is the case that $p \in \mathbb{Q}[x, y]$ and $\deg(p) = 24$. If we have $q = f/g$ for $f, g \in \mathbb{Q}[x, y]$ where

$$q(x, y) = \frac{69x^4y^{20} + 4x^{20} + y^{69}}{6y^6x^6 + 80085x^3},$$

it is the case that $q \in \mathbb{Q}(x, y)$, but $q \notin \mathbb{Q}[x, y]$; also, $\deg(f) = 24$ and $\deg(g) = 12$.

A.3 Further Properties of Polynomials

We assume all ring homomorphisms preserve the multiplicative identity; i.e. if $\phi\colon R \to S$ is a homomorphism where 1_R and 1_S are the multiplicative identities of rings R and S respectively, then $\phi(1_R) = 1_S$.

Theorem A.3.1 (Evaluation Homomorphism). *Let F be a field that is contained in the ring R. Take $\alpha_1, \ldots, \alpha_n \in R$ and let the mapping $\phi\colon F[x_1, \ldots, x_n] \to R$ be given by $\phi(f(x_1, \ldots, x_n)) = f(\alpha_1, \ldots, \alpha_n)$ for some $f \in F[x_1, \ldots, x_n]$. Then, ϕ is a ring homomorphism.*

Proof. Take $f, g \in F[x_1, \ldots, x_n]$. Then,

$$\begin{aligned}\phi(f+g) &= \phi(f(x_1, \ldots, x_n) + g(x_1, \ldots, x_n)) = \phi((f+g)(x_1, \ldots, x_n)) = (f+g)(\alpha_1, \ldots, \alpha_n) \\ &= f(\alpha_1, \ldots, \alpha_n) + g(\alpha_1, \ldots, \alpha_n) = \phi(f(x_1, \ldots, x_n)) + \phi(g(x_1, \ldots, x_n)) = \phi(f) + \phi(g)\end{aligned}$$

$$\begin{aligned}\phi(fg) &= \phi(f(x_1, \ldots, x_n)g(x_1, \ldots, x_n)) = \phi((fg)(x_1, \ldots, x_n)) = (fg)(\alpha_1, \ldots, \alpha_n) \\ &= f(\alpha_1, \ldots, \alpha_n)g(\alpha_1, \ldots, \alpha_n) = \phi(f(x_1, \ldots, x_n))\phi(g(x_1, \ldots, x_n)) = \phi(f)\phi(g)\end{aligned}$$

■

Suppose F is contained in L. Take $f \in F[x]$ and $\alpha \in L$ (note that L is also a ring), and suppose the evaluation homomorphism $\phi\colon F[x] \to L$ is given by $\phi(f(x)) = f(\alpha)$. Then,

$$f(x) = a_n x^n + \ldots + a_1 x + a_0$$

is mapped to

$$f(\alpha) = a_n \alpha^n + \ldots + a_1 \alpha + a_0$$

Theorem A.3.1 is a formalization of the notion of "plugging in" values into polynomial variables. For $f \in F[x]$, whenever we see $f(\alpha)$ for some α in some ring R containing F (in this case L), the evaluation homomorphism is implicitly applied; here, we say that f is *evaluated* at $\alpha \in L$. Also, if $f(\alpha) = 0$, then α is said to be a **root** of f.

For polynomials of one variable over F, note that from Eq. (A.1):

$$\begin{aligned}f(x) &= a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \\ &= a_n(x^n + (a_{n-1}/a_n)x^{n-1} + \ldots + (a_1/a_n)x + (a_0/a_n))\end{aligned}$$

where $a_n \neq 0$. Therefore, if $f(\alpha) = 0$, then

$$\alpha^n + (a_{n-1}/a_n)\alpha^{n-1} + \ldots + (a_1/a_n)\alpha + (a_0/a_n) = 0$$

and redefining the coefficients yields

$$\alpha^n + b_n\alpha^{n-1} + \ldots + b_2\alpha + b_1 = 0$$

We see that this expression corresponds to the following polynomial:

$$g(x) = x^n + b_n x^{n-1} + \ldots + b_2 x + b_1 \tag{A.2}$$

The polynomial given by Eq. (A.2) is called a **monic** polynomial. From any univariate polynomial f as in Eq. (A.1), so long as the leading coefficient is nonzero, we can construct a monic polynomial such that if α is a root of f in then α is also a root of g in Eq. (A.2) since f is a constant multiple of g.

Theorem A.3.2 (Polynomial Division Algorithm). *Take $f, g \in F[x]$; then, $\exists q, r \in F[x]$ such that q and r are unique and*

$$f = qg + r, \quad \text{where } r = 0 \text{ or } \deg(r) < \deg(g)$$

Proof. Let $f = qg + r$ for some $q, r \in F[x]$. Suppose that $q \neq 0$ and $\deg(g) > \deg(f)$. Then,

$$\deg(f) = \max\{\deg(q) + \deg(g), \deg(r)\}$$

It follows that:

$$\deg(q) + \deg(g) \leq \deg(f) \Rightarrow \deg(q) \leq \deg(f) - \deg(g)$$

However, since $\deg(g) > \deg(f)$,

$$\deg(f) - \deg(g) < 0 \Rightarrow \deg(q) < 0$$

It follows that $q \notin F[x]$, a contradiction. We must have that either $q = 0$ or $\deg(g) \leq \deg(f)$. If we assume that $\deg(g) > \deg(f)$, then, we must have that $q = 0$ and $r = f$ where indeed, $\deg(r) < \deg(g)$ by assumption of $\deg(f)$; moreover, we see that assuming $\deg(g) > \deg(f)$ necessarily leads to $r = f$ and $q = 0$ which demonstrates uniqueness.

Suppose that $\deg(g) \leq \deg(f)$. Let

$$f = \sum_{i=0}^{n} a_i x^i, \ (a_i)_{i=0}^n \subset F \text{ and } \ g = \sum_{j=0}^{m} b_j x^j, \ (b_j)_{j=0}^m \subset F$$

where $f \neq 0$, $a_n \neq 0$ and $b_m \neq 0$ so that we have $n = \deg(f)$ and $m = \deg(g)$ where $m \leq n$. Also, let $f^{(0)} = f$ and

$$f^{(1)} = f^{(0)} - \frac{a_n}{b_m} x^{n-m} g$$

Then, we have

$$f^{(1)} = \left(a_{n-1} - \frac{a_n b_{m-1}}{b_m}\right)x^{n-1} + \left(a_{n-2} - \frac{a_n b_{m-2}}{b_m}\right)x^{n-2} + \ldots + \left(a_{n-m} - \frac{a_n b_0}{b_m}\right)x^{n-m} + \ldots$$

For the coefficients of $f^{(1)}$, redefine them such that

$$f^{(1)} = a_{n-1}^{(1)}x^{n-1} + a_{n-2}^{(1)}x^{n-2} + \ldots + a_{n-m}^{(1)}x^{n-m} + \ldots + a_1^{(1)}x + a_0^{(1)}$$

Here, we have that if $f^{(1)} \neq 0$, then $\deg(f^{(1)}) \leq n-1$ since there is the possibility that $a_{n-1}^{(1)} = 0$; therefore, we have that $\deg(f^{(1)}) < \deg(f^{(0)})$ or that $f^{(1)} = 0$. Note that if n=m,

setting $q = \frac{a_n}{b_m}x^{n-m}$ and $r = f^{(1)}$, we have that $\deg(f^{(0)}) = \deg(f) = \deg(g)$, so either

$\deg(r) < \deg(g)$ or $r = 0$. Thus, for the rest of the argument, we may assume $m < n$.
Now, suppose that it is the case that $\deg(f^{(k)}) < \deg(f^{(k-1)})$ or $f^{(k)} = 0$ where $k \in \mathbb{N}$ and

$$f^{(k-1)} = a_{n-k+1}^{(k-1)}x^{n-k+1} + a_{n-k}^{(k-1)}x^{n-k} + \ldots + a_0^{(k-1)}$$

$$f^{(k)} = a_{n-k}^{(k)}x^{n-k} + a_{n-k-1}^{(k)}x^{n-k-1} + \ldots + a_0^{(k)}$$

$$f^{(k)} = f^{(k-1)} - \frac{a_{n-k+1}^{(k-1)}}{b_m}x^{n-m-k+1}g$$

Observe that if $f^{(k)} \neq 0$, then,

$$f^{(k+1)} \coloneqq f^{(k)} - \frac{a_{n-k}^{(k)}}{b_m}x^{n-m-k}g$$

$$\Rightarrow f^{(k+1)} = \left(a_{n-k-1}^{(k)} - \frac{a_{n-k}^{(k)}b_{m-1}}{b_m}\right)x^{n-k-1} + \ldots + a_0^{(k)}$$

Here, we have that $\deg(f^{(k+1)}) \leq n-k-1$, so $\deg(f^{(k+1)}) < \deg(f^{(k)})$ or $f^{(k+1)} = 0$. Thus, by the principle of mathematical induction, for the sequence $(f^{(i)})_{i=0}^{l}$ for some $l \in \mathbb{N}$ such that $\forall i \in \{0, 1, \ldots, l\}$, $f^{(i)} \neq 0$, we may conclude that $\deg(f^{(i)}) < \deg(f^{(i-1)})\ \forall i \in \{1, \ldots, l\}$ and that $\deg(f^{(l+1)}) < \deg(f^{(l)})$ or $f^{(l+1)} = 0$.

Since $m < n$, $\exists k \in \mathbb{N} : m = n - k$, so $\exists k \in \mathbb{N} : \deg(f^{(k)}) \leq m$. If $f^{(k)} \in (f^{(i)})_{i=0}^{l}$, then, of course, $\deg(f^{(l+1)}) < \deg(f^{(k)}) \leq \deg(g)$. Otherwise, if $f^{(k)} \notin (f^{(i)})_{i=0}^{l}$ and $k \neq l+1$, then if $f^{(l+1)} \neq 0$, we may create a larger sequence $(f^{(i)})_{i=0}^{l'}$ until $f^{(l'+1)} = 0$ or until $f^{(k)}$ is in the sequence. Therefore, we may assume that if $f^{(k)} \notin (f^{(i)})_{i=0}^{l}$, then $f^{(l+1)} = 0$.

Note that

$$f^{(l+1)} = f - \left(\frac{a_n}{b_m}x^{n-m} + \frac{a_{n-1}^{(1)}}{b_m}x^{n-m-1} + \ldots + \frac{a_{n-l+1}^{(l-1)}}{b_m}x^{n-m-l+1} + \frac{a_{n-l}^{(l)}}{b_m}x^{n-m-l}\right)g$$

$$\Rightarrow \boxed{f = qg + f^{(l+1)}}$$

where

$$q \coloneqq \left(\frac{a_n}{b_m}x^{n-m} + \frac{a_{n-1}^{(1)}}{b_m}x^{n-m-1} + \ldots + \frac{a_{n-l+1}^{(l-1)}}{b_m}x^{n-m-l+1} + \frac{a_{n-l}^{(l)}}{b_m}x^{n-m-l}\right)$$

Furthermore, letting $r = f^{(l+1)}$, we see that $\exists q, r \in F[x]$ such that

$$f = qg + r \text{ where } r = 0 \text{ or } \deg(r) < \deg(g)$$

• To demonstrate the uniqueness of q and r suppose that $\exists \tilde{q}, \tilde{r} \in F[x]$:

$$f = qg + r = \tilde{q}g + \tilde{r} \text{ where } \tilde{r} = 0 \text{ or } \deg(\tilde{r}) < \deg(g)$$

We also assume the same conditions for r. It follows that

$$g(q - \tilde{q}) = \tilde{r} - r$$

Suppose that $q - \tilde{q} \neq 0$. Then, $\deg(g) \leq \deg(g(q - \tilde{q})) = \deg(\tilde{r} - r)$. However, this yields a contradiction since $\deg(\tilde{r} - r) \leq \deg(\tilde{r}) < \deg(g)$. Therefore we must assume that $q - \tilde{q} = 0$ and hence, $\tilde{q} = q$ and $\tilde{r} = r$ ■

Corollary A.3.3. *Take $f \in F[x]$ and $\alpha \in F$. Then, for some $q \in F[x]$,*

$$f = (x - \alpha)q \iff f(\alpha) = 0$$

Proof. Suppose $f(x) = (x - \alpha)q(x)$ for some $q \in F[x]$. Then, $f(\alpha) = (\alpha - \alpha)q(\alpha) \Rightarrow f(\alpha) = 0$. Now, suppose that $f(\alpha) = 0$. By Theorem A.3.2, there are $q, r \in F[x]$:

$$f(x) = (x - \alpha)q(x) + r(x)$$

such that $\deg(r) < \deg(x - \alpha)$ or $r(x) = 0$. Then, $f(\alpha) = (\alpha - \alpha)q(\alpha) + r(\alpha) \Rightarrow f(\alpha) = r(\alpha)$. Since $f(\alpha) = 0$, we have that $r(\alpha) = 0$. Furthermore, if $r(x) \neq 0$, since $\deg(r) < \deg(x - \alpha)$ where $\deg(x - \alpha) = 1$, this must mean that $\deg(r) = 0$, so $r(x) = c$ for some $c \in F \backslash \{0\}$. However, evaluation of r at $\alpha \in F$ yields $r(\alpha) = c$ which is a contradiction, so it must be that $r(x) = 0$. Thus, $f(x) = (x - \alpha)q(x)$. ■

Corollary A.3.4. *Take nonconstant $f \in F[x]$ and let $n = \deg(f)$ where $n > 0$. Then, f can have at most n roots in the field F.*

Proof. By Corollary A.3.3, if $\alpha_1 \in F$ is a root of f, then $f = (x - \alpha_1)q_1$ where $\deg(q_1) = n - 1$. Similarly, for q_1, if $\alpha_2 \in F$ is a root of q_1 then $q_1 = (x - \alpha_2)q_2$ where $\deg(q_2) = n - 2$. Continuing this process, we conclude that f factorizes such that

$$f = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_r)q_r$$

where $\deg(q_r) = n - r$. Here, q_r is irreducible in F and hence, has no roots in F. Since $\deg(f) = n$ and each linear factor is of degree one, it must be that $r \leq n$ since there can be at most n linear factors; if $r = n$, then q_r is constant ($\deg(q_r) = 0$) and hence, $q_r \in F \backslash \{0\}$. Moreover, each linear factor yields one root in F, so this process yields r roots of f in F where $r \leq n$. ■

A.4 Symmetric Polynomials

Definition A.4.1. *A polynomial $f \in F[x_1, \ldots, x_n]$ is deemed to be* ***symmetric*** *if for all $\sigma \in S_n$ (where S_n is the symmetric group on n letters):*

$$f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

Definition A.4.2. *Let $l = F[x_1, \ldots, x_n]$. Then, $\widetilde{f} \in l[x]$ given by*

$$\widetilde{f} = (x - x_1) \cdots (x - x_n) \tag{A.3}$$

is the ***universal polynomial*** *of degree n.*

Definition A.4.3. *The polynomials $\sigma_1, \sigma_2, \ldots, \sigma_n \in F[x_1, x_2, \ldots, x_n]$ defined by*

$$\begin{aligned}
\sigma_1 &= x_1 + \ldots + x_n \\
\sigma_2 &= \sum_{1 \le i < j \le n} x_i x_j \\
&\vdots \\
\sigma_r &= \sum_{1 \le k_1 < \ldots < k_r \le n} x_{k_1} x_{k_2} \cdots x_{k_r} \\
&\vdots \\
\sigma_n &= x_1 x_2 \cdots x_n
\end{aligned}$$

are called the ***elementary symmetric polynomials****.*

Remark. *One can show that the elementary symmetric polynomials are indeed symmetric.*

Suppose $f \in F[x]$ is a monic polynomial given by

$$f(x) = x^n + a_1 x^{n-1} \ldots + a_{n-1} x + a_n$$

with roots $\alpha_1, \ldots, \alpha_n$ in a larger field L. Then, f can be expressed as the product of linear factors involving its roots over L, so we find in L that

$$x^n + a_1 x^{n-1} \ldots + a_{n-1} x + a_n = (x - \alpha_1) \cdots (x - \alpha_n)$$

Then, after multiplying out the right-hand side, we will find that

$$\begin{aligned}
a_1 &= -(\alpha_1 + \ldots + \alpha_n) \\
a_2 &= \sum_{1 \le i < j \le n} \alpha_i \alpha_j \\
&\vdots \\
a_r &= (-1)^r \sum_{1 \le k_1 < \ldots < k_r \le n} \alpha_{k_1} \alpha_{k_2} \cdots \alpha_{k_r} \\
&\vdots \\
a_n &= (-1)^n \alpha_1 \alpha_2 \cdots \alpha_n
\end{aligned} \tag{A.4}$$

Thus, we find that each a_r is $(-1)^r$ times the elementary symmetric polynomial σ_r evaluated at the roots of f.

Now, we state the Fundamental Theorem of Symmetric Polynomials without proof. It is presented here, because it plays a big role in the solution process of solving for the roots of solvable polynomials.

Theorem A.4.4. *Any symmetric polynomial $f \in F[x_1, \ldots, x_n]$ is expressible in terms of the elementary symmetric polynomials; i.e. f can be written as a polynomial in $\sigma_1, \ldots, \sigma_n$ with coefficients in F and hence, $f \in F[\sigma_1, \ldots, \sigma_n]$.*

Corollary A.4.5. *Suppose $f \in F[x]$ is a nonconstant monic polynomial where $\deg(f) = n$ with roots $\alpha_1, \ldots, \alpha_n$ in a larger field L and that $p \in F[x_1, \ldots, x_n]$ is a symmetric polynomial. Then, using evaluation, $p(\alpha_1, \ldots, \alpha_n) \in F$.*

Proof. Since p is symmetric, it follows by Theorem A.4.4 that p is a polynomial in $\sigma_1, \ldots, \sigma_n$. Evaluation of p at $\alpha_1, \ldots, \alpha_n$ yields that $p(\alpha_1, \ldots, \alpha_n)$ is a polynomial in $\sigma_1(\alpha_1, \ldots, \alpha_n), \ldots$ $\ldots, \sigma_n(\alpha_1, \ldots, \alpha_n)$ with coefficients in F. By Eq. (A.4), we see that each $\sigma_r(\alpha_1, \ldots, \alpha_n)$ is a coefficient of f up to a factor of $(-1)^r$ so each $\sigma_r(\alpha_1, \ldots, \alpha_n)$ is in F. Thus, it follows that $p(\alpha_1, \ldots, \alpha_n) \in F$. ■

For each $k \in \{1, \ldots, n\}$, $\sigma_k \in F[x_1, \ldots, x_n]$, so

$$F[\sigma_1, \ldots, \sigma_n] \subset F[x_1, \ldots, x_n]$$

and hence,

$$F(\sigma_1, \ldots, \sigma_n) \subset F(x_1, \ldots, x_n) \tag{A.5}$$

Definition A.4.6. *Define $\Delta \in F[x_1, \ldots, x_n]$ by*

$$\Delta = \prod_{1 \le i < j \le n} (x_i - x_j)^2 \tag{A.6}$$

*where $n \geq 2$. Here, Δ is called the **discriminant**.*

Note that the discriminant consists of the product of the square of the differences between x_i and x_j for $i, j \in \{1, \ldots, n\}$ such that $i < j$; thus, the discriminant consists of the product of squared differences between unique variable pairs out of n variables and hence has $\binom{n}{2} = \frac{1}{2}n(n-1)$ factors. Additionally, $(x_i - x_j)^2 = -(x_i - x_j)(x_j - x_i)$, so the discriminant may be expressed as

$$\Delta = (-1)^{\frac{1}{2}n(n-1)} \prod_{i=1}^{n} \prod_{\substack{j=1 \\ j \neq i}}^{n} (x_i - x_j)$$

In this regard, we can see that the action of all permutations of S_n on Δ will yield the same product, so we can conclude that Δ is symmetric and hence $\Delta \in F[\sigma_1, \ldots, \sigma_n]$ by Theorem A.4.4.

Eq. (A.6) indicates that $\sqrt{\Delta} \in F[x_1, \ldots, x_n]$; we define it as

$$\sqrt{\Delta} = \prod_{1 \leq i < j \leq n} (x_i - x_j)$$

Proposition A.4.7. *If $\sigma \in S_n$, then*

$$\sigma\sqrt{\Delta} = \operatorname{sgn}(\sigma)\sqrt{\Delta}$$

Recall that $\operatorname{sgn}(\sigma) = 1$ if σ consists of the product of an even number of transpositions and $\operatorname{sgn}(\sigma) = -1$ if σ consists of the product of an odd number of transpositions.

Proof. Let $\tau \in S_n$ be the transposition given by $(l\ m)$ where $1 \leq l < m \leq n$. Then,

$$\begin{aligned}\tau\sqrt{\Delta} = (l\ m) \prod_{1 \leq i < j \leq n} (x_i - x_j) &= (l\ m)(x_l - x_m) \prod_{\substack{1 \leq i < j \leq n \\ (i,j) \neq (l,m)}} (x_i - x_j) \\ = (x_m - x_l) \prod_{\substack{1 \leq i < j \leq n \\ (i,j) \neq (l,m)}} (x_i - x_j) &= -(x_l - x_m) \prod_{\substack{1 \leq i < j \leq n \\ (i,j) \neq (l,m)}} (x_i - x_j) = -\sqrt{\Delta}\end{aligned}$$

Thus, $\tau\sqrt{\Delta} = -\sqrt{\Delta}$; the successive application of another transposition ρ tacks on another factor of -1, so that $\tau\rho\sqrt{\Delta} = -(-\sqrt{\Delta}) = \sqrt{\Delta}$. Now, suppose $\sigma \in S_n$ is given by a product of $k \in \mathbb{N}$ transpositions in S_n where $\sigma = \tau_1 \cdots \tau_k$. Since the action of each transposition on $\sqrt{\Delta}$ tacks on a factor of -1 it follows that

$$\sigma\sqrt{\Delta} = \tau_1 \cdots \tau_k\sqrt{\Delta} = (-1)^k\sqrt{\Delta}$$

where we see that if k is even, $\sigma\sqrt{\Delta} = \sqrt{\Delta}$ and if k is odd, $\sigma\sqrt{\Delta} = -\sqrt{\Delta}$. Therefore, $\sigma\sqrt{\Delta} = \operatorname{sgn}(\sigma)\sqrt{\Delta}$. ■

Let $k = F[\sigma_1, \ldots, \sigma_n]$. Upon expansion of the universal polynomial $\widetilde{f}$, we can find that

$$\widetilde{f} = x^n - \sigma_1 x^{n-1} + \ldots + (-1)^r \sigma_r x^{n-r} + \ldots + (-1)^n \sigma_n \tag{A.7}$$

In this regard, we see that $\widetilde{f} \in k[x]$.

Suppose $f \in F[x]$ is a monic polynomial given by

$$f = x^n + a_n x^{n-1} + \ldots + a_k x^{n-k} + \ldots + a_n$$

with roots $\alpha_1, \ldots, \alpha_n$ in a field L containing F. Then, let $\phi\colon F[\sigma_1, \ldots, \sigma_n] \to F$ be the evaluation homomorphism given by $\phi(\sigma_k) = (-1)^k a_k$. Since $\Delta \in F[\sigma_1, \ldots, \sigma_n]$, $\Delta = \Delta(\sigma_1, \ldots, \sigma_n)$. Now, define the discriminant of f as

$$\Delta(f) = \phi(\Delta) = \Delta(-a_1 \ldots, (-1)^k a_k, \ldots, (-1)^n a_n)$$

Now, let λ be the evaluation given by $\lambda\colon x_k \mapsto \alpha_k$. Then,

$$\lambda(\Delta) = \lambda\Bigg(\prod_{1 \le i < j \le n} (x_i - x_j)^2\Bigg) = \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j)^2$$

Furthermore, note that

$$\lambda(\Delta) = \lambda(\Delta(\sigma_1(x_1, \ldots, x_n), \ldots, \sigma_n(x_1, \ldots, x_n))) = \Delta(\sigma_1(\alpha_1, \ldots, \alpha_n), \ldots, \sigma_n(\alpha_1, \ldots, \alpha_n))$$

and by Eq. (A.4), we find that

$$\lambda(\Delta) = \Delta(-a_1 \ldots, (-1)^k a_k, \ldots, (-1)^n a_n)$$

so it follows that

$$\Delta(f) = \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j)^2 \tag{A.8}$$

References

[1] Peter Pesic. "Abel's Proof: An Essay on the Sources and Meaning of Mathematical Unsolvability". In: (2004).

[2] Jim Brown. "Abel and the Unsolvability of the Quintic". In: (2007).

[3] Peter M. Neumann. "The Mathematical Writings of Evariste Galois". In: (2011).

[4] D. S. Dummit. "Solving Solvable Quintics". In: *Mathematics of Computation* 57.195 (1991), pp. 387–401. ISSN: 00255718, 10886842. URL: http://www.jstor.org/stable/2938681.

[5] D.A. Cox. "Galois Theory". In: Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts (2012).

[6] Richard Ganaye. "Github Repository: Cox Galois Theory Exercises". In: *GitHub* (2020). URL: https://github.com/RichardGanaye/Cox-Galoi-Theory-Exercises.

[7] John B. Fraleigh and Victor J. Katz. "A First Course in Abstract Algebra: 7th. Ed." In: (2003).

[8] Vipul Naik. "Groupprops, the Group Properties Wiki". In: *Groupprops* (2006). URL: https://groupprops.subwiki.org/wiki/Main_Page.