Spring 2013

Galois Theory and the Hilbert Irreducibility Theorem

Damien Adams
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Adams, Damien, "Galois Theory and the Hilbert Irreducibility Theorem" (2013). *Master's Theses*. 4256.
DOI: <https://doi.org/10.31979/etd.y7e3-xbdv>
https://scholarworks.sjsu.edu/etd_theses/4256

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

GALOIS THEORY AND THE HILBERT IRREDUCIBILITY THEOREM

A Thesis

Presented to

The Faculty of the Department of Mathematics

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Damien Adams

May 2013

© 2013

Damien Adams

ALL RIGHTS RESERVED

The Designated Thesis Committee Approves the Thesis Titled

GALOIS THEORY AND THE HILBERT IRREDUCIBILITY THEOREM

by

Damien Adams

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

SAN JOSÉ STATE UNIVERSITY

May 2013

Dr. Timothy Hsu	Department of Mathematics
Dr. Roger Alperin	Department of Mathematics
Dr. Daniel Goldston	Department of Mathematics

ABSTRACT

GALOIS THEORY AND THE HILBERT IRREDUCIBILITY THEOREM

by Damien Adams

We study abstract algebra and Hilbert's Irreducibility Theorem. We give an exposition of Galois theory and Hilbert's Irreducibility Theorem: given any irreducible polynomial $f(t_1, t_2, \dots, t_n, x)$ over the rational numbers, there are an infinite number of rational n -tuples $(a_1, a_2, \dots, a_n)$ such that $f(a_1, a_2, \dots, a_n, x)$ is irreducible over the rational numbers.

We take a preliminary look at linear algebra, symmetric groups, extension fields, splitting fields, and the Chinese Remainder Theorem. We follow this by studying normal extension fields and Galois theory, proving the fundamental theorem and some immediate consequences. We expand on Galois theory by exploring subnormal series of subgroups and define solvability with group property P , ultimately proving Galois' Theorem. Beyond this, we study symmetric functions and large extension fields with Galois group S_n .

We detour into complex analysis, proving a few of Cauchy's theorems, the identity theorem, which is a key to proving Hilbert's Irreducibility Theorem, and meromorphic functions. We study affine plane curves, regular values, and the Density Lemma — which bounds the rational outputs a non-rational meromorphic function has for rational inputs. Ultimately, we prove the Hilbert Irreducibility Theorem and apply it to symmetric functions to construct fields whose Galois group is S_n .

ACKNOWLEDGEMENTS

The 22 months it took me to write this have been the most difficult of my collegiate career. Consequently, they are the most rewarding. I began this thesis as a purely algebraic one, and it ended up with some interpolation, combinatorics, analysis, and a lot of complex analysis. I would not have been able to have come as far as I have without the help of Dr. Timothy Hsu. He was encouraging, understanding, inspiring, and an excellent well of knowledge. I also want to give my most gracious praise to my wife, Brittney Cathey-Adams, whom has shown me infinite patience and encouragement with the hundreds of hours I spent at my computer in solitude. Lastly, and of course not least, I want to thank my mom, dad, sister, in-laws, and all of my closest friends for their continued support and love.

TABLE OF CONTENTS

CHAPTER

1	INTRODUCTION	1
1.1	Motivation	1
1.2	Outline	1
2	A BIT OF BACKGROUND	3
2.1	Assumptions	3
2.2	Linear Algebra	3
2.3	Symmetric Groups	6
2.4	Subrings	7
2.5	Extension Fields	7
2.6	Minimal Polynomials	8
2.7	Simple Extensions	9
2.8	Multiple Zeros	13
2.9	GCD and LCM	13
2.10	Gauss' Lemma	17
2.11	Chinese Remainder Theorem	19
2.12	Splitting Fields	20
3	THE FUNDAMENTAL THEOREM OF GALOIS THEORY	25
3.1	Fundamentals of Galois Groups	25
3.2	Normal Extensions	27
3.3	The Fundamental Theorem of Galois Theory	30

4	SOLVABILITY	35
4.1	Solvable Groups and Polynomials	35
4.2	Roots of Unity	39
4.3	Galois' Theorem Part I	41
4.4	Galois' Theorem Part II	46
5	RATIONAL POLYNOMIALS AND SYMMETRIC GALOIS GROUPS	52
5.1	Symmetric Functions	52
5.2	Large Extensions with Galois Group S_n	57
6	COMPLEX ANALYSIS	64
6.1	Domains	64
6.2	Complex Integration & Uniform Convergence	64
6.3	Holomorphic Functions and the Theorems of Cauchy	67
6.4	Power Series	69
6.5	Taylor Series	73
6.6	Laurent Series and Singularities	74
6.7	Meromorphic Functions	78
6.8	Complex Exponential	80
7	LEMMAS FOR THE HILBERT IRREDUCIBILITY THEOREM	82
7.1	Affine Plane Curves and Regular Values	82
7.2	Interpolation and the Mean Value Theorem	92
7.3	The Density Lemma	93
8	HILBERT'S IRREDUCIBILITY THEOREM	100
8.1	Hilbert's Irreducibility Theorem	100
8.2	The Symmetric Group as a Galois Group Over $\mathbb{Q}$	106

LIST OF FIGURES

Figure

4.1	Tower of Fields with $t = 1$	44
4.2	Diagram of Fields with $t > 1$	45

CHAPTER 1

INTRODUCTION

1.1 Motivation

In grade school, we learn the quadratic formula — a formula that solves any complex quadratic equation. We later learn of cubic formulas, usually Cardano's formula, and a quartic formula that is several pages long. Ultimately, in a second or third course of abstract algebra, we learn that there is no quintic formula. The proof is usually given in the form of a contradiction. We present an alternative approach using Hilbert's Irreducibility Theorem, constructing degree n polynomials with $n > 4$ whose Galois groups are S_n , proving that these polynomials cannot be solved by a general formula.

1.2 Outline

In order to attain our motivated goal, we must develop several topics from abstract algebra, Galois theory, and complex analysis. Assuming that the reader is familiar with an introduction to abstract algebra, we begin in Chapter 2 by covering several preliminary concepts from linear algebra and abstract algebra. We follow in Chapter 3 by giving a proof of the fundamental theorem of Galois theory, which is not usually given in a second course in undergraduate abstract algebra. Moreover, we will conclude our exploration of Galois theory in Chapter 4 with Galois' Theorem, a theorem that will be critical in Chapter 8.

Continuing with abstract algebra, we explore symmetric functions and their Galois groups in Chapter 5. This will be key in constructing polynomials over some

extension field of the rational numbers whose Galois groups happen to be symmetric groups. When paired with Galois' Theorem, this will show that these polynomials are not solvable by radicals and therefore cannot be solved by a general formula.

A critical issue in the development of Hilbert's Irreducibility Theorem is the possibility of a function having several zeros very close together. We use polynomials with coefficients that are meromorphic functions, which we explore at length in Chapter 6. We also cover the Identity Theorem. In Chapter 7, we use the Density Lemma to limit the number of rational values an algebraic non-rational meromorphic function might have.

We finish off in Chapter 8 by proving Hilbert's Irreducibility Theorem in two variables, learn how to transform multivariable polynomials into two-variable polynomials (due to Kronecker), and then prove the theorem in an arbitrary finite number of variables, using the actual strategy that Hilbert used. Once we have achieved this, we can do what we set out to do — construct a polynomial whose Galois group is symmetric.

CHAPTER 2

A BIT OF BACKGROUND

2.1 Assumptions

Unless otherwise stated, we will assume the following:

- All fields have characteristic 0.
- The degree of the zero polynomial is undefined.

2.2 Linear Algebra

We follow Jacobson [Jac85] and Strang [Str03] in this section.

Definition 2.2.1. Let F be a field and $n \in \mathbb{Z}_{>0}$. If $A \in \text{GL}_n(F)$, then we define the **determinant of A** as

$$\det A = \sum_{\sigma} (\text{sgn } \sigma) a_{1i_1} a_{2i_2} \cdots a_{ni_n}, \quad (2.1)$$

where the summation is taken over all permutations σ of $1, 2, \dots, n$ and

$$\text{sgn } \sigma = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}. \quad (2.2)$$

Remark 2.2.2. Notice that if F is a field, $n \in \mathbb{Z}_{>0}$, and $A \in \text{GL}_n(F)$, then $\det A$ is a polynomial in the entries of A .

Definition 2.2.3. Let F be a field, $n \in \mathbb{Z}_{>0}$, and $A \in \text{GL}(n, F)$. The **cofactor of the ij th entry of A** , C_{ij} , is $(-1)^{i+j}$ times the determinant of the matrix found by removing the i th row and j th column of A . The **cofactor matrix of A** , C , is the matrix with entries C_{ij} for $1 \leq i, j \leq n$.

Theorem 2.2.4. *Let F be a field, $n \in \mathbb{Z}_{>0}$, $A \in \text{GL}(n, F)$, and C be the cofactor matrix of A . If $\det A \neq 0$, then*

$$A^{-1} = \frac{C^T}{\det A}. \quad (2.3)$$

Proof. See Strang [Str03, 5C]. □

Definition 2.2.5. Let F be a field. A **Vandermonde matrix** with entries in F is an $n \times n$ matrix V of the form

$$V = \begin{bmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_m & a_m^2 & \dots & a_m^{n-1} \end{bmatrix}, \quad (2.4)$$

where $a_1, a_2, \dots, a_m \in F$.

Theorem 2.2.6. *Let F be a field. If V is the Vandermonde matrix with entries in F given by*

$$V = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{n-1} \\ 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \dots & x_{n-1}^{n-1} \end{bmatrix}, \quad (2.5)$$

then we have

$$\det V = \prod_{0 \leq i < j \leq n-1} (a_j - a_i). \quad (2.6)$$

Proof. By Definition 2.2.1, $\det V$ is the sum of monomials of degree

$$0 + 1 + \dots + (n-1) = \binom{n}{2} \quad (2.7)$$

in the ring $D = F[x_0, x_1, \dots, x_{n-1}]$. If we set $x_i = x_j$, then V has two identical rows.

Thus, $\det V = 0$, and $(x_j - x_i)$ divides $\det V$.

There are $\binom{n}{2}$ possible $(x_j - x_i)$, all irreducible in the UFD D , so

$$\det V = c \prod_{0 \leq i < j \leq (n-1)} (x_j - x_i) \quad (2.8)$$

for some $c \in F$.

We claim that $c = 1$. We compare the coefficients for the x_{n-1}^{n-1} term of $\det V$.

Let us note that

$$\det V = \sum (\det V_{(n-1)(n-1)}) x_{n-1}^{n-1}. \quad (2.9)$$

We will induct on n . For $n = 1$,

$$\det V = 1. \quad (2.10)$$

For $n > 1$,

$$\prod_{0 \leq i < j \leq n-1} (x_j - x_i) = \left(\prod_{0 \leq i < j \leq n-2} (x_j - x_i) \right) x_{n-1}^{n-1}. \quad (2.11)$$

Since the coefficient of the x_{n-1}^{n-1} term in $\det V$ is 1, $c = 1$. $\square$

Lemma 2.2.7. *Let*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= 0 \end{aligned} \quad (2.12)$$

be a system of $n < m$ homogeneous linear equations with $a_{ij} \in F$. Then there exists a solution $(c_1, c_2, \dots, c_m) \neq 0$ with $c_i \in F$ for $1 \leq i \leq m$.

Proof. Consider the set of vectors

$$\left\{ \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}, \dots, \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} \right\}. \quad (2.13)$$

These are m vectors in F^n , where $n < m$. Therefore, the set in (2.13) is linearly dependent. If we look at the vector representation of (2.12), we have

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} x_m = 0. \quad (2.14)$$

Because of the linear dependency of (2.13), we are guaranteed a nontrivial solution $(c_1, c_2, \dots, c_m) \in F^n$ of (2.14). $\square$

Theorem 2.2.8 (Cramer's Rule). *Let F be a field, $n \in \mathbb{Z}_{>0}$, $A \in \text{GL}_n(F)$ such that $\det A \neq 0$, b be an $n \times 1$ F -vector. If $B_i \in \text{GL}_n(F)$ is obtained from A by replacing the i th column of A with the column vector b , then $Ax = b$ has the unique solution $x_i = \frac{\det B_i}{\det A}$ for $i = 1, 2, \dots, n$.*

Proof. See Strang [p. 259][Str03]. $\square$

2.3 Symmetric Groups

We follow Gallian [Gal10] in this section.

Definition 2.3.1. For all $n \in \mathbb{Z}_{>1}$, a permutation in S_n is called a **transposition** if it is a 2-cycle.

Theorem 2.3.2. *For all $n \in \mathbb{Z}_{>1}$, every permutation in S_n is a product of transpositions.*

Proof. See Gallian [Gal10, Theorem 5.4]. $\square$

2.4 Subrings

We follow Gallian [Gal10] in this section.

Theorem 2.4.1. *Let R be a ring, I be an index set, S_i be a subring of R for all $i \in I$. Then $\bigcap_{i \in I} S_i$ is a subring of R .*

Proof. Recall that

$$\bigcap_{i \in I} S_i = \{r \in R \mid r \in S_i \text{ for all } i \in I\}. \quad (2.15)$$

Let $R' = \bigcap_{i \in I} S_i$. First, $R' \neq \emptyset$ because $0 \in S_i$ for all $i \in I$, so $0 \in R'$. Let $a, b \in R'$.

By the subring test, it suffices to show that $a - b, ab \in R'$. Because $a, b \in R'$,

$a, b \in S_i$ for all $i \in I$. Because each S_i is a subring of R , $a - b, ab \in S_i$ for all $i \in I$.

Hence, $a - b, ab \in R'$. Therefore, R' is a subring of R . $\square$

2.5 Extension Fields

We follow Gallian [Gal10] and Hadlock [Had78] in this section.

Definition 2.5.1. If E is a field, and F is a subfield of E , then E is called an **extension field** of F . Write E/F to denote E as an extension field of F .

Theorem 2.5.2. *Let F be a field. Then F contains a subfield isomorphic to $\mathbb{Q}$. It follows that all fields extend the rational numbers.*

Proof. See Gallian [Gal10, Theorem 15.5], where F has characteristic 0. $\square$

Definition 2.5.3. Let F be a field, and let E be an extension of F . We say that E **has degree n as an extension over F** if the dimension of E as a vector space over F is n , and denote this by $[E : F] = n$.

Definition 2.5.4. Let E be an extension of a field F . If $[E : F] = n$ for some positive integer n , then E is called a **finite extension** of F .

Definition 2.5.5. Let F be a field, E/F be an extension of F . If $a_1, a_2, \dots, a_n \in E$, we define $F(a_1, a_2, \dots, a_n)$ to be the smallest subfield of E containing both F and the set $\{a_1, a_2, \dots, a_n\}$.

Definition 2.5.6. If E is an extension of a field F , and K is a subfield of E containing F , then K is called an **intermediate field** of E/F .

Theorem 2.5.7. *If E is a finite extension of a field F and K is an intermediate field of E/F , then K is a finite extension of F and $[E : F] = [E : K][K : F]$.*

Proof. See Gallian [Gal10, Theorem 21.5]. □

Definition 2.5.8. Let E be an extension of a field F and $a \in E$. If a is a zero of some nonzero polynomial over F , then a is **algebraic** over F .

Definition 2.5.9. Let E be an extension field of a field F . If every element of E is algebraic over F , then E is called an **algebraic extension** of F .

Theorem 2.5.10. *Let F be a field, and let E be a finite extension of F . Then E is an algebraic extension of F .*

Proof. See Gallian [Gal10, Theorem 21.4]. □

2.6 Minimal Polynomials

We follow Gallian [Gal10] in this section.

Theorem 2.6.1. *If a is algebraic over a field F , then there is a unique monic irreducible polynomial $p(x) \in F[x]$ such that $p(a) = 0$.*

Proof. See Gallian [Gal10, Theorem 21.2]. □

Definition 2.6.2. The polynomial in Theorem 2.6.1 is called the **minimal polynomial** for a over F .

Theorem 2.6.3. *Let F be a field, E be an extension of F , $a \in E$ be algebraic over F , and $p(x) \in F[x]$ be the minimal polynomial for a over F . If $f(x) \in F[x]$ has the property that $f(a) = 0$, then $p(x)$ divides $f(x)$ in $F[x]$.*

Proof. See Gallian [Gal10, Theorem 21.3]. □

Definition 2.6.4. Let F be a field, E be an extension of F , $a \in E$ be algebraic over F . Then the **degree of a over F** is the degree of the minimal polynomial for a over F .

Corollary 2.6.5. *Let F be a field, E be an extension of F , $a \in E$ be algebraic over F , and $p(x) \in F[x]$ be the minimal polynomial for a over F . The following are equivalent:*

- (i) $f(x) \in F[x]$ is irreducible over F and $f(a) = 0$.
- (ii) $f(x) = cp(x)$ for some nonzero $c \in F$.

Proof. Suppose $f(x) \in F[x]$ is irreducible over F and a is a zero of $f(x)$. By Theorem 2.6.3, $p(x)$ divides $f(x)$, so $f(x) = q(x)p(x)$ for some $q(x) \in F[x]$. Since $f(x)$ is irreducible over F , $q(x)$ must have degree 0. Therefore, $f(x) = cp(x)$ for some nonzero $c \in F$.

Conversely, suppose that $g(x) = dp(x)$ for some nonzero $d \in F$. Then $p(x) = d^{-1}g(x)$. Since $p(x)$ is irreducible over F , then $g(x)$ must also be irreducible over F . Moreover, since $p(a) = 0$ by definition, $d^{-1}g(a) = 0$, so $g(a) = 0$. □

2.7 Simple Extensions

We will follow Gallian [Gal10] for the following section.

Definition 2.7.1. Let F be a field. An extension E of F is a **simple extension** if $E = F(a)$ for some $a \in E$.

Theorem 2.7.2. *Let F be a field, and let $p(x) \in F[x]$ be monic irreducible over F with $\deg p(x) = n$. If a and b are zeros of $p(x)$ in some extension E of F , then*

- (i) *The map $\varphi_a : F[x]/\langle p(x) \rangle \rightarrow F(a)$ such that $\varphi_a(f(x) + \langle p(x) \rangle) = f(a)$ is a ring isomorphism;*
- (ii) *The set $\{1, a, a^2, \dots, a^{n-1}\}$ is a basis for $F(a)$ over F ; and*
- (iii) *There exists an isomorphism $\psi : F(a) \rightarrow F(b)$ such that $\psi(a) = b$ and $\psi(c) = c$ for all $c \in F$.*

Proof. (i) Define the homomorphism $\varphi : F[x] \rightarrow F(a)$ by $\varphi(f(x)) = f(a)$.

Because $\varphi(p(x)) = 0$, the ideal $\langle p(x) \rangle$ is a subset of $\text{Ker}(\varphi)$. Because $p(x)$ is irreducible over F , $\langle p(x) \rangle$ is a maximal ideal in $F[x]$, and $\varphi(1) = 1 \neq 0$. Hence, $\langle p(x) \rangle = \text{Ker}(\varphi)$. By the first isomorphism theorem, the map $\varphi_a : F[x]/\langle p(x) \rangle \rightarrow F(a)$ such that $\varphi_a(f(x) + \langle p(x) \rangle) = f(a)$ is a ring isomorphism.

- (ii) By part (i), we know that $F[x]/\langle p(x) \rangle \cong F(a)$. If $g(x) + \langle p(x) \rangle \in F[x]/\langle p(x) \rangle$, then the division algorithm guarantees that we may choose $g(x)$ so that $\deg g(x) < \deg p(x)$. Thus, we can write

$$g(x) + \langle p(x) \rangle = c_{n-1}x^{n-1} + \dots + c_1x + c_0 + \langle p(x) \rangle, \quad (2.16)$$

for some $c_0, c_1, \dots, c_{n-1} \in F$. It follows that an arbitrary element of $F(a)$ has the form

$$c_{n-1}a^{n-1} + \dots + c_1a + c_0. \quad (2.17)$$

Therefore,

$$B = \{1, a, \dots, a^{n-1}\} \quad (2.18)$$

spans $F(a)$.

Suppose $\sum_{i=0}^{n-1} c_i a^i = 0$ for some $c_i \in F$ not all zero. Then

$\sum_{i=0}^{n-1} c_i x^i + \langle p(x) \rangle = \langle p(x) \rangle$ in $F[x]/\langle p(x) \rangle$. Then write

$$-c_{n-1}x^{n-1} + \langle p(x) \rangle = \sum_{i=0}^{n-2} c_i x^i + \langle p(x) \rangle. \quad (2.19)$$

Because $\deg p(x) = n$, $-c_{n-1}x^{n-1} + \langle p(x) \rangle, \sum_{i=0}^{n-2} c_i x^i + \langle p(x) \rangle \neq \langle p(x) \rangle$.

However, $\deg \sum_{i=0}^{n-2} c_i x^i < \deg -c_{n-1}x^{n-1}$. It follows that $c_i = 0$ for all

$0 \leq i \leq n-1$, and B is F -linearly independent.

(iii) By (i), we can construct a ring isomorphism $\varphi_b : F[x]/\langle p(x) \rangle \rightarrow F(b)$ such that

$\varphi_b(f(x) + \langle p(x) \rangle) = f(b)$. Because φ_a is an isomorphism, we have an

isomorphism $\varphi_a^{-1} : F(a) \rightarrow F[x]/\langle p(x) \rangle$ such that $\varphi_a^{-1}(f(a)) = f(x) + \langle p(x) \rangle$.

Then the composition of ring isomorphisms $\varphi_b \varphi_a^{-1} : F(a) \rightarrow F(b)$ maps $a \mapsto b$

and $c \mapsto c$ for all $c \in F$.

□

Theorem 2.7.3. *If a and b are algebraic over a characteristic 0 field F , then there is an element $c \in F(a, b)$ such that $F(a, b) = F(c)$.*

Proof. See Gallian [Gal10, Theorem 21.6].

□

Theorem 2.7.4. *If F is a field, then the set of all algebraic elements over F is a field.*

Proof. Let A be the set of all elements that are algebraic over F . Consider the polynomial $x - a \in F[x]$. Since $a \in F$ is a zero of this polynomial, $a \in A$. Hence, $F \subseteq A$.

Let $b, c \in A$. Then $F(b, c) \subseteq A$ is an algebraic extension of F . Since

$b, c \in F(b, c), b + c, bc, bc^{-1} \in A$. Hence, A is a field.

□

Corollary 2.7.5. *Any finite extension is simple.*

Proof. Let F be a field, E be a finite extension of F , $[E : F] = n$ for some positive integer n . Because E has degree n as a vector space over F , choose a basis $\{1, a_1, a_2, \dots, a_{n-1}\}$ for E over F , where $a_1, a_2, \dots, a_{n-1} \in E$. Then for any $k \in E$, $k = c_0 + c_1 a_1 + c_2 a_2 + \dots + c_{n-1} a_{n-1}$ for some $c_0, c_1, \dots, c_{n-1} \in F$. That is, $k \in F(a_1, a_2, \dots, a_{n-1})$. Thus, $E \subseteq F(a_1, a_2, \dots, a_{n-1})$. But $F(a_1, a_2, \dots, a_{n-1})$ is the smallest field containing F and $\{a_1, a_2, \dots, a_{n-1}\}$, so $F(a_1, a_2, \dots, a_{n-1}) \subseteq E$. Hence, $E = F(a_1, a_2, \dots, a_{n-1})$.

We will proceed by induction on n . If $n = 1$, then we have the trivial case. Suppose $n > 1$. Then $E = F(a_1, a_2, \dots, a_{n-1})$. Because $\text{char } F = 0$, we can apply Theorem 2.7.3,

$$E = (F(a_1, a_2, \dots, a_{n-3}))(a_{n-2}, a_{n-1}) = F(a_1, a_2, \dots, a_{n-3}, b) \quad (2.20)$$

for some $b \in E$. By our inductive hypothesis, this extension is simple. $\square$

Theorem 2.7.6. *Let $E = F(a)$, $\varphi, \psi : E \rightarrow R$ be ring homomorphisms such that $\varphi(x) = \psi(x)$ for all $x \in F$ and $\varphi(a) = \psi(a)$. Then $\varphi = \psi$.*

Proof. Let n denote the degree of a over F . By Theorem 2.7.2, $\{1, a, a^2, \dots, a^{n-1}\}$ is a basis for E over F . Let $\beta = b_0 + b_1 a + \dots + b_{n-1} a^{n-1}$ be an arbitrary element of E . By hypothesis, $\varphi(b_i) = \psi(b_i)$ and $\varphi(a)^i = \psi(a)^i$ for all $0 \leq i \leq n-1$. Then

$$\begin{aligned} \varphi(\beta) &= \varphi(b_0 + b_1 a + \dots + b_{n-1} a^{n-1}) \\ &= \varphi(b_0) + \varphi(b_1) \varphi(a) + \dots + \varphi(b_{n-1}) \varphi(a)^{n-1} \\ &= \psi(b_0) + \psi(b_1) \psi(a) + \dots + \psi(b_{n-1}) \psi(a)^{n-1} \\ &= \psi(b_0 + b_1 a + \dots + b_{n-1} a^{n-1}) \\ &= \psi(\beta). \end{aligned} \quad (2.21)$$

Since $\beta \in E$ was arbitrary, $\varphi = \psi$. $\square$

2.8 Multiple Zeros

We will follow Gallian [Gal10] for the following section.

Definition 2.8.1. Let F be a field and

$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \in F[x]$. We define the **derivative of $f(x)$** , denoted by $f'(x)$, as the polynomial

$$n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + 2 a_2 x + a_1 \in F[x]. \quad (2.22)$$

Theorem 2.8.2. *Let F be a field. A polynomial $f(x) \in F[x]$ has a multiple zero in some extension E if and only if $f(x)$ and $f'(x)$ have a common factor of positive degree in $F[x]$.*

Proof. See Gallian [Gal10, Theorem 20.5]. □

Theorem 2.8.3. *If F is a field and $f(x) \in F[x]$ is irreducible over F , then $f(x)$ has no multiple zeros.*

Proof. Assume $f(x) \in F[x]$ is irreducible over F . By definition, $\deg f(x) \geq 1$. If $f(x)$ has a multiple zero, then by Theorem 2.8.2, $\gcd(f(x), f'(x)) = q(x)$ for some $q(x) \in F[x]$ of positive degree. Because $f(x)$ is irreducible over F , $q(x) = c f(x)$ for some $c \in F$. Then $f(x)$ divides $f'(x)$. However, either $\deg f'(x) < \deg f(x)$ or $f'(x) = 0$. However, $f'(x)$ cannot be the zero polynomial because $\text{char } F = 0$ and $\deg f(x) \geq 1$. Therefore, $f(x)$ has no multiple zeros. □

2.9 GCD and LCM

We follow Gallian [Gal10] in this section.

Lemma 2.9.1. *Let D be an integral domain, $a, b \in D$. Then $\langle a \rangle = \langle b \rangle$ iff $a = bu$ for some unit $u \in D$.*

Proof. Suppose $\langle a \rangle = \langle b \rangle$. Then $a \in \langle b \rangle$, so $a = bu$ for some $u \in D$. Similarly, $b = av$ for some $v \in D$. Thus, $a = (av)u = a(vu)$. Because integral domains have the cancellation property, $1 = vu$. Hence, v and u are units.

Suppose $a = bu$ for some unit $u \in D$. Then $av = b$ for some unit $v \in D$. Let $a' \in \langle a \rangle$. Then $a' = ax$ for some $x \in D$. Since $a = bu$, $a' = bux = b(y)$ for $y = ux \in D$. Hence, $a' \in \langle b \rangle$. Similarly, let $b' \in \langle b \rangle$. Then $b' = by$ for some $y \in D$, so $b' = avy = ax$ for $x = vy \in D$. Hence, $b' \in \langle a \rangle$. Therefore, $\langle a \rangle = \langle b \rangle$. $\square$

Theorem 2.9.2. *Let D be a principal ideal domain, $a, b \in D \setminus \{0\}$. There exists some $d \in D$ such that $\langle d \rangle = \langle a \rangle + \langle b \rangle$, and so*

(i) *We have that d divides a and b ;*

(ii) *If $c \in D$ divides both a and b , then c divides d ; and*

(iii) *There exist $x, y \in D$ such that $ax + by = d$.*

Moreover, d is unique up to associates.

Proof. Let $I = \langle a \rangle + \langle b \rangle$. The sum of two ideals is an ideal, so I is an ideal of D .

Because D is a principal ideal domain, I is generated by some element, $d \in D$. That is,

$$\{ax + by | x, y \in D\} = \langle a \rangle + \langle b \rangle = I = \langle d \rangle = \{dz | z \in D\}. \quad (2.23)$$

Because $a, b \in I = \langle d \rangle$, $a = dz_1$ and $b = dz_2$ for some $z_1, z_2 \in D$. By definition, d divides both a and b . Because $d \in I = \langle a \rangle + \langle b \rangle$,

$$d = ax + by \quad (2.24)$$

for some $x, y \in D$. Suppose $c \in D$ divides both a and b . Then $a = cr$ and $b = cs$ for some $r, s \in D$. Substitute these into (2.24) to get $d = crx + csy$. Since c divides the right side, c must also divide d .

By Lemma 2.9.1, d is unique up to associates. $\square$

Definition 2.9.3. Let D be a principal ideal domain, $a_1, a_2, \dots, a_n \in D$. If $d \in D$ such that $\langle d \rangle = \langle a_1 \rangle + \langle a_2 \rangle + \dots + \langle a_n \rangle$, then d is called the **greatest common divisor** of $a_1, a_2, \dots, a_n$, denoted by $\gcd(a_1, a_2, \dots, a_n)$, and d is unique up to associates.

Lemma 2.9.4. Let D be a principal ideal domain, $a, b, r, s \in D$. If $ra + sb = 1$, then $\gcd(a, b) = 1$.

Proof. Suppose $d = \gcd(a, b)$. Because $ra + sb = 1$ and d divides both a and b , then d divides 1. Therefore, d is a unit. By Lemma 2.9.1, we can assume $d = 1$. $\square$

Theorem 2.9.5. Let D be a principal ideal domain, $a, b \in D \setminus \{0\}$. There exists some $m \in D$ such that $\langle m \rangle = \langle a \rangle \cap \langle b \rangle$, and so

(i) We have that a divides m and b divides m ; and

(ii) If $c \in D$ such that a and b both divide c , then m divides c .

Moreover, m is unique up to associates.

Proof. Let $I = \langle a \rangle \cap \langle b \rangle$. The intersection of two ideals is an ideal, so I is an ideal of D . Because D is a principal ideal domain, I is generated by some element $m \in D$. Because $m \in \langle a \rangle \cap \langle b \rangle$, $m \in \langle a \rangle$ and $m \in \langle b \rangle$. That is, $m = az_1$ and $m = bz_2$ for some $z_1, z_2 \in D$.

Suppose $c \in D$ such that $c = ax$ and $c = by$ for some $x, y \in D$. Then $c \in \langle a \rangle \cap \langle b \rangle = \langle m \rangle$, so $c = mz$ for some $z \in D$.

By Lemma 2.9.1, m is unique up to associates. $\square$

Definition 2.9.6. Let D be a principal ideal domain, $a_1, a_2, \dots, a_n \in D \setminus \{0\}$. If $m \in D$ such that $\langle m \rangle = \langle a_1 \rangle \cap \langle a_2 \rangle \cap \dots \cap \langle a_n \rangle$, then m is called the **least common**

multiple of $a_1, a_2, \dots, a_n$, denoted by $\text{lcm}(a_1, a_2, \dots, a_n)$, and m is unique up to associates.

Lemma 2.9.7. *Let D be a PID, $a, b \in D$. Then*

$$\gcd(a, b) \cdot \text{lcm}(a, b) = ab. \quad (2.25)$$

Proof. Let $d = \gcd(a, b)$, $m = \text{lcm}(a, b)$. Then there exists some $j, k, \ell \in D$ such that $dj = ab$, $ak = m$, and $b\ell = m$. By Theorem 2.9.2, there exist $x, y \in D$ such that $d = ax + by$. It follows that

$$\begin{aligned} dm &= axm + bym \\ &= ax(b\ell) + by(ak) \\ &= abx\ell + abyk \\ &= ab(x\ell + yk) \\ &= dj(x\ell + yk), \end{aligned} \quad (2.26)$$

so by cancellation, $m = j(x\ell + yk)$. It follows that j divides m .

On the other hand, there exists some $p, q \in D$ such that $dp = a$ and $dq = b$. Then $dpb = ab = dj$ and $dqa = ab = dj$. That is, $j = pb = qa$. It follows that a and b both divide j . By Theorem 2.9.5, m divides j . Therefore, $j = m$, and $dm = ab$. $\square$

Lemma 2.9.8. *Let D be a PID, $a_1, a_2, \dots, a_n, b \in D$ such that $\gcd(a_i, b) = 1$ for $1 \leq i \leq n$. Then $\gcd(a_1 a_2 \cdots a_n, b) = 1$.*

Proof. Induct on n . If $n = 1$, then $\gcd(a_1, b) = 1$ by assumption. Suppose the theorem holds for $k - 1 > 1$. Then $\gcd(a_1 a_2 \cdots a_{k-1}, b) = 1$. Let $a_1 a_2 \cdots a_{k-1} = A_{k-1}$. Then $\gcd(A_{k-1} a_k, b) = 1$ by the inductive hypothesis. $\square$

2.10 Gauss' Lemma

We follow Gallian [Gal10] in this section.

Theorem 2.10.1 (Gauss' Lemma). *Let D be a principal ideal domain. If $f, g \in D[x]$ such that*

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0, \\ g(x) &= b_m x^m + b_{m-1} x^{m-1} + \cdots + b_0, \\ f(x)g(x) &= c_{m+n} x^{m+n} + c_{m+n-1} x^{m+n-1} + \cdots + c_0, \end{aligned} \tag{2.27}$$

with $\gcd(a_0, a_1, \dots, a_n) = 1 = \gcd(b_0, b_1, \dots, b_m)$, then $\gcd(c_0, c_1, \dots, c_{m+n}) = 1$.

Proof. Suppose $d = \gcd(c_0, c_1, \dots, c_{m+n}) \neq 1$. Let $p \in D$ be irreducible such that p divides d , and let $\bar{f} = f \pmod{p}$, $\bar{g} = g \pmod{p}$, and $\overline{fg} = fg \pmod{p}$. Choose $\bar{f}, \bar{g}, \overline{fg} \in (D/\langle p \rangle)[x]$. Since p is irreducible in D , $\langle p \rangle$ is a maximal ideal, so $D/\langle p \rangle$ is a field. Because p divides d , and d divides every term of fg , $\overline{fg} = 0$. It follows that $\bar{f} = 0$ or $\bar{g} = 0$ in $(D/\langle p \rangle)[x]$, since $D/\langle p \rangle$ is a field and $(D/\langle p(x) \rangle)[x]$ is an integral domain. Without loss of generality, say $\bar{f} = 0$. Then p divides each term of f , contradicting the assumption that $\gcd(a_0, a_1, \dots, a_n) = 1$. Therefore, $\gcd(c_0, c_1, \dots, c_{m+n}) = 1$. $\square$

Lemma 2.10.2. *Let D be an integral domain, $b_i \in D$, and $r \in D$ such that $r \mid b_i$ for $0 \leq i \leq n$. Then $r = \gcd(b_0, b_1, \dots, b_n)$ if and only if $1 = \gcd(\frac{b_0}{r}, \frac{b_1}{r}, \dots, \frac{b_n}{r})$.*

Proof. Assume $r = \gcd(b_0, b_1, \dots, b_n)$. Let $s = \gcd(\frac{b_0}{r}, \frac{b_1}{r}, \dots, \frac{b_n}{r})$. Then s divides $\frac{b_i}{r}$ for all $0 \leq i \leq n$. So rs divides b_i for $0 \leq i \leq n$. By Theorem 2.9.2, rs must divide r , so $r = rst$ for some $t \in D$. Cancel r on each side to get $1 = st$, so s, t are units. By Lemma 2.9.1, we can assume that $s = 1$.

Conversely, assume $1 = \gcd(\frac{b_0}{r}, \frac{b_1}{r}, \dots, \frac{b_n}{r})$. Suppose $s = \gcd(b_0, b_1, \dots, b_n)$. Then r divides s , so $s = ru$ for some $u \in D$. Since ru divides b_i , u divides $\frac{b_i}{r}$ for all

$0 \leq i \leq n$. By Theorem 2.9.2, u divides 1, so u is a unit. By Lemma 2.9.1, we may take $r = s$. $\square$

Lemma 2.10.3. *Let D be a principal ideal domain, K be the fraction field of D , $f \in K[x]$ such that*

$$f = \frac{a_n}{A_n}x^n + \frac{a_{n-1}}{A_{n-1}}x^{n-1} + \cdots + \frac{a_0}{A_0}, \quad (2.28)$$

where $a_i, A_i \in D$ and $A_i \neq 0$ for $0 \leq i \leq n$. If $A = \text{lcm}(A_0, A_1, \dots, A_n)$, $b_i = \frac{Aa_i}{A_i}$ for $0 \leq i \leq n$, and $r = \text{gcd}(b_0, b_1, \dots, b_n)$, then we can write $Af = rf_1$, for some $f_1 \in D[x]$, where f_1 has the property that the greatest common divisor of the coefficients of f_1 is 1.

Proof. Write $f_1(x) = \frac{b_n}{r}x^n + \frac{b_{n-1}}{r}x^{n-1} + \cdots + \frac{b_0}{r}$. By Lemma 2.10.2, $\text{gcd}(\frac{b_0}{r}, \frac{b_1}{r}, \dots, \frac{b_n}{r}) = 1$. Then

$$rf_1 = b_nx^n + b_{n-1}x^{n-1} + \cdots + b_0 = \frac{Aa_n}{A_n}x^n + \frac{Aa_{n-1}}{A_{n-1}}x^{n-1} + \cdots + \frac{Aa_0}{A_0} = Af. \quad (2.29)$$

$\square$

Theorem 2.10.4. *Let D be a principal ideal domain, K be the fraction field of D , and $f \in D[x]$. If $f = gh$ for some $g, h \in K[x]$, then there exists $s \in K$ such that $f = (sg)(s^{-1}h)$, where $sg, s^{-1}h \in D[x]$.*

Proof. Write

$$\begin{aligned} f &= a_nx^n + a_{n-1}x^{n-1} + \cdots + a_0, \\ g &= \frac{b_m}{B_m}x^m + \frac{b_{m-1}}{B_{m-1}}x^{m-1} + \cdots + \frac{b_0}{B_0}, \\ h &= \frac{c_\ell}{C_\ell}x^\ell + \frac{c_{\ell-1}}{C_{\ell-1}}x^{\ell-1} + \cdots + \frac{c_0}{C_0}, \end{aligned} \quad (2.30)$$

where $a_i, b_j, c_k, B_j, C_k \in D$ and $B_j, C_k \neq 0$ for $0 \leq i \leq n, 0 \leq j \leq m$, and

$0 \leq k \leq \ell$. Without loss of generality, assume $\text{gcd}(a_0, a_1, \dots, a_n) = 1$, for if it is not, then by Lemma 2.10.2, we can divide by $\text{gcd}(a_0, a_1, \dots, a_n)$ to make it so. Let

$B = \text{lcm}(B_0, B_1, \dots, B_m)$, $C = \text{lcm}(C_0, C_1, \dots, C_\ell)$. Then $BCf = (Bg)(Ch)$, where $Bg, Ch \in D[x]$. Let $r_1 = \gcd(\frac{Bb_0}{B_0}, \frac{Bb_1}{B_1}, \dots, \frac{Bb_m}{B_m})$ and $r_2 = \gcd(\frac{Cc_0}{C_0}, \frac{Cc_1}{C_1}, \dots, \frac{Cc_\ell}{C_\ell})$. Then we can write $Bg = r_1g_1$ and $Ch = r_2h_1$, where by Lemma 2.10.3, $g_1, h_1 \in D[x]$ have the property that the greatest common divisor of the coefficients of g_1 , and respectively h_1 , is 1. Moreover, $BCf = r_1r_2g_1h_1$. By Lemma 2.10.2, the greatest common divisor of the coefficients of BCf is BC , since we are assuming that $\gcd(a_0, a_1, \dots, a_n) = 1$; by Theorem 2.10.1 and Lemma 2.10.2, the greatest common divisor of the coefficients of $r_1r_2g_1h_1$ is r_1r_2 . It follows from Theorem 2.9.2 that $BC = ur_1r_2$, where u is a unit in D . Therefore, $f = (u^{-1}g_1)h_1$, where $u^{-1}g_1, h_1 \in D[x]$. Let $s = \frac{B}{ur_1} = \frac{r_2}{C}$. Then $f = (sg)(s^{-1}h)$, where $sg, s^{-1}h \in D[x]$. $\square$

2.11 Chinese Remainder Theorem

Theorem 2.11.1 (Chinese Remainder Theorem). *Let D be a PID. Given*

$p_0, p_1, \dots, p_n \in D$ such that $\gcd(p_j, p_k) = 1$ for $0 \leq j < k \leq n$ and $b_i \in D/\langle p_i \rangle$ for $0 \leq i \leq n$, there exists a unique $f \in D/\langle p_0p_1 \cdots p_n \rangle$ such that $f = b_i \pmod{p_i}$ for $0 \leq i \leq n$.

Proof. We proceed by induction on n . For the base case, $f = b_0$ suffices and is unique.

For $n = 1$, we have $\gcd(p_0, p_1) = 1$. From Theorem 2.9.2, $\langle p_0 \rangle + \langle p_1 \rangle = \langle 1 \rangle = D$, and there exists $x_0, x_1 \in D$ such that $p_0x_0 - p_1x_1 = b_1 - b_0$. Then $f = b_0 + p_0x_0 = b_1 + p_1x_1$. Then $f = b_0 \pmod{p_0}$ and $f = b_1 \pmod{p_1}$.

Suppose the theorem holds for $n = k$. By the inductive hypothesis, there exists a unique $g \in D/\langle p_0p_1 \cdots p_{k-1} \rangle$ such that $g = b_i \pmod{p_i}$ for $0 \leq i \leq k$. Also by the $n = 1$ case, there exists a unique $f \in D/\langle (p_0p_1 \cdots p_{k-1})p_k \rangle$ such that $f = b_k \pmod{p_k}$ and $f = g \pmod{(p_0p_1 \cdots p_{k-1})}$. Since $g = b_i \pmod{p_i}$ for $0 \leq i \leq k$ and

$f = g \pmod{p_i}$, $f = b_i \pmod{p_i}$ for $0 \leq i \leq k+1$. Therefore, the theorem holds for $n = k+1$.

Suppose f_1, f_2 are two polynomials that satisfy $f_1 = b_i \pmod{p_i} = f_2$ for $0 \leq i \leq n$. Then $f_1 - f_2 = 0 \pmod{p_i}$ for $0 \leq i \leq n$. This is only possible if $f_1 = f_2$. □

Theorem 2.11.2. *Let F be a field, $a_0, a_1, \dots, a_n \in F$ be distinct, and $b_0, b_1, \dots, b_n \in F$. Then there exists a unique $f(x) \in F[x]$ such that $\deg f(x) \leq n$ and $f(a_i) = b_i$ for $0 \leq i \leq n$.*

Proof. By the remainder theorem, $f(a_i) = b_i$ if and only if division of $f(x)$ by $(x - a_i)$ yields the remainder b_i . This is true if and only if $f(x) = b_i \pmod{(x - a_i)}$. By Theorem 2.11.1, there is a unique $f(x) \in F[x]/\langle (x - a_0)(x - a_1) \cdots (x - a_n) \rangle$ such that $f(x) = b_i \pmod{(x - a_i)}$ for $0 \leq i \leq n$. By the division algorithm, $f(x) \in F[x]/\langle (x - a_0)(x - a_1) \cdots (x - a_n) \rangle$ has a unique representative such that $\deg f(x) \leq n$. □

Corollary 2.11.3. *Let $g(t) \in \mathbb{C}[t]$ with $\deg g(t) = m \in \mathbb{Z}_{>0}$. If $g(t_i) \in \mathbb{Q}$ for $m+1$ distinct $t_i \in \mathbb{Q}$, where $0 \leq i \leq m$, then $g(t) \in \mathbb{Q}[t]$.*

Proof. By assumption, there are $m+1$ distinct $t_0, t_1, \dots, t_m \in \mathbb{Q}$ and $T_0, T_1, \dots, T_m \in \mathbb{Q}$ such that $g(t_i) = T_i$ for $0 \leq i \leq m$. By Theorem 2.11.2, there exists $f(t) \in \mathbb{Q}[t]$ with $\deg f(t) \leq m$ such that $f(t_i) = T_i$ for $0 \leq i \leq m$. Because $g(t)$ is the unique polynomial in $\mathbb{C}[t]$ such that $g(t_i) = T_i$ for $0 \leq i \leq m+1$, it must be that $g(t) = f(t) \in \mathbb{Q}[t]$. □

2.12 Splitting Fields

We will follow Gallian [Gal10] for the following section.

Definition 2.12.1. Let E be an extension of a field F , and let $f(x) \in F[x]$. We say that $f(x)$ **splits in E** if $f(x)$ can be factored as a product of linear factors in $E[x]$.

Definition 2.12.2. Let E be an extension of a field F and $f(x) \in F[x]$. We say that E is a **splitting field for $f(x)$ over F** if $f(x)$ splits in E but in no proper intermediate field of E/F .

Theorem 2.12.3. *Let F be a field, and let $f(x) \in F[x]$ have positive degree. Then there exists an extension E of F where $f(x)$ splits.*

Proof. See Gallian [Gal10, Theorem 20.2]. □

Theorem 2.12.4. *Let F be a field, E be an extension of F . If $f(x) \in F[x]$ factors as $b(x - a_1)(x - a_2) \cdots (x - a_n)$ in $E[x]$, then $F(a_1, a_2, \dots, a_n)$ is a splitting field for $f(x)$ over F in E .*

Proof. Let $f(x) = b(x - a_1)(x - a_2) \cdots (x - a_n) \in F[x]$, where $a_1, a_2, \dots, a_n \in E$ and $b \in F$. Certainly, $f(x)$ splits in $F(a_1, a_2, \dots, a_n)$. Suppose $f(x)$ splits in some intermediate field L of E/F . Then $a_1, a_2, \dots, a_n \in L$. It follows that $F(a_1, a_2, \dots, a_n) \subseteq L$, by definition. Then $F(a_1, a_2, \dots, a_n)$ is the smallest subfield of E containing F and all of the zeros of $f(x)$ and is therefore the splitting field for $f(x)$ over F . □

Theorem 2.12.5. *Let F be a field, $\varphi : F \rightarrow \tilde{F}$ be an isomorphism of fields, $f(x) \in F[x]$ be monic irreducible over F with $\deg f(x) = n$, and let $\tilde{f}(x) = \varphi(f(x))$. If r is a zero of $f(x)$ in some extension K of F and $\tilde{r}$ is a zero of $\tilde{f}(x)$ in some extension $\tilde{K}$ of $\tilde{F}$, then there is a unique isomorphism $\tilde{\varphi} : F(r) \rightarrow \tilde{F}(\tilde{r})$ such that $\tilde{\varphi}(p(r)) = \tilde{p}(\tilde{r})$ for any polynomial $p(x) \in F[x]$.*

Proof. Because $F \cong \tilde{F}$, $F[x] \cong \tilde{F}[x]$. Since $\varphi(f(x)) = \tilde{f}(x)$, $\langle f(x) \rangle \cong \langle \tilde{f}(x) \rangle$. By Theorem 2.7.2 and the First Isomorphism Theorem, we have isomorphisms

$\psi : F(r) \rightarrow F[x]/\langle f(x) \rangle$, $\varphi : F[x]/\langle f(x) \rangle \rightarrow \tilde{F}[x]/\langle \tilde{f}(x) \rangle$, and $\sigma : \tilde{F}[x]/\langle \tilde{f}(x) \rangle \rightarrow \tilde{F}(\tilde{r})$ such that

$$\begin{aligned}\psi(p(r)) &= p(x) + \langle f(x) \rangle, \\ \varphi(p(x) + \langle f(x) \rangle) &= \tilde{p}(x) + \langle \tilde{f}(x) \rangle, \\ \sigma(\tilde{p}(x) + \langle \tilde{f}(x) \rangle) &= \tilde{p}(\tilde{r}).\end{aligned}\tag{2.31}$$

Let $g(r) \in F(r)$ be arbitrary, and define $\tilde{\varphi} : F(r) \rightarrow \tilde{F}(\tilde{r})$ by $\tilde{\varphi} = \sigma\varphi\psi$. It follows that

$$\begin{aligned}\tilde{\varphi}(g(r)) &= \sigma\varphi\psi(g(r)) \\ &= \sigma\varphi(g(r) + \langle f(x) \rangle) \\ &= \sigma(\tilde{g}(r) + \langle \tilde{f}(x) \rangle) \\ &= \tilde{g}(\tilde{r}).\end{aligned}\tag{2.32}$$

Finally, suppose $\tilde{\tau}$ is another isomorphism from F to $\tilde{F}$ that satisfies $\tilde{\tau}(g(r)) = \tilde{g}(\tilde{r})$. Then $\tilde{\tau}^{-1}\tilde{\varphi}(g(r)) = g(r)$ for all $g(r) \in F(r)$. By Theorem 2.7.6, $\tilde{\tau}^{-1}\tilde{\varphi} = id$, so $\tilde{\tau} = \tilde{\varphi}$. □

Corollary 2.12.6. *Let F be a field, $\varphi : F \rightarrow \tilde{F}$ be an isomorphism of fields, $f(x) \in F[x]$ be monic irreducible over F with $\deg f(x) = n$, and $\tilde{f}(x) = \varphi(f(x))$. Let K be a splitting field for $f(x)$ over F and $\tilde{K}$ be a splitting field for $\tilde{f}(x)$ over $\tilde{F}$. If $r \in K$ is a zero of $f(x)$, then for each zero $\tilde{r}$ of $\tilde{f}(x)$, there exists a unique isomorphism $\tau : F(r) \rightarrow L$ extending φ , where L is some subfield of $\tilde{K}$. Moreover, any such τ is unique and sends r to one of the n zeros $\tilde{r}$ of $\tilde{f}(x)$.*

Proof. By Corollary 2.6.5, $f(x)$ is the minimal polynomial for r over F . Let $\varphi(f(x)) = \tilde{f}(x)$. By Theorem 2.12.5, if $\tilde{r} \in \tilde{K}$ is a zero of $\tilde{f}(x)$, then there is a unique isomorphism $\tau : F(r) \rightarrow \tilde{F}(\tilde{r})$ extending φ such that $\tau(r) = \tilde{r}$.

Because $f(x)$ is irreducible over F , by Theorem 2.8.3, $f(x)$ has no multiple zeros. So $\tilde{f}(x)$ has n distinct zeros. Because τ extends φ , $\tau(k) = \tilde{k}$ for all $k \in F$.

Since $r \notin F$, $\tau(r) \notin \tilde{F}$. Now, $f(x) = (x - r)g(x) \in F(r)[x]$. So $\tau(f(x)) = (x - \tilde{r})\tilde{g}(x) = \tilde{f}(x)$. Thus, τ maps r to a zero of $\tilde{f}(x)$, there are exactly n possible choices for τ . $\square$

Theorem 2.12.7. *Let $\varphi : F \rightarrow \tilde{F}$ be an isomorphism of fields, $f(x) \in F[x]$ be of positive degree, and $\varphi(f(x)) = \tilde{f}(x)$. Let K be a splitting field of $f(x)$ over F , and $\tilde{K}$ be a splitting field of $\tilde{f}(x)$ over $\tilde{F}$. Then there are exactly $[K : F]$ isomorphisms $\psi : K \rightarrow \tilde{K}$ extending φ .*

Proof. We will induct on $n = [K : F]$. If $n = 1$, then $f(x)$ splits in $K = F$. Then $\tilde{K} = \tilde{F}$ and $\psi = \varphi$.

Suppose $n > 1$. Then $K \neq F$, $f(x)$ does not split in F , and $\deg f(x) > 1$. Let $g(x) \in F[x]$ such that $g(x)$ is monic irreducible over F , divides $f(x)$, and $\deg g(x) = m > 1$. Let $\varphi(g(x)) = \tilde{g}(x)$. Because isomorphisms preserve irreducibility, $\tilde{g}(x)$ is a monic irreducible factor of $\tilde{f}(x)$ over $\tilde{F}$.

Let $a \in K$ so that $g(a) = 0$. Because $g(x)$ is irreducible, Theorem 2.8.3 states there are m zeros of $g(x)$ in K . Then $[F(a) : F] = m$. By Theorem 2.12.5, there is a unique isomorphism $\tilde{\varphi} : F(a) \rightarrow \tilde{F}(\tilde{a})$ extending φ such that $\tilde{\varphi}(a) = \tilde{a}$, where $\tilde{a}$ is a zero of $\tilde{g}(x)$. By Corollary 2.12.6, there are exactly m such isomorphisms.

We now notice that K is a splitting field for $f(x)$ over $F(a)$, and $[K : F(a)] < n$. By our inductive hypothesis, there are $[K : F(a)]$ isomorphisms $\psi : K \rightarrow \tilde{K}$ extending $\tilde{\varphi}$, where $\tilde{K}$ is a splitting field for $\tilde{\varphi}(f(x)) = \tilde{f}(x)$ over $\tilde{F}(\tilde{a})$. By counting, there are $[K : F(a)]m = [K : F(a)][F(a) : F]$ different choices for ψ . By the degree theorem, there are

$$[K : F(a)][F(a) : F] = [K : F] = n \quad (2.33)$$

extensions ψ of φ . $\square$

Corollary 2.12.8. *Let F be a field, and let $f(x) \in F[x]$. Any two splitting fields of $f(x)$ over F are isomorphic.*

Proof. In Theorem 2.12.7, take $F = \tilde{F}$ and φ to be the identity map. Then all $\psi : K \rightarrow \tilde{K}$ extending φ are isomorphic to each other. □

Corollary 2.12.9. *Let F be a field, and let E be the splitting field of some polynomial over F . Then $|\text{Gal}(E/F)| = [E : F]$.*

Proof. In Theorem 2.12.7, take $F = \tilde{F}$, $K = \tilde{K} = E$, and φ to be the identity map. Then there are $[E : F]$ isomorphisms $\psi : E \rightarrow E$ extending the identity. By Definition 3.2.2 $[E : F] = |\text{Gal}(E/F)|$. □

CHAPTER 3

THE FUNDAMENTAL THEOREM OF GALOIS THEORY

3.1 Fundamentals of Galois Groups

We follow Hadlock [Had78] in this section.

Definition 3.1.1. Let E be an extension of a field F . We define

$$\text{Gal}(E/F) = \{\varphi \in \text{Aut}(E) \mid \varphi(c) = c \text{ for all } c \in F\} \quad (3.1)$$

to be the **Galois group of E over F** .

Lemma 3.1.2. *Let F be a field, E be an extension of F , $r \in E$ with $\deg r = n$, and let $f(x) \in F[x]$ be the minimal polynomial for r over F . Then there are exactly n distinct zeros of $f(x)$ in some extension L of E .*

Proof. Since $\deg r = n$, $\deg f(x) = n$. By definition, $f(x)$ is irreducible over F , so by Theorem 2.8.3, $f(x)$ has no multiple zeros. By Theorem 2.12.3, $f(x)$ splits into linear factors in some extension L of E . Because the degree of $f(x)$ is n , and since $f(x)$ has no multiple zeros, $f(x)$ has n distinct roots in L . □

Definition 3.1.3. Let F be a field, r be an algebraic element of some extension E of F , $f(x) \in F[x]$ be the minimal polynomial for r over F , and $r = r_1, r_2, \dots, r_n$ be all of the zeros of $f(x)$ in some extension L of E . Then $r_1, r_2, \dots, r_n$ are the **conjugates of r in L** .

Lemma 3.1.4. *Let E be an extension of a field F . If $\varphi \in \text{Gal}(E/F)$, $p(x) \in F[x]$, then for all $a \in E$, $\varphi(p(a)) = p(\varphi(a))$.*

Proof. Write $p(a) = c_0 + c_1a + \cdots + c_k a^k$ for $c_0, c_1, \dots, c_k \in F$. Because φ fixes F , $\varphi(c_i) = c_i$. Then

$$\begin{aligned}
 \varphi(p(a)) &= \varphi(c_0 + c_1a + \cdots + c_k a^k) \\
 &= \varphi(c_0) + \varphi(c_1)\varphi(a) + \cdots + \varphi(c_k)\varphi(a)^k \\
 &= c_0 + c_1\varphi(a) + \cdots + c_k\varphi(a)^k \\
 &= p(\varphi(a)).
 \end{aligned} \tag{3.2}$$

□

Theorem 3.1.5. *Let F be a field and let $E = F(r)$, where r is algebraic over F , with conjugates $r = r_1, r_2, \dots, r_n$ in some extension L containing E . Then for each $\varphi \in \text{Gal}(E/F)$, $\varphi(r) = r_i$ for some i such that $1 \leq i \leq n$. Moreover, for each $r_i \in E$, there is exactly one $\varphi \in \text{Gal}(E/F)$ satisfying $\varphi(r) = r_i$, and $|\text{Gal}(E/F)|$ is the number of conjugates of r in E .*

Proof. Suppose $\varphi \in \text{Gal}(E/F)$ and $f(x) \in F[x]$ is the minimal polynomial for r over F . By Lemma 3.1.4, $0 = \varphi(0) = \varphi(f(r)) = f(\varphi(r))$. The roots of $f(x)$ in L are precisely the conjugates of r , so $\varphi(r) = r_i$ for some i such that $1 \leq i \leq n$.

Suppose there exists two automorphisms φ and ψ in $\text{Gal}(E/F)$ satisfying $\varphi(r) = r_i$ and $\psi(r) = r_i$ for some i such that $1 \leq i \leq n$. Since $\varphi, \psi \in \text{Gal}(E/F)$, $\varphi(c) = \psi(c) = c$ for all $c \in F$. By assumption, $\varphi(r) = \psi(r)$. Hence, by Theorem 2.7.6, $\varphi = \psi$. Finally, since there is a one-to-one correspondence between automorphisms of $\text{Gal}(E/F)$ and the conjugates of r in E , $|\text{Gal}(E/F)|$ is equal to the number of conjugates of r in E . □

Corollary 3.1.6. *If E is any finite extension of F , then $|\text{Gal}(E/F)| \leq [E : F]$.*

Proof. Let $n = [E : F]$, $E = F(r)$, where r is algebraic over F with conjugates $r = r_1, r_2, \dots, r_m$ in E . We observe that $m \leq n$, since by Lemma 3.1.2 there are

exactly n conjugates of r in some extension L of E . By Theorem 3.1.5, there is exactly one $\varphi \in \text{Gal}(E/F)$ mapping r to r_i for some i such that $1 \leq i \leq m$.

Therefore, $|\text{Gal}(E/F)| \leq [E : F]$. $\square$

3.2 Normal Extensions

We follow Gallian [Gal10], Hadlock [Had78], and Jacobson [Jac85] in this section.

Definition 3.2.1. Let E be an extension of F . If every irreducible polynomial over F that has one root in E has all of its roots in E , then E is called a **normal extension** of F .

Definition 3.2.2. Let E be an extension of a field F , and let H be a subgroup of $\text{Gal}(E/F)$. We define the **fixed field of H** to be

$$E_H := \{a \in E \mid \varphi(a) = a \text{ for all } \varphi \in H\}. \quad (3.3)$$

Lemma 3.2.3. *Let E be an extension of a field F . If $K_1 \subseteq K_2$ are intermediate fields of E/F , then $\text{Gal}(E/K_1) \geq \text{Gal}(E/K_2)$.*

Proof. Let $\varphi \in \text{Gal}(E/K_2)$. Then $\varphi(a) = a$ for all $a \in K_2$. Since $a \in K_1$ implies $a \in K_2$, $\varphi(a) = a$ for all $a \in K_1$. Therefore, $\varphi \in \text{Gal}(E/K_1)$. $\square$

Lemma 3.2.4. *Let E be an extension of a field F , and let F' be the fixed field of $\text{Gal}(E/F)$. Then $\text{Gal}(E/F) = \text{Gal}(E/F')$.*

Proof. Let $a \in F$. For all $\varphi \in \text{Gal}(E/F)$, $\varphi(a) = a$. Because F' is the fixed field of $\text{Gal}(E/F)$ and $a \in F$, $F \subseteq F'$. By Lemma 3.2.3, $\text{Gal}(E/F') \leq \text{Gal}(E/F)$.

Conversely, let $\varphi \in \text{Gal}(E/F')$. Then by assumption, for all $a \in F'$, $\varphi(a) = a$. So $\varphi \in \text{Gal}(E/F)$, and $\text{Gal}(E/F) \leq \text{Gal}(E/F')$. Therefore,

$$\text{Gal}(E/F) = \text{Gal}(E/F'). \quad \square$$

Theorem 3.2.5. *Let E be a finite extension of a field F . Then the following are equivalent:*

(i) E is a normal extension of F .

(ii) $|\text{Gal}(E/F)| = [E : F]$.

(iii) E is a splitting field of some irreducible polynomial over F .

(iv) F is the fixed field of $\text{Gal}(E/F)$.

Proof. (i $\Rightarrow$ ii) By Corollary 2.7.5, write $E = F(r)$, where r is algebraic over F with $\deg r = n$ and has conjugates $r = r_1, r_2, \dots, r_n$ in some extension L of E . Because E is normal as a field over F , $r_i \in E$ for $1 \leq i \leq n$. By Theorem 3.1.5, because there are n conjugates of r in E , $n = [E : F] = |\text{Gal}(E/F)|$.

(ii $\Rightarrow$ iii) Let $n = |\text{Gal}(E/F)| = [E : F]$, and, by Corollary 2.7.5, write $E = F(r)$, where $r \in E$ has minimal polynomial $f(x) \in F[x]$ with zeros $r = r_1, r_2, \dots, r_n$. By Theorem 3.1.5, there are exactly $|\text{Gal}(E/F)|$ conjugates of r in E . Thus, $r_1, r_2, \dots, r_n \in E$. So $E = F(r) \subseteq F(r_1, r_2, \dots, r_n) \subseteq E$. Therefore, by Theorem 2.12.4, E is a splitting field for $f(x)$ over F .

(iii $\Rightarrow$ iv) Let us suppose that $g(x)$ is monic irreducible over F . Let E be a splitting field for $g(x)$ over F and F' be the fixed field of $\text{Gal}(E/F)$. We notice that E is the splitting field of $f(x)$ taken as a polynomial in $F'[x]$. By Corollary 2.12.9,

$$[E : F] = |\text{Gal}(E/F)| \text{ and } [E : F'] = |\text{Gal}(E/F')|. \quad (3.4)$$

By Lemma 3.2.4, $\text{Gal}(E/F) = \text{Gal}(E/F')$. Hence, $[E : F] = [E : F']$. By the degree theorem,

$$[E : F] = [E : F'] [F' : F], \quad (3.5)$$

so $[F' : F] = 1$, and $F = F'$. Therefore, F is the fixed field of $\text{Gal}(E/F)$.

(iv $\Rightarrow$ i) Let E be an extension of F , and assume F is the fixed field of $\text{Gal}(E/F)$. Let $h(x) \in F[x]$ be an irreducible polynomial over F with some zero $t \in E$. Suppose

$$S = \{t = t_1, t_2, \dots, t_\ell\} \quad (3.6)$$

is the orbit of t under the action of $\text{Gal}(E/F)$. The set S is comprised of distinct elements. Let $\varphi \in \text{Gal}(E/F)$. Because $\varphi(0) = 0$,

$$0 = \varphi(h(t)) = h(\varphi(t)) = h(t_i), \quad (3.7)$$

for some i such that $1 \leq i \leq \ell$. Then for $\varphi \in \text{Gal}(E/F)$, $\{\varphi(t_1), \varphi(t_2), \dots, \varphi(t_\ell)\}$ is a permutation of S , and $(x - t_i)$ divides $h(x)$. Define

$$u(x) = (x - t_1)(x - t_2) \cdots (x - t_\ell) \in E[x]. \quad (3.8)$$

Since $t_1, t_2, \dots, t_\ell$ are distinct, and because $\text{Gal}(E/F)$ cannot map t to an element of F , $u(x)$ has no multiple zeros. Thus, $u(x)$ divides $h(x)$. Now, let us apply φ to $u(x)$ to get

$$\begin{aligned} \varphi(u(x)) &= \varphi((x - t_1)(x - t_2) \cdots (x - t_\ell)) \\ &= (x - \varphi(t_1))(x - \varphi(t_2)) \cdots (x - \varphi(t_\ell)) \\ &= (x - t_1)(x - t_2) \cdots (x - t_\ell) \quad \text{after reordering} \\ &= u(x). \end{aligned} \quad (3.9)$$

We notice that $u(x)$ is fixed by φ , and so it is fixed by $\text{Gal}(E/F)$. Hence, $u(x) \in F[x]$, since F is the fixed field of $\text{Gal}(E/F)$. Because $h(x)$ is monic irreducible over F by assumption, and because $u(x)$ divides $h(x)$, $u(x) = h(x)$. Therefore, $t_1, t_2, \dots, t_\ell \in E$ whenever $t \in E$, and E is a normal extension of F . $\square$

Corollary 3.2.6. *Let E be a finite extension of a field F . The extension E/F is normal if and only if E is the splitting field of some not necessarily irreducible polynomial over F .*

Proof. Suppose E is a normal extension of F . By Theorem 3.2.5, E is the splitting field of a polynomial over F . Conversely, suppose that E is the splitting field of a polynomial over F . By Corollary 2.12.9, $[E : F] = |\text{Gal}(E/F)|$, and by Theorem 3.2.5, E is a normal extension of F . $\square$

Corollary 3.2.7. *Let E be a finite normal extension of a field F . If K is an intermediate field of E/F , then E is a normal extension of K .*

Proof. By Theorem 3.2.5 and Corollary 3.2.6, E is a splitting field of some polynomial, say, $f(x) \in F[x]$ over F . It follows that E is a splitting field for $f(x)$ over K . Then E/K is a normal extension. $\square$

3.3 The Fundamental Theorem of Galois Theory

We follow Hadlock [Had78] and Artin [Art91] in this section.

Lemma 3.3.1. *Let E be a normal extension of a field F and K_1 and K_2 be normal intermediate fields of E/F . If $\text{Gal}(E/K_1) = \text{Gal}(E/K_2)$, then $K_1 = K_2$.*

Proof. By Theorem 3.2.5, $E_{\text{Gal}(E/K_1)} = K_1$ and $E_{\text{Gal}(E/K_2)} = K_2$. Because $\text{Gal}(E/K_1) = \text{Gal}(E/K_2)$,

$$K_1 = E_{\text{Gal}(E/K_1)} = E_{\text{Gal}(E/K_2)} = K_2. \quad (3.10)$$

$\square$

Lemma 3.3.2. *Let E be a finite normal extension of a field F , and let K be an intermediate extension of E/F . Let $\varphi \in \text{Gal}(E/F)$. Then*

$\text{Gal}(E/\varphi(K)) = \varphi \text{Gal}(E/K) \varphi^{-1}$. Moreover, if $\text{Gal}(E/\varphi(K)) = \text{Gal}(E/K)$, then $\varphi(K) = K$.

Proof. Let $\varphi(K) = K'$, $\psi \in \text{Gal}(E/K)$, and let $k' \in K'$. By definition, $k' = \varphi(k)$ for some $k \in K$. Then

$$\varphi\psi\varphi^{-1}(k') = \varphi\psi(k) = \varphi(k) = k'. \quad (3.11)$$

Then $\varphi \text{Gal}(E/K) \varphi^{-1} \subseteq \text{Gal}(E/K')$. Let $\sigma \in \text{Gal}(E/K')$. Then by symmetry, $\sigma \in \varphi \text{Gal}(E/K) \varphi^{-1}$, so $\text{Gal}(E/K') = \varphi \text{Gal}(E/K) \varphi^{-1}$.

Assume that $\text{Gal}(E/K') = \text{Gal}(E/K)$. By Lemma 3.3.1, $K' = K$. $\square$

Theorem 3.3.3 (Fundamental Theorem of Galois Theory). *Let E be a finite normal extension of a field F . Let S be the set of all subgroups of $\text{Gal}(E/F)$, and let $\mathcal{F}$ be the set of all intermediate fields of E/F . Define $\Phi : S \rightarrow \mathcal{F}$ and $\Psi : \mathcal{F} \rightarrow S$ by*

$$\Phi(H) = E_H, \text{ and} \quad (3.12)$$

$$\Psi(K) = \text{Gal}(E/K). \quad (3.13)$$

Then Φ and Ψ are inverse functions of each other and are therefore bijections.

Furthermore, if K and L are intermediate fields of E/F , then:

(i) *We have that $K \subseteq L$ if and only if $|\text{Gal}(E/K)| \geq |\text{Gal}(E/L)|$.*

(ii) *We have that $[E : K] = |\text{Gal}(E/K)|$, and therefore,*

$$[K : F] = |\text{Gal}(E/F) : \text{Gal}(E/K)| = \frac{|\text{Gal}(E/F)|}{|\text{Gal}(E/K)|}. \quad (3.14)$$

(iii) *For $f(x) \in F[x]$, $\text{Gal}(E/F)$ permutes the zeros of $f(x)$ in E . If $f(x)$ is irreducible over F , then $\text{Gal}(E/F)$ acts transitively on the set of zeros of $f(x)$ in E .*

(iv) The field K is a normal extension of F if and only if $\text{Gal}(E/K)$ is normal in $\text{Gal}(E/F)$. In that case,

$$\text{Gal}(K/F) \cong \text{Gal}(E/F) / \text{Gal}(E/K). \quad (3.15)$$

Proof. Let us show that Φ and Ψ are inverse functions. If K is an intermediate field of E/F such that $\text{Gal}(E/K)$ is a subgroup of $\text{Gal}(E/F)$, then by Corollary 3.2.7, E/K is normal. Then by Theorem 3.2.5, K is the fixed field of $\text{Gal}(E/K)$. Thus,

$$\Phi(\Psi(K)) = \Phi(\text{Gal}(E/K)) = E_{\text{Gal}(E/K)} = K. \quad (3.16)$$

Hence, $\Phi = \Psi^{-1}$.

Now suppose $\text{Gal}(E/K)$ is a subgroup of $\text{Gal}(E/F)$. Then K is an intermediate field of E/F . Since E/F is normal, by Corollary 3.2.7, E/K is normal. By Theorem 3.2.5, $K = E_{\text{Gal}(E/K)}$, so

$$\Psi(\Phi(\text{Gal}(E/K))) = \Psi(E_{\text{Gal}(E/K)}) = \Psi(K) = \text{Gal}(E/K). \quad (3.17)$$

Hence, $\Psi = \Phi^{-1}$, and Φ and Ψ are inverses of each other. Therefore, Φ and Ψ are bijective functions.

Now, let K and L be intermediate fields of E/F .

(i) By Lemma 3.2.3, if $K \subseteq L$, then $\text{Gal}(E/K) \geq \text{Gal}(E/L)$. For the converse, suppose that $\text{Gal}(E/K) \geq \text{Gal}(E/L)$. Let $k \in K$, $\sigma \in \text{Gal}(E/L)$. Then $\sigma \in \text{Gal}(E/K)$, and $\sigma(k) = k$. Since k is fixed by $\sigma \in \text{Gal}(E/L)$ by Theorem 3.2.5 and Corollary 3.2.7, $k \in L$. Therefore, $K \subseteq L$.

(ii) By Corollary 3.2.7, E is a normal extension of K . By Theorem 3.2.5,

$[E : K] = |\text{Gal}(E/K)|$. By the degree theorem, $[E : F] = [E : K][K : F]$, so

$|\text{Gal}(E/F)| = |\text{Gal}(E/K)| [K : F]$. It follows that $[K : F] = \frac{|\text{Gal}(E/F)|}{|\text{Gal}(E/K)|}$. It also

follows by Lagrange's Theorem that $[K : F]$ is precisely the index of $|\text{Gal}(E/K)|$ in $|\text{Gal}(E/F)|$.

- (iii) Let $f(x) \in F[x]$, and let $X = \{a_1, a_2, \dots, a_n\} \subset E$ be the set of all zeros of $f(x)$ in E . Since E is normal over F , there are no zeros of $f(x)$ not in X . By Theorem 3.1.5, $\text{Gal}(E/F)$ permutes X .

Now suppose $f(x)$ is monic irreducible over F . Because $f(x)$ is irreducible, there are no repeated elements in X . By Theorem 3.1.5, for $1 \leq i, j \leq n$ with $i \neq j$, there exists $\sigma \in \text{Gal}(E/F)$ such that $\sigma(a_i) = a_j$.

- (iv) Suppose that $\text{Gal}(E/K)$ is normal in $\text{Gal}(E/F)$. Then for all $\varphi \in \text{Gal}(E/F)$, $\text{Gal}(E/K) = \varphi \text{Gal}(E/K) \varphi^{-1}$. Then by Lemma 3.3.2, $\text{Gal}(E/K) = \text{Gal}(E/\varphi K)$, and $K = \varphi K$. Hence, φ sends K into itself. Define a homomorphism $\Pi : \text{Gal}(E/F) \rightarrow \text{Gal}(K/F)$ by $\varphi \mapsto \varphi|_K$. Now,

$$\text{Ker } \Pi = \{\psi \mid \psi(k) = k \text{ for all } k \in K\} = \text{Gal}(E/K). \quad (3.18)$$

By the first isomorphism theorem, $\text{Gal}(E/F)/\text{Gal}(E/K)$ is isomorphic to a subgroup of $\text{Gal}(K/F)$. By part (ii) above, we have

$$[K : F] = |\text{Gal}(E/F) : \text{Gal}(E/K)| = |\text{Gal}(E/F)/\text{Gal}(E/K)| \leq |\text{Gal}(K/F)|. \quad (3.19)$$

From Corollary 3.1.6, $|\text{Gal}(K/F)| \leq [K : F] = |\text{Gal}(E/F)/\text{Gal}(E/K)|$.

Therefore, $\text{Gal}(E/F)/\text{Gal}(E/K) \cong \text{Gal}(K/F)$, and K is a normal extension of F .

Now suppose that K is normal over F . By Theorem 3.2.5, K is a splitting field of some $g(x) \in F[x]$ irreducible over F , and the zeros of $g(x)$ in K are the zeros of $g(x)$ in E . By Theorem 3.1.5, $\text{Gal}(E/K)$ permutes the zeros of $g(x)$ in

E and thus the zeros of $g(x)$ in K . So if $\psi \in \text{Gal}(K/F)$, then $\psi(K) = K$. By Lemma 3.3.2,

$$\text{Gal}(E/K) = \text{Gal}(E/\psi(K)) = \psi \text{Gal}(E/K) \psi^{-1}. \quad (3.20)$$

Therefore, $\text{Gal}(K/F)$ is normal in $\text{Gal}(E/F)$.

□

Corollary 3.3.4. *If E is a normal extension of F , then there is a one-to-one correspondence between intermediate fields K of E/F and the subgroups of $\text{Gal}(E/F)$.*

CHAPTER 4

SOLVABILITY

4.1 Solvable Groups and Polynomials

We follow Gallian [Gal10] and Jacobson [Jac85] in this section.

Definition 4.1.1. Let G be a group with subgroups $H_0, H_1, \dots, H_k$. If

$$\{e\} = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_k = G, \quad (4.1)$$

then (4.1) is said to be a **subnormal series of subgroups**.

From this point on, a **group property** P will be any property invariant under isomorphism.

Definition 4.1.2. Let G be a group and

$$\{e\} = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_k = G \quad (4.2)$$

be a subnormal series of subgroups. If for all $0 \leq i < k$ the factor group H_{i+1}/H_i has property P , then (4.2) is said to have **P -factors**.

Definition 4.1.3. If G is a group, then G is said to be **solvable** if G has a finite subnormal series of subgroups with Abelian factors.

Theorem 4.1.4. *A homomorphic image of a solvable group is solvable.*

Proof. Suppose G is a solvable group, L is a group, and $\varphi : G \rightarrow L$ is a surjective homomorphism. Since G is solvable, it has a subnormal series

$$\{e\} = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_n = G \quad (4.3)$$

with Abelian factors. Consider

$$\{e\} = \varphi(H_0) \subseteq \varphi(H_1) \subseteq \cdots \subseteq \varphi(H_n) = L. \quad (4.4)$$

Since surjective homomorphisms preserve normality, and since $H_i \triangleleft H_{i+1}$ for $0 \leq i \leq n-1$, $\varphi(H_i) \triangleleft \varphi(H_{i+1})$.

Define $\psi : H_{i+1}/H_i \rightarrow \varphi(H_{i+1})/\varphi(H_i)$ by $\psi(aH_i) = \varphi(a)\varphi(H_i)$. We claim that ψ is a well-defined, onto homomorphism. Suppose $aH_i = a'H_i$ for some $a, a' \in H_{i+1}$. Then $a' = ah_i$ for some $h_i \in H_i$, and

$$\begin{aligned} \psi(a'H_i) &= \varphi(a')\varphi(H_i) \\ &= \varphi(ah_i)\varphi(H_i) \\ &= \varphi(a)\varphi(h_i)\varphi(H_i) \\ &= \varphi(a)\varphi(H_i) \\ &= \varphi(aH_i). \end{aligned} \quad (4.5)$$

Hence, ψ is well-defined. Let $aH_i, bH_i \in H_{i+1}/H_i$. Consider

$$\begin{aligned} \psi(aH_i bH_i) &= \psi(abH_i) \\ &= \varphi(ab)\varphi(H_i) \\ &= \varphi(a)\varphi(b)\varphi(H_i) \\ &= \varphi(a)\varphi(H_i)\varphi(b)\varphi(H_i) \\ &= \psi(aH_i)\psi(bH_i). \end{aligned} \quad (4.6)$$

Hence, ψ is a homomorphism. Now suppose $\varphi(a)\varphi(H_i) \in \varphi(H_{i+1})/\varphi(H_i)$. Then for $aH_i \in H_{i+1}/H_i$, $\psi(aH_i) = \varphi(a)\varphi(H_i)$. It follows that ψ is onto.

Now, since H_{i+1}/H_i is Abelian by assumption and a quotient of an Abelian group is Abelian, $\varphi(H_{i+1})/\varphi(H_i)$ is Abelian. Therefore, $\varphi(G)$ is a solvable group. □

Theorem 4.1.5. *Let G be a group and $N \triangleleft G$. If N and G/N have subnormal series with P -factors, then G has a subnormal series with P -factors.*

Proof. Let

$$\{e\} = N_0 \triangleleft N_1 \triangleleft \cdots \triangleleft N_t = N \quad (4.7)$$

be a subnormal series for N with P -factors. Moreover, let

$$\{e\} = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_s = G/N \quad (4.8)$$

be a subnormal series for G/N with P -factors. Let $\varphi : G \rightarrow G/N$ be the natural homomorphism. Because preimages preserve normality, we have a series of preimages

$$N = \varphi^{-1}(H_0) \triangleleft \varphi^{-1}(H_1) \triangleleft \cdots \triangleleft \varphi^{-1}(H_s) = \varphi^{-1}(G/N) = G. \quad (4.9)$$

By the Third Isomorphism Theorem,

$$\varphi^{-1}(H_{i+1})/\varphi^{-1}(H_i) \cong H_{i+1}/H_i, \quad (4.10)$$

so (4.9) is a subnormal series with P -factors.

Consider the series

$$\{e\} = N_0 \subset N_1 \subset \cdots \subset N_t = N = \varphi^{-1}(H_0) \triangleleft \varphi^{-1}(H_1) \triangleleft \cdots \triangleleft \varphi^{-1}(H_s) = G. \quad (4.11)$$

By (4.7) and (4.9), the series in (4.11) has P -factors. Therefore, G has a subnormal series with P -factors. $\square$

Corollary 4.1.6. *Any finite Abelian group has a subnormal series with prime order factors.*

Proof. Let G be a finite Abelian group. We will induct on $n = |G|$. If $n = 1$, then the corollary is trivial.

Suppose $n = p_1 p_2 \cdots p_k \geq 2$. By Cauchy's Theorem, G contains an element of order p_1 , and hence a cyclic subgroup H of order p_1 . Since every subgroup of an Abelian group is normal, G/H is a group. Because $|G/H| = \frac{|G|}{|H|}$, the inductive hypothesis and Theorem 4.1.5 imply G/H has a subnormal series with prime factors. □

Lemma 4.1.7. *The following are equivalent*

- (i) G is a finite solvable group.
- (ii) G has a subnormal series of subgroups

$$\{e\} = H_0 \triangleleft H_1 \triangleleft H_2 \triangleleft \cdots \triangleleft H_k = G \quad (4.12)$$

with prime order factors.

Proof. Assume $\frac{|H_{i+1}|}{|H_i|}$ is prime; then $H_{i+1}/H_i \cong \mathbb{Z}_p$. Because $\mathbb{Z}_p$ is Abelian, so too is H_{i+1}/H_i .

Conversely, suppose G has a subnormal series with Abelian factors and assume for all $0 \leq i < k$, H_{i+1}/H_i is Abelian. We will induct on the number of groups in the subnormal series of G , k . If $k = 1$, then the lemma follows from Corollary 4.1.6. Suppose $k > 1$. By our inductive hypothesis, H_{k-1} has a subnormal series with Abelian prime order factors. Consider the factor group G/H_{k-1} . Since G/H_{k-1} is Abelian, G/H_{k-1} has a subnormal series with prime order factors. By Theorem 4.1.5, G has a subnormal series with prime order factors. □

Definition 4.1.8. If $F_0, F_1, \dots, F_k$ for some $k \in \mathbb{Z}_{>0}$ are fields such that

$$F_0 \subseteq F_1 \subseteq \cdots \subseteq F_k, \quad (4.13)$$

then (4.13) is called a **tower of fields**.

Definition 4.1.9. Let F be a field, and let $f(x) \in F[x]$. We say that $f(x)$ is **solvable by radicals over F** if $f(x)$ splits in some extension $F(a_1, a_2, \dots, a_n)/F$ and there exist $k_1, k_2, \dots, k_n \in \mathbb{Z}_{>0}$ such that $a_1^{k_1} \in F$ and $a_i^{k_i} \in F(a_1, \dots, a_{i-1})$ for $i = 2, \dots, n$.

The goal of this chapter is to prove the following theorem.

Theorem 4.1.10. *[Galois' Theorem] A polynomial is solvable by radicals if and only if its Galois group is solvable.*

4.2 Roots of Unity

We follow Gallian [Gal10], Jacobson [Jac89], and Saff and Snider [Sni03] in this section.

Definition 4.2.1. A field F is **algebraically closed** if every monic polynomial $f(x) \in F[x]$ with $\deg f(x) > 0$ has a zero in F .

Theorem 4.2.2. *[The Fundamental Theorem of Algebra] Every nonconstant polynomial in $\mathbb{C}[x]$ has at least one zero in $\mathbb{C}$.*

Proof. See Saff and Snider [Sni03, Theorem 22]. □

Theorem 4.2.3. *The field $\mathbb{C}$ is algebraically closed.*

Theorem 4.2.4. *Any finite extension of a subfield of $\mathbb{C}$ is isomorphic to a subfield of $\mathbb{C}$.*

Proof. Let F be a subfield of $\mathbb{C}$, E be a finite extension of F . By Corollary 2.7.5, $E = F(a)$ for some $a \in E$, where a has minimal polynomial $f(x) \in F[x]$. By Theorem 4.2.2, $f(x)$ has a zero b in a subfield L of $\mathbb{C}$. It follows that

$$E = F(a) \cong F(b) \subseteq L \subseteq \mathbb{C}. \quad (4.14)$$

Hence, E is isomorphic to an intermediate field of $\mathbb{C}/F$. $\square$

Hereafter, since we mainly study finite extensions of $\mathbb{Q}$, any field is assumed to be a subfield of $\mathbb{C}$ unless otherwise stated.

Definition 4.2.5. Let $n \in \mathbb{Z}_{>0}$. An **n th root of unity** is any number $\omega \in \mathbb{C}$ such that $\omega^n - 1 = 0$. If $\omega^k \neq 1$ for all $0 < k < n$, then ω is a **primitive n th root of unity**.

Theorem 4.2.6. Let $n \in \mathbb{Z}_{>0}$. If $\omega = \cos(\frac{2\pi}{n}) + i \sin(\frac{2\pi}{n})$, then ω is a primitive n th root of unity.

Proof. Let

$$\omega = e^{\frac{2\pi i}{n}} = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right). \quad (4.15)$$

By DeMoivre's Theorem,

$$\omega^k = e^{\frac{2k\pi i}{n}} = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right) \quad (4.16)$$

for all $k \in \mathbb{Z}_{>0}$. It follows that $\omega^n = 1$, and if $1 \leq k < n$, then $\omega^k \neq 1$. $\square$

Theorem 4.2.7. Let $n \in \mathbb{Z}_{>0}$. The sum of all n th roots of unity is zero.

Proof. Let ω be a primitive n th root of unity. Then

$$X = \{1, \omega, \omega^2, \dots, \omega^{n-1}\} \quad (4.17)$$

is the set of all n th roots of unity. Now, $\omega \neq 1$ and $\omega^n = 1$, so

$$\omega^{n-1} + \omega^{n-2} + \dots + \omega + 1 = \frac{\omega^n - 1}{\omega - 1} = 0. \quad (4.18)$$

$\square$

Theorem 4.2.8. Let $n \in \mathbb{Z}_{>0}$, ω be a primitive n th root of unity, and F be a subfield of $\mathbb{C}$ containing ω . Then $F(\omega)$ is the splitting field for $x^n - 1$ over F .

Proof. Because

$$X = \{1, \omega, \omega^2, \dots, \omega^{n-1}\} \quad (4.19)$$

is the set of all zeros of $x^n - 1$, and $X = \langle \omega \rangle$ as a multiplicative group, we have from Theorem 2.12.4 that $F(\omega)$ is a splitting field for $x^n - 1$ over F . $\square$

4.3 Galois' Theorem Part I

We follow Gallian [Gal10] in this section.

Theorem 4.3.1. *Let F be a subfield of $\mathbb{C}$, $a \in F$, $n \in \mathbb{Z}_{>0}$, and $f(x) = x^n - a \in F[x]$. If E is the splitting field for $f(x)$ over F , then $G = \text{Gal}(E/F)$ is solvable. In particular, G has a normal subgroup N such that N is isomorphic to a subgroup of $\mathbb{Z}_n$, and G/N is isomorphic to a subgroup of the multiplicative group*

$$U(n) = \{k \in \mathbb{Z}_n \mid \gcd(k, n) = 1\} \cong \text{Aut } \mathbb{Z}_n. \quad (4.20)$$

Proof. If $a = 0$, then $E = F$, and we are done. Assume $a \neq 0$. Let ω be a primitive n th root of unity and $b \in E$ such that $f(b) = 0$. Since $b \neq 0$, the zeros of $f(x)$ in E make up the set

$$X = \{b, \omega b, \omega^2 b, \dots, \omega^{n-1} b\}. \quad (4.21)$$

Let us first suppose that $\omega \in F$. Then $E = F(b)$. Let $\varphi \in G$. Now, by Theorem 3.1.5, φ is completely determined by where φ maps b . By the Fundamental Theorem of Galois Theory, $\varphi(b) \in X$, so $\varphi(b) = \omega^i b$ for some $1 \leq i < n$. Suppose $\psi \in G$. Then $\psi(b) = \omega^j b$ for some $1 \leq j < n$. Since $\omega \in F$,

$\varphi(\omega) = \omega = \psi(\omega)$. We observe that

$$\begin{aligned}
 (\varphi\psi)(b) &= \varphi(\psi(b)) \\
 &= \varphi(\omega^j b) \\
 &= \varphi(\omega)^j \varphi(b) \\
 &= \omega^j (\omega^i b) \\
 &= \omega^{i+j} b.
 \end{aligned} \tag{4.22}$$

Define the map $\Phi : G \rightarrow \mathbb{Z}_n$ by $\Phi(\sigma) = k$ if for $\sigma \in G$, $\sigma(b) = \omega^k b$. By (4.22), $\Phi(\varphi\psi) = i + j = \Phi(\varphi) + \Phi(\psi)$, so Φ is a homomorphism. Suppose $\Phi(\varphi) = \Phi(\sigma)$. Then $i = k$, so $\omega^i = \omega^k$, and $\omega^i b = \omega^k b$. It follows that $\varphi(b) = \sigma(b)$, and by Theorem 3.1.5, $\varphi = \sigma$. Therefore, Φ is an injective homomorphism. By the first isomorphism theorem, G is isomorphic to $\Phi(G) \leq \mathbb{Z}_n$. Because $\mathbb{Z}_n$ is Abelian, G is Abelian.

On the other hand, suppose that F does not contain a primitive n th root of unity. So $\omega \notin F$. Since $b \neq 0$, $b^{-1} \in E$. Moreover, since $f(b) = 0$ and $f(\omega b) = 0$, $(\omega b)(b^{-1}) \in E$. Hence, $F(\omega) \subseteq E$. Notice, too, that $F(\omega)$ is the splitting field for $x^n - 1$ over F by Theorem 4.2.8.

Let $\alpha, \beta \in \text{Gal}(F(\omega)/F)$, and suppose $\alpha(\omega) = \omega^i, \beta(\omega) = \omega^j$. We observe that

$$\begin{aligned}
 (\alpha\beta)(\omega) &= \alpha(\beta(\omega)) \\
 &= \alpha(\omega^j) \\
 &= (\alpha(\omega))^j \\
 &= (\omega^i)^j \\
 &= \omega^{ij}.
 \end{aligned} \tag{4.23}$$

Define a map $\Psi : \text{Gal}(F(\omega)/F) \rightarrow U(n)$ by $\Psi(\alpha) = i$ if $\alpha(\omega) = \omega^i$. Suppose $\sigma \in \text{Gal}(F(\omega)/F)$ such that $\sigma(\omega) = \omega^k$ for some $k \in \mathbb{Z}_n$. Since $\text{Gal}(F(\omega)/F)$ is a group, there exists some $\tau \in \text{Gal}(F(\omega)/F)$ such that $\tau\sigma = \text{id}$ and $\tau(\omega) = \omega^\ell$. That

is,

$$\omega^{k\ell} = (\tau\sigma)(\omega) = id(\omega) = \omega^1. \quad (4.24)$$

Then $k\ell = 1 \pmod{n}$, and k is invertible. Therefore, Ψ is well-defined.

By (4.23), Ψ is a homomorphism. Suppose $\Psi(\alpha) = \Psi(\gamma)$ for some $\alpha, \gamma \in \text{Gal}(F(\omega)/F)$, where $\alpha(\omega) = \omega^i$ and $\gamma(\omega) = \omega^k$. Then $i = k \pmod{n}$, $\omega^i = \omega^k$, and $\alpha(\omega) = \gamma(\omega)$. Therefore, by Theorem 3.1.5, $\alpha = \gamma$, and Ψ is an injective homomorphism. Because $U(n)$ is Abelian, $\text{Gal}(F(\omega)/F)$ is Abelian.

Notice that E is a splitting field for $f(x)$ over $F(\omega)$. Then by the first case, $\text{Gal}(E/F(\omega))$ is Abelian. By the Fundamental Theorem of Galois Theory,

$$\{e\} \triangleleft \text{Gal}(E/F(\omega)) \triangleleft G, \quad (4.25)$$

and

$$G/\text{Gal}(E/F(\omega)) \cong \text{Gal}(F(\omega)/F). \quad (4.26)$$

Since, G and $\text{Gal}(F(\omega)/F)$ are Abelian, G is solvable. $\square$

Theorem 4.3.2. *Let F be a subfield of $\mathbb{C}$ and $f(x) \in F[x]$. Suppose $f(x)$ splits in $F(a_1, a_2, \dots, a_t)$, where $a_1^{n_1} \in F$ and $a_i^{n_i} \in F(a_1, \dots, a_{i-1})$ for $i = 2, 3, \dots, t$ where n_1, n_i are nonnegative integers. If E is the splitting field for $f(x)$ over F in $F(a_1, a_2, \dots, a_t)$, then $\text{Gal}(E/F)$ is solvable.*

Proof. We will proceed by induction on t .

Suppose $t = 1$. Then $F \subseteq E \subseteq F(a_1)$.

Let L be the splitting field for $x^{n_1} - a_1^{n_1}$ over F . As shown in Figure 4.3, we have $F \subseteq E \subseteq F(a_1) \subseteq L$. By Theorem 3.3.3,

$$\text{Gal}(E/F) \cong \text{Gal}(L/F)/\text{Gal}(L/E). \quad (4.27)$$

By Theorem 4.3.1, the group $\text{Gal}(L/F)$ is solvable. From Theorem 4.1.4, the group $\text{Gal}(L/F)/\text{Gal}(L/E)$ is solvable, and so $\text{Gal}(E/F)$ is solvable.

Figure 4.1: Tower of Fields with $t = 1$.

Suppose $t > 1$.

Let $a_1^{n_1} \in F$, $g(x) = x^{n_1} - a_1^{n_1}$, L be the splitting field for $g(x)$ over F , and K be the splitting field for $g(x)$ over E , as shown in Figure 4.3. By Theorem 4.3.1, $\text{Gal}(L/F)$ is solvable. Notice that $F \subseteq L \subseteq K$ and $F \subseteq E \subseteq K$. Since $f(x) \in F[x]$ has splitting field E , K is the splitting field for $g(x)f(x)$ over F , K is a splitting field for $f(x)$ over L , and $L \subseteq K \subseteq L(a_2, a_3, \dots, a_t)$, as shown in Figure 4.3. Because $a_1 \in L$, $f(x) \in L[x]$ splits in $L(a_2, a_3, \dots, a_t)$. Since K is the splitting field for $f(x)$ over L , our inductive hypothesis states that $\text{Gal}(K/L)$ is solvable. By Theorem 3.3.3,

$$\text{Gal}(L/F) \cong \text{Gal}(K/F) / \text{Gal}(K/L). \quad (4.28)$$

Since $\text{Gal}(L/F)$ and $\text{Gal}(K/L)$ are solvable, Theorem 4.1.5 implies that $\text{Gal}(K/F)$ is solvable. By part (iv) of Theorem 3.3.3, since E/F is normal, $\text{Gal}(K/E)$ is a normal subgroup of $\text{Gal}(K/F)$. Hence, Theorem 4.1.4 states that $\text{Gal}(K/F) / \text{Gal}(K/E)$ is solvable. Then by Theorem 3.3.3,

$$\text{Gal}(E/F) \cong \text{Gal}(K/F) / \text{Gal}(K/E), \quad (4.29)$$

and $\text{Gal}(E/F)$ is therefore solvable. □

Figure 4.2: Diagram of Fields with $t > 1$.

4.4 Galois' Theorem Part II

We follow Hadlock [Had78] in this section.

Definition 4.4.1. Let $n \in \mathbb{Z}_{>0}$. We define the n th cyclotomic polynomial over $\mathbb{Q}$ as

$$\Phi_n(x) = \frac{x^n - 1}{x - 1} = x^{n-1} + x^{n-2} + \cdots + x + 1. \quad (4.30)$$

Theorem 4.4.2. If $p \in \mathbb{Z}_{>0}$ is prime, then the p th cyclotomic polynomial is irreducible over $\mathbb{Q}$.

Proof. See Gallian [Gal10, Theorem 17.4]. □

Lemma 4.4.3. Let F be a field and $f(x) \in F[x]$ have splitting field E . If $\text{Gal}(E/F)$ is solvable, then there is a tower of fields

$$F = F_0 \subset F_1 \subset \cdots \subset F_k = E \quad (4.31)$$

such that for each i such that $0 \leq i < k$, F_{i+1} is a normal extension of F_i and $[F_{i+1} : F_i]$ is prime.

Proof. By Corollary 4.1.6, there is a subnormal series of subgroups of $\text{Gal}(E/F)$

$$\{e\} = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_N = \text{Gal}(E/F), \quad (4.32)$$

such that for each j such that $0 < j \leq N$, $|H_j/H_{j-1}|$ is prime. By Corollary 3.3.4, there is a corresponding tower of fields

$$E = F_N \supset F_{N-1} \supset \cdots \supset F_0 = F, \quad (4.33)$$

where $E_{H_i} = F_{N-i}$ with $0 \leq i \leq N$. By Theorem 3.2.5, E is normal over F , so E is normal over F_i . Hence, $H_i = \text{Gal}(E/F_{N-i})$. Because $H_{i+1} \triangleleft H_i$, F_{i+1}/F_i is normal by Theorem 3.3.3 applied to E/F_i .

We now observe the degree of the extension F_{i+1}/F_i is

$$[F_{i+1} : F_i] = \frac{[E : F_i]}{[E : F_{i+1}]} = \frac{|\text{Gal}(E/F_i)|}{|\text{Gal}(E/F_{i+1})|} = \frac{|H_i|}{|H_{i+1}|}. \quad (4.34)$$

Since the last quotient is prime by assumption, $[F_{i+1} : F_i]$ is prime. $\square$

Lemma 4.4.4. *Let $n \in \mathbb{Z}_{>0}$. Fix a prime integer p and let ω be a primitive n th root of unity. If K is a field containing ω and $\tilde{K}$ is a normal extension of K with $[\tilde{K} : K] = p$, then there exist $a_1, a_2, \dots, a_{p-1} \in \mathbb{C}$ such that for each $1 \leq i \leq (p-1)$, $(a_i)^p \in K$ and $\tilde{K} \subseteq K(a_1, a_2, \dots, a_{p-1})$.*

Proof. Since $\tilde{K}/K$ is finite, we can write $\tilde{K} = K(r)$ for some $r \in \tilde{K}$ of degree p over K . Because $\tilde{K}/K$ is normal, all of the conjugates of r are in $\tilde{K}$.

Because $\tilde{K}/K$ is normal, Theorem 3.2.5 states $[\tilde{K} : K] = |\text{Gal}(\tilde{K}/K)| = p$. Since $\text{Gal}(\tilde{K}/K)$ has prime order, it is cyclic. Then for some $\varphi \in \text{Gal}(\tilde{K}/K)$,

$$\langle \varphi \rangle = \text{Gal}(\tilde{K}/K) = \{id, \varphi, \varphi^2, \dots, \varphi^{p-1}\}. \quad (4.35)$$

Define $r_i = \varphi^i(r)$ for $0 \leq i \leq p-1$. Because the φ^i for $0 \leq i \leq p-1$ are distinct and $[\tilde{K} : K] = p$, the set $R = \{r_0, r_1, \dots, r_{p-1}\}$ is the set of the p distinct conjugates of r .

Let ω be a primitive p th root of unity. Notice that $\varphi(\omega) = \omega$ because $\omega \in K$ by assumption, and φ fixes K . Define $a_0, a_1, \dots, a_{p-1}$ by

$$\begin{aligned} a_0 &= r_0 + r_1 + r_2 + \dots + r_{p-1}, \\ a_1 &= r_0 + \omega r_1 + \omega^2 r_2 + \dots + \omega^{(p-1)} r_{p-1}, \\ &\vdots \\ a_j &= r_0 + \omega^j r_1 + \omega^{2j} r_2 + \dots + \omega^{(p-1)j} r_{p-1}, \\ &\vdots \\ a_{p-1} &= r_0 + \omega^{p-1} r_1 + \omega^{2(p-1)} r_2 + \dots + \omega^{(p-1)(p-1)} r_{p-1}. \end{aligned} \quad (4.36)$$

We observe that

$$\begin{aligned}
\varphi(a_0) &= \varphi(r_0 + r_1 + \cdots + r_{p-1}) \\
&= \varphi(r_0) + \varphi(r_1) + \cdots + \varphi(r_{p-1}) \\
&= \varphi(r) + \varphi(\varphi(r)) + \cdots + \varphi(\varphi^{p-1}(r)) \\
&= \varphi(r) + \varphi^2(r) + \cdots + \varphi^p(r) \\
&= r_1 + r_2 + \cdots + r_{p-1} + r_0 \\
&= r_0 + r_1 + \cdots + r_{p-1} \\
&= a_0,
\end{aligned} \tag{4.37}$$

so $a_0 \in K$. Now, for each $1 \leq j < p$,

$$\begin{aligned}
\varphi((a_j)^p) &= (\varphi(a_j))^p \\
&= (\varphi(r_0) + \omega^j \varphi(r_1) + \cdots + \omega^{(p-1)j} \varphi(r_{p-1}))^p \\
&= (r_1 + \omega^j r_2 + \cdots + \omega^{(p-1)j} r_0)^p \\
&= (\omega^{-j}(r_1 \omega^j + r_2 \omega^{2j} + \cdots + \omega^{pj} r_0))^p \\
&= (\omega^{-j}(r_0 + r_1 \omega^j + r_2 \omega^{2j} + \cdots + \omega^{(p-1)j} r_{p-1}))^p \\
&= (\omega^{-j} a_j)^p \\
&= \omega^{-pj} (a_j)^p \\
&= (a_j)^p.
\end{aligned} \tag{4.38}$$

Hence, $\varphi^i((a_j)^p) = (a_j)^p$ for all i . Since $\varphi^i \in \text{Gal}(\tilde{K}/K)$ and by Theorem 3.2.5, $(a_j)^p \in K$.

Consider the system

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{p-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{p-1} \\ 1 & \omega^2 & (\omega^2)^2 & \dots & (\omega^2)^{p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{p-1} & (\omega^{p-1})^2 & \dots & (\omega^{p-1})^{p-1} \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_{p-1} \end{bmatrix} = V \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ \vdots \\ r_{p-1} \end{bmatrix}. \quad (4.39)$$

By Theorem 2.2.6,

$$\det V = \prod_{0 \leq i < j \leq p-1} (\omega^j - \omega^i). \quad (4.40)$$

Because $\omega^j \neq \omega^i$ for all i, j such that $0 \leq i < j \leq p-1$, $\det V \neq 0$. Therefore, since

$$\begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{p-1} \end{bmatrix} = V^{-1} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p-1} \end{bmatrix}, \text{ for each } i, r_i \in K(a_1, a_2, \dots, a_{p-1}), \text{ and}$$

$$\tilde{K} = K(r) \subset K(a_1, a_2, \dots, a_{p-1}). \quad (4.41)$$

□

Lemma 4.4.5. *Fix a prime integer p , and let K be a field. If $\tilde{K}$ is a normal extension of K such that $[\tilde{K} : K] = p$ and ω is a p th root of unity in $\tilde{K}$, then $\tilde{K}(\omega)$ is a normal extension of $K(\omega)$ such that $[\tilde{K}(\omega) : K(\omega)] = p$.*

Proof. Since $\tilde{K}/K$ is finite, we can write $\tilde{K} = K(r)$ for some $r \in \tilde{K}$ algebraic over K , and let $g(x) \in K[x]$ be the minimal polynomial of r over K . By assumption, $[\tilde{K} : K] = p$. By Theorem 4.4.2, since p is prime, the p th cyclotomic polynomial is irreducible over $\mathbb{Q}$, ω is a zero of this polynomial, and $\deg_{\mathbb{Q}} \omega = p-1$. Since ω is a p th root of unity and

$$m = \deg_K \omega \leq \deg_{\mathbb{Q}} \omega = p-1, \quad (4.42)$$

$[K(\omega) : K] = m$. We have from the degree theorem that

$$[\tilde{K}(\omega) : K] = [\tilde{K}(\omega) : \tilde{K}][\tilde{K} : K] = [\tilde{K}(\omega) : \tilde{K}]p, \quad (4.43)$$

and

$$[\tilde{K}(\omega) : K] = [\tilde{K}(\omega) : K(\omega)][K(\omega) : K] = [\tilde{K}(\omega) : K(\omega)]m. \quad (4.44)$$

Because $\gcd(p, m) = 1$, p divides $[\tilde{K}(\omega) : K(\omega)]$, so $[\tilde{K}(\omega) : K(\omega)] \geq p$. On the other hand, because $\tilde{K} = K(r)$,

$$[\tilde{K}(\omega) : K(\omega)] = [K(r, \omega) : K(\omega)] \leq [K(r) : K] = [\tilde{K} : K] = p. \quad (4.45)$$

Hence, $[\tilde{K}(\omega) : K(\omega)] \leq p$. Therefore, $[\tilde{K}(\omega) : K(\omega)] = p$.

By Theorem 4.2.8, the splitting field for $(x^p - 1)$ over K is $K(\omega)$. Let $f(x) = g(x)(x^p - 1)$. Then the splitting field for $f(x)$ over K is $K(r, \omega) = \tilde{K}(\omega)$. By Theorem 3.2.5, $\tilde{K}(\omega)$ is normal over K . By Corollary 3.2.7, $\tilde{K}(\omega)$ is normal over $K(\omega)$. $\square$

Theorem 4.4.6. *Let F be a field, $f(x) \in F[x]$, and E be the splitting field for $f(x)$ over F . If $\text{Gal}(E/F)$ is solvable, then $f(x)$ is solvable by radicals.*

Proof. By Lemma 4.4.3, there exists a finite tower of fields

$$F = F_0 \subset F_1 \subset \cdots \subset F_k = E \quad (4.46)$$

such that for each $0 \leq j \leq k$, F_{j+1}/F_j is normal and $[F_{j+1} : F_j] = p$, where p is a prime integer. Let ω be a primitive p th root of unity. By Lemmas 4.4.4 and 4.4.5,

$$F_{j+1} \subseteq F_j(\omega_j, a_{1j}, a_{2j}, \dots, a_{(p-1)j}), \quad (4.47)$$

where, for each i such that $1 \leq i < p$, $(a_i)^p \in F_j(\omega)$. Consider the tower of fields obtained by successively adjoining to F the elements

$$\omega_0, a_{10}, a_{20}, \dots, a_{(p-1)0}, \omega_1, a_{11}, a_{21}, \dots, a_{(p-1)1}, \dots, a_{(p-1)k}. \quad (4.48)$$

This is a finite tower of radical extensions such that the ultimate extension contains $F_k = E$. Therefore, $f(x)$ is solvable by radicals. $\square$

By combining Theorems 4.3.2 and 4.4.6, we obtain Galois' Theorem.

Theorem 4.4.7 (Galois' Theorem). *A polynomial is solvable by radicals over a subfield of $\mathbb{C}$ if and only if its Galois group is solvable.*

Proof. Q.E.D. $\square$

CHAPTER 5

RATIONAL POLYNOMIALS AND SYMMETRIC GALOIS GROUPS

5.1 Symmetric Functions

We follow Hadlock [Had78] in this section.

We continue the assumption that any field F is a subfield of $\mathbb{C}$.

Definition 5.1.1. Let F be a field and $P(x_1, x_2, \dots, x_n) \in F[x_1, x_2, \dots, x_n]$, where $n \in \mathbb{Z}_{>0}$. We say that $P(x_1, x_2, \dots, x_n)$ is a **symmetric polynomial** if

$$P(x_1, x_2, \dots, x_n) = P(x_{\varphi(1)}, x_{\varphi(2)}, \dots, x_{\varphi(n)}) \quad (5.1)$$

for all $\varphi \in S_n$.

Definition 5.1.2. For $n \in \mathbb{Z}_{>0}$, let $[n]$ denote the set $\{1, 2, \dots, n\}$, and $\binom{[n]}{k}$ denote the k -subsets of $[n]$. Then for $I = \{i_1, i_2, \dots, i_k\} \in \binom{[n]}{k}$, we define $x_I = x_{i_1}x_{i_2} \dots x_{i_k}$. Define symmetric polynomials $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_n$ on the n variables $x_1, x_2, \dots, x_n$ by

$$\begin{aligned} \sigma_0 &= 1 \\ \sigma_1 &= x_1 + x_2 + \dots + x_n \\ \sigma_2 &= \sum_{1 \leq i < j \leq n} x_i x_j \\ &\vdots \\ \sigma_k &= \sum_{I \in \binom{[n]}{k}} x_I \\ &\vdots \\ \sigma_n &= x_1 x_2 \dots x_n. \end{aligned} \quad (5.2)$$

We call the symmetric functions $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_n$ the **elementary symmetric functions on n variables**.

Lemma 5.1.3. *Given $n \in \mathbb{Z}_{>0}$ and $0 \leq k \leq n$, let the function F_n adjoin to any subset of $[n-1]$ the element n . Then $\binom{[n]}{k}$ is the disjoint union of $\binom{[n-1]}{k}$ and $F_n \left(\binom{[n-1]}{k-1} \right)$.*

Proof. By definition $\binom{[n-1]}{k}$ contains all of the k -subsets of $\binom{[n]}{k}$ that do not contain n in any term. Moreover, $F_n \left(\binom{[n-1]}{k-1} \right)$ contains all of the k -subsets of $\binom{[n]}{k}$ that have n in every term. Because these two sets form a partition of $\binom{[n]}{k}$ into k -subsets containing n and k -subsets not containing n , the disjoint union of the two is precisely $\binom{[n]}{k}$. □

Lemma 5.1.4. *Let F be a field and $n \in \mathbb{Z}_{>0}$, and suppose*

$f(x) = \prod_{i=1}^n (x - a_i) \in F[x]$ has splitting field $E = F(a_1, a_2, \dots, a_n)$ over F . Let $b_i = (-1)^i \sigma_i(a_1, a_2, \dots, a_n)$ where σ_i is the i th elementary symmetric function on n objects, $0 \leq i \leq n$. If $g(x) = \sum_{i=0}^n b_i x^{n-i}$, then $f(x) = g(x)$.

Proof. We will induct on n . If $n = 1$, then $f_1(x) = x - a_1$, $b_1 = -a_1$, $g_1(x) = x - a_1$, and $f_1(x) = g_1(x)$.

Suppose the lemma holds for $n = k - 1$. Consider $f_k(x) = \prod_{i=1}^k (x - a_i)$, and let $b_i = (-1)^i \sigma_i(a_1, a_2, \dots, a_k)$ and $B_j = (-1)^j \overline{\sigma}_j(a_1, a_2, \dots, a_{k-1})$ where $\overline{\sigma}_j$ is the j th elementary symmetric function on $k - 1$ elements, $0 \leq j \leq k - 1$. By our inductive

hypothesis, $f_{k-1}(x) = g_{k-1}(x) = \sum_{j=0}^{k-1} B_j x^{k-1-j}$. Then

$$\begin{aligned}
f_k(x) &= (f_{k-1}(x))(x - a_k) \\
&= \left(\sum_{j=0}^{k-1} B_j x^{(k-1)-j} \right) (x - a_k) \\
&= \sum_{j=0}^{k-1} (-1)^j \overline{\sigma_j}(a_1, a_2, \dots, a_{k-1}) x^{(k-1)-j} (x - a_k) \\
&= \sum_{j=0}^{k-1} (-1)^j \sum_{I \in \binom{[k-1]}{j}} a_I x^{(k-1)-j} (x - a_k) \\
&= \sum_{j=0}^{k-1} \left((-1)^j \sum_{I \in \binom{[k-1]}{j}} a_I x^{k-j} \right) + \sum_{j=0}^{k-1} \left((-1)^{j+1} \sum_{I \in \binom{[k-1]}{j}} a_I a_k x^{k-1-j} \right) \\
&= \sum_{\ell=0}^{k-1} \left((-1)^\ell \sum_{I \in \binom{[k-1]}{\ell}} a_I x^{k-\ell} \right) + \sum_{\ell=1}^k \left((-1)^\ell \sum_{I \in \binom{[k-1]}{\ell-1}} a_I a_k x^{k-\ell} \right) \\
&= x^k + \left(\sum_{\ell=1}^{k-1} (-1)^\ell \sum_{I \in \binom{[k-1]}{\ell}} a_I x^{k-\ell} \right) + \left(\sum_{\ell=1}^{k-1} (-1)^\ell \sum_{I \in \binom{[k-1]}{\ell-1}} a_I a_k x^{k-\ell} \right) \\
&\quad + (-1)^k a_1 a_2 \cdots a_{k-1} a_k \\
&= x^k + \left(\sum_{\ell=1}^{k-1} (-1)^\ell \sum_{I \in \binom{[k]}{\ell}} a_I x^{k-\ell} \right) + (-1)^k a_1 a_2 \cdots a_k \quad \text{by Lemma 5.1.3} \\
&= \sum_{\ell=0}^k (-1)^\ell \sum_{I \in \binom{[k]}{\ell}} a_I x^{k-\ell} \\
&= \sum_{\ell=0}^k b_\ell x^{k-\ell} \\
&= g_k(x).
\end{aligned} \tag{5.3}$$

□

Definition 5.1.5. Let $n \in \mathbb{Z}_{>0}$ and $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n), \beta = (\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{Z}_{\geq 0}^n$.

We define the **lexicographic ordering** on $\mathbb{Z}_{\geq 0}^n$ by declaring that $\alpha > \beta$ if and only if $\alpha_i - \beta_i > 0$ in the leftmost nonzero entry of $\alpha - \beta$, and $\alpha \geq \beta$ if and only if $\alpha > \beta$ or $\alpha = \beta$.

Lemma 5.1.6. *Let $n \in \mathbb{Z}_{>0}$. The lexicographic ordering is a total ordering on $\mathbb{Z}_{\geq 0}^n$.*

Proof. Let $\alpha, \beta, \gamma \in \mathbb{Z}_{\geq 0}^n$. Suppose $\alpha \leq \beta$ and $\beta \leq \alpha$. Because $\alpha \leq \beta$, $\alpha_i \leq \beta_i$ in the first place that they differ. Suppose they differ first in the k th place. Then since $\alpha \leq \beta$, $\alpha(k) \leq \beta(k)$, and because $\beta \leq \alpha$, $\beta(k) \leq \alpha(k)$. Hence, $\alpha = \beta$, and the lexicographic ordering is antisymmetric.

Now suppose $\alpha \leq \beta \leq \gamma$ in lexicographic ordering. If $\alpha = \beta$ or $\beta = \gamma$, then $\alpha \leq \gamma$, so we can assume that $\alpha < \beta < \gamma$. Suppose α and β differ first in the k th position and β and γ differ first in the j th position. If $k > j$, then $\alpha_j = \beta_j < \gamma_j$, and $\alpha_i = \beta_i = \gamma_i$ for all $1 \leq i < j$. Hence, $\alpha < \gamma$. If $k < j$, then $\alpha_k < \beta_k = \gamma_k$, and $\alpha_i = \beta_i = \gamma_i$ for all $1 \leq i < k$. Hence, $\alpha < \gamma$. If $k = j$, then $\alpha_k < \beta_k < \gamma_k$, and $\alpha < \gamma$. Hence, the lexicographic ordering is transitive.

It is left to be shown that if α and β are arbitrary elements of $\mathbb{Z}_{\geq 0}^n$, then either $\alpha \leq \beta$ or $\beta \leq \alpha$. This holds when $\alpha = \beta$, so suppose they differ first in the k th place. If $\alpha_k > \beta_k$, then $\alpha > \beta$. Similarly, if $\alpha_k < \beta_k$, then $\alpha < \beta$. Therefore, $\alpha \leq \beta$ or $\beta \leq \alpha$, and the lexicographic ordering is a total ordering. $\square$

Theorem 5.1.7. *Let F be a field, $n \in \mathbb{Z}_{>0}$, and $P(x_1, x_2, \dots, x_n) \in F[x_1, x_2, \dots, x_n]$ be a symmetric polynomial. Then P can be written as a polynomial $Q \in F[\sigma_1, \sigma_2, \dots, \sigma_n]$, and if P has integer coefficients, then Q has integer coefficients.*

Proof. Let $M = \deg P$. For this proof, we will consider each monomial term of the

symmetric polynomial P as an element $(k_1, k_2, \dots, k_n)$ of

$$\mathbb{Z}_{\geq 0}^n(M) = \{(k_1, k_2, \dots, k_n) | k_i \in \mathbb{Z}_{\geq 0}, \sum_{i=1}^n k_i \leq M\}. \quad (5.4)$$

Let us call the monomial term of P with the greatest n -tuple representation in lexicographic order the *highest* term of P .

Let N be the number of n -tuples of the finite set $\mathbb{Z}_{\geq 0}^n(M)$ less than or equal to the highest term of $P(x_1, x_2, \dots, x_n)$. We will induct on N . In the base case, if the highest term is $(0, 0, \dots, 0)$, then all exponents of x_i in P are 0. Hence, P is a constant polynomial, and $P = Q \in F[\sigma_1, \sigma_2, \dots, \sigma_n]$. If $P \in \mathbb{Z}$, then $Q \in \mathbb{Z}$.

Suppose the highest term of P is $ax_1^{i_1}x_2^{i_2} \cdots x_n^{i_n}$, where $a \in F$. Because P is symmetric, by definition of lexicographic order, it follows that $i_1 \geq i_2 \geq \cdots \geq i_n$. Let

$$Q(x_1, x_2, \dots, x_n) = a\sigma_1^{i_1-i_2}\sigma_2^{i_2-i_3} \cdots \sigma_{n-1}^{i_{n-1}-i_n}\sigma_n^{i_n} = a \left(\prod_{j=1}^{n-1} \sigma_j^{i_j-i_{j+1}} \right) \sigma_n^{i_n}. \quad (5.5)$$

Since $i_j > 0$ for all j , $i_j - i_{j+1} \leq i_j$ for $1 \leq j \leq n-1$. Substituting the $\sigma_i(x_1, x_2, \dots, x_n)$ into Q , we have

$$\begin{aligned} Q_1(x_1, x_2, \dots, x_n) &= a(x_1 + x_2 + \cdots + x_n)^{i_1-i_2} \cdot \\ &\quad (x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n)^{i_2-i_3} \cdots \\ &\quad (x_1x_2 \cdots x_n)^{i_n}. \end{aligned} \quad (5.6)$$

After expanding, we can see from symmetry that each term of Q_1 has the same degree. By our ordering, the highest term of Q_1 in $F[x_1, x_2, \dots, x_n]$ is the term with the most x_1 's, then the most x_2 's, and so on. Thus, the highest term of Q_1 is

$$ax_1^{i_1-i_2}(x_1x_2)^{i_2-i_3}(x_1x_2x_3)^{i_3-i_4} \cdots (x_1x_2 \cdots x_n)^{i_n} = ax_1^{i_1}x_2^{i_2} \cdots x_n^{i_n}. \quad (5.7)$$

Let $R = P - Q_1$. Because the highest term of P is the same as the highest term of Q_1 , either $R = 0$ or R has a lesser highest term than P . If $R = 0$, then $P = Q_1$, and

we're done. If $R \neq 0$, then by our inductive hypothesis, R can be written as a polynomial $S \in F[\sigma_1, \sigma_2, \dots, \sigma_n]$, where S has integer coefficients if R does. Because Q and S are both polynomials in $F[\sigma_1, \sigma_2, \dots, \sigma_n]$ and both have integer coefficients if P does, $Q + S$ satisfies the same conditions. $\square$

5.2 Large Extensions with Galois Group S_n

We follow Hadlock [Had78] in this section.

Definition 5.2.1. Let F be a subfield of $\mathbb{C}$. A collection of elements $a_1, a_2, \dots, a_n \in \mathbb{C}$ are **algebraically independent over F** if for all nonzero $p(x_1, x_2, \dots, x_n) \in F[x_1, x_2, \dots, x_n]$, $p(a_1, a_2, \dots, a_n) \neq 0$.

Lemma 5.2.2. *The set of algebraic elements over a countable field is countable.*

Proof. Let F be a countable field, $n \in \mathbb{Z}_{>0}$, and for $\alpha = (a_0, a_1, \dots, a_{n-1}) \in F^n$, let

$$f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 \in F[x]. \quad (5.8)$$

Let E be a splitting field for $f(x)$ over F , and let A_α be the set of all zeros of $f(x)$ in E . Observe that $|A_\alpha| \leq n$. Let

$$A_n = \bigcup_{\alpha \in F^n} A_\alpha. \quad (5.9)$$

We can express the set of all algebraic elements over F as

$$\mathcal{A} = \bigcup_{n \in \mathbb{Z}_{>0}} A_n. \quad (5.10)$$

Because A_n is a countable union of finite sets and $\mathcal{A}$ is a countable union of countable unions of finite sets, $\mathcal{A}$ is countable. $\square$

Lemma 5.2.3. *If F is a countable field, then for every $n \in \mathbb{Z}_{>0}$ there exist n algebraically independent elements over F .*

Proof. Fix $n \in \mathbb{Z}_{>0}$. By Lemma 5.2.2, the set of elements algebraic over F must also be countable. Because $\mathbb{C}$ is an uncountable field, there exists a transcendental element $a_1 \in C$ over F . Because F is countable, so too is the set

$$F(a_1) = \left\{ \frac{f(a_1)}{g(a_1)} \left| f(x), g(x) \in F[x], g(x) \neq 0 \right. \right\}. \quad (5.11)$$

By parallel argument, we can construct a_2 transcendental over $F(a_1)$, a_3 transcendental over $F(a_1, a_2)$, and by iteration, a_n transcendental over $F(a_1, a_2, \dots, a_{n-1})$.

Let $S = \{a_1, a_2, \dots, a_n\}$. Suppose that S is algebraically dependent. Then there exists a nonzero polynomial $p(x_1, x_2, \dots, x_n) \in F[x_1, x_2, \dots, x_n]$ such that

$$p(a_1, a_2, \dots, a_n) = 0. \quad (5.12)$$

Choose the smallest N such that $1 \leq N \leq n$ and

$p(a_1, a_2, \dots, a_N, a_{N+1}, \dots, a_n) = p(a_1, a_2, \dots, a_N, 0, \dots, 0)$. Consider p as a polynomial in $F(a_1, a_2, \dots, a_{N-1})[x]$ by letting a_N be the indeterminant x . Clearly, $p(a_N) = p(a_1, a_2, \dots, a_N) = 0$, and a_N is algebraic over $F(a_1, a_2, \dots, a_{N-1})$. This contradicts the assumption that a_N is transcendental over $F(a_1, a_2, \dots, a_N)$.

Therefore, S is algebraically independent. □

Theorem 5.2.4. *If D, D' are integral domains with field of fractions E, E' respectively, then an isomorphism $\sigma : D \rightarrow D'$ induces an isomorphism*

$$\hat{\sigma} : E \rightarrow E' \quad \text{by} \quad (5.13)$$

$$\hat{\sigma} \left(\frac{a}{b} \right) = \frac{\sigma(a)}{\sigma(b)}. \quad (5.14)$$

Proof. Suppose $\frac{a}{b}, \frac{c}{d} \in E$. Note that $b, d \neq 0$, so $\sigma(b), \sigma(d) \neq 0$. Consider

$$\begin{aligned}
 \hat{\sigma} \left(\frac{a}{b} + \frac{c}{d} \right) &= \hat{\sigma} \left(\frac{ad + bc}{bd} \right) \\
 &= \frac{\sigma(ad + bc)}{\sigma(bd)} \\
 &= \frac{\sigma(ad) + \sigma(bc)}{\sigma(bd)} \\
 &= \frac{\sigma(ad)}{\sigma(bd)} + \frac{\sigma(bc)}{\sigma(bd)} \\
 &= \frac{\sigma(a)\sigma(d)}{\sigma(b)\sigma(d)} + \frac{\sigma(b)\sigma(c)}{\sigma(b)\sigma(d)} \\
 &= \frac{\sigma(a)}{\sigma(d)} + \frac{\sigma(c)}{\sigma(d)} \\
 &= \hat{\sigma} \left(\frac{a}{b} \right) + \hat{\sigma} \left(\frac{c}{d} \right).
 \end{aligned} \tag{5.15}$$

Moreover,

$$\begin{aligned}
 \hat{\sigma} \left(\frac{a}{b} \cdot \frac{c}{d} \right) &= \hat{\sigma} \left(\frac{ac}{bd} \right) \\
 &= \frac{\sigma(ac)}{\sigma(bd)} \\
 &= \frac{\sigma(a)\sigma(c)}{\sigma(b)\sigma(d)} \\
 &= \frac{\sigma(a)}{\sigma(b)} \cdot \frac{\sigma(c)}{\sigma(d)} \\
 &= \hat{\sigma} \left(\frac{a}{b} \right) \cdot \hat{\sigma} \left(\frac{c}{d} \right).
 \end{aligned} \tag{5.16}$$

Hence, $\hat{\sigma}$ is a homomorphism of rings.

Suppose $\frac{a}{b} = \frac{c}{d}$. Then $ad = bc$. Apply σ to get $\sigma(ad) = \sigma(bc)$. Because σ is a homomorphism,

$$\begin{aligned}
 \sigma(a)\sigma(d) &= \sigma(ad) = \sigma(bc) = \sigma(b)\sigma(c) \\
 \frac{\sigma(a)}{\sigma(b)} &= \frac{\sigma(c)}{\sigma(d)} \\
 \hat{\sigma} \left(\frac{a}{b} \right) &= \hat{\sigma} \left(\frac{c}{d} \right).
 \end{aligned} \tag{5.17}$$

Hence, $\hat{\sigma}$ is well-defined.

Suppose $\frac{a}{b} \in \text{Ker } \hat{\sigma}$. Then

$$\begin{aligned} 0 &= \hat{\sigma} \left(\frac{a}{b} \right) \\ &= \frac{\sigma(a)}{\sigma(b)}. \end{aligned} \tag{5.18}$$

It follows that $0 = \sigma(a)$, and thus $a = 0$ and $\frac{a}{b} = 0$. Hence, the kernel is trivial, and $\hat{\sigma}$ is a monomorphism.

Suppose $\hat{\sigma} \left(\frac{c}{d} \right) \in E'$. Then $\frac{\sigma(c)}{\sigma(d)} \in E'$. Because σ is an isomorphism, there exist $c, d \in D$ such that $c \mapsto \sigma(c)$ and $d \mapsto \sigma(d)$, with $d \neq 0$, else $\sigma(d) = 0$. Hence we have that

$$\frac{c}{d} \xrightarrow{\sigma} \frac{\sigma(c)}{\sigma(d)} = \hat{\sigma} \left(\frac{c}{d} \right), \tag{5.19}$$

and $\hat{\sigma}$ is an epimorphism. Therefore, $\hat{\sigma} : E \rightarrow E'$ is an isomorphism. $\square$

Corollary 5.2.5. *If $a_1, a_2, \dots, a_n \in \mathbb{C}$ are algebraically independent over a field F , then the map*

$$\varphi : F(x_1, x_2, \dots, x_n) \rightarrow F(a_1, a_2, \dots, a_n), \tag{5.20}$$

$$\varphi \left(\frac{f(x_1, x_2, \dots, x_n)}{g(x_1, x_2, \dots, x_n)} \right) = \frac{f(a_1, a_2, \dots, a_n)}{g(a_1, a_2, \dots, a_n)} \tag{5.21}$$

is an isomorphism of fields.

Proof. Because $a_1, a_2, \dots, a_n$ are algebraically independent,

$F[x_1, x_2, \dots, x_n] \cong F[a_1, a_2, \dots, a_n]$ as integral domains. By Theorem 5.2.4, φ is an isomorphism. $\square$

Corollary 5.2.6. *Let F be a field, $a_1, a_2, \dots, a_n \in \mathbb{C}$ be algebraically independent over F , $E = F(a_1, a_2, \dots, a_n)$. Then each permutation of $a_1, a_2, \dots, a_n$ induces an automorphism $\tau \in \text{Aut}(E/F)$.*

Proof. Consider the polynomial ring $F[x_1, x_2, \dots, x_n]$. Let $\sigma \in S_n$. Then let

$$\hat{\sigma} : F[x_1, x_2, \dots, x_n] \rightarrow F[x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}] \quad (5.22)$$

$$\hat{\sigma}(f(x_1, x_2, \dots, x_n)) = f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) \quad (5.23)$$

is an automorphism of $F[x_1, x_2, \dots, x_n]$ fixing F . By Theorem 5.2.4, $\hat{\sigma}$ is an automorphism of $F(x_1, x_2, \dots, x_n)$. By Corollary 5.2.5, $\hat{\sigma}$ is an automorphism of $F(a_1, a_2, \dots, a_n)$. Therefore, each permutation of $a_1, a_2, \dots, a_n$ induces an automorphism of $\text{Aut}(E/F)$. $\square$

For the remainder of this section, for $1 \leq i \leq n$, we let $a_1, a_2, \dots, a_n$ be algebraically independent over $\mathbb{Q}$, $E = \mathbb{Q}(a_1, a_2, \dots, a_n)$, $b_i = (-1)^i \sigma_i(a_1, a_2, \dots, a_n)$, where σ_i the i th symmetric function on n variables, and $K = \mathbb{Q}(b_1, b_2, \dots, b_n)$.

Theorem 5.2.7. *Let $f(x) = x^n + b_1x^{n-1} + \dots + b_n$. The splitting field for $f(x)$ over $\mathbb{Q}$ is E , $f(x)$ is irreducible over K , and $\text{Gal}(E/K) = S_n$.*

Proof. By Lemma 5.1.4, $f(x) = x^n + b_1x^{n-1} + \dots + b_n = \prod_{i=1}^n (x - a_i)$, and since $\mathbb{Q} \subseteq K \subseteq E$, E is the splitting field for $f(x)$ over $\mathbb{Q}$. Because the elements of $\text{Gal}(E/\mathbb{Q})$ permute the n zeros of $f(x)$ by Theorem 3.1.5, $\text{Gal}(E/\mathbb{Q})$ is a subgroup of S_n . By Corollary 5.2.6, each permutation of the a_i 's induces an automorphism of $\text{Gal}(E/\mathbb{Q})$. That is, $\text{Gal}(E/\mathbb{Q}) \cong S_n$. Let $\sigma \in \text{Gal}(E/\mathbb{Q})$. Because permutations fix symmetric functions, $\sigma(b_i) = b_i$ for all $1 \leq i \leq n$. Hence, σ fixes every element of K . That is, $\text{Gal}(E/\mathbb{Q}) = \text{Gal}(E/K) \cong S_n$. Since E is the splitting field for $f(x)$ over K and $\text{Gal}(E/K) \cong S_n$, $f(x)$ is irreducible over K by Theorem 3.2.5. $\square$

Lemma 5.2.8. *There exist $m_1, m_2, \dots, m_n \in \mathbb{Z}$ such that if*

$$\alpha = m_1a_1 + m_2a_2 + \dots + m_na_n, \quad (5.24)$$

then $\tau(\alpha) \in S_n$ is distinct for all $\tau \in S_n$, and $E = K(\alpha)$. In particular, E is a simple extension of K . Moreover, there exist $p_i \in \mathbb{Q}[x]$ such that $a_i = p_i(\alpha)$.

Proof. By Theorem 5.2.7, E is a splitting field for $f(x)$ over $\mathbb{Q}$, and $\text{Gal}(E/K) = S_n$. By the Fundamental Theorem of Galois Theory,

$$[E : K] = |\text{Gal}(E/K)| = |S_n| = n!. \quad (5.25)$$

Since $[E : K]$ is finite, $E = K(m_1a_1 + m_2a_2 + \cdots + m_na_n)$ for some $m_1, m_2, \dots, m_n \in \mathbb{Z}$ by Corollary 2.7.5. In particular, since $E = K(\alpha)$, $\deg_K \alpha = [E : K] = n!$. It follows that α has $n!$ conjugates in E , since $f(x)$ is irreducible over K and $\tau \in \text{Gal}(E/K)$ permutes the zeros of $f(x)$. Hence, τ permutes $a_1, a_2, \dots, a_n$, and $\tau(a_i) = a_{\tau(i)}$ for each $1 \leq i \leq n$. Thus, $\tau(\alpha)$ is a conjugate of α in E . By Lemma 3.1.2, the conjugates are distinct.

Furthermore, since $E = K(\alpha)$ is algebraic over K , a_i is the root of some polynomial $p_i \in K[x]$ for $1 \leq i \leq n$. Because $a_1, a_2, \dots, a_n \mid b_i$, b_i does not divide any coefficient of p_i . Therefore, $p_i \in \mathbb{Q}[x]$. $\square$

Theorem 5.2.9. *Let $\alpha = m_1a_1 + m_2a_2 + \cdots + m_na_n$ for some $m_1, m_2, \dots, m_n \in \mathbb{Z}$ such that $E = K(\alpha)$. If*

$$g(x) = \prod_{\tau \in S_n} (x - \tau(\alpha)), \quad (5.26)$$

then $g(x)$ is irreducible over K , and E is a splitting field for $g(x)$ over K .

Proof. Let $\psi \in S_n$. Then

$$\begin{aligned} \psi(g(x)) &= \psi\left(\prod_{\tau \in S_n} (x - \tau(\alpha))\right) \\ &= \prod_{\tau \in S_n} \psi(x - \tau(\alpha)) \\ &= \prod_{\tau \in S_n} (x - \psi(\tau(\alpha))) \\ &= \prod_{\tau' \in S_n} (x - \tau'(\alpha)) \\ &= g(x). \end{aligned} \quad (5.27)$$

Hence, for all $\psi \in S_n$, $\psi(g(x)) = g(x)$. It follows that the coefficients of $g(x)$ are symmetric polynomials on $a_1, a_2, \dots, a_n$. Since $a_1, a_2, \dots, a_n$ are the zeros of $f(x) \in K[x]$ and $g(x)$ is fixed by τ , $g(x) \in K[x]$. Because the coefficients of $g(x)$ are symmetric polynomials on $a_1, a_2, \dots, a_n$, the splitting field for $g(x)$ over K is E . By Theorem 5.2.7, $\text{Gal}(E/K) = S_n$, so $g(x)$ is irreducible over K by Theorem 3.2.5. $\square$

CHAPTER 6

COMPLEX ANALYSIS

We omit many proofs in this chapter, for the details are not related to our goal.

6.1 Domains

We follow Munkres [Mun75] and Saff and Snider [Sni03] in this section.

Definition 6.1.1. An open subset $S \subseteq \mathbb{C}$ is called **path-connected** if every pair of points $z_1, z_2 \in S$ can be joined by a polygonal path that lies entirely in S .

Definition 6.1.2. An open subset $S \subseteq \mathbb{C}$ is called **connected** if it is not the disjoint union of two nonempty open sets.

Theorem 6.1.3. *Every path-connected set is a connected set.*

Proof. See Munkres [Mun75, p. 155]. □

Definition 6.1.4. A path-connected open subset of $\mathbb{C}$ is called a **domain**.

6.2 Complex Integration & Uniform Convergence

We follow Rudinf [Rud76] and Saff and Snider [Sni03] in this section.

Definition 6.2.1. Let $z_0 \in \mathbb{C}$, $U \subseteq \mathbb{C}$ be a neighborhood of z_0 , $f : U \rightarrow \mathbb{C}$. We say that $L \in \mathbb{C}$ is the **limit of $f(z)$ as z approaches z_0** if for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f(z) - L| < \epsilon \quad \text{whenever} \quad 0 < |z - z_0| < \delta, \quad (6.1)$$

and we write

$$\lim_{z \rightarrow z_0} f(z) = L. \quad (6.2)$$

Definition 6.2.2. If $U \subseteq \mathbb{C}$ is a neighborhood of $z_0 \in \mathbb{C}$, $f : U \rightarrow \mathbb{C}$ is called **continuous at z_0** if

$$\lim_{z \rightarrow z_0} f(z) = f(z_0). \quad (6.3)$$

Definition 6.2.3. A **contour** is a finite sequence of directed smooth curves $(\gamma_1, \gamma_2, \dots, \gamma_n)$, where $\gamma_i : [0, 1] \rightarrow \mathbb{C}$ for $i = 1, 2, \dots, n$, and $\gamma_i(1) = \gamma_{i+1}(0)$ for $i = 1, 2, \dots, n-1$.

Definition 6.2.4. A contour $\Gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$ is called a **loop** if $\gamma_1(0) = \gamma_n(1)$. If the image of Γ is a single point, we call Γ a **constant loop**. A loop Γ is a **simple loop** if for $1 \leq i \leq j \leq n$, $\gamma_i(z_1) \neq \gamma_j(z_2)$ unless $i = j$ and $z_1 = z_2$, or $i = 1, j = n, z_1 = 0$, and $z_2 = 1$.

Definition 6.2.5. If every loop in an open set $\Omega \subseteq \mathbb{C}$ can be continuously deformed to a constant loop, then Ω is called a **simply connected domain**.

Definition 6.2.6. Let $\Omega \subseteq \mathbb{C}$, $\gamma : [0, 1] \rightarrow \Omega$ be a smooth curve, $f : \Omega \rightarrow \mathbb{C}$ be a function continuous on γ . We define the **complex integral of $f(z)$ along γ** to be

$$\int_{\gamma} f(z) dz = \int_a^b f(\gamma(t)) \gamma'(t) dt. \quad (6.4)$$

Definition 6.2.7. Let $\Omega \subseteq \mathbb{C}$, $\gamma : [0, 1] \rightarrow \Omega$ be a smooth curve. A parametrization of γ given by $z(t)$ is an **admissible parametrization** if $z(t)$ is a continuous complex-valued function on the real interval $[a, b]$ such that

- (i) $z(t)$ has a continuous derivative on $[a, b]$, and
- (ii) $z'(t) \neq 0$ on $a \leq t \leq b$.

Theorem 6.2.8. Let $\Omega \subseteq \mathbb{C}$, $\gamma : [0, 1] \rightarrow \Omega$ be a smooth curve, $f : \Omega \rightarrow \mathbb{C}$ be a function continuous on γ . If $z_1(t)$ and $z_2(t)$ are any two admissible parametrizations of γ consistent with its direction, where $a, c \leq t \leq b, d$, then

$$\int_a^b f(z_1(t))z_1'(t)dt = \int_c^d f(z_2(t))z_2'(t)dt. \quad (6.5)$$

Proof. See Saff and Snider [Sni03, Theorem 4, p.165]. $\square$

Definition 6.2.9. Let $\Gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$ be a contour in $\mathbb{C}$. If f is a complex function continuous on Γ , then we define $\int_{\Gamma} f(z)dz$, the **complex integral of f along Γ** , to be

$$\int_{\Gamma} f(z)dz = \int_{\gamma_1} f(z)dz + \int_{\gamma_2} f(z)dz + \dots + \int_{\gamma_n} f(z)dz. \quad (6.6)$$

Definition 6.2.10. Let Ω be a simply connected domain, $(f_n : \Omega \rightarrow \mathbb{C})$ be a sequence of functions, and $n \in \mathbb{Z}_{>0}$. The sequence (f_n) **converges uniformly** to $f : \Omega \rightarrow \mathbb{C}$ in Ω if for every $\epsilon > 0$ there exists $N \in \mathbb{Z}_{\geq 0}$ such that $n \geq N$ implies

$$|f_n(z) - f(z)| < \epsilon \quad (6.7)$$

for all $z \in \Omega$.

Theorem 6.2.11. Suppose Ω is a domain, Γ is a contour in Ω , $f : \Omega \rightarrow \mathbb{C}$ is continuous, and $\{f_n\}$ converges uniformly to f in Ω . Then $\lim_{n \rightarrow \infty} \int_{\Gamma} f_n dz$ exists and

$$\lim_{n \rightarrow \infty} \int_{\Gamma} f_n dz = \int_{\Gamma} \left(\lim_{n \rightarrow \infty} f_n \right) dz = \int_{\Gamma} f dz. \quad (6.8)$$

Proof. For the real integral case, see Rudin [Rud76, Theorem 7.16]. The complex case follows from the real case and Definition 6.2.6. $\square$

Corollary 6.2.12. Let Ω be a domain and Γ be a contour in Ω . If $f_n : \Omega \rightarrow \mathbb{C}$ is continuous for $n \in \mathbb{Z}_{>0}$ and $\sum_{n=1}^{\infty} f_n(z)$ converges uniformly to $f(z)$ in Ω , then

$$\int_{\Gamma} f(z)dz = \sum_{n=1}^{\infty} \int_{\Gamma} f_n(z)dz. \quad (6.9)$$

$\square$

6.3 Holomorphic Functions and the Theorems of Cauchy

We follow Saff and Snider [Sni03] in this section.

Definition 6.3.1. Let $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, and call $\hat{\mathbb{C}}$ the **extended complex plane**.

Definition 6.3.2. In the extended complex plane, a **neighborhood of infinity** is $\{z \in \mathbb{C} \mid |z| > R\}$ for some $R \in \mathbb{R}_{\geq 0}$.

Definition 6.3.3. Let $\Omega \subseteq \mathbb{C}$ be a domain. A function $f(z) : \Omega \rightarrow \mathbb{C}$ is called **holomorphic at a** if

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = f'(a) \quad (6.10)$$

exists.

Definition 6.3.4. Let $\Omega \subseteq \mathbb{C}$ be a domain. A function $f(z) : \Omega \rightarrow \mathbb{C}$ is called **holomorphic on Ω** if it is holomorphic at all $a \in \Omega$; that is, $f(z)$ is called holomorphic if it is complex differentiable.

Definition 6.3.5. Let $\Omega \subseteq \hat{\mathbb{C}}$ be a domain, $f : \Omega \rightarrow \mathbb{C}$, and $\{t \in \mathbb{C} \mid |t| > R\} \subseteq \Omega$ for some $R \in \mathbb{R}_{\geq 0}$. The function $f(t)$ is **holomorphic at infinity** if

$$\lim_{z \rightarrow 0} f\left(\frac{1}{z}\right) = L, \text{ and}$$

$$g(z) = \begin{cases} f\left(\frac{1}{z}\right) & \text{if } z \neq 0 \\ L & \text{if } z = 0 \end{cases} \quad (6.11)$$

is holomorphic at 0.

Theorem 6.3.6. Let $\Omega \subseteq \hat{\mathbb{C}}$ be a domain, $a \in \Omega$. The set $\mathcal{H}(a)$ of functions holomorphic at a is a subring of $\mathcal{F}(\Omega)$, the ring of functions from Ω to $\mathbb{C}$.

Proof. Let $f(z), g(z) \in \mathcal{H}(a)$. Recall that because f, g are holomorphic at a , f and g are differentiable at a , and by the rules of differentiation, extending the usual proofs

over $\mathbb{R}$, αf , $f + g$, fg are all differentiable at a , where $\alpha \in \Omega$. Therefore, $\alpha f, f + g, fg \in \mathcal{H}(a)$, and $\mathcal{H}(a)$ is a subring of $\mathcal{F}(\Omega)$. $\square$

Theorem 6.3.7. *Let $\Omega \subseteq \hat{\mathbb{C}}$ be a domain, $a \in \Omega$. If $f \in \mathcal{H}(a)$ with $f(a) \neq 0$, then $\frac{1}{f} \in \mathcal{H}(a)$.*

Proof. Let $f \in \mathcal{H}(a)$ such that $f(a) \neq 0$. Then $\frac{1}{f}$ is differentiable at a by extending the usual proof over $\mathbb{R}$ of the quotient rule. Therefore, $\frac{1}{f} \in \mathcal{H}(a)$. $\square$

Theorem 6.3.8. *Let $\Omega \subseteq \hat{\mathbb{C}}$ be a domain. Then $\mathcal{H}(\Omega)$ is a subring of $\mathcal{F}(\Omega)$.*

Proof. By Theorem 6.3.6, $\mathcal{H}(a)$ is a subring of $\mathcal{F}(\Omega)$ for all $a \in \Omega$. By Theorem 2.4.1, $\bigcap_{a \in \Omega} \mathcal{H}(a)$ is a subring of $\mathcal{F}(\Omega)$, and by definition, $\bigcap_{a \in \Omega} \mathcal{H}(a) = \mathcal{H}(\Omega)$. $\square$

Theorem 6.3.9. *Let $\Omega \subseteq \mathbb{C}$ be an open set, $f : \Omega \rightarrow \mathbb{C}$ be holomorphic in Ω . If Γ_0 and Γ_1 are two loops that can be continuously deformed into one another in Ω , then*

$$\int_{\Gamma_0} f(z)dz = \int_{\Gamma_1} f(z)dz. \quad (6.12)$$

Proof. See Saff and Snider [Sni03, Theorem 8, p.186]. $\square$

Theorem 6.3.10 (Cauchy's Theorem). *Let Ω be a simply connected domain, $f(z) : \Omega \rightarrow \mathbb{C}$ be holomorphic. If Γ is any loop in Ω , then*

$$\int_{\Gamma} f(z)dz = 0. \quad (6.13)$$

Proof. Since Ω is a simply connected domain, we can continuously deform Γ to a constant loop. By Theorem 6.3.9,

$$\int_{\Gamma} f(z)dz = \int_{\Gamma_0} f(z)dz \quad (6.14)$$

for any loop Γ_0 that Γ deforms to. We observe, then, that the integral of $f(z)$ over a constant loop is 0. $\square$

Theorem 6.3.11 (Morera's Theorem). *If Ω is a simply connected domain, $f : \Omega \rightarrow \mathbb{C}$ is continuous, and $\int_{\Gamma} f(z)dz = 0$ for all loops Γ in Ω , then f is holomorphic in Ω .*

Proof. See Saff and Snider [Sni03, Theorem 18, p. 210]. □

Theorem 6.3.12 (Cauchy's Formula). *Let Ω be a simply connected domain, $f(z) : \Omega \rightarrow \mathbb{C}$ be holomorphic, Γ be a positively oriented simple loop in Ω , and z_0 be a point in Ω inside Γ . Then all derivatives of f exist on Ω ,*

$$f(z_0) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{z - z_0} dz, \quad (6.15)$$

and for $k \geq 1$,

$$f^{(k)}(z_0) = \frac{k!}{2\pi i} \int_{\Gamma} \frac{f(z)}{(z - z_0)^{k+1}} dz. \quad (6.16)$$

Proof. See Saff and Snider [Sni03, Theorem 19, p. 211]. □

6.4 Power Series

We follow Rudin [Rud76] and Saff and Snider [Sni03] in this section.

Definition 6.4.1. Given a fixed point $z_0 \in \mathbb{C}$, a series of the form

$$\sum_{i=0}^{\infty} c_i (z - z_0)^i, \quad (6.17)$$

where $c_i \in \mathbb{C}$ for all $i \in \mathbb{Z}_{\geq 0}$ is called a **power series with center z_0** .

Theorem 6.4.2. *For any power series $p(z) = \sum_{i=0}^{\infty} c_i (z - z_0)^i$, there exists some $R \in \mathbb{R}_{>0}$, called the radius of convergence, such that $p(z)$ converges in $|z - z_0| < R$ and converges uniformly in any closed subdisk $|z - z_0| \leq R' < R$.*

Proof. See Saff and Snider [Sni03, Theorem 7, p. 264]. □

Lemma 6.4.3. *Let Ω be a simply connected domain. If $(f_n : \Omega \rightarrow \mathbb{C})$ is a sequence of continuous functions converging uniformly to $f : \Omega \rightarrow \mathbb{C}$, then f is continuous on Ω .*

Proof. Let $z_0 \in \Omega$, $\epsilon > 0$. Because (f_n) is uniformly convergent, we can choose $N \in \mathbb{Z}_{>0}$ such that

$$|f(z) - f_N(z)| < \frac{\epsilon}{3} \quad (6.18)$$

for any $z \in \Omega$. Because f_N is continuous, there exists $\delta_N > 0$ such that

$$|f_N(z_0) - f_N(z)| < \frac{\epsilon}{3} \quad (6.19)$$

for any $z \in \Omega$ such that $|z_0 - z| < \delta_N$. It follows that for any $z \in \Omega$ such that $|z_0 - z| < \delta_N$, $|f(z_0) - f(z)| < \epsilon$ because, by the triangle inequality,

$$\begin{aligned} |f(z_0) - f(z)| &\leq |f(z_0) - f_N(z_0)| + |f_N(z_0) - f_N(z)| + |f_N(z) - f(z)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon. \end{aligned} \quad (6.20)$$

Hence, f is continuous at z_0 . Because z_0 was arbitrary in Ω , f is continuous in Ω . □

Theorem 6.4.4. *If Ω is a simply connected domain and $(f_n : \Omega \rightarrow \mathbb{C})$ is a sequence of holomorphic functions converging uniformly to a function $f : \Omega \rightarrow \mathbb{C}$, then f is holomorphic in Ω .*

Proof. By Lemma 6.4.3, f is continuous in Ω . Let Γ be a loop in Ω . By Theorem 6.2.11 and Cauchy's Theorem (Theorem 6.3.10),

$$\int_{\Gamma} f(z) dz = \int_{\Gamma} \left(\lim_{n \rightarrow \infty} f_n(z) \right) dz = \lim_{n \rightarrow \infty} \int_{\Gamma} f_n(z) dz = \lim_{n \rightarrow \infty} 0 = 0. \quad (6.21)$$

By Morera's Theorem (Theorem 6.3.11), f is holomorphic in Ω . □

Theorem 6.4.5. *If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ has radius of convergence $R \in \mathbb{R}_{>0}$, then $f(z)$ is holomorphic for $|z| < R$ and*

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1} \quad (6.22)$$

for $|z| < R$.

Proof. Let Ω be the open disk $|z| < R$, z_0 be a point in Ω , Γ be the circle of radius $R_1 = \frac{|z_0|+R}{2}$. By Theorem 6.4.4, $f(z)$ is holomorphic for $|z| < R$ and converges uniformly on $|z| \leq R_2 = \frac{R_1+R}{2}$. By Cauchy's Formula (Theorem 6.3.12),

$$\begin{aligned} f'(z_0) &= \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{(z - z_0)^2} dz \\ &= \frac{1}{2\pi i} \int_{\Gamma} \frac{1}{(z - z_0)^2} \sum_{n=0}^{\infty} a_n z^n dz \\ &= \sum_{n=0}^{\infty} (a_n) \left(\frac{1}{2\pi i} \int_{\Gamma} \frac{z^n}{(z - z_0)^2} dz \right) && \text{by Theorem 6.4.2 and Corollary 6.2.12} \\ &= \sum_{n=0}^{\infty} (a_n) \frac{d}{dz} (z^n) \Big|_{z=z_0} && \text{by Cauchy's Formula} \\ &= \sum_{n=0}^{\infty} a_n (n z_0^{n-1}) \\ &= \sum_{n=0}^{\infty} n a_n z_0^{n-1}. \end{aligned} \quad (6.23)$$

Since z_0 is any point inside the radius of convergence of $f(z)$, we have that

$$f'(z) = \sum_{n=0}^{\infty} n a_n z^{n-1} \quad (6.24)$$

for $|z| < R$. □

Corollary 6.4.6. *Let Ω be a domain containing $\{z \in \mathbb{C} \mid |z| < R\}$ for some*

$R \in \mathbb{R}_{>0}$. *If $f : \Omega \rightarrow \mathbb{C}$ satisfies $0 = f(z) = \sum_{n=0}^{\infty} a_n z^n$ inside $|z| < R$, then $a_n = 0$ for all $n \in \mathbb{Z}_{>0}$.*

Proof. By the hypothesis and Theorem 6.4.5, $f^{(k)}(z)$ is holomorphic inside $|z| < R$ for all $k \in \mathbb{Z}_{>0}$, and

$$\begin{aligned}
 0 &= f(z) = \sum_{n=0}^{\infty} a_n z^n, \\
 0 &= f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}, \\
 &\vdots \\
 0 &= f^{(k)}(z) = \sum_{n=k}^{\infty} n(n-1)(n-2) \cdots (n-(k-1)) a_n z^{n-k}, \\
 &\vdots
 \end{aligned} \tag{6.25}$$

Evaluate each of these at $z = 0$ to obtain

$$\begin{aligned}
 0 &= a_0 \\
 0 &= a_1 \\
 &\vdots \\
 0 &= a_k \\
 &\vdots
 \end{aligned} \tag{6.26}$$

Therefore, $a_n = 0$ for all $n \in \mathbb{Z}_{>0}$. □

Theorem 6.4.7 (Identity Theorem). *Let $\Omega \subseteq \mathbb{C}$ be a simply connected domain containing $\{z \in \mathbb{C} \mid |z| < R\}$ for some $R \in \mathbb{R}_{>0}$. If $f(z), g(z) : \Omega \rightarrow \mathbb{C}$ such that*

$$f(z) = \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} b_n z^n = g(z) \tag{6.27}$$

for $|z| < R \in \mathbb{R}_{>0}$, then $a_n = b_n$ for $n \in \mathbb{Z}_{>0}$.

Proof. For $|z| < R$, where $R > 0$, we have

$$\begin{aligned}
 0 &= f(z) - g(z) \\
 &= \sum_{n=0}^{\infty} a_n z^n - \sum_{n=0}^{\infty} b_n z^n \\
 &= \sum_{n=0}^{\infty} (a_n - b_n) z^n.
 \end{aligned} \tag{6.28}$$

By Corollary 6.4.6, $a_n - b_n = 0$ for $n \in \mathbb{Z}_{>0}$. Therefore, $a_n = b_n$. $\square$

Theorem 6.4.8. *Suppose*

$$(i) \sum_{n=0}^{\infty} a_n \text{ converges absolutely,}$$

$$(ii) \sum_{n=0}^{\infty} a_n = A,$$

$$(iii) \sum_{n=0}^{\infty} b_n = B,$$

$$(iv) c_n = \sum_{k=0}^n a_k b_{n-k} \text{ where } n \in \mathbb{Z}_{\geq 0}.$$

Then $\sum_{n=0}^{\infty} c_n$ converges absolutely and is equal to AB .

Proof. See Rudin [Rud76, Theorem 3.50]. $\square$

6.5 Taylor Series

We follow Saff and Snider [Sni03] in this section.

Definition 6.5.1. Let $f(z) : \Omega \rightarrow \mathbb{C}$ be holomorphic at $z_0 \in \mathbb{C}$, where Ω is some neighborhood of z_0 . Then the series

$$f(z_0) + f'(z_0)(z - z_0) + \frac{f''(z_0)}{2!}(z - z_0)^2 + \cdots = \sum_{i=0}^{\infty} \frac{f^{(i)}(z_0)}{i!}(z - z_0)^i \tag{6.29}$$

is called the **Taylor series for f about z_0** .

Theorem 6.5.2. Let $R \in \mathbb{R}_{>0}$, $z_0 \in \mathbb{C}$, $f(z) : \Omega \rightarrow \mathbb{C}$ be holomorphic in the disk

$$\Omega = \{z \in \mathbb{C} \mid |z - z_0| < R\}. \quad (6.30)$$

Then the Taylor series for f about z_0 converges to $f(z)$ for all $z \in \Omega$. Furthermore, the convergence of the series is uniform in any closed subdisk

$$|z - z_0| \leq R' < R. \quad (6.31)$$

Proof. See Saff and Snider [Sni03, Theorem 3, p. 243]. $\square$

Theorem 6.5.3. If

$$\sum_{n=0}^{\infty} c_n (z - z_0)^n \quad (6.32)$$

converges to $f(z)$ in some neighborhood of z_0 , then for all $n \in \mathbb{Z}_{>0}$,

$$c_n = \frac{f^{(n)}(z_0)}{n!}. \quad (6.33)$$

That is, (6.32) is actually the Taylor expansion of $f(z)$ about z_0 .

Proof. By Theorem 6.4.7, c_n is the coefficient of the n th summand of the Taylor expansion of $f(z)$. Therefore, (6.32) is the Taylor expansion of $f(z)$ about z_0 . $\square$

6.6 Laurent Series and Singularities

We follow Saff and Snider [Sni03] in this section.

Definition 6.6.1. Let $\Omega \subseteq \mathbb{C}$ be a domain, $f : \Omega \rightarrow \mathbb{C}$, $z_0 \in \Omega$, and $n \in \mathbb{Z}_{\leq 0}$. A series $f(z) = \sum_{k=n}^{\infty} a_k (z - z_0)^k$ that converges in $\{z \in \mathbb{C} \mid 0 < |z - z_0| < R\} \subseteq \Omega$ is called a **Laurent series** for f in a punctured neighborhood of z_0 .

Definition 6.6.2. Let $\Omega \subseteq \hat{\mathbb{C}}$ be an open set and $f : \Omega \rightarrow \mathbb{C}$. An **isolated singularity** of f is either a point $z_0 \in \Omega$ such that f is holomorphic in $0 < |z - z_0| < R$ for some $R \in \mathbb{R}_{>0}$ but not holomorphic at z_0 for $z \in \mathbb{C}$, or such that f is holomorphic in $|z - z_0| > R$ for some $R \in \mathbb{R}_{>0}$ if $z_0 = \infty$.

Definition 6.6.3. Let $\Omega \subseteq \mathbb{C}$ be an open set, $f : \Omega \rightarrow \mathbb{C}$. If $\lim_{z \rightarrow z_0} f(z) = L$ and

$$f_1(z) = \begin{cases} f(z) & \text{if } z \neq z_0 \\ L & \text{if } z = z_0 \end{cases} \quad (6.34)$$

is holomorphic in a neighborhood of z_0 , then z_0 is called a **removable singularity** of f .

Remark 6.6.4. Henceforth, we will remove any singularity if we can without comment.

Definition 6.6.5. Let $\Omega \subseteq \mathbb{C}$ be an open set. A function $f(z) : \Omega \rightarrow \mathbb{C}$ has a **zero of order m** at z_0 if and only if $f(z) = (z - z_0)^m g(z)$ in some neighborhood of z_0 , where $g(z)$ is holomorphic at z_0 and $g(z_0) \neq 0$.

Definition 6.6.6. Let $\Omega \subseteq \mathbb{C}$ be an open set. A function $f(z) : \Omega \rightarrow \mathbb{C}$ has a **pole of order m** at z_0 if and only if $f(z) = \frac{g(z)}{(z - z_0)^m}$ in some punctured neighborhood of z_0 , where $g(z)$ is holomorphic at z_0 and $g(z_0) \neq 0$.

Definition 6.6.7. Let $\Omega \subseteq \hat{\mathbb{C}}$ be a domain, $f : \Omega \rightarrow \mathbb{C}$. If $f(\frac{1}{z})$ has a pole at $z_0 = 0$, then we say that $f(z)$ has a **pole at ∞** . If in a neighborhood of z_0 $f(\frac{1}{z})$ has a zero at $z_0 = 0$, then we say that $f(z)$ has a **zero at ∞** .

Theorem 6.6.8. Let $\Omega \subseteq \mathbb{C}$ be a domain, $f : \Omega \rightarrow \mathbb{C}$, $z_0 \in \Omega$. Then f is either holomorphic in a neighborhood of z_0 or has a pole of order less than or equal to m if and only if f has a convergent Laurent series $\sum_{k=-m}^{\infty} a_k(z - z_0)^k$ in some punctured neighborhood of z_0 .

Proof. Assume f is either holomorphic or has a pole of order less than or equal to m . If f is holomorphic, then it has a convergent Laurent series in some punctured neighborhood of z_0 . If f has a pole of order $n \leq m$, then $f(z) = \frac{g(z)}{(z - z_0)^n}$ for some

$g(z)$ holomorphic at z_0 and $g(z_0) \neq 0$. Then $f(z)(z - z_0)^n = g(z)$. By Theorem 6.5.2, $g(z) = \sum_{k=0}^{\infty} a_k(z - z_0)^k$ in some neighborhood of z_0 . It follows that $f(z) = \sum_{k=-n}^{\infty} a_{k+n}(z - z_0)^k$ in some punctured neighborhood of z_0 .

Conversely, assume that f has a convergent Laurent series $\sum_{k=-m}^{\infty} a_k(z - z_0)^k$ in some punctured neighborhood of z_0 . Then

$$f(z) = \sum_{k=-m}^{\infty} a_k(z - z_0)^k = \frac{\sum_{k=0}^{\infty} a_{k-m}(z - z_0)^k}{(z - z_0)^m}. \quad (6.35)$$

Since $\sum_{k=0}^{\infty} a_{k-m}(z - z_0)^k$ is holomorphic in a neighborhood of z_0 , $f(z)$ has a pole of order less than or equal to m at z_0 or $f(z)$ is holomorphic if

$$a_{-m} = a_{1-m} = \cdots = a_{-1} = 0. \quad \square$$

Lemma 6.6.9. *Let $\Omega \subseteq \mathbb{C}$ be a domain, $f : \Omega \rightarrow \mathbb{C}$ be a function, $z_0 \in \Omega$. Then*

(i) *If f has a zero of order m at z_0 , then $\frac{1}{f}$ has a pole of order m at z_0 , and*

(ii) *If f has a pole of order m at z_0 , then $\frac{1}{f}$ has a removable singularity at z_0 .*

Proof. Suppose f has a zero of order m at z_0 . Then $f(z) = (z - z_0)^m g(z)$ for some $g(z) : \Omega \rightarrow \mathbb{C}$ holomorphic at z_0 such that $g(z_0) \neq 0$. Then $\frac{1}{f}(z) = \frac{h(z)}{(z - z_0)^m}$, where $h(z) = \frac{1}{g(z)}$ is holomorphic at z_0 and $h(z_0) \neq 0$. Therefore, $\frac{1}{f}$ has a pole of order m at z_0 .

Suppose f has a pole of order m at z_0 . Then $f(z) = \frac{g(z)}{(z - z_0)^m}$ in some punctured neighborhood of z_0 , where $g(z)$ is holomorphic at z_0 and $g(z_0) \neq 0$. Let $H(z) = (z - z_0)^m h(z)$, where $h(z) = \frac{1}{g(z)}$ is holomorphic in a neighborhood of z_0 .

Then

$$\lim_{z \rightarrow z_0} (z - z_0)^m h(z) = \left(\lim_{z \rightarrow z_0} (z - z_0)^m \right) \left(\lim_{z \rightarrow z_0} h(z) \right) = H(z_0). \quad (6.36)$$

Since $H(z) = \frac{1}{f}(z)$, $\frac{1}{f}(z)$ has a removable singularity at z_0 . $\square$

Lemma 6.6.10. *Let $\Omega \subseteq \mathbb{C}$ be an open set, $z_0 \in \Omega$, $f(z), g(z), h(z), \tilde{h}(z) : \Omega \rightarrow \mathbb{C}$ such that f has a pole of degree m at z_0 , g has a pole of degree n at z_0 , h has a zero of degree n at z_0 , and $\tilde{h}$ is holomorphic in a neighborhood of z_0 . Then*

(i) fg has a pole of order $m + n$ at z_0 ,

(ii) If $n < m$, then fh has a pole of order $m - n$, or if $n \geq m$, then a zero of degree $n - m$ at z_0 ,

(iii) $f + \tilde{h}$ has a pole of order m at z_0 , and

(iv) $f + g$ has a pole of order less than or equal to $M = \max(m, n)$.

Proof. By definition,

$$\begin{aligned} f(z) &= \frac{f_1(z)}{(z - z_0)^m}, \\ g(z) &= \frac{g_1(z)}{(z - z_0)^n}, \\ h(z) &= h_1(z)(z - z_0)^n, \end{aligned} \tag{6.37}$$

in a punctured neighborhood Ω of z_0 , where $f_1(z_0), g_1(z_0), h_1(z_0) \neq 0$ and f_1, g_1, h_1 are holomorphic in some neighborhood of z_0 . Then, in Ω ,

$$\begin{aligned} (fg)(z) &= \frac{f_1(z)g_1(z)}{(z - z_0)^{m+n}}, \\ (fh)(z) &= \frac{f_1(z)h_1(z)(z - z_0)^n}{(z - z_0)^m}, \\ (f + \tilde{h})(z) &= \frac{f_1(z) + \tilde{h}(z)(z - z_0)^m}{(z - z_0)^m}, \\ (f + g)(z) &= \frac{f_1(z)(z - z_0)^{M-n} + g_1(z)(z - z_0)^{M-m}}{(z - z_0)^M}, \end{aligned} \tag{6.38}$$

where $(f_1g_1)(z_0), (f_1h_1)(z_0), (f_1 + \tilde{h})(z_0) \neq 0$. Then fg has a pole of order $m + n$, and $f + \tilde{h}$ has a pole of order m . If $n < m$, then fh has a pole of order $m - n$, and if $n > m$, then fh has a zero of order $n - m$. Lastly, if

$f_1(z_0)(0)^{M-n} + g_1(z_0)(0)^{M-m} = 0$, then $f + g$ is holomorphic, and if $f_1(z_0)(0)^{M-n} + g_1(z_0)(0)^{M-m} \neq 0$, then $f + g$ has a pole of order less than or equal to M . $\square$

Remark 6.6.11. By Lemma 6.6.9 and Theorem 6.3.7, given any domain $\Omega \subseteq \mathbb{C}$ and function $f : \Omega \rightarrow \mathbb{C}$, if f has only poles and zeros, then $\frac{1}{f}$ has only removable singularities and poles.

6.7 Meromorphic Functions

Definition 6.7.1. Let $\Omega \subseteq \hat{\mathbb{C}}$ be an open set. A function $f(z) : \Omega \rightarrow \mathbb{C}$ is called **meromorphic** on Ω if it is holomorphic on Ω except at a set of isolated poles.

Theorem 6.7.2. Let $\Omega \subseteq \mathbb{C}$ be a domain, $f : \Omega \rightarrow \mathbb{C}$ be meromorphic on Ω . Either $f(z) = 0$ for all $z \in \Omega$, or f has only isolated zeros and poles on Ω .

Proof. Let

$$\begin{aligned} A &= \{z_0 \in \Omega \mid \text{Either } f(z_0) \neq 0, z_0 \text{ is an isolated zero, or } z_0 \text{ is an isolated pole}\}, \\ B &= \{z_0 \in \Omega \mid \text{There exists } \epsilon > 0 \text{ such that } f(z) = 0 \text{ for all } |z - z_0| < \epsilon\}. \end{aligned} \quad (6.39)$$

Suppose $z_0 \in A \cap B$. Because $z_0 \in B$, $f(z_0) = 0$. Because $z_0 \in A$, z_0 is an isolated zero. However, in some neighborhood about z_0 , $f(z)$ is identically zero, contradicting the assumption that z_0 is an isolated zero. Hence, $A \cap B = \emptyset$.

It follows that Ω is the disjoint union of the open sets A and B . Because Ω is a connected set by Theorem 6.1.3, $\Omega = A$ or $\Omega = B$. Therefore, if $\Omega = B$, then $f(z) = 0$, and if $\Omega = A$, then $f(z)$ has only isolated zeros and poles on Ω . $\square$

Lemma 6.7.3. Let $\Omega \subseteq \hat{\mathbb{C}}$ be an open set, $f(z) : \Omega \rightarrow \mathbb{C}$ be a meromorphic function not identically zero. Then $\frac{1}{f(z)}$ is a meromorphic function.

Proof. Because $f(z)$ is meromorphic on Ω , $f(z)$ is holomorphic on Ω except at a set of isolated poles. By Theorem 6.3.7, $\frac{1}{f(z)}$ is a holomorphic function. If X is the set of isolated zeros of $f(z)$, then by Lemma 6.6.9, X is the set of isolated poles of $\frac{1}{f(z)}$. Therefore, $\frac{1}{f(z)}$ is meromorphic. $\square$

Theorem 6.7.4. *Let F be a field. If a function $f(t)$ is meromorphic in a neighborhood of infinity and $g \in F[t]$, then $f(t)g(t)$ is equal to $\sum_{k=-m}^{\infty} \frac{a_k}{t^k}$ for some $m \in \mathbb{Z}_{>0}$ in some neighborhood of infinity.*

Proof. For some $R \in \mathbb{R}_{\geq 0}$, for all $z > R$,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\frac{1}{f(z+h)} - \frac{1}{f(z)}}{h} &= \lim_{h \rightarrow 0} \frac{f(z) - f(z+h)}{f(z+h)f(z)h} \\ &= \left(\lim_{h \rightarrow 0} -\frac{f(z+h) - f(z)}{h} \right) \left(\lim_{h \rightarrow 0} \frac{1}{f(z+h)f(z)} \right) \\ &= \lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h} \cdot \frac{-1}{f(z)^2}. \end{aligned} \quad (6.40)$$

Because $f(z)$ is not identically zero, $\frac{-1}{f(z)^2}$ exists except at a set of isolated points.

Moreover, $\lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h}$ exists for all $z \in \Omega \setminus X$. Hence, $\lim_{h \rightarrow 0} \frac{\frac{1}{f(z+h)} - \frac{1}{f(z)}}{h}$ exists except at a set of isolated points. Therefore, $\frac{1}{f(z)}$ is meromorphic. $\square$

Definition 6.7.5. Let $\Omega \subseteq \hat{\mathbb{C}}$ be a domain. We define $\mathcal{H}_t(\Omega)$ to be the **set of all holomorphic functions on Ω in the variable t** .

Definition 6.7.6. Let $\Omega \subseteq \hat{\mathbb{C}}$ be a domain. We define $\mathcal{M}_t(\Omega)$ to be the **set of all meromorphic functions on Ω in the variable t** .

Remark 6.7.7. Theorem 6.6.8 implies that if $\Omega \subseteq \hat{\mathbb{C}}$, $f \in \mathcal{M}_t(\Omega)$, then for all $a \in \Omega$, f has a convergent Laurent series in some neighborhood of a .

Theorem 6.7.8. *If $\Omega \subseteq \hat{\mathbb{C}}$ is a domain, then $\mathcal{M}_t(\Omega)$ is a field.*

Proof. By Theorem 6.3.6 and Lemma 6.6.10, $\mathcal{M}_t(\Omega)$ is a ring. Because $\mathbb{C}$ is a field, $\mathcal{M}_t(\Omega)$ is a commutative ring with unity, where 1 is the constant function $f(t) = 1$. Let $g \in \mathcal{M}_t(\Omega)$, $g \neq 0$. By Theorem 6.7.2, g has isolated zeros and poles, so by Lemma 6.7.3, $\frac{1}{g}$ is meromorphic. Since g is arbitrary, $\mathcal{M}_t(\Omega)$ is a field. $\square$

6.8 Complex Exponential

Definition 6.8.1. Let $z \in \mathbb{C}$ have the form $z = re^{i\theta}$, where $r, \theta \in \mathbb{R}$. We call $-\pi < \theta' \leq \pi$ such that $z = re^{i\theta'}$ the **principal argument of z** and write $\text{Arg } z = \theta'$.

Definition 6.8.2. Let $z \in \mathbb{C}$. The **principal logarithm of z** is defined as

$$\text{Log } z = \log |z| + i \text{Arg } z. \quad (6.41)$$

Definition 6.8.3. Let $z, s \in \mathbb{C}$. The **complex exponential z^s** is defined by

$$z^s = (e^{\text{Log } z})^s = e^{s \text{Log } z}. \quad (6.42)$$

Lemma 6.8.4. Let $z, w \in \mathbb{C}$. Then either $-\pi < \text{Arg } z + \text{Arg } w \leq \pi$ if and only if $\text{Arg}(zw) = \text{Arg } z + \text{Arg } w$ or z and w differ by $\pm 2\pi i$.

Proof. Write $z = re^{i\theta}$, $w = se^{i\varphi}$, $\theta = \text{Arg } z$, $\varphi = \text{Arg } w$.

Assume $-\pi < \text{Arg } z + \text{Arg } w \leq \pi$. Then $-\pi < \theta + \varphi \leq \pi$. It follows that

$$\begin{aligned} \text{Arg}(zw) &= \text{Arg}(re^{i\theta}se^{i\varphi}) \\ &= \text{Arg}(rse^{i(\theta+\varphi)}) \\ &= \theta + \varphi \\ &= \text{Arg } z + \text{Arg } w, \end{aligned} \quad (6.43)$$

where the third equation follows because $-\pi < \theta + \varphi \leq \pi$.

Now assume $\text{Arg}(zw) = \text{Arg } z + \text{Arg } w$. Suppose $\text{Arg } z + \text{Arg } w > \pi$ or $\text{Arg } z + \text{Arg } w \leq -\pi$. By definition, $-\pi < \text{Arg}(zw) \leq \pi$, so $\text{Arg } z + \text{Arg } w \neq \text{Arg}(zw)$. Therefore, $-\pi < \text{Arg } z + \text{Arg } w \leq \pi$. $\square$

Theorem 6.8.5. *Let $z, w \in \mathbb{C}$. Then either $\text{Log}(zw) = \text{Log } z + \text{Log } w$ if and only if $-\pi < \text{Arg } z + \text{Arg } w \leq \pi$ or z and w differ by $\pm 2\pi i$.*

Proof. Assume $-\pi < \text{Arg } z + \text{Arg } w \leq \pi$. Then

$$\begin{aligned}
 \text{Log}(zw) &= \log |zw| + i \text{Arg}(zw) \\
 &= \log |z| + \log |w| + i(\text{Arg } z + \text{Arg } w) \\
 &= \log |z| + i \text{Arg } z + \log |w| + i \text{Arg } w \\
 &= \text{Log } z + \text{Log } w,
 \end{aligned} \tag{6.44}$$

where the second equation follows by Lemma 6.8.4.

Suppose $\text{Log}(zw) = \text{Log } z + \text{Log } w$. If $-\pi \geq \text{Arg } z + \text{Arg } w$ or $\text{Arg } z + \text{Arg } w > \pi$, then $\text{Arg}(zw) \neq \text{Arg } z + \text{Arg } w$. From the above calculation, $\text{Log}(zw) \neq \text{Log } z + \text{Log } w$. Therefore, $-\pi < \text{Arg } z + \text{Arg } w \leq \pi$. $\square$

Corollary 6.8.6. *Given $z, w \in \mathbb{C}$, $(zw)^{\frac{1}{2}} = z^{\frac{1}{2}}w^{\frac{1}{2}}$ if $-\pi < \text{Arg } z + \text{Arg } w \leq \pi$.*

Proof. Assume $-\pi < \theta + \varphi \leq \pi$. By Theorem 6.8.5, $\text{Log}(zw) = \text{Log } z + \text{Log } w$. In particular,

$$(zw)^{\frac{1}{2}} = e^{\frac{1}{2} \text{Log } zw} = e^{\frac{1}{2} \text{Log } z} e^{\frac{1}{2} \text{Log } w} = z^{\frac{1}{2}} w^{\frac{1}{2}}. \tag{6.45}$$

$\square$

CHAPTER 7

LEMMAS FOR THE HILBERT IRREDUCIBILITY THEOREM

7.1 Affine Plane Curves and Regular Values

We follow Gallian [Gal10], Hadlock [Had78], and Wilf [Wil94] in this section.

Recall our assumption that every field F is a subfield of $\mathbb{C}$.

Definition 7.1.1. Let F be a field, $f(t, x) \in F[t, x]$. The set

$$C = \{(t, x) | f(t, x) = 0\} \subseteq F^2 \quad (7.1)$$

is called an **affine plane curve**.

Remark 7.1.2. Let F be a field, $D = F[t]$, $f(t, x) \in F[t, x]$. Then we can group together terms containing the same powers of x to rewrite $f(t, x)$ as

$$f(t, x) = \sum_{i=0}^n a_i(t)x^i \in D[x], \quad (7.2)$$

where $a_i(t) \in D$. Note that if $t_0 \in F$ with $a_n(t_0) \neq 0$, then there are exactly n zeros of the polynomial $f(t_0, x) \in F[x]$ in some extension of F , counting multiplicities.

Definition 7.1.3. Let F be a field, $n \in \mathbb{Z}_{>0}$, and

$f(x_1, x_2, \dots, x_n) \in F[x_1, x_2, \dots, x_n]$. We define the **partial derivative of** $f(x_1, x_2, \dots, x_n)$ **with respect to** x_i , denoted by $f_{x_i}(x_1, x_2, \dots, x_n)$, as the derivative of $f(x_1, x_2, \dots, x_n)$ taken as a polynomial over the ring $F[x_1, \dots, \widehat{x_i}, \dots, x_n]$.

Definition 7.1.4. Let F be a field, $f(t, x) \in F[t, x]$ such that $f(t, x) = \sum_{i=0}^n a_i(t)x^i$.

If $t_0 \in F$ such that $a_n(t_0) \neq 0$ and there are n distinct zeros of $f(t_0, x) \in F[x]$ in some extension of F , then t_0 is called a **regular value of** f in the variable t .

Lemma 7.1.5. *Let F be a field, $D = F[t]$. If $f(t, x) \in F[t, x]$ is irreducible in $F[t, x]$, then all but a finite number of values of $t_0 \in F$ are regular values.*

Proof. Let $f(t, x) = \sum_{i=0}^n a_i(t)x^i$ with $a_n(t) \in D \setminus \{0\}$. There exist only finitely many t_0 such that $a_n(t_0) = 0$, so we may assume $a_n(t_0) \neq 0$. Let $K = F(t)$ and $r(t, x) = \gcd(f(t, x), f_x(t, x))$ in the PID $K[x]$. Write $r(t, x) = \sum_{i=0}^m \frac{r_i(t)}{s_i(t)} x^i$, where $r_i(t), s_i(t) \in F[t]$ and $s_i(t) \neq 0$ for $0 \leq i \leq m$. By Theorem 2.9.2, $r(t, x)$ divides $f(t, x)$ in $K[x]$, so $f(t, x) = r(t, x)s(t, x)$ for some $s(t, x) \in K[x]$. By Theorem 2.10.4, there exist $R(t, x), S(t, x) \in D[x]$ such that $f(t, x) = R(t, x)S(t, x)$ and $R(t, x) = s(t)r(t, x)$ for some $s(t) \in K$. Because $f(t, x)$ is irreducible in $D[x]$, either $\deg_x R(t, x) = \deg_x f(t, x)$ or $\deg_x R(t, x) = 0$. By Theorem 2.9.2, $r(t, x)$ divides $f_x(t, x)$ in $K[x]$, so

$$\deg_x R(t, x) = \deg_x r(t, x) \leq \deg_x f_x(t, x) < \deg_x f(t, x). \quad (7.3)$$

Hence, $\deg_x R(t, x) = 0$, and $R(t, x) \in F \setminus \{0\}$. Thus, $r(t, x) \in K \setminus \{0\}$, and since K is a field, we can assume without loss of generality that $r(t, x) = 1$.

By Theorem 2.9.2, $r(t, x)$ is a $K[x]$ -linear combination of $f(t, x)$ and $f_x(t, x)$, so there exist $p_1(t, x), p_2(t, x) \in D[x]$ and $q_1(t), q_2(t) \in D$ such that

$$\frac{p_1(t, x)}{q_1(t)} f(t, x) + \frac{p_2(t, x)}{q_2(t)} f_x(t, x) = 1, \quad (7.4)$$

where $p_1(t, x), p_2(t, x)$ are not both zero and $q_1(t), q_2(t)$ have finitely many zeros in F . Suppose $t_0 \in F$ such that $q_1(t_0), q_2(t_0) \neq 0$. Then by Lemma 2.9.4, $\gcd(f(t_0, x), f_x(t_0, x))$ in $F[x]$ is 1. Hence, all but finitely many $t_0 \in F$ satisfy $\gcd(f(t_0, x), f_x(t_0, x)) = 1$. Therefore, by Theorem 2.8.2, $f(t_0, x) \in F[x]$ has distinct zeros for all but finitely many $t_0 \in F$, and all but finitely many $t_0 \in F$ are regular. □

Definition 7.1.6. If $a \in \mathbb{R}_{>0}$, we say that

$$D_a = \{z \in \mathbb{C} \mid |z| \leq a\} \quad (7.5)$$

is the **disk of radius a centered at the origin**.

Lemma 7.1.7. If $x(t) = \sum_{k=0}^{\infty} b_k t^k$ is a power series in t absolutely convergent in D_R for some $R \in \mathbb{R}_{>0}$, then

$$\sum_{k=0}^{\infty} \left(\sum_{\substack{r_1 + \dots + r_j = k \\ r_i \geq 0}} b_{r_1} b_{r_2} \dots b_{r_j} \right) t^k \quad (7.6)$$

converges absolutely to $(x(t))^j$ inside D_R , where $0 \leq r_n \leq k$ for $0 \leq n \leq j$.

Proof. We will induct on j . Assume $j = 1$. Then $(x(t))^j = \sum_{k=0}^{\infty} b_k t^k$ by hypothesis, and

$$\sum_{K=0}^{\infty} \left(\sum_{\substack{r_1 = K \\ r_1 \geq 0}} b_{r_1} \right) t^K = \sum_{k=0}^{\infty} b_k t^k, \quad (7.7)$$

so our base case is true.

Assume the lemma is true for j . Then by absolute convergence and Theorem

6.4.8, inside D_R , we have the absolute convergence of

$$\begin{aligned}
(x(t))^{j+1} &= (x(t))^j x(t) \\
&= \left(\sum_{k=0}^{\infty} \left(\sum_{\substack{r_1+r_2+\dots+r_j=k \\ r_i \geq 0}} b_{r_1} b_{r_2} \cdots b_{r_j} \right) t^k \right) \left(\sum_{k=0}^{\infty} b_k t^k \right) \\
&= \sum_{k=0}^{\infty} \left(\sum_{n=0}^k \left(\sum_{\substack{r_1+r_2+\dots+r_j=n \\ r_i \geq 0}} b_{r_1} b_{r_2} \cdots b_{r_j} \right) t^n (b_{k-n} t^{k-n}) \right) \\
&= \sum_{k=0}^{\infty} \left(\sum_{n=0}^k \sum_{\substack{r_1+r_2+\dots+r_j=n \\ r_i \geq 0}} b_{r_1} b_{r_2} \cdots b_{r_j} b_{k-n} t^k \right) \tag{7.8} \\
&= \sum_{k=0}^{\infty} \left(\sum_{r_{j+1}=0}^k \sum_{\substack{r_1+r_2+\dots+r_j=k-r_{j+1} \\ r_i \geq 0}} b_{r_1} b_{r_2} \cdots b_{r_j} b_{r_{j+1}} t^k \right) \\
&= \sum_{k=0}^{\infty} \left(\sum_{\substack{r_1+r_2+\dots+r_{j+1}=k \\ r_i \geq 0}} b_{r_1} b_{r_2} \cdots b_{r_{j+1}} \right) t^k,
\end{aligned}$$

where the next to last equality holds by letting $r_{j+1} = k - n$, which is always nonnegative. By induction, the lemma follows. $\square$

Corollary 7.1.8. *If $x(t) = \sum_{k=0}^{\infty} b_k t^k$ is a power series in t absolutely convergent in D_R with $R \in \mathbb{R}_{>0}$ and $b_0 = 0$, then for $j \geq 2$,*

$$x(t)^j = \sum_{k=0}^{\infty} c_k t^k \tag{7.9}$$

converges in D_R with

$$c_k = \sum_{\substack{r_1+\dots+r_j=k \\ 1 \leq r_i \leq k-1}} b_{r_1} b_{r_2} \cdots b_{r_j}, \tag{7.10}$$

which is an integral polynomial with nonnegative coefficients in $b_1, b_2, \dots, b_{k-1}$.

Proof. Let $j \geq 2$. By Lemma 7.1.7, since $b_0 = 0$, if

$$c_k = \sum_{\substack{r_1 + \dots + r_j = k \\ r_i \geq 0}} b_{r_1} b_{r_2} \dots b_{r_j} = \sum_{\substack{r_1 + \dots + r_j = k \\ r_i \geq 1}} b_{r_1} b_{r_2} \dots b_{r_j}, \quad (7.11)$$

then $\sum_{k=0}^{\infty} c_k t^k$ converges to $(x(t))^j$ in D_R . Now, because $r_i \geq 1$ and $j \geq 2$, $r_i \leq k - 1$ for $1 \leq i \leq j$. It follows that c_k is a polynomial in $b_1, b_2, \dots, b_{k-1}$ with nonnegative integral coefficients. $\square$

Lemma 7.1.9. *Suppose Ω_t, Ω_x are neighborhoods of 0 such that the double power series*

$$f(t, x) = a_{10}t - x + \sum_{2 \leq i+j} a_{ij} t^i x^j \quad (7.12)$$

converges absolutely for $t \in \Omega_t$ and $x \in \Omega_x$, and $f(t, x)$ is holomorphic in each variable (holding the other fixed) for $t \in \Omega_t$ and $x \in \Omega_x$. For $k \geq 1$, define $p_k(a_{ij}) \in \mathbb{Z}[a_{ij}]$ recursively by

$$\begin{aligned} p_1(a_{ij}) &= a_{10}, \\ p_k(a_{ij}) &= a_{k0} + \sum_{n=1}^{k-1} a_{n1} p_{k-n}(a_{ij}) \\ &\quad + \sum_{j=2}^k \sum_{n=j}^k \left(\sum_{\substack{r_1 + r_2 + \dots + r_j = n \\ 1 \leq r_i \leq n-1 < k}} p_{r_1}(a_{ij}) p_{r_2}(a_{ij}) \dots p_{r_j}(a_{ij}) \right) a_{(k-n)j}. \end{aligned} \quad (7.13)$$

Then the $p_k(a_{ij})$ have the following properties:

(i) *For $k \geq 1$, the coefficients of $p_k(a_{ij})$ are nonnegative, and for $i + j > k$, the coefficient of a_{ij} in $p_k(a_{ij})$ is 0.*

(ii) *If there exists holomorphic $x(t) \in \mathcal{H}_t(\Omega_t)$ such that*

(a) $x(t) \in \Omega_x$ for $t \in \Omega_t$,

(b) $f(t, x(t)) = 0$ for $t \in \Omega_t$,

(c) $x(0) = 0$, and

(d) $x(t) = \sum_{k=1}^{\infty} b_k t^k$ converges absolutely,

then $b_k = p_k(a_{ij})$.

(iii) If $x(t) = \sum_{k=1}^{\infty} b_k t^k$ converges absolutely in Ω_t , $b_k = p_k(a_{ij})$, and $x(t) \in \Omega_x$ for $t \in \Omega_t$, then $x(0) = 0$ and $f(t, x(t)) = 0$ for $t \in \Omega_t$.

Proof. (i) We will induct on k . The case where $k = 1$ is true by definition of $p_1(a_{ij})$. Suppose (i) holds for $p_r(a_{ij})$, $1 \leq r < k$. Consider (7.13). By our inductive hypothesis, the coefficients of the summands of (7.13) meet the claim, as well. It follows by induction that the coefficients of $p_k(a_{ij})$ meet the claim.

(ii) Assume that $x(t)$ satisfies hypotheses (a), (b), (c), and (d) for some $b_0, b_1, \dots \in \mathbb{C}$. Then for all $t \in \Omega_t$,

$$\begin{aligned} 0 &= f(t, x(t)) \\ &= a_{10}t - x(t) + \sum_{2 \leq i+j} a_{ij} t^i (x(t))^j \\ &= a_{10}t - \left(\sum_{k=1}^{\infty} b_k t^k \right) + \sum_{2 \leq i+j} a_{ij} t^i \left(\sum_{k=1}^{\infty} b_k t^k \right)^j. \end{aligned} \tag{7.14}$$

It follows from Corollary 7.1.8 and absolute convergence that

$$\begin{aligned} \sum_{k=1}^{\infty} b_k t^k &= a_{10}t + \sum_{2 \leq i+j} a_{ij} t^i \left(\sum_{k=1}^{\infty} b_k t^k \right)^j \\ &= a_{10}t + \sum_{k=2}^{\infty} a_{k0} t^k + \left(\sum_{i=1}^{\infty} a_{i1} t^i \right) \left(\sum_{k=1}^{\infty} b_k t^k \right) \\ &\quad + \sum_{j=2}^{\infty} \left(\sum_{i=0}^{\infty} a_{ij} t^i \right) \left(\sum_{k=0}^{\infty} \left(\sum_{\substack{r_1+r_2+\dots+r_j=k \\ 1 \leq r_i \leq k-1}} b_{r_1} b_{r_2} \dots b_{r_j} \right) t^k \right). \end{aligned} \tag{7.15}$$

Examining the double summation in the last equation,

$$\left(\sum_{i=1}^{\infty} a_{i1} t^i \right) \left(\sum_{k=1}^{\infty} b_k t^k \right) = \sum_{k=2}^{\infty} \left(\sum_{n=1}^{k-1} a_{n1} b_{k-n} \right) t^k \quad (7.16)$$

by Theorem 6.4.8. Also, observe that $k - n < k$ since $n \geq 1$.

Examining the quadruple summation, by Theorem 6.4.8 and absolute convergence,

$$\begin{aligned} & \sum_{j=2}^{\infty} \left(\sum_{i=0}^{\infty} a_{ij} t^i \right) \left(\sum_{k=0}^{\infty} \left(\sum_{\substack{r_1+r_2+\dots+r_j=k \\ 1 \leq r_i \leq k-1}} b_{r_1} b_{r_2} \dots b_{r_j} \right) t^k \right) \\ &= \sum_{j=2}^{\infty} \left(\sum_{k=0}^{\infty} \left(\sum_{n=0}^k \left(\sum_{\substack{r_1+r_2+\dots+r_j=n \\ 1 \leq r_i \leq n-1}} b_{r_1} b_{r_2} \dots b_{r_j} \right) a_{(k-n)j} \right) t^k \right) \\ &= \sum_{k=0}^{\infty} \left(\sum_{j=2}^{\infty} \sum_{n=0}^k \left(\sum_{\substack{r_1+r_2+\dots+r_j=n \\ 1 \leq r_i \leq n-1}} b_{r_1} b_{r_2} \dots b_{r_j} \right) a_{(k-n)j} \right) t^k \\ &= \sum_{k=2}^{\infty} \left(\sum_{j=2}^k \sum_{n=j}^k \left(\sum_{\substack{r_1+r_2+\dots+r_j=n \\ 1 \leq r_i \leq n-1 < k}} b_{r_1} b_{r_2} \dots b_{r_j} \right) a_{(k-n)j} \right) t^k, \end{aligned} \quad (7.17)$$

where the last equality holds because for $n < j$, we have $r_i \geq 1$ and

$r_1 + r_2 + \dots + r_j > n$. Furthermore, k begins at 2 since j begins at 2 and

terminates at k .

By the Identity Theorem, $b_1 = a_{10}$, and for $k > 1$,

$$b_k = a_{k0} + \sum_{n=1}^{k-1} a_{n1} b_{k-n} + \sum_{j=2}^k \sum_{n=j}^k \left(\sum_{\substack{r_1+r_2+\dots+r_j=n \\ r_i \geq 1}} b_{r_1} b_{r_2} \dots b_{r_j} \right) a_{(k-n)j}. \quad (7.18)$$

Therefore, $b_k = p_k(a_{ij})$, where $p_k(a_{ij})$ is as claimed.

(iii) Assume $\sum_{k=1}^{\infty} b_k t^k$ converges absolutely in Ω_t , $b_k = p_k(a_{ij})$, and $x(t) \in \Omega_x$ for $t \in \Omega_t$. Trivially, $x(0) = 0$. Consider

$$\begin{aligned} f(t, x(t)) &= a_{10}t - x(t) + \sum_{2 \leq i+j} a_{ij} t^i (x(t))^j \\ &= a_{10}t - \left(\sum_{k=1}^{\infty} b_k t^k \right) + \sum_{2 \leq i+j} a_{ij} t^i \left(\sum_{k=1}^{\infty} b_k t^k \right)^j. \end{aligned} \quad (7.19)$$

Because $\sum_{k=1}^{\infty} b_k t^k$ converges absolutely and $b_k = p_k(a_{ij})$, by the argument from (ii) backwards,

$$\sum_{k=1}^{\infty} b_k t^k = a_{10}t + \sum_{2 \leq i+j} a_{ij} t^i \left(\sum_{k=1}^{\infty} b_k t^k \right)^j. \quad (7.20)$$

Therefore, $f(t, x(t)) = 0$.

□

Theorem 7.1.10. *If $f(t, x) \in \mathbb{C}[t, x]$, $t_0 \in \mathbb{C}$ is a regular value of f , and $f(t_0, x_0) = 0$ for some $x_0 \in \mathbb{C}$, then there exists a neighborhood Ω_t of t_0 and $x(t) \in \mathcal{H}_t(\Omega_t)$ with $x(t_0) = x_0$ and $f(t, x(t)) = 0$ for all $t \in \Omega_t$.*

Proof. Let $A > 0$. Consider

$$f_A(t, x) = At - x + A \frac{x^2}{1-x} + At \frac{x}{1-x} + A \left(\frac{t^2}{1-t} \right) \left(\frac{1}{1-x} \right). \quad (7.21)$$

For $t, x \in D_1$, by geometric series,

$$f_A(t, x) = At - x + At^0 \sum_{j=2}^{\infty} x^j + At^1 \sum_{j=1}^{\infty} x^j + A \sum_{j=2}^{\infty} t^j \left(\sum_{j=0}^{\infty} x^j \right). \quad (7.22)$$

By the absolute convergence of geometric series in D_1 ,

$$f_A(t, x) = At - x + A \sum_{i+j \geq 2} t^i x^j \quad (7.23)$$

for $t, x \in D_1$. But also

$$\begin{aligned} (1-x)f_A(t, x) &= At - Atx - x + x^2 + Ax^2 + Atx + A \left(\frac{t^2}{1-t} \right) \\ &= (A+1)x^2 - x + \frac{At}{1-t}. \end{aligned} \quad (7.24)$$

By the quadratic formula, $(1-x)f_A(t, x) = 0$ if and only if

$$x = x_A(t) = \frac{1 \pm \sqrt{1 - 4(A+1) \left(\frac{At}{1-t}\right)}}{2(A+1)}. \quad (7.25)$$

So if $x_A(t) \neq 1$, $f_A(t, x_A(t)) = 0$. By Corollary 6.8.6, for $t \in D_r$, where $r = \frac{1}{4A^2+4A+1}$, $x_A(t) = \frac{1 - \frac{\sqrt{1-(1-4A^2-4A)t}}{\sqrt{1-t}}}{2A+2}$. Moreover, from (7.21), $|x_A(t)| < 1$.

By Theorem 6.5.2, if $|u| < 1$, then the expression $\sqrt{1-u}$ can be expressed as a power series in u . In particular, for $t \in D_r$, $x_A(t)$ can be expressed as a convergent power series at 0; in fact, by Lemma 7.1.9 we must have $x_A(t) = \sum_{k=1}^{\infty} A_k t^k$, where $A_k = p_k(A)$.

Now consider $f(t, x) \in \mathbb{C}[t, x]$ such that $f(t_0, x_0) = 0$ and t_0 is a regular value of f . Since $g(t - t_0, x - x_0)$ is a linear transformation of $f(t, x)$, we can assume that $t_0 = x_0 = 0$. Now, $f(0, 0) = 0$ by assumption, so write

$$f(t, x) = a_{10}t + a_{01}x + \sum_{2 \leq i+j \leq N} a_{ij}t^i x^j, \quad (7.26)$$

where N is the total degree of $f(t, x)$. Observe that the polynomial $f(t, x)$ satisfies the hypotheses of Lemma 7.1.9 for any neighborhoods Ω_t and Ω_x in $\mathbb{C}$. By differentiation, $f_x(0, 0) = a_{01}$. Because 0 is a regular value of f , $\gcd(f(0, x), f_x(0, x)) = 1$, so $a_{01} \neq 0$. By scaling f , assume $a_{01} = -1$. Hence,

$$f(t, x) = a_{10}t - x + \sum_{2 \leq i+j \leq N} a_{ij}t^i x^j. \quad (7.27)$$

Let $A = \max |a_{ij}|$, $b_k = p_k(a_{ij})$ where p_k is defined in Lemma 7.1.9, and $x(t) = \sum_{k=1}^{\infty} b_k t^k$. Since p_k has nonnegative integer coefficients,

$$A_k = p_k(A) \geq |p_k(a_{ij})| = |b_k| \quad (7.28)$$

for $k \geq 1$, so, for each k , $A_k \geq |b_k|$, and for

$$\sum_{k=1}^{\infty} |b_k t^k| \leq \sum_{k=1}^{\infty} |A_k t^k| < \infty. \quad (7.29)$$

Therefore, $x(t)$ has a positive radius of convergence, so by Lemma 7.1.9, $x(t)$ satisfies the theorem. $\square$

Corollary 7.1.11. *Let $f(t, x) \in \mathbb{C}[t, x]$, $n = \deg_x f(t, x)$, $t_0 \in \mathbb{C}$ be a regular value of f , and $x_1(t), x_2(t), \dots, x_n(t) \in \mathcal{H}_t(\Omega)$ be distinct roots at t_0 . Then there exists a neighborhood Ω of t_0 such that $x_1(t), x_2(t), \dots, x_n(t)$ are each continuous and pairwise not equal on Ω and $f(t, x)$ factors in $(\mathcal{H}_t(\Omega))[x]$ and $f(t, x_i(t)) = 0$ for all $t \in \Omega$, $1 \leq i \leq n$.*

Proof. By Theorem 7.1.10, there exist respective $x_1(t), x_2(t), \dots, x_n(t) \in \mathcal{H}_t(\Omega)$ such that $f(t, x_i(t)) = 0$ for all $t \in \Omega$, $1 \leq i \leq n$.

Let $\epsilon = \frac{1}{2} \min |x_i(t_0) - x_j(t_0)|$. Since $x_i(t) \in \mathcal{H}_t(\Omega)$ are continuous and distinct at t_0 , there exist $\delta_i > 0$ such that $x_i(N_{\delta_i}(t_0)) \subseteq N_\epsilon(x_i(t_0))$ for $1 \leq i \leq n$. Hence, the $x_i(t)$ are continuous and pairwise not equal on $\Omega = D_\delta$, where $\delta = \min_{1 \leq i \leq n} \delta_i$. $\square$

Theorem 7.1.12. *Let $f(t, x) \in \mathbb{Q}[t, x]$ be irreducible over $\mathbb{Q}$, t_0 a regular value of f . Then there exists a neighborhood Ω of t_0 such that f can be expressed as a product of linear factors in $\mathcal{M}_t(\Omega)[x]$.*

Proof. Write

$$f(t, x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0, \quad (7.30)$$

where $a_i \in \mathbb{Q}[t]$ for $0 \leq i \leq n$. By Corollary 7.1.11, there exists a neighborhood Ω of t_0 such that there exist $x_1(t), x_2(t), \dots, x_n(t) \in \mathcal{M}_t(\Omega)$ with $f(t, x_i(t)) = 0$ for $1 \leq i \leq n$ and for all $t \in \Omega$. Moreover, $a_n(t) \neq 0$, or else $\deg_x f(t, x) < n$.

It follows that

$$f(t, x) = a_n(t) \prod_{i=1}^n (x - x_i(t)) \quad (7.31)$$

is in $\mathcal{M}_t(\Omega)[x]$. $\square$

7.2 Interpolation and the Mean Value Theorem

We follow Hadlock [Had78] in this section.

Lemma 7.2.1. *Let $t_0, t_1, \dots, t_m \in \mathbb{R}$ where $t_0 < t_1 < \dots < t_m$ for some $m \in \mathbb{Z}_{>0}$ and $z : [t_0, t_m] \rightarrow \mathbb{C}$ be a function that is m times differentiable on $[t_0, t_m]$. There exists a unique polynomial $y(t)$ with $\deg y \leq m$ such that $y(t_i) = z(t_i)$ for $0 \leq i \leq m$. Moreover, there exists $T \in (t_0, t_m)$ such that $y^{(m)}(T) = z^{(m)}(T)$.*

Proof. Let $y(t)$ be as in Theorem 2.11.2. We will prove by induction on k that for $0 \leq k \leq m$ there exist $m - k + 1$ distinct points $t_0^{(k)} < t_1^{(k)} < \dots < t_{m-k}^{(k)}$ such that $y^{(k)}(t_i^{(k)}) = z^{(k)}(t_i^{(k)})$ for $0 \leq i \leq m - k$, and for $k \geq 1$, $t_0 < t_0^{(k)} < \dots < t_{m-k}^{(k)} < t_m$.

For $k = 0$, the lemma holds by Theorem 2.11.2.

Suppose the lemma holds for $0 \leq k < m$. Consider $Y_k(t) = y^{(k)}(t) - z^{(k)}(t)$. For $0 \leq i \leq m - k - 1$, since $Y_k(t_i^{(k)}) = Y_k(t_{i+1}^{(k)}) = 0$, by the Mean Value Theorem, there exists $t_i^{(k+1)} < t_{i+1}^{(k+1)} < t_{i+1}^{(k)}$ such that $y^{(k+1)}(t_i^{(k+1)}) = z^{(k+1)}(t_i^{(k+1)})$. The lemma follows by induction. $\square$

Lemma 7.2.2. *Let $t_0, t_1, \dots, t_m \in \mathbb{R}$ where $t_0 < t_1 < \dots < t_m$ for some $m \in \mathbb{Z}_{>0}$, $z : [t_0, t_m] \rightarrow \mathbb{C}$,*

$$\begin{aligned}
 V_m &= \begin{bmatrix} 1 & t_0 & t_0^2 & \dots & t_0^{m-1} & t_0^m \\ 1 & t_1 & t_1^2 & \dots & t_1^{m-1} & t_1^m \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^{m-1} & t_m^m \end{bmatrix}, \\
 W_m &= \begin{bmatrix} 1 & t_0 & t_0^2 & \dots & t_0^{m-1} & z(t_0) \\ 1 & t_1 & t_1^2 & \dots & t_1^{m-1} & z(t_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & t_m & t_m^2 & \dots & t_m^{m-1} & z(t_m) \end{bmatrix},
 \end{aligned} \tag{7.32}$$

$V = \det V_m$, and $W = \det W_m$. If $z(t)$ is m times differentiable on (t_0, t_m) , then there exists some $T \in (t_0, t_m)$ such that

$$\frac{z^{(m)}(T)}{m!} = \frac{W}{V}. \quad (7.33)$$

Proof. By Lemma 7.2.1, choose $y(t)$ to be the unique polynomial of degree less than or equal to m with $y(t_i) = z(t_i)$ and $T \in (t_0, t_m)$ such that $y^{(m)}(T) = z^{(m)}(T)$.

Write $y(t) = a_0 + a_1t + \cdots + a_mt^m$. Since $y(t_i) = z(t_i)$ for $i = 0, 1, \dots, m$,

$$\begin{aligned} a_0 + a_1t_0 + \cdots + a_{m-1}t_0^{m-1} + a_mt_0^m &= z(t_0) \\ a_0 + a_1t_1 + \cdots + a_{m-1}t_1^{m-1} + a_mt_1^m &= z(t_1) \\ &\vdots \\ a_0 + a_1t_m + \cdots + a_{m-1}t_m^{m-1} + a_mt_m^m &= z(t_m). \end{aligned} \quad (7.34)$$

By Theorem 2.2.6, $V = \prod_{i>j} (t_i - t_j)$. Because $t_i < t_j$ for all $0 \leq i < j \leq m$, $V \neq 0$.

Therefore, by Theorem 2.2.8, $a_m = \frac{W}{V}$. Since $\deg y(t) \leq m$,

$$y^{(m)}(t) = m!a_m = \frac{m!W}{V}. \quad (7.35)$$

Therefore,

$$\frac{z^{(m)}(T)}{m!} = \frac{y^{(m)}(T)}{m!} = \frac{W}{V}. \quad (7.36)$$

□

7.3 The Density Lemma

Theorem 7.3.1. *Let Ω be a domain in $\hat{\mathbb{C}}$. The field $\mathbb{C}(t)$ is isomorphic to a subfield of $\mathcal{M}_t(\Omega)$.*

Proof. Define a homomorphism $\varphi : \mathbb{C}(t) \rightarrow \mathcal{M}_t(\Omega)$ as follows: For $a(t) \in \mathbb{C}(t)$, define $\varphi_a : \Omega \rightarrow \mathbb{C}$ by $\varphi_a(t) = a(t)$.

By Theorem 6.7.2, if $\varphi_a(t) = 0$ as a function, then $a(t) = 0$ in $\mathbb{C}(t)$, so $|\ker \varphi| = 1$, and φ is a monomorphism. By the first isomorphism theorem, $\mathbb{C}(t)$ is isomorphic to a subfield of $\mathcal{M}_t(\Omega)$. $\square$

Corollary 7.3.2. *For every $\Omega \subseteq \mathbb{C}$, $\mathbb{Q}(t)$ is isomorphic to a subfield of $\mathcal{M}_t(\Omega)$.*

Proof. Notice that $\mathbb{Q}(t)$ is a subfield of $\mathbb{C}(t)$. By Theorem 7.3.1, $\mathbb{C}(t)$ is isomorphic to a subfield of $\mathcal{M}_t(\Omega)$. The corollary follows. $\square$

Remark 7.3.3. The previous corollary implies that we may think of $\mathcal{M}_t(\Omega)$ as an extension field of $\mathbb{Q}(t)$.

Theorem 7.3.4 (The Density Lemma). *Let Ω be a domain, $T_0 \in \mathbb{R}_{>0}$ such that $\{t \in \mathbb{C} \mid |t| \geq T_0\} \subseteq \Omega$, and $y(t) \in \mathcal{M}_t(\Omega)$ such that $y(t)$ is algebraic over $\mathbb{Q}(t)$ but $y(t) \notin \mathbb{Q}(t)$. For any $\epsilon > 0$, there exist $T(\epsilon) > T_0$ and $N(\epsilon) \in \mathbb{Z}_{>0}$ such that for any $N_1, N_2 \in \mathbb{Z}$, $N_1 > T(\epsilon)$, $N_2 \geq N(\epsilon)$, if*

$$\rho(N_1, N_2) = \{t_0 \in \mathbb{Z} \mid y(t_0) \in \mathbb{Q} \text{ and } N_1 \leq t_0 < N_1 + N_2\}, \quad (7.37)$$

then we have

$$\frac{|\rho(N_1, N_2)|}{N_2} \leq \epsilon. \quad (7.38)$$

The Density Lemma shows that the proportion of rational outputs coming from integral inputs that an algebraic but non-rational meromorphic function might have in a neighborhood of infinity tends to zero. We will examine three cases examining how far apart the rational values $y(t_0)$ must be given that $t_0 \in \mathbb{Z}$. We first will transform $y(t)$ into a Laurent series and examine two trivial cases. Lastly, we will consider an interval of the real number line. We will subdivide this interval into several equal subintervals, and a “leftover” interval, examining what proportion of rational values $y(t_0)$ might attain in the entire interval.

Proof. Say the minimal polynomial $D(u) \in (\mathbb{Q}(t))[u]$ of $y(t)$ over $\mathbb{Q}(t)$ is

$$D(u) = d_m(t)u^m + d_{m-1}(t)u^{m-1} + \cdots + d_0(t), \quad (7.39)$$

with $D(y(t)) = 0$ for all $t \in \Omega$. The denominator of $d_k(t)$ is not zero in $\mathbb{Q}[t]$ for

$0 \leq k \leq m$, so without loss of generality $D(u) \in (\mathbb{Z}[t])[u]$. Let

$z(t) = d_m(t)y(t) \in \mathcal{M}_t(\Omega)$, and let $d(u) = d_m(t)^{m-1}D(u) \in (\mathbb{Z}[t])[u]$. Then

$$\begin{aligned} 0 &= d(y(t)) \\ &= d_m(t)^{m-1}d_m(t)y(t)^m + d_m(t)^{m-1}d_{m-1}(t)y(t)^{m-1} + \cdots + d_m(t)^{m-1}d_0(t) \\ &= (d_m(t)y(t))^m + d_{m-1}(t)(d_m(t)y(t))^{m-1} + d_{m-2}(t)d_m(t)(d_m(t)y(t))^{m-2} + \cdots \\ &\quad + d_0(t)d_m(t)^{m-1} \\ &= z(t)^m + b_{m-1}(t)z(t)^{m-1} + \cdots + b_0(t), \end{aligned} \quad (7.40)$$

where $b_k(t) = d_k(t)d_m(t)^{m-1-k} \in \mathbb{Z}[t]$ for $0 \leq k \leq m-1$.

Choose $T_1 \geq T_0$ to be greater than all of the roots of $d_m(t)$. Let $t_0 \in \mathbb{Z}$ such that $t_0 > T_1$. Suppose $z(t_0) \in \mathbb{Z}$. Since $z(t) = d_m(t)y(t)$ and $d_m(t_0) \neq 0$, $y(t_0) = \frac{z(t_0)}{d_m(t_0)}$, so $y(t_0) \in \mathbb{Q}$. On the other hand, suppose $y(t_0) \in \mathbb{Q}$. Now,

$$0 = d(y(t_0)) = z(t_0)^m + b_{m-1}(t_0)z(t_0)^{m-1} + \cdots + b_0(t_0). \quad (7.41)$$

By the Rational Root Theorem, $z(t_0)$ is a ratio of the factors of $b_0(t_0) \in \mathbb{Z}$ to factors of 1, so $z(t_0) \in \mathbb{Z}$. Therefore, $z(t_0) \in \mathbb{Z}$ if and only if $y(t_0) \in \mathbb{Q}$. Hence, for $N_1, N_2 \in \mathbb{Z}_{>0}$,

$$\rho(N_1, N_2) = \{t_0 \in \mathbb{Z} \mid z(t_0) \in \mathbb{Z} \text{ and } N_1 \leq t_0 < N_1 + N_2\}. \quad (7.42)$$

Since $z(t)$ is meromorphic at ∞ , it follows by Theorem 6.7.4 that $z(t)$ has a Laurent series in $\frac{1}{t}$ that converges in some neighborhood $\Omega_1 = \{t \mid |t| > T_2 \in \mathbb{R}_{>0}\}$

of ∞ , $T_2 \geq T_1$. Write

$$z(t) = c_k t^k + c_{k-1} t^{k-1} + \cdots + c_1 t + c_0 + \frac{c_{-1}}{t} + \frac{c_{-2}}{t^2} + \cdots, \quad (7.43)$$

which is holomorphic in Ω_1 except at ∞ .

We first consider two trivial cases.

- (i) Suppose $z(t) \in \mathbb{R}[t]$. Because $y(t) = \frac{z(t)}{d_m(t)} \in \mathbb{R}(t)$ is not in $\mathbb{Q}(t)$, $z(t)$ must have an irrational coefficient. By the contrapositive of Corollary 2.11.3, there are a finite number of $t_0 \in \mathbb{Z}$ such that $z(t_0) \in \mathbb{Z}$, and the theorem follows.
- (ii) Suppose $c_i \notin \mathbb{R}$ for some $i \leq k$. Suppose c_i is the coefficient of the highest term of $z(t)$ such that $c_i \notin \mathbb{R}$. Then

$$\lim_{t \rightarrow \infty} \left(\operatorname{Im} \frac{z(t)}{t^i} \right) = \operatorname{Im} c_i. \quad (7.44)$$

It follows that for some $T_3 \geq T_2$, $z(t)$ is not real for all $t \geq T_3$. Hence, there are only a finite number of integers $t_0 \geq T_2$ such that $z(t_0) \in \mathbb{Z}$, and the theorem follows.

Since the theorem holds in the trivial cases (i) and (ii), we now assume $c_i \in \mathbb{R}$ for all $i \leq k$ with $c_\ell \neq 0$ for some $\ell < 0$. We can repeatedly differentiate $z(t)$ so that

$$z^{(m)}(t) = \frac{p_1}{t^q} + \frac{p_2}{t^{q+1}} + \cdots, \quad (7.45)$$

where $m = \max(k+1, 1) \geq 1$, $p_1 \neq 0$, $q \in \mathbb{Z}_{>0}$. Now,

$$\lim_{t \rightarrow \infty} t^q z^{(m)}(t) = p_1. \quad (7.46)$$

It follows that there exists $T_3 \geq T_2$ such that

$$0 < \frac{1}{2} \left(\frac{|p_1|}{t^q} \right) \leq |z^{(m)}(t)| \leq \frac{2|p_1|}{t^q} \quad (7.47)$$

for all $t \geq T_3$.

Suppose we have $m + 1$ integers $t_0 < t_1 < \cdots < t_m$ such that $T_2 \leq t_0$ and $z(t_i) \in \mathbb{Z}$ for $0 \leq i \leq m$. Let V_m, W_m, V, W be as in Lemma 7.2.2. By Lemma 7.2.2, there exists T such that $t_0 \leq T \leq t_m$ and $\frac{z^{(m)}(T)}{m!} = \frac{W}{V}$. Furthermore, by (7.47),

$$\frac{2|p_1|}{m!t_0^q} \geq \frac{2|p_1|}{m!T^q} \geq \frac{|z^{(m)}(T)|}{m!} = \frac{W}{V}. \quad (7.48)$$

But since $\frac{|z^{(m)}(T)|}{m!} > 0$ and $V, W \in \mathbb{Z}$,

$$\frac{2|p_1|}{m!t_0^q} \geq \frac{1}{V}. \quad (7.49)$$

Taking reciprocals,

$$\frac{m!}{2|p_1|} t_0^q \leq V = \prod_{j>k} (t_j - t_k) < (t_m - t_0)^{\frac{m(m+1)}{2}}. \quad (7.50)$$

Let $\alpha = \left(\frac{m!}{2|p_1|}\right)^{\frac{2}{m(m+1)}}$ and $\beta = \frac{2q}{m(m+1)}$. Then

$$0 < \alpha t_0^\beta < t_m - t_0. \quad (7.51)$$

Let $\epsilon > 0$. Take $T(\epsilon)$ such that $T(\epsilon) > T_3$ and $\alpha(T(\epsilon))^\beta \geq \frac{2m}{\epsilon} + 1$. Let $N = N(\epsilon) = \lceil \frac{2m}{\epsilon} \rceil$.

By (7.51), if there are $m + 1$ integers $T(\epsilon) < t_0 < t_1 < \cdots < t_m$ such that $z(t_i) \in \mathbb{Z}$ for $0 \leq i \leq m$, then

$$t_m - t_0 > \alpha t_0^\beta > \alpha(T(\epsilon))^\beta \geq \frac{2m}{\epsilon} + 1, \quad (7.52)$$

so $t_m - t_0 > N$. Equivalently, if $a > T(\epsilon)$ and $b \leq N$, then there are at most m integers $t_0 < t_1 < \cdots < t_{m-1}$ such that $t_i \in \rho(a, b)$ for $0 \leq i \leq (m-1)$.

Choose $N_1 > T(\epsilon)$, $N_2 \in \mathbb{Z}$ such that $N_2 \geq N$. Let $r = \lfloor \frac{N_2}{N} \rfloor$ and $r' = \frac{N_2}{N} - r$. Consider the interval $[N_1, N_1 + N_2) = [N_1, N_1 + (r + r')N)$. Subdivide this interval

into r intervals of length N and one interval of length $< N$. Then

$$\begin{aligned} \rho(N_1, N_2) &= \rho(N_1, N) \cup \rho(N_1 + N, N) \cup \rho(N_1 + 2N, N) \cup \cdots \\ &\cup \rho(N_1 + (r-1)N, N) \cup \rho(N_1 + rN, r'N). \end{aligned} \quad (7.53)$$

There are at most m integers in each $\rho(N_1 + jN, N)$, $0 \leq j \leq r-1$, and in $\rho(N_1 + rN, r'N)$. Hence,

$$\frac{|\rho(N_1, N_2)|}{N_2} \leq \frac{(r+1)m}{rN} \leq \frac{2m}{N} \leq \frac{2m}{\left(\frac{2m}{\epsilon}\right)} = \epsilon. \quad (7.54)$$

The Density Lemma follows. $\square$

Corollary 7.3.5. *For any domain Ω containing some $t_0 \in \mathbb{Q}$,*

$y_1(t), y_2(t), \dots, y_M(t) \in \mathcal{M}_t(\Omega)$ each algebraic over $\mathbb{Q}$ with $y_i(t) \notin \mathbb{Q}[t]$, there exist infinitely many $t_1 \in \mathbb{Q}$ such that $y_i(t_1) \notin \mathbb{Q} \cap \Omega$ for all $1 \leq i \leq M$.

Proof. Let $f, g : \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$ be defined by $f(t) = t_0 + \frac{1}{t}$, $g(t) = \frac{1}{t-t_0}$, and let $z_i(t) = y_i(f(t))$ for all $1 \leq i \leq M$. Because

$$\begin{aligned} f(g(t)) &= t_0 + \frac{1}{\frac{1}{t-t_0}} = t_0 + t - t_0 = t, \\ g(f(t)) &= \frac{1}{\left(t_0 + \frac{1}{t}\right) - t_0} = t, \end{aligned} \quad (7.55)$$

f and g are inverse functions. If $t \in \mathbb{Q} \cup \{\infty\}$, then we see that $f(t), g(t) \in \mathbb{Q} \cup \{\infty\}$.

As f and g are inverse functions, if $t \notin \mathbb{Q} \cup \{\infty\}$, then $f(t), g(t) \notin (\mathbb{Q} \cup \{\infty\})(t)$.

Now, we claim that since $y_i(t)$ is algebraic over $\mathbb{Q}(t)$, $z_i(t)$ is also algebraic over $\mathbb{Q}(t)$ for all $1 \leq i \leq M$. By definition, there exists $r(s) \in (\mathbb{Q}(t))[s]$ such that $r(y_i(t)) = 0$ for all $t \in \Omega$. Write

$$r(s) = r_n(t)s^n + r_{n-1}(t)s^{n-1} + \cdots + r_1(t)s + r_0(t), \quad (7.56)$$

so

$$0 = r(y_i(t)) = r_n(t)(y_i(t))^n + r_{n-1}(t)(y_i(t))^{n-1} + \cdots + r_1(t)(y_i(t)) + r_0(t). \quad (7.57)$$

Because $r(y_i(t))$ is equivalently zero for all $t \in \Omega$, it follows that $r(y_i(t_0 + \frac{1}{t})) = 0$. Moreover, $r(y_i(f(t))) = 0$, so $y_i(f(t))$ is algebraic over $\mathbb{Q}(t)$, and $z_i(t)$ is algebraic over $\mathbb{Q}(t)$. This proves the claim.

Let $\Omega_2 = g(\Omega)$. Since Ω contains t_0 and g sends t_0 to ∞ , Ω_2 is a neighborhood of infinity. It follows that $\{t \in \mathbb{C} \mid |t| \geq T_0\} \subseteq \Omega_2$ for some $T_0 \in \mathbb{R}$.

We also claim that $t_1 \in \mathbb{Q}$ such that $z_i(t_1) \notin \mathbb{Q}$ if and only if $t_2 = t_0 + \frac{1}{t_1} \in \mathbb{Q}$ such that $y_i(t_2) \notin \mathbb{Q}$. Since

$$z_i(t_1) = y_i(f(t_1)) = y_i(t_0 + \frac{1}{t_1}) = y_i(t_2), \quad (7.58)$$

$y_i(t_2) \notin \mathbb{Q}$ if $t_2 \in \mathbb{Q}$ if and only if $z_i(t_1) \notin \mathbb{Q}$ if $t_1 \in \mathbb{Q}$. Moreover, by the chain rule, $y_i(t_0 + \frac{1}{t}) \in \mathcal{M}_t(\Omega_2)$, so $z_i(t) \in \mathcal{M}_t(\Omega_2)$. Hence, we can transform $y_i(t)$ to $z_i(t)$ and we can assume $t_0 = \infty$.

Choose $\epsilon \leq \frac{1}{2M}$. Let $T^{(i)}(\epsilon), N^{(i)}(\epsilon), N_1^{(i)}(\epsilon), N_2^{(i)}(\epsilon)$ be as in Theorem 7.3.4 for each $z_i(t)$, and let $N = \max N^{(i)}$, $N_1 = \max N_1^{(i)}(\epsilon)$, $N_2 = \max N_2^{(i)}(\epsilon)$, and $T(\epsilon) = \max T^{(i)}(\epsilon)$. Let

$$\begin{aligned} \rho_i(N_1, N_2) &= \{t_1 \in \mathbb{Z} \mid y_i(t_0) \in \mathbb{Q} \text{ and } N_1 \leq t_1 < N_1 + N_2\}, \\ \rho(N_1, N_2) &= \{t_1 \in \mathbb{Z} \mid y_i(t_0) \in \mathbb{Q} \text{ for some } 1 \leq i \leq M \text{ and } N_1 \leq t_1 < N_1 + N_2\}. \end{aligned} \quad (7.59)$$

By the Density Lemma,

$$\frac{|\rho(N_1, N_2)|}{N_2} \leq \sum_{i=1}^M \left(\frac{|\rho_i(N_1, N_2)|}{N_2} \right) \leq \frac{1}{2}. \quad (7.60)$$

So, at least half of the rational numbers in $[N_1, N_1 + N_2)$ are mapped simultaneously to irrational numbers by $z_i(t)$ for all $1 \leq i \leq M$. That is, there are infinitely many $t_1 \in \mathbb{Q}$ such that $z_i(t_1) \notin \mathbb{Q}$ for all $1 \leq i \leq M$. The corollary follows. $\square$

CHAPTER 8

HILBERT'S IRREDUCIBILITY THEOREM

8.1 Hilbert's Irreducibility Theorem

We base our arguments on Hadlock [Had78] in this section.

Definition 8.1.1. For $f(t, x) \in (\mathbb{Q}(t))[x]$, we say that t_0 is a **Hilbert value** of $f(t, x)$ if $t_0 \in \mathbb{Q}$ is a regular value of $f(t, x)$ and $f(t_0, x) \in \mathbb{Q}[x]$ is irreducible over $\mathbb{Q}$.

We continue with Hilbert's Irreducibility Theorem on two variables.

Theorem 8.1.2. *[Hilbert's Irreducibility Theorem on Two Variables] Let $f(t, x) \in \mathbb{Q}[t, x]$ be irreducible over $\mathbb{Q}$. Then there exist infinitely many Hilbert values of $f(t, x)$. If $f_1(t, x), f_2(t, x), \dots, f_M(t, x)$ are M such polynomials, then there exist infinitely many $t_0 \in \mathbb{Q}$ such that t_0 is a Hilbert value of $f_i(t, x)$ for each $1 \leq i \leq M$.*

Proof. Write

$$f(t, x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0, \quad (8.1)$$

where $a_i \in \mathbb{Q}[t]$ for $0 \leq i \leq n$. By Theorem 7.1.12, there exists a neighborhood Ω of a regular value $t_0 \in \mathbb{Q}$ of f such that $f(t, x)$ can be expressed as a product of linear factors in $\mathcal{M}_t(\Omega)[x]$, so

$$f(t, x) = a_n(t) \prod_{i=1}^n (x - x_i(t)) \in \mathcal{M}_t(\Omega)[x]. \quad (8.2)$$

It follows that if, for some $t_0 \in \mathbb{Q}$, $f(t_0, x) = h_1(x)h_2(x)$ for some $h_1(x), h_2(x) \in \mathbb{C}[x]$, then because $\mathbb{C}[x]$ is a unique factorization domain, we must have

$$h_1(x) = a_{n_1}(t) \prod_{i \in S} (x - x_i(t_0)), \quad h_2(x) = a_{n_2}(t) \prod_{i \notin S} (x - x_i(t_0)) \quad (8.3)$$

for some $S \subseteq [n]$ where $a_n(t) = a_{n_1}(t)a_{n_2}(t)$ for some $a_{n_1}(t), a_{n_2}(t) \in \mathcal{M}_t(\Omega)$. Since $f(t, x)$ is irreducible over $\mathbb{Q}$, each factorization of $f(t, x)$ of the form

$$a_n(t) \left(\prod_{i \in S} (x - x_i(t)) \right) \left(\prod_{i \notin S} (x - x_i(t)) \right) \quad (8.4)$$

must contain one coefficient in either $\prod_{i \in S} (x - x_i(t))$ or $\prod_{i \notin S} (x - x_i(t))$ that is not in $\mathbb{Q}(t)$. Because there are n different linear factors, there can only be a finite number of ways to factor $f(t, x)$ in this manner.

Let $y(t) \in \mathcal{M}_t(\Omega) \setminus \mathbb{Q}[t]$ be one of the non-rational coefficient polynomials obtained above. Since $x_i(t) \in \mathcal{M}_t(\Omega)[t]$ and is algebraic over $\mathbb{Q}(t)$ for $1 \leq i \leq n$, $y(t) \in \mathcal{M}_t(\Omega)[t]$ and is algebraic over $\mathbb{Q}(t)$.

Because there are a finite number of nontrivial factorizations of $f(t, x)$ in $\mathcal{M}_t(\Omega)[x]$ over $\mathbb{C}$, there are a finite number of non-rational coefficient functions $y(t)$ obtained from those factorizations; call them $y_1(t), y_2(t), \dots, y_M(t)$. By Corollary 7.3.5, for some neighborhood $\Omega_1 \subseteq \Omega$ of t_0 , there exist infinitely many $t_0 \in (\Omega_1 \cap \mathbb{Q})$ such that $y_j(t_0) \notin \mathbb{Q}$ for $1 \leq j \leq M$. Hence, there exist $t_0 \in \mathbb{Q}$ such that $y_j(t_0) \notin \mathbb{Q}$ for $1 \leq j \leq M$. Therefore, there are infinitely many $t_0 \in \mathbb{Q}$ such that $f(t_0, x)$ is irreducible over $\mathbb{Q}$. $\square$

Lemma 8.1.3 (Kronecker's Criterion). *Let $d \in \mathbb{Z}$, $d \geq 2$, $D = \mathbb{Q}[u_0]$ or $\mathbb{Q}$,*

$$P_d = \{g \in D[u_1, \dots, u_n] \mid \text{the maximum exponent of any } u_1, u_2, \dots, u_n \text{ is strictly less than } d\}, \quad (8.5)$$

$$K_d = \{\hat{g} \in D[y] \mid \deg_y \hat{g} \leq d^n - 1\}.$$

(i) *The map $\hat{\cdot}: P_d \rightarrow K_d$ defined by the $\mathbb{Q}$ -linear extension of*

$$u_1^{i_1} u_2^{i_2} \cdots u_n^{i_n} \mapsto y^{i_1 + di_2 + d^2 i_3 + \cdots + d^{n-1} i_n} \quad (8.6)$$

is a bijection;

(ii) If $G, H \in P_d$ such that $GH \in P_d$, then $\widehat{GH} = \widehat{G}\widehat{H}$; and

(iii) The following are equivalent:

(a) The polynomial $f \in P_d$ is irreducible over $\mathbb{Q}$.

(b) If $\widehat{f}$ factors nontrivially in K_d such that $\widehat{f} = \widehat{g}\widehat{h}$, then $gh \notin P_d$, where $g, h \in P_d$.

To expand on (ii) and (iii), we will present examples with $d = 4$ and $n = 2$.

Consider $G(u_1, u_2), H(u_1, u_2) \in P_4$ where

$$G = u_1^2 + u_1u_2 + u_2^2, \quad H = u_1 - u_2. \quad (8.7)$$

Then $GH = u_1^3 - u_2^3$. Using the map defined in (i),

$$\begin{aligned} \widehat{G} &= y^2 + y^{1+4(1)} + y^{4(2)} &&= y^2 + y^5 + y^8, \\ \widehat{H} &= y^1 - y^{4(1)} &&= y^1 - y^4, \\ \widehat{G}\widehat{H} &= (y^2 + y^5 + y^8)(y^1 - y^4) &&= y^3 - y^{12}, \text{ and} \\ \widehat{GH} &= y^3 - y^{4(3)} &&= y^3 - y^{12}. \end{aligned} \quad (8.8)$$

Hence, $\widehat{G}\widehat{H} = \widehat{GH}$.

Now consider $f = u_1^3 + u_2^2$, which is irreducible in $\mathbb{Q}[u_1, u_2]$. Then

$$\begin{aligned} \widehat{f} &= y^3 + y^8 \\ &= y^3(1 + y^5), \text{ or} \\ &= y^2(y + y^6), \text{ or} \\ &= y(y^2 + y^7). \end{aligned} \quad (8.9)$$

From the first equation, letting $\widehat{g} = y^3, \widehat{h} = 1 + y^5$, we obtain $g = u_1^3$ and $h = 1 + u_1u_2$. It follows that $gh = u_1^3 + u_1^4u_2$ which is not an element of P_4 . From the second equation, letting $\widehat{g} = y^2, \widehat{h} = y + y^6$, we obtain $g = u_1^2$ and $h = u_1 + u_1^2u_2$.

It follows that $gh = u_1^3 + u_1^4 u_2$ which is not an element of P_4 . From the third equation, letting $\widehat{g} = y$, $\widehat{h} = y^2 + y^7$, we obtain $g = u_1$ and $h = u_1^2 + u_1^3 u_2$. It follows that $gh = u_1^3 + u_1^4 u_2$ which is not an element of P_4 . Therefore, any nontrivial factorization in K_4 $\widehat{f} = \widehat{g}\widehat{h}$ with $g, h \in P_4$ forces $gh \notin P_4$.

Proof. (i) The map $\widehat{}$ is a bijection because every positive integer less than $d^n - 1$ can be written uniquely as an n -digit number in base d , given leading zeros.

(ii) Let $u^{(a_1, a_2, \dots, a_n)} = u_1^{a_1} u_2^{a_2} \cdots u_n^{a_n}$, where $(a_1, a_2, \dots, a_n) \in \mathbb{Z}_{\geq 0}^n$. Let

$$\begin{aligned} G &= \sum_{\ell=1}^N q_\ell u^{(i_{1\ell}, i_{2\ell}, \dots, i_{n\ell})}, \\ H &= \sum_{m=1}^M r_m u^{(j_{1m}, j_{2m}, \dots, j_{nm})}, \\ GH &= \sum_{\ell=1}^N \sum_{m=1}^M q_\ell r_m u^{(i_{1\ell}+j_{1m}, i_{2\ell}+j_{2m}, \dots, i_{n\ell}+j_{nm})} \end{aligned} \quad (8.10)$$

where $q_\ell, r_\ell \in \mathbb{Q}[u_0]$ and such that $i_{k\ell}, j_{km} < d$ for $1 \leq k \leq n$, $1 \leq \ell \leq N$, and $1 \leq m \leq M$. Moreover,

$$\begin{aligned} \widehat{G} &= \sum_{\ell=1}^N q_\ell y^{i_{1\ell} + di_{2\ell} + \cdots + d^{n-1}i_{n\ell}}, \\ \widehat{H} &= \sum_{m=1}^M r_m y^{j_{1m} + dj_{2m} + \cdots + d^{n-1}j_{nm}}, \\ \widehat{GH} &= \sum_{\ell=1}^N \sum_{m=1}^M q_\ell r_m y^{((q_{1\ell}+r_{1m})+d(q_{2\ell}+r_{2m})+\cdots+d^{n-1}(q_{n\ell}+r_{nm}))}. \end{aligned} \quad (8.11)$$

Let $i_{K_\alpha} + j_{K_{\alpha'}} = \max i_{k\ell} + j_{km}$. Suppose $i_{K_\alpha} + j_{K_{\alpha'}} \geq d$. Since $GH \in P_d$, either $q_\alpha r_{\alpha'} = 0$ or for some $1 \leq \beta \leq N$ and $1 \leq \beta' \leq M$,

$$q_\alpha r_{\alpha'} u^{(i_{1\alpha}+j_{1\alpha'}, i_{2\alpha}+j_{2\alpha'}, \dots, i_{n\alpha}+j_{n\alpha'})} = -q_\beta r_{\beta'} u^{(i_{1\beta}+j_{1\beta'}, i_{2\beta}+j_{2\beta'}, \dots, i_{n\beta}+j_{n\beta'})}. \quad (8.12)$$

If $q_\alpha r_{\alpha'} = 0$, then either $q_\alpha = 0$ or $r_{\alpha'} = 0$. It follows that $i_{K_\alpha} = 0$ or $j_{K_{\alpha'}} = 0$.

This forces $i_{K_\alpha} \geq d$ or $j_{K_{\alpha'}} \geq d$, contradicting the assumption that $G, H \in P_d$.

On the other hand, (8.12) implies that $q_\alpha r_{\alpha'} = -q_\beta r_{\beta'}$. It follows that if $i_{K_\alpha} + j_{K_{\alpha'}} \geq d$, then the coefficient of that term in GH is 0, so $GH \in P_d$. Hence, for each term with a nonzero coefficient, $i_{k_\ell} + j_{k_m} < d$ for $1 \leq k \leq n$, $1 \leq \ell \leq N$, and $1 \leq m \leq M$, $\widehat{G}\widehat{H} = \widehat{GH} \in K_d$.

If $i_{K_\alpha} + j_{K_{\alpha'}} < d$, then $i_{k_\ell} + j_{k_m} < d$ for $1 \leq k \leq n$ and $1 \leq \ell \leq N$ and $1 \leq m \leq M$, $\widehat{G}\widehat{H} = \widehat{GH} \in K_d$.

(iii) Assume f is reducible. Then $f = gh$ for some $g, h \in P_d$, both of positive degree over D , since the degree of each variable in f is the sum of the degrees of the respective variables in g and h . Then $\widehat{f} = \widehat{gh} = \widehat{g}\widehat{h}$, by (ii), is a nontrivial factorization in K_d . Hence, (b) is false, since $gd = f \in P_d$.

Assume f is irreducible. Suppose $\widehat{f} = \widehat{g}\widehat{h}$ for some $\widehat{g}, \widehat{h} \in K_d$ of positive degree over D . Suppose $g, h, gh \in P_d$. Then $\widehat{gh} = \widehat{g}\widehat{h} = \widehat{f}$. Since $\widehat{g}, \widehat{h}$ have positive degree over D in $\mathbb{Q}[u_0, \dots, u_n]$, g, h have positive degree over D , contradicting the irreducibility of f . Therefore, since $\widehat{}$ is one-to-one, any factorization $\widehat{f} = \widehat{g}\widehat{h}$ must lead back to a product $gh \notin P_d$, where $g, h \in P_d$.

□

Main Theorem (Hilbert's Irreducibility Theorem). *Let $f \in \mathbb{Q}[t_1, t_2, \dots, t_n, x]$ be irreducible over $\mathbb{Q}$. Then there exist an infinite number of n -tuples $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{Q}^n$ such that $f(\alpha_1, \alpha_2, \dots, \alpha_n, x)$ is irreducible in $\mathbb{Q}[x]$.*

Proof. Induct on n . Theorem 8.1.2 implies the case $n = 1$.

Suppose the theorem holds for $n - 1$ in $\mathbb{Q}[t_1, t_2, \dots, t_{n-1}]$. Let $u_{i-1} = t_i$ for $1 \leq i \leq n$, $u_n = x$. Then $f(t_1, t_2, \dots, t_n, x) = f(u_0, u_1, \dots, u_n)$. Let $d - 1$ be the maximum of the exponents of $u_1, u_2, \dots, u_n$ of f and let $\widehat{}$, K_d , and P_d be as in Kronecker's Criterion. By induction, it suffices to show that there exist infinitely

many $\alpha \in \mathbb{Q}$ such that $f(\alpha, u_1, u_2, \dots, u_n)$ is irreducible over $\mathbb{Q}$, which can be broken into two cases: $\widehat{f}(u_0, y)$ is either reducible or irreducible in $\mathbb{Q}[u_0, y]$.

Suppose $\widehat{f}(u_0, y)$ is reducible in $\mathbb{Q}[u_0, y]$. Then $\widehat{f}(u_0, y)$ is a product of N irreducible polynomials

$$\widehat{f}(u_0, y) = \prod_{i=1}^N g_i(u_0, y), \quad (8.13)$$

for some $2 \leq N < n$, where $g_i(u_0, y)$ is irreducible in $\mathbb{Q}[u_0, y]$. Let

$$A = \{\alpha \in \mathbb{Q} \mid g_i(\alpha, y) \text{ is irreducible in } \mathbb{Q}[y] \text{ for each } 1 \leq i \leq N\}, \quad (8.14)$$

and let $\alpha \in A$. Suppose

$$\widehat{f}(\alpha, y) = \left(\prod_{i \in S} g_i(\alpha, y) \right) \left(\prod_{i \notin S} g_i(\alpha, y) \right), \quad (8.15)$$

where $\emptyset \neq S \subset [N]$. Since $\prod_{i \in S} g_i(\alpha, y), \prod_{i \notin S} g_i(\alpha, y) \in K_d$, by Lemma 8.1.3, there exist $q, r \in P_d$ such that

$$\begin{aligned} \widehat{f} &= \widehat{q}\widehat{r} \\ \widehat{q}(\alpha, y) &= \prod_{i \in S} g_i(\alpha, y) \\ \widehat{r}(\alpha, y) &= \prod_{i \notin S} g_i(\alpha, y). \end{aligned} \quad (8.16)$$

The polynomial $f(u_0, u_1, \dots, u_n)$ is irreducible over $\mathbb{Q}$, so by Lemma 8.1.3

$$q(u_0, u_1, \dots, u_n)r(u_0, u_1, \dots, u_n) \notin P_d. \quad (8.17)$$

Hence, any factorization of $f(\alpha, u_1, u_2, \dots, u_n)$ must occur from some $q, r \in P_d$ and $\alpha \in A$. Because there are a finite number of $\emptyset \neq S \subset [N]$, the factorization $\widehat{f} = \widehat{q}\widehat{r}$ leads back to $qr \notin P_d$, where $q, r \in P_d$, only a finite number of times when $\widehat{f}(\alpha, y)$ is reducible.

By Theorem 8.1.2, A is an infinite set. Let M be the maximum of all of the exponents of $u_1, u_2, \dots, u_n$ in qr in (8.17), u_{m_S} the variable having this exponent,

a_{m_S} the coefficient of u_{m_S} . Then $a_{m_S} \in \mathbb{Q}[u_0]$. If $a_{m_S}(\alpha) \neq 0$, then

$$q(\alpha, u_1, u_2, \dots, u_n)r(\alpha, u_1, u_2, \dots, u_n) \notin P_d, \quad (8.18)$$

and $q(\alpha, u_1, u_2, \dots, u_n)r(\alpha, u_1, u_2, \dots, u_n) \neq f(\alpha, u_1, u_2, \dots, u_n)$.

Let $A_{a_{m_S}} = \{\alpha \in \mathbb{Q} \mid a_{m_S}(\alpha) = 0\}$. By Lemma 8.1.3,

$$\mathcal{A} = A \setminus \left(\bigcup_{\emptyset \neq S \subseteq [n]} A_{a_{m_S}} \right) \quad (8.19)$$

is the set of all $\alpha \in \mathbb{Q}$ such that $f(\alpha, u_1, \dots, u_n)$ is irreducible over $\mathbb{Q}$. Since

$\bigcup_{\emptyset \neq S \subseteq [n]} A_{a_{m_S}}$ is a finite set, $\mathcal{A}$ is an infinite set.

If $\widehat{f}(u_0, y)$ is irreducible in $\mathbb{Q}[u_0, y]$, then there are an infinite number of $\alpha \in \mathbb{Q}$ such that $\widehat{f}(\alpha, y)$ is irreducible in $\mathbb{Q}[y]$, by Theorem 8.1.2. By Kronecker's Criterion, $f(\alpha, u_1, u_2, \dots, u_n)$ is irreducible over $\mathbb{Q}$ for each such α . $\square$

8.2 The Symmetric Group as a Galois Group Over $\mathbb{Q}$

We follow Hadlock [Had78] in this section.

For this section, we will let $s_1, s_2, \dots, s_n$ be variables for $n \in \mathbb{Z}_{>0}$ and t_i be $\pm$ the i th symmetric function on $s_1, s_2, \dots, s_n$. We will construct a polynomial with roots $s_1, s_2, \dots, s_n$ so that the coefficients of that polynomial are the t_i .

Theorem 8.2.1. *Let $n \in \mathbb{Z}_{>0}$ be fixed,*

$$t_i = (-1)^i \sigma_i(s_1, s_2, \dots, s_n) \in \mathbb{Z}[s_1, s_2, \dots, s_n], \quad (8.20)$$

where σ_i is the i th symmetric function on the n variables s_i , $i = 1, 2, \dots, n$,

$\alpha = \sum_{j=1}^n m_j s_j$ where the m_j are as in Lemma 5.2.8. Let

$F(x), G(x) \in \mathbb{Q}[s_1, s_2, \dots, s_n, x]$ such that

$$\begin{aligned} F_s(s_1, s_2, \dots, s_n, x) &= \prod_{i=1}^n (x - s_i), \\ F_t(t_1, t_2, \dots, t_n, x) &= x^n + t_1 x^{n-1} + \dots + t_n, \\ G(t_1, t_2, \dots, t_n, x) &= \prod_{\tau \in S_n} (x - \tau(\alpha)), \end{aligned} \tag{8.21}$$

where $\tau \in S_n$ acts on the s_i on the left. Then F_t and G are irreducible in $\mathbb{Q}[t_1, t_2, \dots, t_n, x]$.

Proof. By Lemma 5.1.4, $F_s = F_t$. Let $a_1, a_2, \dots, a_n \in \mathbb{C}$ be algebraically independent over $\mathbb{Q}$, $b_i = (-1)^i \sigma_i(a_1, a_2, \dots, a_n)$ for $i = 1, 2, \dots, n$. Suppose $F_t = RS$ for some $R, S \in \mathbb{Q}[t_1, t_2, \dots, t_n, x]$, with $\deg_x R, \deg_x S > 0$. Then

$$F_t(b_1, b_2, \dots, b_n, x) = R(b_1, b_2, \dots, b_n, x)S(b_1, b_2, \dots, b_n, x). \tag{8.22}$$

This contradicts Theorem 5.2.7, so F is irreducible.

Suppose $G = rs$ for some $r, s \in \mathbb{Q}[t_1, t_2, \dots, t_n, x]$, with $\deg_x r, \deg_x s > 0$.

Then

$$G(b_1, b_2, \dots, b_n, x) = r(b_1, b_2, \dots, b_n, x)s(b_1, b_2, \dots, b_n, x). \tag{8.23}$$

This contradicts Theorem 5.2.9, so G is irreducible. □

Theorem 8.2.2. For all $n \in \mathbb{Z}_{>0}$, there exists $f(x) \in \mathbb{Q}[x]$ with splitting field E such that $\text{Gal}(E/\mathbb{Q}) \cong S_n$.

Proof. Fix $n \in \mathbb{Z}_{>0}$ and let $s_i, t_i, \sigma_i, m_i, \alpha, F_t, G$ be as in Theorem 8.2.1. By Theorem 8.2.1, F_t, G are irreducible in $\mathbb{Q}[t_1, t_2, \dots, t_n, x]$.

By Hilbert's Irreducibility Theorem, choose $\beta_1, \beta_2, \dots, \beta_n \in \mathbb{Q}$ such that $\overline{G}(x) = G(\beta_1, \beta_2, \dots, \beta_n, x)$ is irreducible over $\mathbb{Q}$, and let

$\overline{F}(x) = F_t(\beta_1, \beta_2, \dots, \beta_n, x) \in \mathbb{Q}[x]$. Call the zeros of $\overline{F}$ $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$. By Lemma 5.1.4, $\beta_i = (-1)^i \sigma_i(\alpha_1, \alpha_2, \dots, \alpha_n)$. Now,

$$\overline{\alpha} = \sum_{i=1}^n m_i \alpha_i \in E = \mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n), \quad (8.24)$$

the splitting field for $\overline{F}$ over $\mathbb{Q}$. Hence, $\mathbb{Q}(\overline{\alpha}) \subseteq \mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n)$. By Theorem 5.2.9,

$$\overline{G}(x) = \prod_{\tau \in S_n} (x - \tau(\overline{\alpha})) \in \mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n). \quad (8.25)$$

It follows that $\overline{G}(x)$ splits over $\mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n)$, and $\mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n) \subseteq \mathbb{Q}(\overline{\alpha})$.

Therefore,

$$\mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n) = \mathbb{Q}(\overline{\alpha}) = E. \quad (8.26)$$

Because G is irreducible over $\mathbb{Q}$ and $\deg G = n!$, $[E : \mathbb{Q}] \geq n!$. By Theorem 5.2.7, the splitting field for $F_t(x)$ over $\mathbb{Q}(t_1, t_2, \dots, t_n)$ is $\mathbb{Q}(s_1, s_2, \dots, s_n)$. By Theorem 5.2.9, the splitting field for $G(x)$ over $\mathbb{Q}(t_1, t_2, \dots, t_n)$ is $\mathbb{Q}(s_1, s_2, \dots, s_n)$. It follows that F_t and G have the same splitting field over $\mathbb{Q}(t_1, t_2, \dots, t_n)$, so considering $\text{Gal}(E/\mathbb{Q})$ as the Galois group for $\overline{F}$ over $\mathbb{Q}$, $|\text{Gal}(E/\mathbb{Q})| \geq n!$. On the other hand, $\deg \overline{F} = n$, so $\text{Gal}(E/\mathbb{Q}) \leq S_n$ by Theorem 3.1.5. Therefore, $|\text{Gal}(E/\mathbb{Q})| = n!$ and $\text{Gal}(E/\mathbb{Q}) = S_n$. □

BIBLIOGRAPHY

- [Art91] M. Artin, *Algebra*, Prentice-Hall, 1991.
- [Gal10] Joseph A. Gallian, *Contemporary abstract algebra*, 7 ed., Brooks/Cole, Cengage Learning, 2010.
- [Had78] Charles Robert Hadlock, *Field theory and its classical problems*, The Mathematical Association of America, Inc., 1978.
- [Jac85] Nathan Jacobson, *Basic algebra I*, 2 ed., Dover Publications, Inc., 1985.
- [Jac89] ———, *Basic algebra II*, 2 ed., Dover Publications, Inc., 1989.
- [Mun75] James Raymond Munkres, *Topology: a first course*, Prentice-Hall, Inc., 1975.
- [Rud76] Walter Rudin, *Principles of mathematical analysis*, 3 ed., McGraw-Hill, Inc., 1976.
- [Sni03] E.B. Saff & A.D. Snider, *Fundamentals of complex analysis with applications to engineering and science*, 3 ed., Pearson Education, Inc., 2003.
- [Str03] Gilbert Strang, *Introduction to linear algebra*, 3 ed., Wellesley-Cambridge Press, 2003.
- [Wil94] Herbert S. Wilf, *Generatingfunctionology*, 2 ed., Academic Press, Inc., 1994.