
GROUP, RING, AND FIELD THEORY: A COMPLETE GUIDE

ABSTRACT ALGEBRA

AUGUST 24, 2021

ELIJAH THOMPSON,
PHYSICS AND MATH HONORS

Solo Pursuit of Learning



Contents

I	Group Theory	5
1	§§Basic Definitions and Examples: Groups	6
1.1	§Initial Definitions	6
1.2	§The Group of Symmetries	7
	§Notation: Cycles	8
1.3	§General Properties	9
2	§§Group Homomorphisms	12
2.1	§Basic Definitions and Examples: Group Homomorphisms	12
	§Group Isomorphisms	15
2.2	§Automorphisms	17
3	§§Subgroups	19
3.1	§Basic Definitions and Examples: Subgroups	19
	§Center	20
3.2	§Cyclic Subgroups	21
3.3	§Dihedral Groups	24
3.4	§Lattice Subgroups of a Group	28
4	§§Free Groups	31
4.1	§Basic Definitions and Examples: Free Groups	31
4.2	§Presentations	33
5	§§Quotient Groups	35
5.1	§Cosets	35
5.2	§Lagrange's Theorem and Applications	37
	§Classification of Groups of Order $2p$ for p a prime	39
5.3	§The Alternating Group	40
5.4	§The Quotient Group Definition and Construction	41
5.5	§Isomorphism Theorems and Correspondence	44
6	§§Group Actions	48
6.1	§Basic Definitions and Examples: Group Actions	48
	§Application to Cycle Decompositions	51
6.2	§Counting and Combinatorial Formulas	52
6.3	§Conjugacy Actions and Actions on Subgroups	55
	§Conjugation in Special groups	57
	§Right Group Actions	57
6.4	§P-Groups	58
6.5	§Sylow's Theorem	59

Applications of Sylow's Theorem	63
7 §§Product Groups	67
7.1 §Basic Definitions and Examples: Product Groups	67
7.2 §Semi-Direct Products	70
Classifications of Certain Finite Groups	74
8 §§Nilpotent and Solvable Groups	76
8.1 § p -Groups	76
8.2 §Nilpotent Groups	77
8.3 §Composition Series and Solvable Groups	79
The Hölder Program	80
II Ring Theory	82
9 §§Basic Definitions and Examples: Rings	83
9.1 §Initial Definitions and Examples	83
§Integral Domains	86
§Subrings	87
9.2 §Ring Homomorphisms	89
9.3 §Domains and Fields of Fractions	91
9.4 §Special Definitions and Facts	94
9.5 §The Gaussian Integers	95
10 §§Ideals and Quotient Rings	97
10.1 §Basic Definitions and Examples: Ideals	97
§Simple Ideals	99
§Maximal and Prime Ideals	100
10.2 §Ideal Arithmetic and the Chinese Remainder Theorem	101
10.3 §Adjunctions	105
10.4 §Isomorphism Theorems and Correspondence	105
11 §§Adjunction of Elements	107
12 §§Unique Factorization Domains	108
12.1 §Basic Definitions and Examples: UFDs	108
12.2 §Unique Factorization in $F[x]$	108
13 §§Principal Ideal Domains	110
13.1 §Basic Definitions and Examples: PIDs	110
14 §§Euclidean Domains	111
14.1 §Basic Definitions and Examples: Euclidean Domains	111
15 §§Polynomial Rings	114
15.1 §Basic Definitions and Examples: Polynomial Rings	114
15.2 §Division Algorithm	118
§Solutions to Polynomials	121
15.3 §Substitution Principle	121

15.4	§Roots and Factorization	123
	§Polynomials over \mathbb{Q} and \mathbb{Z}	125
	§Parallels between the Integers and Polynomials over a Field	129
15.5	§Polynomials over a Field	129
	§Field Extensions	131
15.6	§GCD of Polynomials	132
III	Field Theory	134
16	§§Basic Definitions and Examples: Fields	135
17	§§Field Extensions	137
17.1	§Initial Definitions and Examples	137
17.2	§Algebraic Extensions	140
17.3	§Classical Straightedge and Compass Constructions	146
17.4	§Splitting Fields	148
	Cyclotomic Fields	150
17.5	§Separable and Inseparable Extensions	151
17.6	Cyclotomic Polynomials and Extensions	155
18	§§Galois Theory	157
18.1	§Basics Definitions and Examples: Galois Theory	157
	The Galois Correspondence	159
18.2	§The Fundamental Theorem of Galois Theory	162
18.3	§Finite Fields	169
18.4	§Composite Extensions and Simple Extensions	170
18.5	§Cyclotomic Extensions and Abelian Extensions	171
18.6	§Galois Group of Polynomials	172
	Polynomials of Degree 2	175
	Polynomials of Degree 3	175
	Polynomials of Degree 4	176
18.7	§Solvable and Radical Extensions	177
IV	Modules	181
19	§§General Definitions and Examples	182
19.1	§Basic Definitions and Examples: Modules	182
19.2	§Module Homomorphisms	186
	§Isomorphism Theorems for Modules	190
	§Evaluation Bijections	192
19.3	§Submodules	193
19.4	§Free Modules and Generators	193
20	§§Linear Transformations	197
21	§§Matrix Theory for Free Modules	198
22	§§Modules over PIDs	199

23 §§Tensor Products	200
23.1 §Module Tensor Products	200
Motivation/Special Case	200
General Construction	203
Appendices	211
.1 §Semi-Groups and Monoids	212

Part I

Group Theory

Chapter 1

§§Basic Definitions and Examples: Groups

1.1.0 §Initial Definitions

Definition 1.1.1. A binary operation on a non-empty set S is a map

$$\beta : S \times S \rightarrow S$$

where $S \times S := \{(a, b) : a, b \in S\}$, and $(a, b) \mapsto a * b = \beta(a, b)$.

Example 1.1.1.

1. $\begin{array}{l} S \times S \rightarrow S \\ (a, b) \mapsto a \end{array}$ known as projection to the first factor
2. $\begin{array}{l} \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z} \\ (a, b) \mapsto 0 = a * b \end{array}$ known as the zero map
3. $\begin{array}{l} \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z} \\ (a, b) \mapsto a + b \end{array}$ addition

Remark 1.1.1. There can be many different binary operations on a given set.

Definition 1.1.2 (Group). A pair (G, \star) where G is a set and \star is a binary operation on G is called a group if:

G1. (Associativity) For all $a, b, c \in G$,

$$(a \star b) \star c = a \star (b \star c)$$

G2. (Identity Element) There exists $e \in G$ such that for all $a \in G$

$$a \star e = e \star a = a$$

G3. (**Inverses**) For all $a \in G$ there exists $b \in G$ such that

$$a \star b = b \star a = e$$

In this case we write $b = a^{-1}$

Remark 1.1.2. The operation \star can be denoted in many ways: \cdot , $+$, juxtaposition. The identity element e is sometimes denoted 1_G , 1 , e_G , and 0 .

Example 1.1.2.

1. $(\mathbb{Z}, +)$ is a group with $e = 0$ and the inverse of a is denoted $-a$
2. $(\mathbb{R}_{>0}, \cdot)$ where \cdot is multiplication. \cdot is a binary operation on $\mathbb{R}_{>0}$ because for all $a, b \in \mathbb{R}_{>0}$ we have $a, b > 0$ so $a \cdot b > 0$ and $a \cdot b \in \mathbb{R}_{>0}$. Moreover, $(\mathbb{R}_{>0}, \cdot)$ is a group with identity 1.
3. (Non-example) $(\mathbb{Z}, -)$, $(a, b) \mapsto a - b$. This is not a group. Indeed, although $-$ is a binary operation on \mathbb{Z} , there is no identity element and it's not associative.
4. (Non-example) $(\mathbb{Z} \setminus \{0\}, \cdot)$ is associative and has identity 1, but does not have inverses for all $a \in \mathbb{Z}$.

Definition 1.1.3. A group (G, \star) is called **abelian** if $a \star b = b \star a$ for all $a, b \in G$.

Remark 1.1.3. Abelian groups are also known as **commutative groups**.

Example 1.1.3. $\text{GL}_n(\mathbb{R})$ is the general linear group of dimension $n \geq 1$, defined by

$$\text{GL}_n(\mathbb{R}) := \left(\{A \in M_{n \times n}(\mathbb{R}) : \det(A) \neq 0\}, \underbrace{\circ}_{\text{matrix product}} \right) \quad (1.1.1)$$

Exercise 1.1.4. $\text{GL}_n(\mathbb{R})$ is a group with identity $I_n = (\delta_{ij})$, but it is non-abelian for $n \geq 2$.

Exercise 1.1.5. If (G, \star) is a group, then (G, \star') is also group with

$$a \star' b := b \star a, \forall a, b \in G$$

Proof. Note that for all $a, b \in G$, $a \star' b = b \star a \in G$ since \star is a binary operation on G by assumption, so \star' is also a binary operation on G . (cont.) ■

1.2.0 §The Group of Symmetries

Definition 1.2.1 (Symmetric Group). Let X be a non-empty set. Then, define

$$S_X := \{\sigma : X \rightarrow X : \sigma \text{ is a bijection}\} \quad (1.2.1)$$

Such a σ is called a **permutation** of X . It follows that (S_X, \circ) is a group where

$$\begin{aligned} \circ : S_X \times S_X &\rightarrow S_X \\ (\sigma, \tau) &\mapsto \sigma \circ \tau \end{aligned} \tag{1.2.2}$$

is function composition. The group is also commonly denoted as $\text{Sym}(X)$.

Proof. Let X be a non-empty set, and define S_X as above. (cont.) ■

Definition 1.2.2. If $X = \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$, then $S_X = S_{\{1, 2, \dots, n\}}$ is denoted S_n and is called the symmetric group of degree n or symmetric group on n letters.

Example 1.2.1. Take $n = 3$: $\sigma \in S_3$ can be represented as a $2 \times n$ matrix by

$$\begin{pmatrix} 1 & 2 & 3 \\ \sigma(1) & \sigma(2) & \sigma(3) \end{pmatrix}$$

where $\sigma(1) \sigma(2) \sigma(3)$ is a permutation of 1 2 3.

Example 1.2.2. $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \gamma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, then

$$\sigma \circ \gamma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

Observe $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \sigma$ and $\gamma^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

The identity permutation is denoted $\text{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$

Note 1.2.1. We can also have written

$$\begin{pmatrix} 2 & 1 & 3 \\ \sigma(2) & \sigma(1) & \sigma(3) \end{pmatrix}$$

instead

Example 1.2.3. $n = 2$: $S_2 = \{\text{Id}, \tau\}$, where

$$\text{Id} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \tau = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \tau^2 = \text{Id}$$

§Notation: Cycles

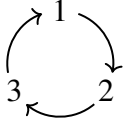
The cycle notation is more compact than the matrix-type notation, although this does come with some ambiguity:

Example 1.2.4.

1. $(1\ 2) \in S_3$ means $1 \mapsto 2, 2 \mapsto 1$, and $3 \mapsto 3$ (a transposition).

↳ Note this is the same as $(2\ 1)$ which is where ambiguity can arise.

2. $(1\ 2\ 3) \in S_3$, means $1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1$

↳ Visual -  (Note $(1\ 2\ 3) = (2\ 3\ 1) = (3\ 1\ 2)$)

Definition 1.2.3. Let $k_1, k_2, \dots, k_r \in \{1, 2, \dots, n\}$ be distinct ($r \leq n$). Then the permutation in S_n which sends $k_1 \mapsto k_2 \mapsto k_3 \mapsto \dots \mapsto k_r \mapsto k_1$ and fixes all other numbers is denoted

$$(k_1\ k_2\ \dots\ k_r) \quad (1.2.3)$$

which is a cycle of length r or an r -cycle.

Remark 1.2.2. The only 1-cycle is the identity permutation in S_n

$$\mapsto (1) = (2) = \dots = (n)$$

Definition 1.2.4. Cycles of length ≥ 2 in S_n are called disjoint if they move disjoint sets of numbers

↳

Example 1.2.5. $(1\ 2), (3\ 4)$ are disjoint in S_4 , but $(1\ 2), (3\ 1)$ are not.

Remark 1.2.3. Every permutation in $S_n (\neq \text{Id})$, can be written as a product of disjoint cycles of length ≥ 2 . Moreover, this factorization is unique up to the order of the factors.

Proof. (Left to the reader) ■

Example 1.2.6. In S_5 $(1\ 4\ 5) \circ (2\ 3) = (2\ 3) \circ (1\ 4\ 5)$ as they are disjoint

↳ (so we can't hope for full unicity of the factorization)

1.3.0 §General Properties

Definition 1.3.1. For a group (G, \star) , the number of elements (cardinal) of G is denoted $|G|$ and called the order of the group (can be infinite).

Example 1.3.1. 1. $(\mathbb{Z}, +)$ has infinite order

2. (S_n, \circ) has order $n!$ (permutations)

Proposition 1.3.1. *Let (G, \star) be a group. Then we have the following properties:*

1. *The identity element is unique.*

Proof. If $e, i \in G$ are identity elements, then $e = e \star i = i$ as $e \star g = e$ and $g \star i = g$ for all $g \in G$ by assumption. ■

2. *The inverse of an element is unique.*

Lemma 1.3.2 (Cancellation Lemma). *If $a \star b = a \star c$ or $b \star a = c \star a$, then $b = c$ for all $a, b, c \in G$.*

Proof. Let a^{-1} be an inverse of a . Then

$$b = e \star b = (a^{-1} \star a) \star b = a^{-1} \star a \star c = e \star c = c$$

and similarly for $b \star a = c \star a$. ■

Remark 1.3.1. $a \star b = c \star a$ does not tell us anything in general.

Proof. Take b, c , inverses of a . Then $b \star a = e = c \star a$, so by the cancellation lemma $b = c$. ■

↳ We will denote the inverse of a by a^{-1} for all $a \in G$.

Corollary 1.3.3. *For all $a \in G$, if $b \star a = e$ (or $a \star b = e$) the identity element, then we have that $b = a^{-1}$, so b is the inverse of a .*

Definition 1.3.2. *Let (G, \star) be a group. Let $n \in \mathbb{Z}$, $g \in G$, then*

$$g^n = \begin{cases} e, & \text{if } n = 0 \\ \underbrace{g \star g \star \dots \star g}_{n\text{-fold times}}, & \text{if } n > 0 \\ \underbrace{g^{-1} \star g^{-1} \star \dots \star g^{-1}}_{-n\text{-fold times}}, & \text{if } n < 0 \end{cases} \quad (1.3.1)$$

Proposition 1.3.4. *For all $g \in G$ and all $n, m \in \mathbb{Z}$,*

1. $(g^n)^m = g^{nm}$
2. $g^n \star g^m = g^{n+m}$

Proof. (Left to the reader) ■

Example 1.3.2. Let $G = \mathbb{Z}/2\mathbb{Z}$, $g = [1]$, and the operation be $+$. Then $g^{-1} = -1 \cdot g = [-1] = [1]$, $g^0 = 0 \cdot g = [0]$, $g^1 = 1 \cdot g = [1]$, $g^2 = 2 \cdot g = [1] + [1] = [2] = [0]$, and $g^3 = 3 \cdot g = [3] = [1]$, etc.

Remark 1.3.2. Due to associativity, the placement of parenthesis is unambiguous and unnecessary:

$$\mapsto ((a \star (b \star c)) \star d) = ((a \star b) \star (c \star d))$$

Thus, $g_1 \star g_2 \star \dots \star g_n$ is well-defined for all $n \in \mathbb{N}$.

Proof. (Left to the reader) ■

Note 1.3.3. However, because we don't necessarily have commutivity, it is not true in general that $(g \star h)^n = g^n \star h^n$. Indeed, $(g \star h)^2 = g \star h \star g \star h$, not necessarily $g^2 \star h^2$.

Remark 1.3.4 (Inverse of a product). Let $a, b \in G$, a group. Then

$$1. (a \star b)^{-1} = b^{-1} \star a^{-1}$$

2. More generally

$$(a_1 \star a_2 \star \dots \star a_n)^{-1} = a_n^{-1} \star \dots \star a_2^{-1} \star a_1^{-1} \tag{1.3.2}$$

for $a_i \in G, 1 \leq i \leq n$

Proof. (Left to the reader) ■

Example 1.3.3. In $(\mathbb{Z}, +)$, g^2 for $g = 3$ is $g^2 = 2 \cdot g = 2 \cdot 3 = 6$. In general $g^n = ng$ for additive groups.

Chapter 2

§§Group Homomorphisms

2.1.0 §Basic Definitions and Examples: Group Homomorphisms

Definition 2.1.1 (A). A group homomorphism from a group A to a group B is a map of sets $\phi : A \rightarrow B$ that satisfies the condition:

$$\forall a_1, a_2 \in A, \quad \phi(a_1 \cdot a_2) = \phi(a_1) \cdot \phi(a_2) \quad (2.1.1)$$

We can rewrite the condition on ϕ in terms of the requirement that the following diagram commute:

$$\begin{array}{ccc} A \times A & \xrightarrow{\text{mult}_A} & A \\ \phi \times \phi \downarrow & & \downarrow \phi \\ B \times B & \xrightarrow{\text{mult}_B} & B \end{array}$$

We denote the set of all group homomorphisms from A to B by $\mathbf{Hom}_{\mathbf{Grp}}(A, B)$.

Definition 2.1.2 (B). Let (G, \star) and (M, \circ) be groups. A map $G \xrightarrow{f} M$ is a homomorphism of groups if

$$f(x \star y) = f(x) \circ f(y), \forall x, y \in G \quad (2.1.2)$$

Example 2.1.1.

1. $H \leq G$, then $H \xrightarrow{\iota} G$ the inclusion map is a monomorphism (an injective homomorphism)
2. $(\mathbb{R}, +) \rightarrow \mathbb{R}^\times$
 $x \mapsto 2^x$ is a monomorphism.

3. $S_3 \hookrightarrow S_4$, where $\sigma'(i) = \sigma(i)$ for $i \in \{1, 2, 3\}$ and $\sigma'(4) = 4$. In particular $\sigma \mapsto \sigma'$

$$\begin{pmatrix} 1 & 2 & 3 \\ \sigma(1) & \sigma(2) & \sigma(3) \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 \\ \sigma'(1) & \sigma'(2) & \sigma'(3) & 4 \end{pmatrix}$$

is a monomorphism.

4. $\det : \mathbf{GL}_n(\mathbb{R}) \rightarrow \mathbb{R}^\times$
 $A \mapsto \det(A)$ is an epimorphism (surjective homomorphism)

5. (G, \star) a group, then for $g \in G$,

$$\begin{aligned} \mathbb{Z} &\xrightarrow{\phi_g} G \\ x &\mapsto g^x \end{aligned}$$

is a group homomorphism. It needs not be injective or surjective.

6. $|\cdot| : \mathbb{C}^\times \rightarrow \mathbb{R}^\times$
 $z \mapsto |z|$ is an epimorphism.

Properties 2.1.3. Let $f : (G, \star) \rightarrow (H, \circ)$ be a group homomorphism. Then

1. $f(e_G) = e_H$
2. for all $g \in G$, $f(g^{-1}) = (f(g))^{-1}$
3. The composition of two homomorphisms $G \xrightarrow{f} H \xrightarrow{\phi} K$ is a homomorphism $G \xrightarrow{\phi \circ f} K$.

Proof. (Left to the reader) ■

Remark 2.1.1.

1. If $H_1 \leq G \xrightarrow{f} K$ where f is a homomorphism, then the canonical map

$$f|_{H_1} : H_1 \rightarrow K \tag{2.1.3}$$

is a homomorphism called the restriction of f to H_1 . Moreover, $f|_{H_1} = f \circ \iota_{H_1}$, where $\iota_{H_1} : H_1 \hookrightarrow G$ is the inclusion homomorphism.

2. Let $f : G \rightarrow K$ be a homomorphism. If the image of f , $\text{Im}(f)$, is contained in a subgroup $H_2 \leq K$, then the associated map

$$f' : G \rightarrow H_2 \tag{2.1.4}$$

is a homomorphism of groups.

Example 2.1.2. For $\det : \mathbf{GL}_n(\mathbb{R}) \rightarrow \mathbb{R}^\times$, we have $\mathbf{SL}_n(\mathbb{R}) \leq \mathbf{GL}_n(\mathbb{R})$, so $\mathbf{SL}_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times$ is a homomorphism, and

$$\text{Im}(\det|_{\mathbf{SL}_n(\mathbb{R})}) = \{1\} \leq \mathbb{R}^\times \tag{2.1.5}$$

so

$$\det : \mathbf{SL}_2(\mathbb{R}) \rightarrow \{1\} \tag{2.1.6}$$

is a homomorphism.

Proposition 2.1.1. Let $G \xrightarrow{f} K$ be a group homomorphism. Then

1. Let $H_1 \leq G$, then the image $f(H_1) \leq K$ is a subgroup of K
2. Let $H_2 \leq K$, then $f^{-1}(H_2) \leq G$ is a subgroup of G , called the inverse image of H_2 , or the pre-image of H_2 by f .

Proof. (Left to the reader) ■

Remark 2.1.2. Note that the image of a cyclic subgroup $\langle g \rangle \leq G$ under a group homomorphism $f : G \rightarrow K$ is the cyclic subgroup

$$\langle f(g) \rangle \leq K \quad (2.1.7)$$

Proof. (Left to the reader) ■

Corollary 2.1.2. Let $G \xrightarrow{f} K$ be a group homomorphism.

1. The **image** $f(G) = \text{Im}(f)$ is a subgroup of K
2. The **kernel** $\ker(f) := f^{-1}(\{e_K\})$ is a subgroup of G

Proposition 2.1.3. Let $G \xrightarrow{f} K$ be a group homomorphism. Then f is injective (i.e. a **monomorphism**) if and only if $\ker(f) = \{e_G\}$.

Proof. (Left to the reader) ■

Example 2.1.3.

1. f a homomorphism is an isomorphism if and only if $\ker(f) = \{e_G\}$ and $\text{Im}(f) = K$ (for $G \xrightarrow{f} K$)
2. $H \leq G$, $H \xhookrightarrow{\iota} G$ the inclusion map is a homomorphism with $\text{Im}(\iota) = H$ and $\ker(\iota) = \{e_G\}$
3. The determinant map $\mathbf{GL}_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times$ is a group homomorphism. Moreover, we obtain the subgroup

$$\mathbf{SL}_n(\mathbb{R}) := \{A \in \mathbf{GL}_n(\mathbb{R}) : \det(A) = 1\} = \ker(\det) \quad (2.1.8)$$
4. The map $\mathbb{Z} \xrightarrow{\phi_g} G$ for some fixed $g \in G$ is a group homomorphism with image $\text{Im } \phi_g = \langle g \rangle$, and $\ker(\phi_g) = \{n \in \mathbb{Z} : g^n = e_g\} = S_g$. Thus, ϕ_g is surjective if and only if $\langle g \rangle = G$ and ϕ_g is injective if and only if $o(g) = +\infty$.
5. The modulus map $\mathbb{C}^\times \xrightarrow{|\cdot|} \mathbb{R}^\times$ is a group homomorphism with $\ker(|\cdot|) = \{z \in \mathbb{C}^\times : |z| = 1\} = S^1$ the circle group, and $\text{Im}(|\cdot|) = \mathbb{R}_{>0}$.
6. The map $\mathbb{R} \xrightarrow{\alpha} \mathbb{C}^\times$ $\theta \mapsto \exp(2\pi i \theta)$ is a group homomorphism with $\ker(\alpha) = \mathbb{Z}$ and $\text{Im}(\alpha) = S^1$.

§Group Isomorphisms

Example 2.1.4 (motivating examples).

1. $g_1 := s_{\{1,2,3\}}$ and $g_2 := s_{\{a,b,c\}}$ are essentially the “same” group, but their elements are not the same. indeed “everything we do” in g_1 using the group operation we can do in g_2 by renaming 1 as a , 2 as b , and 3 as c : order ($|g_1| = |g_2|$), subgroups, orders of elements, “equations,” etc.
2. $\mathbb{Z}/2\mathbb{Z} = \{[0], [1]\}$, $g = \{+1, -1\} \leq (\mathbb{R} \setminus \{0\}, \cdot) = \mathbb{R}^\times$ are the “same” group. let us consider their *cayley tables*:

$(\mathbb{Z}/2\mathbb{Z}, +)$	$[0]$	$[1]$	(g, \cdot)	$+1$	-1
$[0]$	$[0]$	$[1]$	$+1$	$+1$	-1
$[1]$	$[1]$	$[0]$	-1	-1	$+1$

$[0]$ plays the role of $+1$ and $[1]$ plays the role of -1 .

Proposition 2.1.4. $\mathbb{Z}/n\mathbb{Z}$ is isomorphic to the group of n th roots of unity:

$$\{z \in \mathbb{C} : z^n = 1\} \quad (2.1.9)$$

Definition 2.1.4. Let (G, \star) and (M, \circ) be groups. A bijective map $G \xrightarrow{f} M$ which is a group homomorphism is called an **isomorphism** of the groups G and M . The groups (G, \star) and (M, \circ) are said to be **isomorphic**, denoted $(G, \star) \cong (M, \circ)$.

Example 2.1.5.

1. $\mathbb{Z} \xrightarrow{f} 2\mathbb{Z}$ is a group isomorphism.
 $n \mapsto 2n$
2. If G is a cyclic group, $G = \langle g \rangle$, and $|G| = n < +\infty$, then

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} &\xrightarrow{\phi} G \\ [m] &\mapsto g^m \end{aligned} \quad (2.1.10)$$

is a group isomorphism. Recall that $g^m = g^{m'}$ if and only if $m \equiv m' \pmod{n}$ ($o(g) = n$), so ϕ is well-defined and injective. By construction ϕ is also surjective. Therefore, $\mathbb{Z}/n\mathbb{Z} \cong G$.

3. If $G = \langle g \rangle$ and $o(g) = +\infty$, then $G \cong \mathbb{Z}$. Indeed, the map

$$\begin{aligned} \mathbb{Z} &\xrightarrow{\phi} G \\ m &\mapsto g^m \end{aligned} \quad (2.1.11)$$

is a group isomorphism.

4. The map

$$\begin{aligned} (\mathbb{R}, +) &\xrightarrow{\exp} (\mathbb{R}_{>0}, \cdot) \\ x &\mapsto e^x \end{aligned} \quad (2.1.12)$$

is a group isomorphism, so $(\mathbb{R}, +) \cong (\mathbb{R}_{>0}, \cdot)$. Another isomorphism is

$$\begin{aligned} (\mathbb{R}, +) &\xrightarrow{\phi} (\mathbb{R}_{>0}, \cdot) \\ x &\mapsto 2^x \end{aligned} \quad (2.1.13)$$

5. For all $h \in G$, where (G, \star) is an arbitrary group,

$$\begin{aligned} (G, \star) &\xrightarrow{\alpha_h} (G, \star) \\ a &\mapsto h^{-1} \star a \star h \end{aligned}$$

is a group isomorphism with inverse $\alpha_{h^{-1}}$.

6. $\begin{aligned} \mathbb{Z} &\xrightarrow{\beta} \mathbb{Z} \\ n &\mapsto -n \end{aligned}$ is a group isomorphism.

Definition 2.1.5. An isomorphism $(G, \star) \rightarrow (G, \star)$ is called an automorphism of (G, \star) .

Proposition 2.1.5. The set of automorphisms of a group G is a subgroup of the symmetric group S_G .

1. The identity map $\begin{aligned} \text{Id} : G &\rightarrow G \\ g &\mapsto g \end{aligned}$ is an isomorphism (hence an automorphism) which acts as an identity for the group
2. If $G \xrightarrow{\phi} H$ is an isomorphism, then the inverse $H \xrightarrow{\phi^{-1}} G$ is an isomorphism
3. If $G \xrightarrow{\phi} H$ and $H \xrightarrow{\psi} K$ are isomorphisms, then the composition $\psi \circ \phi$ is also an isomorphism

Proof. (Left to the reader) ■

Corollary 2.1.6. The set $\text{Aut}(G)$ of automorphisms of G is a group for the composition of maps.

Example 2.1.6 (Non-example). $\mathbb{Q} \not\cong \mathbb{Q}^\times$ because $-1 \in \mathbb{Q}^\times$ has order 2, but \mathbb{Q} does not have any element of order 2. Indeed, if $x \in \mathbb{Q}$ such that $x + x = 2x = 0$, then $x = 0$ so $o(x) = 1$. But, if $\mathbb{Q}^\times \xrightarrow{\phi} \mathbb{Q}$ is an isomorphism, then $o(\phi(-1)) = 2$, which is not possible.

Theorem 2.1.7 (Dihedral Group Isomorphisms). Let $x, y \in G$ such that $G = \langle x, y \rangle (= \langle \{x, y\} \rangle)$ with the relations $x^n = y^2 = e$, $yx = x^{-1}y$, and $n \geq 1$. Then $|G| \leq 2n$. If $|G| = 2n$, then the relation characterizes the group up to isomorphism.

Proof. If $n = 1$, $G \cong \mathbb{Z}/2\mathbb{Z} \cong D_1$. If $n = 2$, $G \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \cong D_2$, mapping $x \mapsto ([1], [0])$ and $y \mapsto ([0], [1])$. Now, for all $n \geq 1$, let $Y_n := \{e, x, x^2, \dots, x^{n-1}, y, xy, \dots, x^{n-1}y\}$. We know

that $G = Y_n$ as sets from our study of the relations on the dihedral group. But $|Y_n| = 2n$ as a set, so $|G| \leq 2n$ as a group. If $|G| = 2n$, then all elements described in Y_n are distinct in G , and the Caley table of the group is fixed by the relations. Thus, the group G is uniquely determined up to isomorphism by the relations and $|G| = 2n$. If $n \geq 3$ we have seen that $\langle x = \phi_{2\pi/n}, y = \psi_0 \rangle = D_n$, x, y satisfy the relations and $|D_n| = 2n$, so $G \rightarrow D_n$ is an isomorphism (taking x in G to x in D_n and y in G to y in D_n). ■

Corollary 2.1.8. $S_3 \cong D_3$ for $x = (1\ 2\ 3), y = (1\ 2)$.

Proof. (Left to the reader) ■

2.2.0 §Automorphisms

Definition 2.2.1. Let G be a group. An isomorphism from G onto itself is called an **automorphism** of G . The group of all automorphisms on G is denoted by $\text{Aut}(G)$.

Proposition 2.2.1. Let H be a normal subgroup of the group G . Then G acts by conjugation on H as automorphisms of H . More specifically, the action of G on H by conjugation is defined for each $g \in G$ by

$$h \mapsto ghg^{-1} \quad \text{for each } h \in H$$

For each $g \in G$, conjugation by g is an automorphism of H . The permutation representation afforded by this action is a homomorphism of G into $\text{Aut}(H)$ with kernel $C_G(H) = \{g \in G : \forall h \in H, ghg^{-1} = h\}$ (The centralizer of H with respect to G). In particular, $G/C_G(H)$ is isomorphic to a subgroup of $\text{Aut}(H)$.

Remark 2.2.1. This proposition implies that a group acts by conjugation on a normal subgroup as *structure preserving permutations*, i.e. automorphisms.

Corollary 2.2.2. If K is any subgroup of the group G and $g \in G$, then $K \cong gKg^{-1}$. Conjugate elements and conjugate subgroups have the same order (as the induced map is an automorphism).

Corollary 2.2.3. For any subgroup H of a group G , the quotient group $N_G(H)/C_G(H)$ is isomorphic to a subgroup of $\text{Aut}(H)$. In particular, $G/Z(G)$ is isomorphic to a subgroup of $\text{Aut}(G)$.

Proof. Since H is a normal subgroup of $N_G(H)$, our previous proposition implies that $N_G(H)$ acts by conjugation on H . Moreover, $C_G(H) \subseteq N_G(H)$, so the kernel of the permutation representation of $N_G(H)$ in $\text{Aut}(H)$ afforded by this action is $C_G(H)$. Hence by the first isomorphism theorem $N_G(H)/C_G(H)$ is isomorphic to a subgroup of $\text{Aut}(H)$.

The second case follows from taking $H = G$, so $N_G(G) = G$ and $C_G(G) = Z(G)$. ■

Definition 2.2.2. Let G be a group and let $g \in G$. Conjugation by g is called an **inner automorphism** of G and the subgroup of $\text{Aut}(G)$ consisting of all inner automorphisms is denoted by $\text{Inn}(G)$.

Definition 2.2.3. A subgroup H of a group G is called **characteristic in G** if and only if every automorphism of G maps H onto itself, i.e., $\sigma(H) = H$ for all $\sigma \in \text{Aut}(G)$.

Proposition 2.2.4. Let H be a subgroup of a group G :

1. If H is characteristic in G then $H \triangleleft G$,
2. If H is the unique subgroup of G of a given order, then H is characteristic in G ,
3. If K is a characteristic subgroup of H and $H \triangleleft G$, then $K \triangleleft G$.

Proof. (To be completed) ■

Corollary 2.2.5. If C is a cyclic group of order n , then every subgroup of C is characteristic in C .

Proposition 2.2.6. The automorphism group of the cyclic group of order n is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^\times$, an abelian group of order $\varphi(n)$ (where φ is the Euler-totient function).

Proof. Let x be a generator of the cyclic group $\mathbb{Z}/n\mathbb{Z}$. If $\psi \in \text{Aut}(\mathbb{Z}/n\mathbb{Z})$, then $\psi(x) = x^a$ for some $a \in \mathbb{Z}$, and the integer a uniquely determines ψ . Denote this automorphism by ψ_a . As usual, since $|x| = n$, the integer a is only defined modulo n . Since ψ_a is an automorphism, x and x^a must have the same order, hence $\gcd(a, n) = 1$. Furthermore, for a relatively prime to n , the map $x \mapsto x^a$ is an automorphism of $\mathbb{Z}/n\mathbb{Z}$. Hence we have a surjective map

$$\begin{aligned} \Psi : \text{Aut}(\mathbb{Z}/n\mathbb{Z}) &\rightarrow (\mathbb{Z}/n\mathbb{Z})^\times \\ \psi_a &\mapsto a \pmod{n} \end{aligned}$$

The map Ψ is a homomorphism because

$$\psi_a \circ \psi_b(x) = \psi_a(x^b) = x^{ab} = \psi_{ab}(x)$$

for all $\psi_a, \psi_b \in \text{Aut}(\mathbb{Z}/n\mathbb{Z})$, so that

$$\Psi(\psi_a \circ \psi_b) = \Psi(\psi_{ab}) = ab \pmod{n} = \Psi(\psi_a)\Psi(\psi_b)$$

Finally, Ψ is injective by construction of the ψ_a , and hence is an isomorphism. ■

Example 2.2.1.

1. If p is an odd prime and $n \in \mathbb{Z}^+$, then the automorphism group of the cyclic group of order p is cyclic of order $p - 1$. More generally, the automorphism group of the cyclic group of order p^n is cyclic of order $p^{n-1}(p - 1)$.
2. Let p be a prime and let V be an abelian group (written additively) with the property that $pv = 0$ for all $v \in V$. If $|V| = p^n$, then V is an n -dimensional vector space over the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. The automorphisms of V are precisely the non-singular linear transformations from V to itself, that is

$$\text{Aut}(V) \cong \text{GL}(V) \cong \text{GL}_n(\mathbb{F}_p)$$

Chapter 3

§§Subgroups

3.1.0 §Basic Definitions and Examples: Subgroups

Definition 3.1.1. A subset H of a group (G, \star) (i.e. $H \subseteq G$) is a subgroup if it satisfies the following properties:

S1 (identity) $e \in H$, where e is the identity in G (so $H \neq \emptyset$)

S2 (closure) If $h_1, h_2 \in H$, then $h_1 \star h_2 \in H$.

S3 (inverses) Of $h \in H$, then $h^{-1} \in H$.

Thus, the group $(H, \star|_H)$ makes sense. We denote this by $H \leq G$, to say H is a subgroup of G (\star is understood from context)

Example 3.1.1.

1. $\{e\} \leq G$ and $G \leq G$, for any group G . These subgroups are called the trivial subgroups of G .

2. $\mathbb{Z} \leq \mathbb{R}$ with addition

3. $l\mathbb{Z} := \{ln : n \in \mathbb{Z}\} \subset \mathbb{Z}$, where l is fixed in \mathbb{Z} . $l\mathbb{Z} \leq \mathbb{Z}$. Indeed

S1 $0 = l \cdot 0 \in l\mathbb{Z}$

S2 $ln + lm = l(n + m) \in l\mathbb{Z}$

S3 $ln + l(-n) = l(n + (-n)) = l \cdot 0 = 0$, so $-(ln) \in l\mathbb{Z}$.

4. $SL_n(\mathbb{R}) := \{A \in Mat_n(\mathbb{R}) : \det(A) = 1\} \leq GL_n(\mathbb{R})$ is a subgroup called the special linear group of degree n .

Proof. (Left to the reader) ■

5. Let $S^1 := \{z \in \mathbb{C} : |z| = 1\}$, then $S^1 \leq \mathbb{C} \setminus \{0\} = \mathbb{C}^\times$, where the operation is multiplication. S^1 is called the circle group

Proposition 3.1.1. *If $H_1 \leq G$, $H_2 \leq G$ are subgroups of G , then $H = H_1 \cap H_2 \leq G$ is a subgroup.*

Proof. (Left to the reader) ■

Corollary 3.1.2. *Let $\{H_\alpha\}_{\alpha \in J}$ be a collection of subgroups of a group G , where J is an indexing set (possibly infinite). Then*

$$\bigcap_{\alpha \in J} H_\alpha \leq G \quad (3.1.1)$$

Proof. (Left to the reader) ■

§Center

Definition 3.1.2. *Let (G, \star) be a group and let $g \in G$. The centralizer of $g \in G$ is*

$$Z(g) := \{x \in G : \underbrace{x \star g = g \star x}_{x \text{ and } g \text{ commute}}\} \quad (3.1.2)$$

Claim 3.1.3. *For all $g \in G$, for a group (G, \star) , $Z(g) \leq G$ is a subgroup of G .*

Proof. (Left to the reader) ■

Example 3.1.2.

1. $Z(2) \leq (\mathbb{Z}, +)$, and actually $Z(2) = \mathbb{Z}$ as \mathbb{Z} is abelian.
2. $Z(I_2 + E_{2,2}) \leq \mathbf{GL}_2(\mathbb{R})$, and in particular

$$Z(I_2 + E_{2,2}) = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \in \mathbf{GL}_2(\mathbb{R}) : ad \neq 0, ad \in \mathbb{R} \right\}$$

Definition 3.1.3. *The center $Z(G)$ of the group G is*

$$Z(G) := \{x \in G : x \star g = g \star x \forall g \in G\} \quad (3.1.3)$$

Proposition 3.1.4. $Z(G) = \bigcap_{g \in G} Z(g)$, so also $Z(G) \leq G$.

Exercise 3.1.3. For $n \geq 2$, $Z(\mathbf{GL}_n(\mathbb{R})) = \{aI_n : a \in \mathbb{R}^\times\}$

Remark 3.1.1. For any group G , $Z(G)$ is an abelian group.

3.2.0 §Cyclic Subgroups

Definition 3.2.1. Let $S \subset G$ a subset of a group G . Then the subgroup of G generated by S is

$$\langle S \rangle := \bigcap \{H \in \mathcal{P}(G) : S \subset H \leq G\} \quad (3.2.1)$$

(The intersection of all the subgroups of G containing S) which is the smallest subgroup of G containing S .

Note 3.2.1. $S \subset G \leq G$, so $S \subset \langle S \rangle$ and $\langle S \rangle$ is well-defined. $\langle S \rangle$ is the smallest subgroup of G containing S with respect to set inclusion.

Example 3.2.1. $\langle \emptyset \rangle = \{e\}$ and $\langle G \rangle = G$ (trivial subgroups)

Proposition 3.2.1. Let $g \in G$, then

$$\langle \{g\} \rangle = \{g^n : n \in \mathbb{Z}\} \quad (3.2.2)$$

and we write $\langle g \rangle := \langle \{g\} \rangle$.

Proof. (Left to the reader) ■

Example 3.2.2.

1. $\langle e \rangle = \{e\}$
2. For $[2] \in \mathbb{Z}/3\mathbb{Z}$, the $\langle [2] \rangle = \{[2], [1], [0]\} = \mathbb{Z}/3\mathbb{Z}$.
3. For $l \in \mathbb{Z}$, $\langle l \rangle = \{nl : n \in \mathbb{Z}\} \leq \mathbb{Z}$. We write $l\mathbb{Z} := \langle l \rangle$.

Definition 3.2.2. A group K such that there exists $g \in K$ with $\langle g \rangle = K$ is called a cyclic group.

↳ i.e) A group generated by single element is called cyclic.

Definition 3.2.3. Then, for all $g \in G$, $\langle g \rangle \leq G$ is cyclic. The order of $|\langle g \rangle|$ of the group $\langle g \rangle$ is called the order of g , and denoted $o(g)$ (could be infinite!).

Example 3.2.3.

1. For $1 \in \mathbb{Z}$, $o(1) = \infty$ given $\langle 1 \rangle = \mathbb{Z}$. Thus, \mathbb{Z} is cyclic of infinite order.
 ↳ Note: The only other generator of \mathbb{Z} is -1 .
 ↳ For $2 \in \mathbb{Z}$ does not generate \mathbb{Z} because $\langle 2 \rangle = 2\mathbb{Z} \subsetneq \mathbb{Z}$
2. For $n \in \mathbb{Z}_{>0}$, $[1]_n \in \mathbb{Z}/n\mathbb{Z}$, $o([1]_n) = n$ given

$$\langle [1]_n \rangle = \{k[1]_n : k \in \mathbb{Z}\} = \mathbb{Z}/n\mathbb{Z}$$

Thus $\mathbb{Z}/n\mathbb{Z}$ is a cyclic group of order n .

3. For any group G , $o(e) = 1$ where e is the identity of G .
4. $\mathbb{Q}^\times = (\mathbb{Q} \setminus \{0\}, \cdot)$ is not cyclic.

Proof. Assume $\mathbb{Q}^\times = \langle \frac{a}{b} \rangle$, then take p relatively prime to a and to b . Then $p \notin \{(\frac{a}{b})^n : n \in \mathbb{Z}\}$, which contradicts the assumption and \mathbb{Q}^\times is not cyclic. ■

Remark 3.2.2. A cyclic group is abelian.

Proof. (Left to the reader) ■

Corollary 3.2.2. If a group is non-abelian it cannot be cyclic.

Proof. Contrapositive of the previous statement. ■

$\hookrightarrow \mathbf{GL}_n(\mathbb{R})$, $n \geq 2$, and S_m , $m \geq 3$, are not cyclic as they are non-abelian.

\mathbb{Z} is cyclic and properties of \mathbb{Z} will have consequences for all cyclic groups (through isomorphisms).

Theorem 3.2.3. Every subgroup H of $(\mathbb{Z}, +)$ is of the form $n\mathbb{Z} = \langle n \rangle$ for some $n \in \mathbb{Z}$.

Proof. (Left to the reader) ■

Corollary 3.2.4 (GCD). Let $a, b \in \mathbb{Z}$ and define

$$a\mathbb{Z} + b\mathbb{Z} := \{an + bm : n, m \in \mathbb{Z}\} \subset \mathbb{Z} \quad (3.2.3)$$

$a\mathbb{Z} + b\mathbb{Z}$ is a subgroup of $(\mathbb{Z}, +)$ generated by $\{a, b\}$. Then, by the previous theorem $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$ for some $d \in \mathbb{Z}$, and if $(a, b) \neq (0, 0)$, we have that $d = \gcd(a, b)$. We choose $d > 0$ to maintain uniqueness.

Proof. (Left to the reader) ■

Corollary 3.2.5. Suppose $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$, $(a, b) \neq (0, 0)$, so $d \neq 0$. Then

1. $d \mid a$ and $d \mid b$.
2. If $e \mid a$ and $e \mid b$, then $e \mid d$.
3. There exist $x, y \in \mathbb{Z}$ such that $d = ax + by$.

Proposition 3.2.6. Take the cyclic subgroup $\langle g \rangle \leq G$ of G . Then define

$$S_g := \{k \in \mathbb{Z} : g^k = e\} \subset \mathbb{Z} \quad (3.2.4)$$

It follows that

1. $S_g \leq \mathbb{Z}$ is a subgroup for all $g \in G$.
2. For $r, s \in \mathbb{Z}$, $g^r = g^s$ if and only if $r - s \in S_g$.
3. If $S_g \neq \{0\}$, then $S_g = n\mathbb{Z}$ for some $n > 0$, and

$$\langle g \rangle = \{e, g, g^2, \dots, g^{n-1}\} \quad (3.2.5)$$

with $o(g) = n$.

4. $S_g = \{0\}$ if and only if $o(g) = \infty$, in which case $g^m = g^n$ if and only if $m = n$.
5. The order of g^l is $\frac{n}{\gcd(l, n)}$ if $o(g) = n$.

Proof. (Left to the reader) ■

Remark 3.2.3. If $o(g) = \infty$, $g^r = g^s$ if and only if $r = s$. If $o(g) = n < \infty$, $g^r = g^s$ if and only if $r \equiv s \pmod n$.

↳ $o(g)$ is the smallest integer $n > 0$ such that $g^n = e$. (If $\nexists n > 0$ such that $g^n = e$, then $o(g) = \infty$)

Corollary 3.2.7. For $g \in G$, g^l is a generator of $\langle g \rangle$ if and only if $\gcd(l, n) = 1$, where $n = o(g)$.

↳ If $o(g) = \infty$ then $\langle g^l \rangle = \langle g \rangle$ if and only if $l \in \{1, -1\}$

Proof. (Left to the reader) ■

Theorem 3.2.8. Every subgroup of a cyclic group is itself cyclic.

Proof. (Left to the reader) ■

Example 3.2.4.

1. For $2 \in (\mathbb{Z}, +)$, $o(2) = \infty$, so $\langle 2 \rangle = 2\mathbb{Z}$, and $|2\mathbb{Z}| = \infty$.
2. For $[2] \in \mathbb{Z}/3\mathbb{Z}$, $o([2]) = 3$, so $\langle [2] \rangle = \mathbb{Z}/3\mathbb{Z}$. Indeed, $\langle [1] \rangle = \mathbb{Z}/3\mathbb{Z}$, so

$$o([2]) = o(2[1]) = \frac{o([1])}{\gcd(o([1]), 2)} = \frac{3}{\gcd(3, 2)} = 3$$

3. For $[2] \in \mathbb{Z}/4\mathbb{Z}$, $o([2]) = 2$, so $\langle [2] \rangle < \mathbb{Z}/4\mathbb{Z}$ is a proper subgroup.

Corollary 3.2.9. Let $G = \langle g \rangle$ be cyclic. Then every subgroup of G is cyclic.

Proof. Consider the epimorphism

$$\begin{aligned} \mathbb{Z} &\xrightarrow{\phi_g} G \\ n &\mapsto g^n \end{aligned} \quad (3.2.6)$$

Let $H \subseteq G$. Then $\phi_g^{-1}(H) = H' \subseteq \mathbb{Z}$ because of the properties of the inverse images of subgroups under group homomorphisms. Thus since ϕ_g is cyclic, $\phi_g(H') = H$, so H is the image of a subgroup H' of \mathbb{Z} . But, all subgroups of \mathbb{Z} are cyclic. Thus $\phi_g(H')$ is cyclic since it is the image of a cyclic subgroup under a group homomorphism. ■

Theorem 3.2.10. *If G is a cyclic group of order $n < +\infty$, then the order of every subgroup H of G divides n . Moreover, for every divisor q of n , there exists a unique subgroup of G of order q .*

Proof. Let $G = \langle g \rangle$, $H \subseteq G$, and $o(g) = n$. By the previous corollary $H = \langle g^l \rangle$ for some $l \geq 0$. It follows that $|H| = \frac{n}{\gcd(n,l)}$. Thus,

$$\frac{n}{\gcd(n,l)} = |H| \mid |G| = n$$

Suppose $|H| = |H'| (= \langle g^{l'} \rangle)$, so $\gcd(l, n) = \gcd(l', n)$. Note that $o(g^{\gcd(l,n)}) = \frac{n}{\gcd(n, \gcd(l,n))} = \frac{n}{\gcd(n,l)} = o(g^l)$ and since $\gcd(l, n) \mid l g^l \in \langle g^{\gcd(l,n)} \rangle$. Hence, we have that

$$H = \langle g^l \rangle = \langle g^{\gcd(l,n)} \rangle = \langle g^{\gcd(l',n)} \rangle = \langle g^{l'} \rangle = H' \quad (3.2.7)$$

Thus we have uniqueness, and that the order of any subgroup must divide that of the group. For existence, suppose $n = qr$ for some integer r . Then $H = \langle g^r \rangle$ is a subgroup of order

$$|H| = o(g^r) = \frac{n}{\gcd(n,r)} = \frac{qr}{r} = q$$

satisfying existence. ■

3.3.0 §Dihedral Groups

Claim 3.3.1. *If S is a non-empty subset of G , then*

$$\langle S \rangle = \{s_1^{k_1} \star s_2^{k_2} \star \dots \star s_m^{k_m} : m \geq 1, s_1, s_2, \dots, s_m \in S, k_1, k_2, \dots, k_m \in \mathbb{Z}\} \quad (3.3.1)$$

Or, equivalently

$$\langle S \rangle = \{r_1^{\alpha_1} \star r_2^{\alpha_2} \star \dots \star r_n^{\alpha_n} : n \geq 1, r_1, r_2, \dots, r_n \in S, \alpha_1, \alpha_2, \dots, \alpha_n \in \{1, -1\}\} \quad (3.3.2)$$

Remark 3.3.1. The s_i 's (and r_i 's) in these two descriptions need not be distinct. This is a generalization of the description of $\langle g \rangle$. This is by no means a unique way to write elements of $\langle S \rangle$.

Example 3.3.1. For $S = \{a, b\}$,

1. If $a \star b = b \star a$, for all $m \geq 1$,

$$s_1^{k_1} \star s_2^{k_2} \star \dots \star s_m^{k_m} = a^k \star b^l \quad (3.3.3)$$

2. In additive notation we get $a^k \star b^l = ka + lb$, so $\langle \{a, b\} \rangle = \mathbb{Z}a + \mathbb{Z}b$. (In general this need not be cyclic)

Remark 3.3.2. If we don't assume $a \star b = b \star a$, then we cannot simplify a general element to the form $a^k \star b^l$ in $\langle S \rangle$.

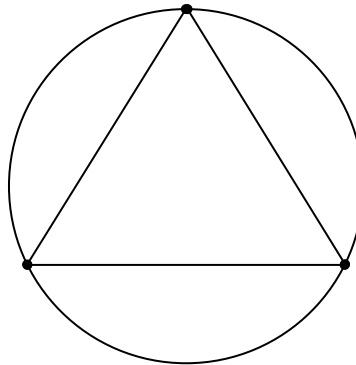
Definition 3.3.1. A polygon X is **regular** if it is **equiangular** (all angles are equal in measure) and **equilateral** (all sides have the same length).

↳

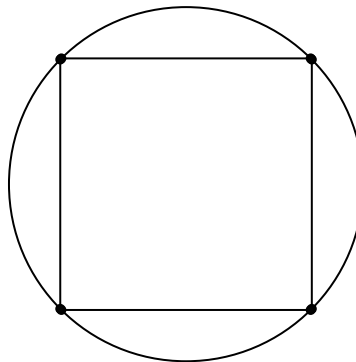
Note 3.3.3. The vertices of such a figure (if it is convex) can always be drawn on a circle called the circumcircle.

Example 3.3.2.

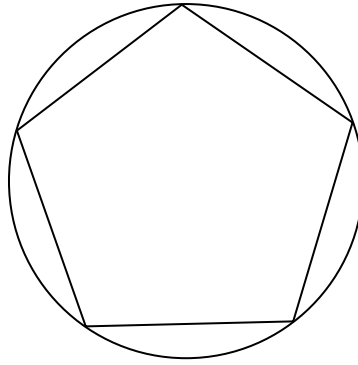
1. 3-sides = equilateral triangle:



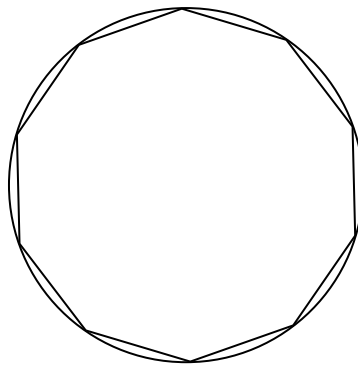
2. 4-sides = square



3. 5-sides = pentagon

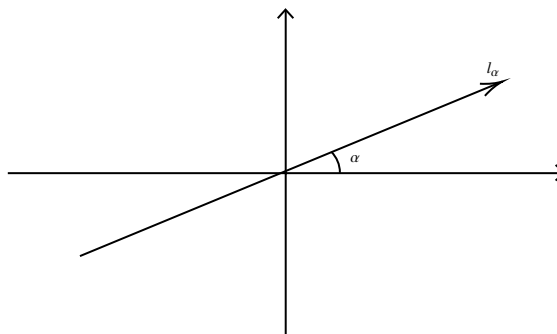


4. n -sides = regular convex n -gon



Definition 3.3.2. Let $n \geq 3$. The **dihedral group** D_n is the symmetry group of the regular convex n -gon. Consider the following maps $\mathbb{R}^2 \rightarrow \mathbb{R}^2$:

1. For all $\alpha \in \mathbb{R}$, ϕ_α is the rotation about the origin of \mathbb{R}^2 of angle α radians.
2. For all $\alpha \in \mathbb{R}$, ψ_α is the reflection with respect to a line l_α going through the origin and forming an angle α radians with the x -axis:



Claim 3.3.2. For all $\alpha, \beta \in \mathbb{R}$,

1. $\phi_\alpha \circ \phi_\beta = \phi_{\alpha+\beta}$ (rotation)
2. $\psi_\alpha \circ \psi_\beta = \phi_{2(\alpha-\beta)}$ (rotation)
3. $\phi_\alpha \circ \psi_\beta = \psi_{\beta+\frac{1}{2}\alpha}$ (reflection)

$$4. \psi_\beta \circ \phi_\alpha = \psi_{\beta - \frac{1}{2}\alpha} \text{ (reflection)}$$

\hookrightarrow (Where \circ is the composition of maps $\mathbb{R}^2 \rightarrow \mathbb{R}^2$).

Note 3.3.4. This is not commutative.

Corollary 3.3.3. The composition of maps induces a binary operation on the set of rigid motions and reflections about the origin, $\mathbb{R}^2 \rightarrow \mathbb{R}^2$.

Proposition 3.3.4. This set is a group with

$$\phi_\alpha^{-1} = \phi_{-\alpha}, \psi_\beta^{-1} = \psi_\beta, \text{ and } \phi_0 = \text{Id} \quad (3.3.4)$$

for all $\alpha, \beta \in \mathbb{R}$. Moreover, since \circ is associative, it is indeed a group.

Definition 3.3.3. This group is called the orthogonal group, $O_2(\mathbb{R})$.

Proposition 3.3.5. Note that $O_2(\mathbb{R}) \leq \mathbf{GL}_2(\mathbb{R})$ since the maps are linear. Indeed:

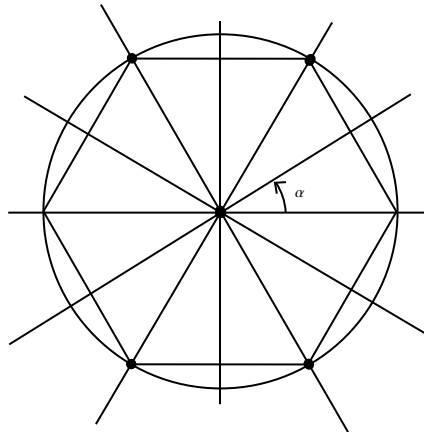
$$\phi_\alpha = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}, \psi_\alpha = \begin{bmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{bmatrix}, \quad (3.3.5)$$

in the standard basis of \mathbb{R}^2 .

Definition 3.3.4. Let $N \geq 3$, the n -th dihedral group D_n is the subgroup of $O_2(\mathbb{R})$ preserving a regular n -gon X_n with a circumcircle centered at the origin of \mathbb{R}^2 :

$$D_n := \{f \in O_2(\mathbb{R}) : f(X_n) = X_n\} \quad (3.3.6)$$

Example 3.3.3. For $n = 6$ we have



We have 6-rotations

$$\{\phi_{2\pi/6}, \phi_{4\pi/6}, \phi_{6\pi/6}, \phi_{8\pi/6}, \phi_{10\pi/6}, \phi_{12\pi/6}\}$$

and 6-reflections

$$\{\psi_{\pi/6}, \psi_{2\pi/6}, \psi_{3\pi/6}, \psi_{4\pi/6}, \psi_{5\pi/6}, \psi_{6\pi/6}\}$$

Corollary 3.3.6. *The order $|D_n|$ is $2n$ (n -rotations and n -reflections). In particular*

$$D_n = \langle \{\phi_{2\pi/n}, \psi_0\} \rangle \quad (3.3.7)$$

Remark 3.3.5. $D_n \leq O_2(\mathbb{R}) \leq \mathbf{GL}_2(\mathbb{R})$

Definition 3.3.5 (Algebraic Description). D_n is the group of order $2n$ generated by 2 elements x, y satisfying the relations

$$x^n = e, y^2 = e, yx = x^{-1}y \quad (3.3.8)$$

(they imply $xyx = y$, and $x^k y x^k = y$ for all $k \in \mathbb{Z}$)

Remark 3.3.6. The elements of D_n are

$$D_n = \{e, x, x^2, \dots, x^{n-1}, y, yx, yx^2, \dots, yx^{n-1}\} \quad (3.3.9)$$

(This is closed under the multiplication using the relations above). In fact, $(x^k y)^{-1} = x^k y$ because $x^k y x^k = y$ implies $x^k y = y x^{-k}$.

Remark 3.3.7. Let $Y \subset X$. Then

$$\{f \in S_X : f(Y) = Y\} \leq S_X \quad (3.3.10)$$

where S_X is the group of symmetries of the set X .

Remark 3.3.8. For the orthogonal group $O_2(\mathbb{R})$ we have the subgroup

$$D_n = \{f \in O_2(\mathbb{R}) \leq S_{\mathbb{R}^2} : f(X_n) = X_n\} \leq O_2(\mathbb{R}) \quad (3.3.11)$$

which is the group of symmetries of the regular convex n -gon, X_n , denoted by D_n .

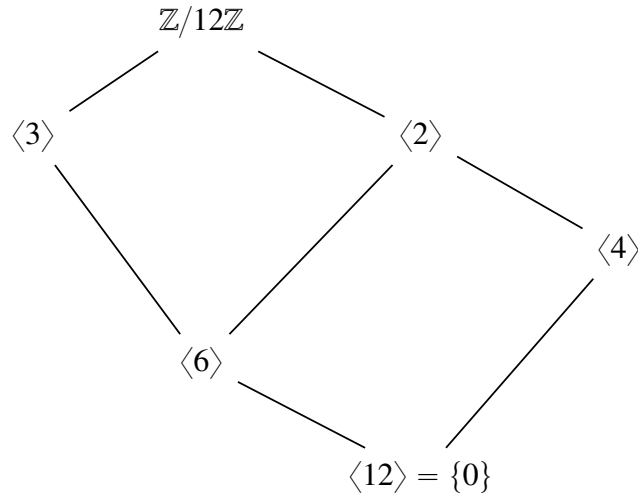
3.4.0 §Lattice Subgroups of a Group

Construction 3.4.1. *Given a finite group G , we plot subgroups of G with $\{e\}$ at the bottom and G at the top. We draw paths upward between subgroups using the rule that an upward line connects a subgroup A to a subgroup B if and only if $A \leq B$, and there are no subgroups properly between A and B .*

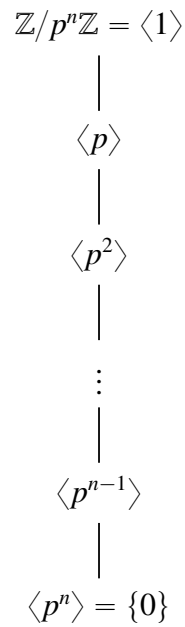
Remark 3.4.1. If $G \cong H$, then G and H have the same lattice structure. That is, group isomorphism induces a one-to-one correspondence between subgroups preserving containment.

Example 3.4.1.

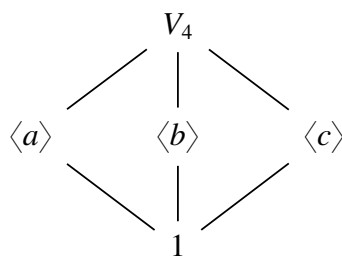
1. For $G = \mathbb{Z}/n\mathbb{Z}$ the lattice of subgroups is the lattice of divisors of \mathbb{Z} . For instance:



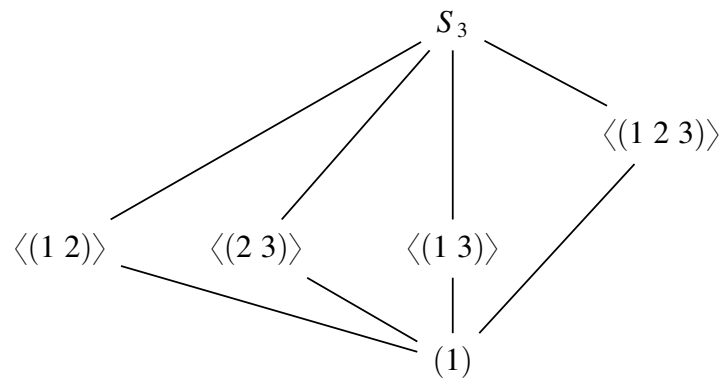
and for a prime p :



2. The Klien 4-group, $V_4 = \langle a, b, c : a^2 = b^2 = c^2 = 1 \rangle$:



3. The symmetric group on 3-letters, S_3 :



Chapter 4

§§Free Groups

4.1.0 §Basic Definitions and Examples: Free Groups

Remark 4.1.1. The idea of a free group, $F(S)$, generated by a set S is that there are no relations satisfied by the elements of S . (S is “free of relations”)

Definition 4.1.1 (Universal Property of Free Groups). Given any set map φ from the set S to the set underlying the group G , there is a unique group homomorphism

$$\Phi : F(S) \rightarrow G \quad (4.1.1)$$

such that $\Phi \circ \iota = \varphi$. That is to say, the following diagram commutes:

$$\begin{array}{ccc} S & \xrightarrow{\iota} & F(S) \\ & \searrow \varphi & \downarrow \exists! \Phi \\ & & G \end{array}$$

Construction 4.1.2. Let S be a set and let S^{-1} be any set disjoint from S such that there is a bijection from S to S^{-1} . For each $s \in S$ denote its corresponding element in S^{-1} by s^{-1} , and for each $t \in S^{-1}$ denote its corresponding element in S by t^{-1} (so $(s^{-1})^{-1} \in S$).

Take a singleton set not contained in $S \cup S^{-1}$ and call it $\{1\}$. Let $1^{-1} = 1$, and for any $x \in S \cup S^{-1} \cup \{1\}$, let $x^1 = x$.

A word on S is defined by a sequence

$$(s_1, s_2, s_3, \dots) \quad (4.1.2)$$

where $s_i \in S \cup S^{-1} \cup \{1\}$ for all i , and there exists $N \in \mathbb{N}$ such that if $i \geq N$, then $s_i = 1$.

The word (s_1, s_2, s_3, \dots) is said to be reduced if

1. $s_{i+1} \neq s_i^{-1}$ for all i with $s_i \neq 1$

2. if $s_k = 1$ for some k , then $s_i = 1$ for all $i \geq k$

The reduced word $(1, 1, 1, \dots)$ is called the **empty word** and is denoted by 1. We write the reduced word $(s_1^{\varepsilon_1}, \dots, s_n^{\varepsilon_n}, 1, 1, \dots)$ with $\varepsilon_i = \pm 1$ as $s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n}$. Note by definition reduced words

$$r_1^{\delta_1} \dots r_m^{\delta_m} \text{ and } s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n} \quad (4.1.3)$$

are equal if and only if $n = m$ and $\delta_i = \varepsilon_i$ for all $1 \leq i \leq n$.

Let $F(S)$ be the set of reduced words on S and embed S into $F(S)$ by

$$s \mapsto (s, 1, 1, \dots) \quad (4.1.4)$$

Note if $S = \emptyset$, $F(S) = \{1\}$.

Operation: Let $r_1^{\delta_1} \dots r_m^{\delta_m}$ and $s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n}$ be reduced words, and assume without loss of generality that $m \leq n$. Let k be the smallest integer in the range $1 \leq k \leq m+1$ such that

$$s_k^{\varepsilon_k} \neq r_{m-k+1}^{-\delta_{m-k+1}} \quad (4.1.5)$$

Then the product of these reduced words is defined as

$$(r_1^{\delta_1} \dots r_m^{\delta_m})(s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n}) := \begin{cases} r_1^{\delta_1} \dots r_{m-k+1}^{\delta_{m-k+1}} s_k^{\varepsilon_k} \dots s_n^{\varepsilon_n}, & \text{if } k \leq m \\ s_{m+1}^{\varepsilon_{m+1}} \dots s_n^{\varepsilon_n}, & \text{if } k = m+1 \leq n \\ 1, & \text{if } k = m+1 \text{ and } m = n \end{cases} \quad (4.1.6)$$

Theorem 4.1.1. $F(S)$ is a group under the above binary operation.

Proof. By construction we note that 1 is an identity element of the binary operation, and that the inverse of a reduced word $s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n}$ is $s_n^{-\varepsilon_n} \dots s_1^{-\varepsilon_1}$. For each $s \in S \cup S^{-1} \cup \{1\}$ define a map $\sigma_s : F(S) \rightarrow F(S)$ by

$$\sigma_s(s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n}) = \begin{cases} s \cdot s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n} & \text{if } s_1^{\varepsilon_1} \neq s^{-1} \\ s_2^{\varepsilon_2} \dots s_n^{\varepsilon_n} & \text{if } s_1^{\varepsilon_1} = s^{-1} \end{cases} \quad (4.1.7)$$

Since $\sigma_{s^{-1}} \circ \sigma_s$ is the identity map on $F(S)$, σ_s is a permutation of $F(S)$. Let $A(F)$ be the subgroup of the symmetric group on $F(S)$ generated by $\{\sigma_s : s \in S\}$. We observe that the map

$$s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n} \mapsto \sigma_{s_1^{\varepsilon_1}} \circ \dots \circ \sigma_{s_n^{\varepsilon_n}} \quad (4.1.8)$$

is a set bijection between $F(S)$ and $A(F)$ which respects their binary operation. Since $A(F)$ is a group, and hence its operation is associative, so is $F(S)$. ■

Theorem 1 (Universal Property of Free Groups).

Let G be a group, S a set, and $S \xrightarrow{\varphi} G$ a set map. Then there exists a unique group homomorphism $\Phi : F(S) \rightarrow G$ such that the diagram commutes

$$\begin{array}{ccc} S & \xrightarrow{\iota} & F(S) \\ & \searrow \varphi & \downarrow \exists! \Phi \\ & & G \end{array}$$

Proof. Choose $\Phi : s_1^{\varepsilon_1} \dots s_n^{\varepsilon_n} \mapsto \varphi(s_1)^{\varepsilon_1} \dots \varphi(s_n)^{\varepsilon_n}$ ■

Corollary 4.1.2. *The free group $F(S)$ is unique up to an isomorphism which is an identity on the set S .*

Proof. Suppose $F(S)$ and $F'(S)$ are two free groups generated by S . Since S is contained in both $F(S)$ and $F'(S)$ we have natural injections

$$S \xhookrightarrow{\iota} F(S), S \xhookrightarrow{\iota'} F'(S) \quad (4.1.9)$$

By the universal property of Free groups there exist unique group homomorphisms $\Phi : F(S) \rightarrow F'(S)$ and $\Phi' : F'(S) \rightarrow F(S)$ such that $\Phi \circ \iota = \iota'$ and $\Phi' \circ \iota' = \iota$, which are both the identity on S . Then, $\Phi' \circ \Phi$ is a map which makes the diagram

$$\begin{array}{ccc} S & \xrightarrow{\iota} & F(S) \\ & \searrow \iota & \downarrow \downarrow ? \\ & & F(S) \end{array}$$

commute. But, $\text{Id}_{F(S)}$ also makes this commute, so by uniqueness $\Phi' \circ \Phi = \text{Id}_{F(S)}$. Similarly, $\Phi \circ \Phi' = \text{Id}_{F'(S)}$, so Φ and Φ' are inverses, and hence bijections. Thus, Φ and Φ' are isomorphisms which are the identity on S , so $F(S) \cong F'(S)$ as claimed. ■

Definition 4.1.3 (Free Group). *The group $F(S)$ is called the **free group** on the set S . A group F is a free group if there is some set S such that $F \cong F(S)$. In this case we call S a set of **free generators** or a **free basis** of F . The cardinality of S is called the **rank** of the free group F .*

Theorem 4.1.3 (Schreier). *Subgroups of a free group are themselves free.*

4.2.0 §Presentations

Remark 4.2.1. For a group G , G is a homomorphic image of a free group. Take $S = G$ and φ as the identity map from G to G . Then by the universal property of free groups there is a surjective group homomorphism from $F(G)$ onto G .

↳ In general, if $S \subseteq G$ such that $G = \langle S \rangle$, then there exists a unique group epimorphism $\varphi : F(S) \rightarrow G$ which is the identity on S .

Definition 4.2.1. *Let S be a subset of a group G such that $G = \langle S \rangle$.*

1. A **presentation** for G is a pair (S, R) where R is a set of words in $F(S)$ such that the **normal closure** of $\langle R \rangle$ in $F(S)$ (the smallest normal subgroup containing $\langle R \rangle$) equals the kernel of the homomorphism $\pi : F(S) \rightarrow G$, where π extends the identity map from S to S . The elements of S are called **generators** and those of R are called **relations** of G .

2. We say G is **finitely generated** if there is a presentation (S, R) such that S is a finite set, and we say G is **finitely presented** if there is a presentation (S, R) where both S and R are finite sets.

Chapter 5

§§Quotient Groups

5.1.0 §Cosets

Definition 5.1.1 (Left Cosets). Let $H \leq G$ be a subgroup, $g \in G$. The left coset of H containing g , or generated by g , is

$$gH := \{g \star h : h \in H\} \subseteq G \quad (5.1.1)$$

↳ Notation depends on the operation: $g + H, g \star H, gH$, etc.

Note 5.1.1. For $H \leq G$ and all $g \in G$, $g = g \star e_G \in gH$ since $e_G \in H$. If $g \in H$, $gH = H$, as $h = g \star (g^{-1} \star h)$ for all $h \in H$. Hence we can have $gH = g'H$ for $g \neq g'$. Note gH is only a subgroup if $g \in H$ so $gH = H$.

Example 5.1.1. Take $G = \mathbb{Z}/4\mathbb{Z}$, $H = \langle [2] \rangle \leq G$. So $H = \{[0], [2]\}$. Then we have the cosets

$$\begin{array}{ll} g & gH = g + H \\ [0] & H \\ [1] & \{[1], [3]\} = [1] + H \\ [2] & H \\ [3] & \{[3], [1]\} = [3] + H = [1] + H \end{array}$$

we have $H = [2] + H$ and $[1] + H = [3] + H$, so we get 2 distinct left cosets and they form a partition of G .

Lemma 5.1.1. The left cosets of $H \leq G$ form a partition of G .

Proof. First, as $g \in gH$ for all $g \in G$ we have that

$$\bigcup_{g \in G} gH = G$$

Then, suppose $gH \cap g'H \neq \emptyset$, and let $gh = g'h' \in gH \cap g'H$. Then $g = g'h'h^{-1} \in g'H$ and $g' = gh'h'^{-1} \in gH$. Then, for all $gh'' \in gH$ and $g'\bar{h} \in g'H$, $gh'' = g'(h'h^{-1}h'') \in g'H$

and $g'\bar{h} = g(hh'^{-1}\bar{h}) \in gH$, so $g'H \supseteq gH$ and $g'H \subseteq gH$. Thus $g'H = gH$, so by proof by contrapositive we have that all distinct cosets are disjoint, completing the proof. ■

Remark 5.1.2. The left cosets of H are the equivalence classes for the equivalence relation

$$a \equiv \iff a^{-1}b \in H \quad (5.1.2)$$

Definition 5.1.2 (Index). The number of left cosets of H in G is called the index of H in G , denoted by

$$|G : H| \text{ or } [G : H] \quad (5.1.3)$$

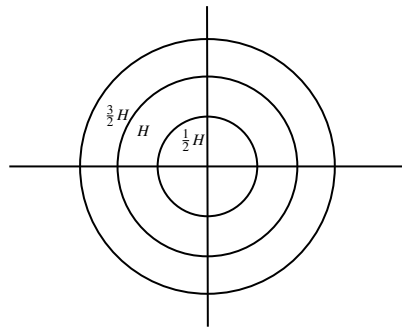
Lemma 5.1.2. For all $g \in G$, G a group, and all $H \leq G$, $|gH| = |H|$.

Proof. Define the map $\phi : H \rightarrow gH$ by $h \mapsto gh$. Then ϕ is a bijection with inverse $\varphi : gH \rightarrow H$ defined by $gh \mapsto g^{-1}gh$. Indeed, we have $\varphi \circ \phi(h) = \varphi(gh) = g^{-1}gh = h$. Thus, by definition of cardinality of sets $|H| = |gH|$. ■

Definition 5.1.3. Let G be a group with $H \leq G$. $g \in G$ is a **representative** of a left coset xH of H in G if and only if $g \in xH$ ($\iff gH = xH$). A complete set of representatives of the left cosets of H in G is a set $S \subseteq G$ such that S contains one, and only one, representative of each left coset of H in G .

Example 5.1.2.

1. For $n \geq 1$, and $H = n\mathbb{Z} \leq \mathbb{Z} = G$, a number $m \in \mathbb{Z}$ is a representative of $a + n\mathbb{Z}$ if and only if $n \mid m - a$. A complete set of representatives would be $\{0, 1, 2, \dots, n-1\}$ of the left cosets of $n\mathbb{Z}$ in \mathbb{Z} .
2. Consider $H = \{g \in \mathbb{C}^\times : |g| = 1\} \leq \mathbb{C}^\times$.



A complete set of representatives is $\mathbb{R}_{>0}$. Indeed, the map

$$\begin{aligned} \mathbb{R}_{>0} &\xrightarrow{f} \{\text{left cosets of } H \text{ in } \mathbb{C}^\times\} \\ r &\mapsto rH \end{aligned} \quad (5.1.4)$$

Indeed, for all $x \in \mathbb{C}^\times$ $x = re^{i\theta}$ so $xH = re^{i\theta}H = rH$, as $e^{i\theta} \in H$, so f is surjective. If $r, r' \in \mathbb{R}_{>0}$ then $r' = rz$ with $|z| = 1$, so $|r'| = |rz| = |r|$. But, for $r, r' \in \mathbb{R}_{>0}$ $|r'| = |r|$ implies $r' = r$. Thus, f is a bijection.

3. Consider $H = \mathbb{R} \leq \mathbb{C}$. Then \mathbb{R} identified with the y axis is a complete set of representative under the map

$$\begin{aligned} \mathbb{R} &\mapsto \{\text{left cosets of } H\} \\ r &\mapsto ir + R \end{aligned} \quad (5.1.5)$$

4. Consider $\mathbb{R}^\times \leq \mathbb{C}^\times$. Then $[0, \pi[$ is a complete set of representatives, with the map

$$\begin{aligned} [0, \pi[&\rightarrow \{\text{left cosets of } \mathbb{R}^\times\} \\ \theta &\mapsto e^{i\theta} \mathbb{R}^\times \end{aligned} \quad (5.1.6)$$

Definition 5.1.4. For a subgroup $H \leq G$, a right coset of H in G is a subset of G of the form

$$Hg := \{hg : h \in H\} \quad (5.1.7)$$

We say that Hg is the right coset generated by g or containing g .

Theorem 5.1.3. The right cosets of H in G form a partition of G .

Proof. (Left to the reader) ■

Remark 5.1.3. The right cosets of H in G are the equivalence classes of the equivalence relation

$$a \equiv b \iff ba^{-1} \in H \quad (5.1.8)$$

Proposition 5.1.4. There is a bijection

$$\begin{aligned} \{\text{left cosets of } H \text{ in } G\} &\rightarrow \{\text{right cosets of } H \text{ in } G\} \\ aH &\mapsto Ha^{-1} \end{aligned} \quad (5.1.9)$$

Proof. (Left to the reader) ■

Corollary 5.1.5. The number of right cosets of H in G is also $|G : H|$, the index.

Remark 5.1.4. When G is abelian, every right coset Hg is a left coset gH .

Example 5.1.3 (Non-example). For $H = \langle y \rangle \leq D_3 = \langle x, y \rangle$, the left and right cosets give two different partitions of D_3 . Indeed, $H = yH$, $xH = xyH$, and $x^2H = x^2yH$ are the left cosets, while $H = yH$, $Hx = Hx^2y$, $Hx^2 = Hxy$ are the right cosets.

5.2.0 §Lagrange's Theorem and Applications

Theorem 2 (Lagrange's Theorem).

If H is a subgroup of a finite group G , then $|H|$ divides $|G|$.

Proof. Let $H \leq G$ for G finite. Then since the cosets of H partition G we have that

$$|G| = \sum_{\text{cosets}} |gH| = \sum_{\text{cosets}} |H| = |G : H| |H| \quad (5.2.1)$$

so by definition $|H|$ divides $|G|$. ■

Example 5.2.1.

1. Let $g \in G$ (a finite group). Then $o(g) \mid |G|$ and thus $g^{|G|} = e_G$.

Proof. (Left to the reader) ■

2. If $|G| = p$ is prime, then G has only the trivial subgroups, $H = \{e_G\}$ and $H = G$, since $|H| \mid |G|$ implies $|H| \in \{1, p\}$. Actually:

Corollary 5.2.1. *If $|G| = p$ a prime, then $G \cong \mathbb{Z}/p\mathbb{Z}$.*

3. If $|G| = p^2$ for a prime p then either $G \cong \mathbb{Z}/p^2\mathbb{Z}$ or $g^p = e$ for all $g \in G$.

Proof. (Left to the reader) ■

Remark 5.2.1. A class $[m] \in \mathbb{Z}/n\mathbb{Z}$ has a multiplicative inverse if and only if $\gcd(m, n) = 1$ (they are relatively prime).

Proof. If $\gcd(m, n) = 1$ then there exist $a, b \in \mathbb{Z}$ such that $am + bn = 1$, so $[a][m] = [1]$ modulo n . On the other hand, if $[a][m] = [1]$ for some $m \in \mathbb{Z}$ then there exists $b \in \mathbb{Z}$ such that $1 = am + bn$. Hence, $m\mathbb{Z} + n\mathbb{Z} = \mathbb{Z}$, so $\gcd(m, n) = 1$. ■

Definition 5.2.1. Fix an integer $n > 1$. Let $(\mathbb{Z}/n\mathbb{Z})^\times$ be the set of these classes with multiplicative inverses from $\mathbb{Z}/n\mathbb{Z}$ with multiplication. Then, it is a group with identity $[1]$.

Definition 5.2.2 (Euler Totient Function). Let $\varphi(n) := |(\mathbb{Z}/n\mathbb{Z})^\times|$ (the Euler totient function), or equivalently

$$\varphi(n) := |\{m \in \{0, 1, \dots, n-1\} : \gcd(m, n) = 1\}| \quad (5.2.2)$$

This is also known as the Euler phi function.

Example 5.2.2. Take p a prime. Then

1. $\varphi(p) = p - 1$
2. $\varphi(p^k) = p^k - p^{k-1}$ for all $k \geq 1$

Theorem 3 (Euler's Theorem).

If a and $n \geq 2$ are relatively prime integers, then

$$a^{\varphi(n)} \equiv 1 \pmod{n} \quad (5.2.3)$$

Proof. (Left to the reader) ■

Theorem 4 (Fermat's Theorem).

If p is a prime then $a^p = a \pmod{p}$ for all $a \in \mathbb{Z}$.

Proof. (Left to the reader) ■

Classification of Groups of Order $2p$ for p a prime

Theorem 5.2.2. *Let G be a group. If $|G| = 2p$, then either G is cyclic ($\cong \mathbb{Z}/2p\mathbb{Z}$) or G is isomorphic to the dihedral group D_p of order $2p$.*

Proof. This proof extends over this subsection and will be completed after stating a few lemmas ■

Lemma 5.2.3. *If G is a group in which $g^2 = e_G$ for all $g \in G$, then G is abelian.*

Proof. Let $x, y \in G$. Then $(xy)^2 = e$ since $xy \in G$, so $xy = (xy)^{-1} = y^{-1}x^{-1} = yx$. Thus G is abelian as claimed. ■

Proof of Theorem for $p = 2$. First, suppose $|G| = 2 \cdot 2 = 2^2 = p^2$. Then, by application of 2 we have G is cyclic or $g^p = g^2 = e_G$ for all $g \in G$. If G is cyclic $G \cong \mathbb{Z}/4\mathbb{Z}$ and we're done. If G is not cyclic, $G = \{e_G, g_2, g_3, g_4\}$. Set $x = g_2$ and $y = g_3$. We have $|G| = 2n$, $x^n = e_G = y^2$ and $yx = xy = x^{-1}y$ since G is abelian by the last Lemma. Thus, we have that $G \cong D_2$, the dihedral group of order 4. ■

Proof of Theorem for $p \neq 2$. If G is cyclic we're done, so suppose G is not cyclic. We must show that $G \cong D_p$. By 2 $o(g) \in \{1, 2, p\}$ for all $g \in G$ (since G is assumed to not be cyclic).

Claim 5.2.4. *G has an element of order p .*

Proof. If $g^2 = e_G$ for all $g \in G$ then G is abelian by the Lemma. Take three distinct elements $\{e_G, a, b\}$ in G , so $\{e_G, a, b, ab\} \leq G$, which is isomorphic to D_2 . But $|D_2| = 4$ and $4 \nmid 2p$ as $p > 2$ is odd. Thus, this is not possible by 2. Hence, there must exist $x \in G$ such that $o(x) = p$. ■

Set $H = \langle x \rangle$, so $|H| = o(x) = p$.

Claim 5.2.5. *If $g \in G$ with $g \notin H$, then $o(g) = 2$.*

Proof. Note $g \neq e_G$ because $e_G \in H$, so $o(g) \neq 1$. We have that $G = H \sqcup gH$ because $|G| = 2p = |H| + |gH|$ and $gH \cap H = \emptyset$ since $g \notin H$ by $g \in gH$. Next, note $g^2 \in gH$ if and only if $g \in H$, so $g^2 \notin gH$, which implies $g^2 \in H$. If $o(g) = p$ then $g = g^{p+1} = (g^2)^{\frac{p+1}{2}} \in H$ since $p+1$ is even. But $g \notin H$, so this is a contradiction. Hence, we must have that $o(g) = 2$. ■

Now, let $y \in G$ such that $y \notin H$, so $o(y) = 2$. Then we have $\langle x, y \rangle \geq H$ and $\langle x, y \rangle \subseteq yH$. But $G = H \sqcup gH$, so $\langle x, y \rangle = G$. Thus $|\langle x, y \rangle| = 2p$. Moreover, $o(x) = p$ and $o(y) = 2$. We want $yx = x^{-1}y$. Note $yx \in yH$ by definition of the left coset, so $(yx)^2 = e_G$. Hence, $yx = (yx)^{-1} = x^{-1}y$ as $y = y^{-1}$. Therefore, G satisfies the criterion for the dihedral group of order $2p$, so $G \cong D_p$. ■

5.3.0 §The Alternating Group

Definition 5.3.1. Let $n \geq 1$, A be an $n \times n$ matrix, and let $\sigma \in S_n$. Define an action of S_n on $\mathbf{GL}_n(\mathbb{R})$ by letting $\sigma(A)$ be the $n \times n$ matrix with the i -th row being the $\sigma^{-1}(i)$ -th row of A . That is

$$\sigma(A)_{\sigma(i)j} = A_{ij} \text{ or } \sigma(A)_{ij} = A_{\sigma^{-1}(i)j} \quad (5.3.1)$$

so σ sends the i -th row of A to the $\sigma(i)$ -th row.

Claim 5.3.1. The map defined by

$$\begin{aligned} S_n &\xrightarrow{f} \mathbf{GL}_n(\mathbb{R}) \\ \sigma &\mapsto \sigma(\text{Id}_n) \end{aligned} \quad (5.3.2)$$

is a well defined group homomorphism.

Proof. First, let $\sigma, \eta \in S_n$. I claim $\sigma(A) = \sigma(\text{Id}_n)A$ for all $A \in \mathbf{GL}_n(\mathbb{R})$. Indeed, observe that

$$\begin{aligned} (\sigma(\text{Id}_n)A)_{ik} &= \sum_{j=1}^n \sigma(\text{Id}_n)_{ij} A_{jk} \\ &= \sum_{j=1}^n \text{Id}_{\sigma^{-1}(i)j} A_{jk} \\ &= \sum_{j=1}^n \delta_{\sigma^{-1}(i)j} A_{jk} \\ &= A_{\sigma^{-1}(i)k} \\ &= \sigma(A)_{ik} \end{aligned}$$

where $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ is the **kroncker delta**. Hence we have that

$$f(\sigma)f(\eta) = \sigma(\text{Id}_n)\eta(\text{Id}_n) = \sigma(\eta(\text{Id}_n)) = (\sigma \circ \eta)(\text{Id}_n) = f(\sigma \circ \eta) \quad (5.3.3)$$

Thus f is multiplicative. Next we want to show $\det(f(\sigma)) \neq 0$, so $f(\sigma) \in \mathbf{GL}_n(\mathbb{R})$ for all $\sigma \in S_n$. For σ a 2-cycle, (i.e. a transposition) we have $\sigma(\text{Id}_n) = E$ an elementary matrix for the elementary operation of type I (exchanging two rows), so $\det(\sigma(\text{Id}_n)) = 1$ or -1 . But, every permutation $\sigma \in S_n$ can be written as a product of transpositions, say $\sigma = \beta_1 \circ \dots \circ \beta_r$. Hence, multiplicativity says $f(\sigma) = f(\beta_1) \dots f(\beta_r)$ where $\det(f(\beta_i)) \in \{1, -1\}$ for all i , so $\det(f(\sigma)) \in \{1, -1\}$. Hence we have that $f(\sigma) \in \mathbf{GL}_n(\mathbb{R})$, completing the proof. ■

Corollary 5.3.2. Due to this result we have a homomorphism

$$\begin{array}{ccc} S_n & \xrightarrow{\det \circ f} & \{-1, 1\} \leq \mathbb{C}^\times \\ & \searrow f & \nearrow \det \\ & \mathbf{GL}_n(\mathbb{R}) & \end{array}$$

Definition 5.3.2 (Parity of Permutations). The permutation $\sigma \in S_n$ is called even if $\text{sign}(\sigma) = 1$ and odd if $\text{sign}(\sigma) = -1$ where we define

$$\text{sign} := \det \circ f \quad (5.3.4)$$

Remark 5.3.1. From the proof above we see that this definition is compatible with the definition of odd or even in terms of the transposition decomposition of a permutation.

Proof. (Left to the reader) ■

Definition 5.3.3. The subgroup of S_n of even permutations, $\ker(\text{sign})$, is called the alternating group of degree n , denoted A_n .

Proposition 5.3.3. For all $n \geq 2$ we have $|A_n| = \frac{n!}{2} = \frac{|S_n|}{2}$.

Proof. Let Odd_n be the subset of all odd permutations. Then $|S_n| = |A_n| + |Odd_n|$. Moreover, the maps

$$\begin{aligned} A_n &\rightarrow Odd_n \\ \sigma &\mapsto (1\ 2) \circ \sigma \end{aligned} \quad (5.3.5)$$

and

$$\begin{aligned} Odd_n &\rightarrow A_n \\ \gamma &\mapsto (1\ 2) \circ \gamma \end{aligned} \quad (5.3.6)$$

are mutual inverses, and hence bijections of sets. Thus $|A_n| = |Odd_n| = \frac{n!}{2}$. ■

Example 5.3.1. $A_2 = \{(1)\}$ and $A_3 = \{(1), (1\ 2\ 3), (1\ 3\ 2)\} \cong \mathbb{Z}/3\mathbb{Z} \cong \langle x \rangle \leq D_3$, where $D_3 \cong S_3$ from before.

5.4.0 §The Quotient Group Definition and Construction

Construction 5.4.1. We want to define a group structure \star on the left cosets of $H \leq G$ in G such that the map

$$\begin{aligned} \pi : G &\rightarrow \{\text{left cosets of } H \text{ in } G\} \\ g &\mapsto gH \end{aligned} \quad (5.4.1)$$

is a group homomorphism. That is we want $\pi(gg') = \pi(g) \star \pi(g')$, so we must define the operation \star by $aH \star bH := abH$. For \star to be well defined we need $\pi(g) = \pi(a)$ and $\pi(g') = \pi(b)$ imply $\pi(gg') = \pi(ab)$. As $\pi(g) = \pi(a)$ and $\pi(g') = \pi(b)$ we have that $gh = a$ and $g'h' = b$ for some $h, h' \in H$. Then, we want $ghg'h' \in gg'H$ with occurs if and only if $hg'h' \in g'H$, which is to say $hg' \in g'H$. In other words, we must have that $h \in g'Hg'^{-1}$ for all $h \in H$. In particular, $H \subseteq g'Hh^{-1}$. But, then $g'^{-1}Hg' \subseteq H$, and as this must hold for all $(g'^{-1})^{-1}Hg'^{-1} = g'Hg'^{-1} \subseteq H$, so $H = g'Hg'^{-1}$. In other words, $H = g^{-1}Hg$ for all $g \in G$ if our operation is to be well defined. Moreover, if \star is well-defined then we obtain a group structure on the left cosets of H in G under \star . This follows from associativity in G and the fact that π is defined to be a group homomorphism.

Definition 5.4.2 (Normal Subgroups). A subgroup $H \leq G$ is a normal subgroup, denoted $H \triangleleft G$, if and only if for all $g \in G$ and all $h \in H$, $g^{-1}hg \in H$.

Note 5.4.1. This is equivalent to $g^{-1}Hg = H$ for all $g \in G$.

Example 5.4.1.

1. $\{e_G\} \triangleleft G$ and $G \triangleleft G$.
2. If G is abelian, $H \leq G \implies H \triangleleft G$.
3. Every subgroup of $Z(G)$, the center of G , is normal in G . Indeed, for all $h \in Z(G)$ and all $g \in G$, $g^{-1}hg = g^{-1}gh = h \in Z(G)$.

Remark 5.4.2. The following are equivalent:

1. $H \triangleleft G$
2. For all $g \in G$, $gH = Hg$
3. Every left coset is a right coset, and vice-versa

Proof. (Left to the reader) ■

Proposition 5.4.1. Let $G \xrightarrow{f} K$ be a group homomorphism, and $H_1 \triangleleft K$. Then $f^{-1}(H_1) \triangleleft G$.

Proof. (Left to the reader) ■

Corollary 5.4.2. For all group homomorphisms $f : G \rightarrow K$, $\ker(f) = f^{-1}(\{e_K\})$ is a normal subgroup of G .

Example 5.4.2. Then $A_n \triangleleft S_n$ as $A_n = \ker(\text{sign})$, $\mathbf{SL}_n(\mathbb{R}) \triangleleft \mathbf{GL}_n(\mathbb{R})$ as $\mathbf{SL}_n(\mathbb{R}) = \ker(\det)$, and for $n = 3$ $S_3 \cong D_3$ so $A_3 \cong \langle x \rangle \triangleleft D_3$.

Example 5.4.3 (Non-example). $H := \langle y \rangle \leq D_3$ is not a normal subgroup of D_3 . Indeed, we have shown previously that the left cosets partition D_3 differently when compared to the right cosets. Alternatively, $x^{-1}yx = yx^2 \notin H$.

Remark 5.4.3. The image of a normal subgroup under a group homomorphism is a subgroup, but not necessarily a normal subgroup.

Example 5.4.4. $\langle y \rangle \hookrightarrow D_3$ is a group homomorphism and $\langle y \rangle \triangleleft \langle y \rangle$ but

$$\iota(\langle y \rangle) = \langle y \rangle \not\triangleleft D_3$$

is not normal.

Notation 5.4.4. For subsets $A, B \subseteq G$, we define the subset product

$$AB := \{ab : a \in A, b \in B\} \tag{5.4.2}$$

Lemma 5.4.3. *If $N \triangleleft G$, then for all $a, b \in G$, $(aN)(bN) = \{anbn' : n, n' \in N\}$ is the left coset abN of N in G .*

Proof. (Left to the reader) ■

Definition 5.4.3 (Quotient Group). *Let $N \triangleleft G$ be a normal subgroup. Then the quotient group G/N of G by N is the set of all left cosets of N in G with the binary operation*

$$(gN) \star (g'N) = gg'N \quad (5.4.3)$$

Note that this is indeed a well-defined group structure as our previous argument shows, and its structure makes the canonical projection

$$\begin{aligned} \pi : G &\rightarrow G/N \\ g &\mapsto gN \end{aligned} \quad (5.4.4)$$

a surjective group homomorphism.

Corollary 5.4.4. *If G is abelian, G/N is abelian and if G is cyclic then G/N is cyclic.*

Proof. (Left to the reader) ■

Example 5.4.5.

1. For all $n \geq 1$, $n\mathbb{Z} \triangleleft \mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$ is as previously defined.
2. $G/G \cong \{e_G\}$
3. $G/\{e_G\} \cong G$
4. $\mathbf{GL}_n(\mathbb{R})/\mathbf{SL}_n(\mathbb{R}) \cong \mathbb{R}^\times$

Claim 5.4.5. *The set $\left\{ \begin{bmatrix} r & \mathbf{0} \\ \mathbf{0} & I_{n-1} \end{bmatrix} : r \in \mathbb{R}^\times \right\}$ is a complete set of representatives of left cosets of $\mathbf{SL}_n(\mathbb{R})$ in $\mathbf{GL}_n(\mathbb{R})$.*

Proof. (Left to the reader) ■

Consider the map

$$\begin{aligned} \mathbb{R}^\times &\xrightarrow{j} \mathbf{GL}_n(\mathbb{R})/\mathbf{SL}_n(\mathbb{R}) \\ r &\mapsto j(r) = \begin{bmatrix} r & \mathbf{0} \\ \mathbf{0} & I_{n-1} \end{bmatrix} \mathbf{SL}_n(\mathbb{R}) \end{aligned} \quad (5.4.5)$$

We claim it is an isomorphism.

Proof. j is a bijection by the first claim. Let $r, r' \in \mathbb{R}^\times$. Then

$$j(r)j(r') = \begin{bmatrix} r & \mathbf{0} \\ \mathbf{0} & I_{n-1} \end{bmatrix} \mathbf{SL}_n(\mathbb{R}) \begin{bmatrix} r' & \mathbf{0} \\ \mathbf{0} & I_{n-1} \end{bmatrix} \mathbf{SL}_n(\mathbb{R}) = \begin{bmatrix} rr' & \mathbf{0} \\ \mathbf{0} & I_{n-1} \end{bmatrix} \mathbf{SL}_n(\mathbb{R}) = j(rr') \quad (5.4.6)$$

so j is a homomorphism. Hence, $\mathbb{R}^\times \cong \mathbf{GL}_n(\mathbb{R})/\mathbf{SL}_n(\mathbb{R})$. ■

5. For $n \geq 2$, $S_n/A_n \cong \{-1, 1\}$ with the isomorphism

$$\begin{aligned} \{-1, 1\} &\xrightarrow{\varphi} S_n/A_n \\ 1 &\mapsto A_n \\ -1 &\mapsto (1\ 2)A_n \end{aligned} \quad (5.4.7)$$

Remark 5.4.5. Consider $H \leq G$ for G not necessarily a finite group, then the number of left cosets of H in G is the index $|G : H|$ by definition. The index can be finite even if H and G are infinite, but it can also be infinite. Next, if $N \triangleleft G$, then the order of the group G/N is $|G : N|$. Moreover, if $|G| < +\infty$, $|G/N| = \frac{|G|}{|N|}$ since $|G| = |N||G : N|$ by 2.

Remark 5.4.6. Given $N \triangleleft G$, one can study G by studying the two groups N and G/N .

Theorem 5.4.6. Let $K \leq Z(G)$ (so $K \triangleleft G$) such that G/K is cyclic. Then G is abelian.

Proof. Let $G/K = \langle gK \rangle$. Then for all $a, b \in G$ there exist $m, n \geq 0$ such that $a = g^m k$ and $b = g^n k'$ for some $k, k' \in K$. Then

$$ab = g^m k g^n k' = g^{m+n} k k' = g^{n+m} k' k = g^n k' g^m k = ba$$

so G is abelian. ■

Remark 5.4.7. However, one can have G/N and N cyclic but G is not cyclic. Similarly, we can have G/N and N abelian but G is not abelian.

Example 5.4.6. For $A_3 \triangleleft S_3$, $S_3/A_3 \cong \mathbb{Z}/2\mathbb{Z}$, where A_3 and $\mathbb{Z}/2\mathbb{Z}$ are cyclic, but S_3 is not even abelian.

Definition 5.4.4 (Simple). A group G is called simple if its only normal subgroups are $\{e_G\}$ and G .

Example 5.4.7.

1. $\mathbb{Z}/p\mathbb{Z}$ is simple for all primes p (no proper subgroups at all by 2).
2. $\{e\}$ is simple
3. A_n is simple for $n \geq 5$.

5.5.0 §Isomorphism Theorems and Correspondence

Theorem 5 (First Isomorphism Theorem of Groups).

Let $f : G \rightarrow G'$ be a group homomorphism and let $N = \ker(f) \triangleleft G$. Then there exists a unique group homomorphism $\bar{f} : G/N \rightarrow G'$ such that the following diagram commutes:

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ \pi \searrow & & \nearrow \exists! \bar{f} \\ & G/\ker(f) & \end{array}$$

In particular, for all $a \in G$ and for all $b \in aN$, $f(b) = f(a)$, and the map

$$\begin{aligned} G/N &\xrightarrow{\bar{f}} G' \\ aN &\mapsto f(a) \end{aligned} \quad (5.5.1)$$

is well defined and satisfies $f = \bar{f} \circ \pi$. Note that we call $\pi : G \rightarrow G/N$ the **canonical map** or **factor map**. Finally, \bar{f} is a group monomorphism.

Proof. First we shall show \bar{f} as defined is a well-defined group monomorphism with $f = \bar{f} \circ \pi$. Let $b \in aN$ so $b = an$ for some $n \in \ker(f)$. Then

$$f(b) = f(an) = f(a)f(n) = f(a)e_{G'} = f(a)$$

so \bar{f} is well defined. Let $bN, b'N \in G/N$. Then

$$\bar{f}(bN \star b'N) = \bar{f}(bb'N) = f(bb') = f(b)f(b') = \bar{f}(bN)\bar{f}(b'N)$$

so \bar{f} is multiplicative. Finally, take $bN \in \ker(\bar{f})$, so $\bar{f}(bN) = f(b) = e_{G'}$. Hence, $b \in \ker(f) = N$, so $bN = N = e_{G/N}$. Thus, $\ker(\bar{f}) = \{e_{G/N}\}$, so \bar{f} is a monomorphism. By construction we have that $f = \bar{f} \circ \pi$. To prove uniqueness suppose g is another such group homomorphism such that $f = g \circ \pi$. Then for all $aN \in G/N$ we have

$$g(aN) = (g \circ \pi)(a) = f(a) = \bar{f}(a)$$

so $g = \bar{f}$. Hence, uniqueness is satisfied and the proof is complete. \blacksquare

Example 5.5.1.

1. The map $\mathbf{GL}_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times$ gives the isomorphism

$$\mathbf{GL}_n(\mathbb{R})/\mathbf{SL}_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times \quad (5.5.2)$$

Indeed, $\mathbf{SL}_n(\mathbb{R}) = \ker(\det)$ and \det is a surjective group homomorphism.

2. For $n \geq 2$, $S_n \xrightarrow{\text{sign}} \{-1, 1\}$ induces the isomorphism

$$S_n/A_n \xrightarrow{\text{sign}} \mathbb{R}^\times \quad (5.5.3)$$

3. The map $\mathbb{C}^\times \xrightarrow{|\cdot|} \mathbb{R}^\times$ is a homomorphism of image $\mathbb{R}_{>0}$ and $\ker(|\cdot|) = \{z \in \mathbb{C}^\times : |z| = 1\} = S^1$ the circle group. Thus, we have the isomorphism

$$\mathbb{C}^\times/S^1 \xrightarrow{|\cdot|} \mathbb{R}_{>0} \quad (5.5.4)$$

Corollary 5.5.1.

1. The **corestriction** of \bar{f} to $\text{Im}(f)$ is an isomorphism

$$\bar{f} : G/N \xrightarrow{\sim} f(G) = \text{Im}(f) \quad (5.5.5)$$

2. For G, G' , finite groups,

$$|G| = |\ker(f)||G/N| = |\ker(f)||\operatorname{Im}(f)| \quad (5.5.6)$$

so $|\ker(f)| \mid |G|$ while $|\operatorname{Im}(f)| \mid |G|$ and $|\operatorname{Im}(f)| \mid |G'|$.

Proof. (Left to the reader) ■

Proposition 5.5.2. Let $\varphi : G \rightarrow G'$ be an epimorphism of groups. If $N \triangleleft G$, then $\varphi(N) \triangleleft G'$.

Proof. (Left to the reader) ■

Theorem 6 (Correspondence Theorem).

Let $\varphi : G \rightarrow G'$ be an epimorphism of groups. Then the preimage by φ induces the bijection

$$\begin{aligned} \{H \leq G : \ker(\varphi) \leq H\} &\leftrightarrow \{H' \leq G'\} \\ H &\mapsto \varphi(H) \\ \varphi^{-1}(H') &\leftarrow H' \end{aligned} \quad (5.5.7)$$

which preserves normality of subgroups. If $H' \leq G'$ and $H \leq G$ are finite groups and correspond to each other, then $|H| = |\ker(\varphi)||H'|$.

Proof. We know that $\varphi^{-1}(H')$ and $\varphi(H)$ are subgroups (respectively, normal subgroups as φ is surjective) if H' and H are. Moreover, $\varphi^{-1}(H') \supseteq \ker(\varphi)$ since $e_{G'} \in H'$. Now, as φ is surjective, $\varphi(\varphi^{-1}(H')) = H'$, and let us show $\varphi^{-1}(\varphi(H)) = H$ if $H \supseteq \ker(\varphi)$. Note

$$\varphi^{-1}(\varphi(H)) \supseteq H$$

by definition. Then let $g \in G$ such that $\varphi(g) \in \varphi(H)$. But then $\varphi(g) = \varphi(h)$ for some $h \in H$, so $\varphi(gh^{-1}) = \varphi(g)\varphi(h)^{-1} = e_{G'}$, which implies $gh^{-1} \in \ker(\varphi)$. Since $\ker(\varphi) \subseteq H$, $gh^{-1} \in H$ so $g = (gh^{-1})h \in H$. Hence

$$\varphi^{-1}(\varphi(H)) \subseteq H$$

so both inclusions hold and $\varphi^{-1}(\varphi(H)) = H$. Thus the correspondence is indeed a bijection. Finally, we have $\varphi|_H : H \rightarrow H' = \varphi(H)$, where $\varphi|_H = \varphi \circ \iota_H$ is a surjective homomorphism. Then, by the corollary of the First Isomorphism Theorem we have that

$$|H| = |\ker(\varphi|_H)||\operatorname{Im}(\varphi|_H)| = |\ker(\varphi)||H'|$$

completing the proof. ■

Remark 5.5.1. The correspondence preserves containment and intersections. That is

$$H' \subseteq K' \iff \varphi^{-1}(H') \subseteq \varphi^{-1}(K') \quad (5.5.8)$$

and

$$\varphi^{-1}(H'_1 \cap H'_2) = \varphi^{-1}(H'_1) \cap \varphi^{-1}(H'_2) \quad (5.5.9)$$

Note that we can apply the correspondence theorem to the canonical epimorphism $\pi : G \rightarrow G/N$ for $N \triangleleft G$, to get a correspondence between subgroups of G/N and subgroups of G containing N .

Example 5.5.2.

1. $\mathbb{Z}/2\mathbb{Z} \cong \{-1, 1\}$ has only two subgroups, $\mathbb{Z}/2\mathbb{Z}$ and $\{1\}$. So, if $A_n \leq H \leq S_n$, then $H = A_n$ or $H = S_n$ because $\text{sign} : S_n \rightarrow \{-1, 1\}$ is an epimorphism and $\ker(\text{sign}) = A_n$.
2. Similarly, for $p\mathbb{Z} \leq \mathbb{Z}$ for p a prime number, if $p\mathbb{Z} \leq H \leq \mathbb{Z}$, then $H = p\mathbb{Z}$ or $H = \mathbb{Z}$ because

$$\begin{aligned} \mathbb{Z} &\rightarrow \mathbb{Z}/p\mathbb{Z} \\ n &\mapsto [n] \end{aligned} \tag{5.5.10}$$

is a surjective group homomorphism, and $\mathbb{Z}/p\mathbb{Z}$ has no proper subgroups.

3. For G finite, let N be a proper normal subgroup which is maximal with respect to normality. That is, if $N \leq N' \triangleleft G$ then $N = N'$ or $N' = G$. By the Correspondance Theorem it follows that G/N is a simple group. An example is $\mathbb{Z}/p\mathbb{Z}$ for a prime p .

Chapter 6

§§Group Actions

6.1.0 §Basic Definitions and Examples: Group Actions

Remark 6.1.1 (Motivation). We want to view elements of a group G as symmetries of a set X . In particular, for every $g \in G$ we want to associate $X \xrightarrow{\alpha_g} X$ a bijection, with $\alpha_{gh} = \alpha_g \circ \alpha_h$ and $\alpha_{e_G} = \text{Id}_X$.

Example 6.1.1.

1. For X_n , the regular convex n -gon, and $G = D_n$, for all $g \in D_n$ we get a bijection $X_n \xrightarrow{g} X_n$.
2. For $X = \{1, 2, \dots, n\}$, and $G = S_n$, for all $\sigma \in S_n$ we get a permutation $X \xrightarrow{\sigma} X$.
3. For $X = \mathbb{R}^n$, and $G = \mathbf{GL}_n(\mathbb{R})$, for all $A \in \mathbf{GL}_n(\mathbb{R})$ we get a bijection

$$\begin{aligned} \mathbb{R}^n &\xrightarrow{L_A} \mathbb{R}^n \\ \vec{v} &\mapsto A\vec{v} \end{aligned} \tag{6.1.1}$$

4. For $X = G$ a group, we have the action by left multiplication, where for all $g \in G$ we get the bijection

$$\begin{aligned} G &\xrightarrow{\ell_g} G \\ x &\mapsto gx \end{aligned} \tag{6.1.2}$$

5. For $X = G$ a group, and $H \leq G$, for all $h \in H$ we have the bijection

$$\begin{aligned} G &\xrightarrow{\ell_h} G \\ g &\mapsto hg \end{aligned} \tag{6.1.3}$$

6. For $X = G$ a group, we have the action by conjugation, where for all $g \in G$, we have the bijection

$$\begin{aligned} G &\xrightarrow{\beta_g} G \\ x &\mapsto gxg^{-1} \end{aligned} \tag{6.1.4}$$

Remark 6.1.2. There are two equivalent ways to formalize the notion of group actions.

Definition 6.1.1 (Group Action). A group action of a group G on a set X is a

1. group homomorphism

$$\begin{aligned} \alpha : G &\rightarrow S_X \\ g &\mapsto \alpha_g \end{aligned} \tag{6.1.5}$$

2. map $a : G \times X \rightarrow X$ such that

- (a) $a(e_G, x) = x$ for all $x \in X$
- (b) $a(gh, x) = a(g, a(h, x))$ for all $g, h \in G$ and all $x \in X$.

Definition 6.1.2. Let G be a group acting on a set X . The data of 1. (or equivalently 2.) in the previous definitions is called an action of G on X and X is called a G -set.

Claim 6.1.1. Definitions 1. and 2. of a group action are equivalent. That is, for any group G and any nonempty set A there is a bijection between the actions of G on A and the group homomorphisms of G into S_A .

Proof. (Left to the reader) ■

Definition 6.1.3. Let G be a group acting on a nonempty set A . Then the homomorphism $\alpha : G \rightarrow S_A$ associated with the action of G on A is called a permutation representation associated to the given action. We say a given action of G on A affords or induces the associated permutation representation of G .

Definition 6.1.4. The kernel of the group homomorphism associated with a group action is $\ker(\alpha) = \{g \in G : \alpha(g) = \text{Id}_X\}$, or equivalently for definition 2. the set $\{g \in G : \forall x \in X, a(g, x) = x\}$. If $\ker(\alpha) = \{e_G\}$, then the action of G on X is said to be faithful.

Remark 6.1.3. Two group elements in G induce the same permutation of the set A if and only if they exist in the same coset of the kernel of the action (i.e., if and only if they are in the same fiber of the permutation representation α).

Moreover, the inherited action of the quotient space $G/\ker(\alpha)$ on A is faithful.

Example 6.1.2.

1. For the left action of G on G , $\ker = \{g \in G : \ell_g = \text{Id}_G\}$. We want $\ell_g(h) = h$ for all $h \in G$, where $\ell_g(h) = gh$, so $g = e_G$. Thus, $\ker = \{e_G\}$, so the action is faithful.
2. For the conjugation action, with the associated group homomorphism

$$\begin{aligned} \beta : G &\rightarrow S_G \\ g &\mapsto \begin{aligned} \beta_g : G &\rightarrow G \\ x &\mapsto gxg^{-1} \end{aligned} \end{aligned} \tag{6.1.6}$$

We claim that $\ker(\beta) = Z(G)$, the center of G .

Definition 6.1.5. The permutation representation afforded by the left multiplication action on the elements of the group G is called the left regular representation of G .

Theorem 7 (Cayley's Theorem).

Every group G is isomorphic to a subgroup of its group of symmetries $\text{Sym}(G)$.

Proof. (Left to the reader) ■

Notation 6.1.4. If $G \times Y \xrightarrow{a} Y$ is a group action, we denote $a(g, y)$ by $g.y$ (or even gy if there is no confusion). Note that we have $(gh).y = g.(h.y)$ and $e_G.y = y$ for all $g, h \in G$ and all $y \in Y$. Moreover, to say that a group G acts on a set Y we write $G \curvearrowright Y$.

Definition 6.1.6. Let $G \times Y \xrightarrow{a} Y$ be an action and let $y \in Y$.

1. The orbit of y under the action by G is the set $O_y = \{g.y : g \in G\} \subseteq Y$ (also denoted $G.y$)
2. The stabilizer of y under the action by G is the set $G_y = \{g \in G : g.y = y\} \subseteq G$.
3. $y \in Y$ is called a fixed point of the action if $G_y = G$, so for all $g \in G$, $g.y = y$.

Definition 6.1.7. Let G be a group and A a nonempty set. The action of G on A is said to be transitive if there is only one orbit, i.e., given any two elements $a, b \in A$, there exists $g \in G$ such that $b = g.a$.

Proposition 6.1.2. For $G \curvearrowright Y$ and all $y \in Y$, $G_y \leq G$.

Proof. (Left to the reader) ■

Example 6.1.3.

1. For the left multiplication action, $G \curvearrowright G$, for all $g \in G$ we have orbits $G.g = O_g = G$, and stabilizers $G_g = \{e_G\}$.
2. For the conjugation action, $G \curvearrowright G$, for all $g \in G$ we have orbits

$$G.g = O_g = \{a \in G : \exists h \in H, hgh^{-1} = a\} \quad (6.1.7)$$

These sets are called the conjugacy classes of G . The stabilizers of the action are

$$G_g = \{a \in G : aga^{-1} = g\} = Z(g) \quad (6.1.8)$$

the centralizer of g in G .

3. The left multiplication action of a subgroup $H \leq G$ on G , $H \curvearrowright G$, for all $g \in G$ the orbit is $H.g = Hg$ the right coset of H . Moreover, the stabilizers still are $H_g = \{e_G\}$.

Lemma 6.1.3. Let $a : G \times Y \rightarrow Y$ be a group action. Then

1. The orbits $G.y = O_y$ of the action form a partition of Y .
2. For all $y \in Y$, the order of the orbit $|G.y| = |O_y|$, is the index $|G : G_y|$ of the stabilizer G_y of y in G . (**Orbit Stabilizer Theorem**)

1. First, note that $y = e_G.y \in G.y$ for all $y \in Y$, so

$$Y = \bigcup_{y \in Y} G.y \quad (6.1.9)$$

Next, let $y, y' \in Y$ and suppose $g.y = g'.y' \in G.y \cap G.y'$, for some $g, g' \in G$. Then we have that $y = g^{-1}.(g'.y') = (g^{-1}g').y' \in G.y'$, so for all $h.y \in G.y$, $h.y = (hg^{-1}g').y' \in G.y'$ so $G.y \subseteq G.y'$. Similarly we have that $G.y \supseteq G.y'$, so $G.y = G.y'$. Hence, the orbits indeed partition Y .

[2.] Define a map

$$\begin{aligned} G/G_y &\xrightarrow{f} G.y \\ aG_y &\mapsto a.y \end{aligned} \quad (6.1.10)$$

where G/G_y denotes the set of left cosets of G_y (not necessarily a group). First, to show the map is well defined suppose $aG_y = bG_y$. Then we have that $a = bg$ for some $g \in G_y$. It follows that

$$a.y = (bg).y = b.(g.y) = b.y$$

since $g \in G_y$, so $f(aG_y) = f(bG_y)$ and the map is well defined. Now suppose $aG_y, cG_y \in G/G_y$ such that $a.y = c.y$. Then $(c^{-1}a).y = c^{-1}.(c.y) = e_G.y = y$, which implies $c^{-1}a \in G_y$. This implies by coset equality that $aG_y = cG_y$, so f is an injection. Finally, if $g.y \in G.y$, we have that $f(gG_y) = g.y$, so f is a surjection. Therefore f is a bijection and we conclude that

$$|G : G_y| = |G/G_y| = |G.y| \quad (6.1.11)$$

as claimed. ■

§Application to Cycle Decompositions

Using the tools we have developed with group actions, we can provide an alternate proof to the fact that any permutation $\sigma \in S_n$ can be decomposed into disjoint cycles.

Proof. Let $A = \{1, 2, \dots, n\}$, let $\sigma \in S_n$, and let $G = \langle \sigma \rangle$. Then consider the action of $\langle \sigma \rangle$ on A . Let O be one of the orbits of this action, and let $x \in O$. Note that there exists a bijection between the elements of O and the left cosets of the stabilizer G_x in G , given explicitly by

$$\sigma^i x \mapsto \sigma^i G_x$$

Since G is cyclic it is abelian, so $G_x \trianglelefteq G$ and G/G_x is cyclic of order d , where d is the order of σ in G/G_x , in particular it is the smallest positive integer for which $\sigma^d \in G_x$. Also, $d = |G : G_x| = |O|$. Thus, the distinct cosets of G_x in G are

$$1G_x, \sigma G_x, \sigma^2 G_x, \dots, \sigma^{d-1} G_x$$

This shows that the distinct elements of \mathcal{O} are $x, \sigma(x), \dots, \sigma^{d-1}(x)$ by our bijection. Ordering the elements of \mathcal{O} in this manner shows that σ cycles the elements of \mathcal{O} , that is, on an orbit of size d , σ acts as a d -cycle. This proves the existence of a cycle decomposition for each $\sigma \in S_n$.

The orbits of $\langle \sigma \rangle$ are uniquely determined by σ . The only choice is in the order the orbits are listed in, which depends on our initial representative from \mathcal{O} . It follows that the cycle decomposition above is unique up to a rearrangement of the cycles and up to a cyclic permutation of the integers within each cycle. ■

6.2.0 §Counting and Combinatorial Formulas

Theorem 8 (Counting Formula).

Suppose $G \curvearrowright Y$. Then for all $y \in Y$ we have the counting formula

$$|G| = |G_y| |G.y| \quad (6.2.1)$$

Proof. (Left to the reader) ■

Theorem 9 (Orbit Decomposition Theorem).

Let Y be a finite set with $G \curvearrowright Y$. Let $Y_f \subseteq Y$ denote the set of fixed points of Y under the action. Let $G.y_1, \dots, G.y_n$ be the distinct non-singular orbits of Y for some integer $n \geq 0$. Then

$$|Y| = |Y_f| + \sum_{i=1}^n |G : G_{y_i}| \quad (6.2.2)$$

Proof. (Left to the reader) ■

Corollary 6.2.1 (Class Equation). For a group G and the conjugation action $G \curvearrowright G$, we have the class equation

$$|G| = \sum_{\text{Conjugacy Classes } C} |C| \quad (6.2.3)$$

Remark 6.2.1. By the counting formula we have that $|C| \mid |G|$ for all conjugacy classes C .

Proposition 6.2.2. The set of fixed elements of the conjugation action of G on G , $G \curvearrowright G$, is the center of G , $Z(G)$.

Proof. (Left to the reader) ■

Proposition 6.2.3. If $H \triangleleft G$, then H is a union of conjugacy classes. Indeed, for all $h \in H$ and all $g \in G$, $g.h = ghg^{-1} \in H$, so $G.h \subseteq H$.

Proof. (Left to the reader) ■

Example 6.2.1.

1. For an abelian group we get the class equation

$$|G| = 1 + 1 + \dots + 1 \quad (6.2.4)$$

2. For $D_3 (\cong S_3)$, $|D_3| = 6$, and $D_3 = \langle x, y \rangle$. First, note that $|G.x| = [D_3 : G_x] = [D_3 : Z(x)] = 6/3 = 2$, where $\langle x \rangle \subseteq Z(x)$ while $y \notin Z(x)$ as $xyx = x^2$, so by 2 $|Z(x)| = 3$ and $Z(x) = \langle x \rangle$. Similarly, $Z(y) = \langle y \rangle$ as $xyx^2 = x^2y \neq y$, so $x, x^2 \notin Z(y)$, so $|G.y| = [D_3 : Z(y)] = 6/2 = 3$. Thus, we have that

$$|D_3| = |D_3.e| + |D_3.x| + |D_3.y| = 1 + 2 + e \quad (6.2.5)$$

3. The class equation of A_5 is

$$|A_5| = 1 + 20 + 12 + 12 + 15 \quad (6.2.6)$$

First note that $|A_5| = \frac{5!}{2} = 60$, and A_5 is composed all even permutations in S_5 . First, $1 = |\{(1)\}| = A_5.(1)|$. Then, for the three cycles in A_5 , there are $\frac{5*4*3}{3} = 20$ three cycles. In S_5 all three cycles are in the same conjugacy class. It follows that for any three cycle $\sigma \in A_5$, $|S_5.\sigma| = 20 = |S_5 : S_{5\sigma}|$, which implies $|S_{5\sigma}| = 6$. Note $\langle \sigma \rangle \subseteq S_{5\sigma}$, but also for the $1 \leq i, j \leq 5$ not moved by σ , $(i \ j), \sigma(i \ j), \sigma^{-1}(i \ j) \in S_{5\sigma}$. But, $(i \ j), \sigma(i \ j), \sigma^{-1}(i \ j) \notin A_5$ as they are odd, but $\langle \sigma \rangle \subseteq A_5$ so we have that $|A_{5\sigma}| = 3$. Thus, $|A_5.\sigma| = [A_5 : A_{5\sigma}] = 60/3 = 20$, which is the entire conjugacy class. Next, for pairs of disjoint transpositions we have $\frac{5*4*3*2}{2*2*2} = 15$, which are all in the same conjugacy class in S_5 . Thus, for a transposition pair τ we have $|S_5.\tau| = 15 = [S_5 : S_{5\tau}]$, so $|S_{5\tau}| = 120/15 = 8$. Individual transpositions of the pair are in the centralizer, but they are not in A_5 as they are odd permutations, so $|A_{5\tau}| \leq 6$. Moreover, the two four cycles composed of the pair of transpositions adjoined are in the centralizer while not being in A_5 , so $|A_{5\tau}| \leq 4$. But, note that if $|A_{5\tau}| \leq 3$ then $|A_5.\tau| \geq 20$, but there are only 15 elements in the conjugacy class of pairs of transpositions in S_5 . Thus, we have that $|A_{5\tau}| = 4$ so $|A_5.\tau| = |A_5 : A_{5\tau}| = 60/4 = 15$. Finally, we have $5!/5 = 24$ five cycles in A_5 . Let α be one such five cycle. Then $|S_5.\alpha| = 24 = [S_5 : S_{5\alpha}]$, which implies that $|S_{5\alpha}| = 5$. But $\langle \alpha \rangle \subseteq S_{5\alpha}$ and $\langle \alpha \rangle \subseteq A_5$ with $|\langle \alpha \rangle| = 5$, so $S_{5\alpha} = A_{5\alpha}$. Thus, we have that $|A_5.\alpha| = [A_5 : A_{5\alpha}] = 60/5 = 12$. Therefore, the conjugacy class is split in two, and we have arrived at our class equation.

Corollary 6.2.4. A_5 is a simple group.

Proof. Let $N \triangleleft A_5$. We must have that $|N| \mid |A_5| = 60$ and $|N| = \sum |C|$ for conjugacy classes C of A_5 , and $\{(1)\}$ is one of them. But by the class equation the only possibilities are $|N| = 1$ and $|N| = 60$. Thus, $N \triangleleft A_5$ implies either $N = \{(1)\}$ or $N = A_5$, so A_5 is a simple group by definition. ■

Proposition 6.2.5. Let $|G| = p^n$ for a prime p (such a group G is called a **p-group**). The center of G is not the trivial subgroup $\{e_G\}$.

Proof. The class equation of G is

$$p^n = |G| = 1 + \sum_i |C_i| \quad (6.2.7)$$

Note $g \in Z(G)$ if and only if $G.g = \{g\}$. So, if $Z(G) = \{e_G\}$ then $|C_i| > 1$ for all i . Thus, $p \mid |C_i|$ for all i since $1 < |C_i| \mid |G| = p^n$. Hence,

$$p \mid \left(|G| - \sum_i |C_i| \right) = |C_1| = 1 \quad (6.2.8)$$

which is a contradiction as this implies $p = 1$, and 1 is not prime. So, $|Z(G)| > 1$ as claimed. ■

Theorem 10 (Cauchy's Theorem).

Let G be a finite group. If $p \mid |G|$ for p a prime, then G has an element of order p .

Proof. We want to define an action on the set

$$G^p = \prod_{i=1}^p G$$

by a cyclic group of order p ($\cong \mathbb{Z}/p\mathbb{Z}$). Let $H = \langle \sigma \rangle$ for $o(\sigma) = p$, and we define the action by

$$\sigma.(g_1, g_2, \dots, g_p) := (g_2, g_3, \dots, g_p, g_1) \quad (6.2.9)$$

This gives a well-defined group action of H on G^p . Moreover, $x \in G^p$ is a fixed point of the action if and only if

$$(g_1, g_2, \dots, g_p) = (g_2, g_3, \dots, g_p, g_1) \quad (6.2.10)$$

so $g_1 = g_2 = \dots = g_p$. We are interested in a subset $Y \subseteq G^p$ defined by

$$Y := \{(g_1, g_2, \dots, g_p) \in G^p : g_1 g_2 \dots g_p = e\} \quad (6.2.11)$$

We see that for all $y \in Y$, $H.y \subseteq Y$. Indeed, if $g_1 \dots g_p = e$, then $e = g_1^{-1} g_1 = g_2 \dots g_p g_1$, which is associated to $\sigma.(g_1, \dots, g_p)$. Hence, if $y \in Y$, then $\sigma.y \in Y$. Thus, we have an action of H on Y taken by corestriction. Next, $|Y| = |G|^{p-1}$. Indeed, choose g_1, g_2, \dots, g_{p-1} freely, which constitutes $|G|^{p-1}$ choices, then choose $G_p = (g_1 g_2 \dots g_{p-1})^{-1}$. Then we have that $(g_1, \dots, g_p) \in Y$. Note that $y \in Y$ is fixed by H , for $y = (g_1, \dots, g_p)$, if and only if $g = g_1 = \dots = g_p$, and $g_1 \dots g_p = g^p = e$. Then, $o(g) \in \{1, p\}$. Note that $(e, \dots, e) \in Y$ is a fixed point. We want to show that $|Y_f| > 1$. Applying the Orbit Decomposition Theorem to the action of H on Y we have that

$$|G|^{p-1} = |Y| = |Y_f| + \sum_{i=1}^n |H : H_{y_i}| \quad (6.2.12)$$

where each $|H : H_{y_i}|$ divides $|H| = p$ and is greater than 1 as they are not fixed points, so in particular $|H : H_{y_i}| = p$ for each i . Thus,

$$|Y_f| = |G|^{p-1} - \sum_{i=1}^n |H : H_{y_i}| \quad (6.2.13)$$

which p divides as $p \mid |G|$ by our initial assumption. Then, $p \mid |Y_f|$ so in particular $|Y_f| > 1$. Hence, there exists $(g, g, \dots, g) \in Y_f$ such that $g \neq e$ and $g^p = e$ as desired. ■

6.3.0 §Conjugacy Actions and Actions on Subgroups

In this section we redefine and make precise certain concepts previously mentioned off hand in examples in relation to conjugation and subgroup actions.

Theorem 6.3.1. *Let G be a group, let H be a subgroup of G , and let G act by left multiplication on the set A of left cosets of H in G . Let π_H be the associated permutation representation afforded by this action. Then*

1. G acts transitively on A
2. the stabilizer in G of the point $1H \in A$ is the subgroup H
3. the kernel of the action (i.e., the kernel of π_H) is $\bigcap_{x \in G} xHx^{-1}$, and $\ker \pi_H$ is the largest normal subgroup of G contained in H .

Proof. (Left to the reader) ■

Definition 6.3.1. *Let G be a group and define $\text{Sub}(G)$ as the set of all subgroups of G . Then G acts on the set $\text{Sub}(G)$ by*

$$a : G \times \text{Sub}(G) \rightarrow \text{Sub}(G); (g, H) \mapsto g \cdot H := gHg^{-1}$$

where

$$gHg^{-1} := \{ghg^{-1} : h \in H\}$$

Observe that the orbits of the action partition the subgroups of G into conjugacy classes.

Definition 6.3.2. *Let $g \in G$, where G is a group. The stabilizer of g under the conjugation action of G on itself is equal to*

$$Z(g) = \{h \in G : hgh^{-1} = g\}$$

and is called the centralizer of g . It follows that the center of G is equal to

$$Z(G) = \bigcap_{g \in G} Z(g)$$

Definition 6.3.3. *Let $g \in G$. The equivalence class of g with respect to the equivalence relation coming from the conjugation action of G on itself is called the conjugacy class of g in G , sometimes denoted $C(g)$; thus*

$$C(g) := \{g' \in G : \exists h \in G; hgh^{-1} = g'\}$$

Definition 6.3.4. *Let $H \leq G$ be a subgroup of G . The stabilizer of H under the action of conjugation on $\text{Sub}(G)$ is*

$$N_G(H) = \{g \in G : gHg^{-1} = H\}$$

and it is called the normalizer of H in G .

Definition 6.3.5. Let G be a group and $S \subseteq G$ a subset of G . Let $g \in G$ and define $gSg^{-1} := \{gsg^{-1} : s \in S\}$. Then G acts on its power set $\mathcal{P}(G)$ of all subsets of itself by defining $g \cdot S = gSg^{-1}$ for any $g \in G$ and $S \in \mathcal{P}(G)$.

Definition 6.3.6. Two subsets S and T of G are said to be conjugate in G if there is some $g \in G$ such that $T = gSg^{-1}$.

Proposition 6.3.2. The number of conjugates of a subset S in G is the index of the normalizer of S , $|G : N_G(S)|$. Moreover, the number of conjugates of an element s of G is the index of the centralizer of s , $|G : Z(s)|$.

Theorem 11 (Class Equation (Alternate Form)).

Let G be a finite group and let g_1, g_2, \dots, g_r be representatives of the distinct conjugacy classes of G not contained in the center $Z(G)$ of G . Then

$$|G| = |Z(G)| + \sum_{i=1}^r |G : Z(g_i)|$$

Proof. Note that for $x \in G$, the conjugacy class of x is the singleton $\{x\}$ if and only if $x \in Z(G)$, since then $gxg^{-1} = x$ for all $g \in G$. Let $Z(G) = \{1, z_2, \dots, z_m\}$, let $\mathcal{K}_1, \dots, \mathcal{K}_r$ be the conjugacy classes of G not contained in the center, and let g_i be a representative of \mathcal{K}_i for each i . Then the full set of conjugacy classes of G is given by

$$\{1\}, \{z_2\}, \dots, \{z_r\}, \mathcal{K}_1, \dots, \mathcal{K}_r$$

Since these partition G we have

$$\begin{aligned} |G| &= \sum_{i=1}^m 1 + \sum_{i=1}^r |\mathcal{K}_i| \\ &= |Z(G)| + \sum_{i=1}^r |G : Z(g_i)| \end{aligned}$$

This proves the class equation. ■

Theorem 6.3.3. If p is a prime and P a group of prime power order p^α for some $\alpha \geq 1$, then P has a nontrivial center: $Z(P) \neq \{1\}$.

Proof. By the class equation

$$|P| = |Z(P)| + \sum_{i=1}^r |P : Z(g_i)|$$

By definition $Z(g_i) \neq P$ for $i \in \{1, 2, \dots, r\}$, so p divides $|P : Z(g_i)|$. Since p also divides $|P|$ it follows that p must divide $|Z(P)|$, hence the center must be nontrivial. ■

Corollary 6.3.4. If $|P| = p^2$ for some prime p , then P is abelian. More precisely, P is isomorphic to either $\mathbb{Z}/p^2\mathbb{Z}$ or $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

Proof. Since $|Z(P)| \neq 1$ by the previous theorem, it follows that $|Z(P)| \in \{p, p^2\}$. Thus $|P/Z(P)| \in \{1, 2\}$, so $P/Z(P)$ is cyclic. Let $P/Z(P) = \langle gZ(P) \rangle$ for some $g \in P$. Let $x, y \in P$, then since $P/Z(P)$ is cyclic there exist $n, m \in \mathbb{Z}$ such that $xZ(P) = g^n Z(P)$ and $yZ(P) = g^m Z(P)$. That is, there exist $z, z' \in Z(P)$ such that $x = g^n z$ and $y = g^m z'$. It follows that

$$xy = g^n z g^m z' = g^{n+m} z z' = g^m g^n z' z = g^m z' g^n z = yx$$

However, x, y were arbitrary elements of P so P must be abelian. Hence $Z(P) = P$. If P has an element of order p^2 , then P is cyclic and $P \cong \mathbb{Z}/p^2\mathbb{Z}$. Assume therefore that every nonidentity element of P has order p . Let x be a non-identity element of P and let $y \in P \setminus \langle x \rangle$. Since $|\langle x, y \rangle| > |\langle x \rangle| = p$, we must have that $P = \langle x, y \rangle$. Both x and y have order p so $\langle x \rangle \times \langle y \rangle \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. It now follows directly that the map $(x^a, y^b) \mapsto x^a y^b$ is an isomorphism from $\langle x \rangle \times \langle y \rangle$ onto P . This completes the proof. ■

§Conjugation in Special groups

Remark 6.3.1. Note that in the matrix group $\mathbf{GL}_n(\mathbb{F})$, conjugation is equivalent to a change of basis: $A \mapsto PAP^{-1}$. An analogous situation arises in S_n .

Proposition 6.3.5. Let σ and τ be elements of the symmetric group S_n , and suppose σ has cycle decomposition

$$(a_1 a_2 \dots a_{k_1})(b_1 b_2 \dots b_{k_2})\dots$$

Then $\tau\sigma\tau^{-1}$ has cycle decomposition:

$$(\tau(a_1) \tau(a_2) \dots \tau(a_{k_1}))(\tau(b_1) \tau(b_2) \dots \tau(b_{k_2}))\dots$$

Proof. (Left to the reader) ■

Definition 6.3.7.

1. If $\sigma \in S_n$ is the product of disjoint cycles of lengths n_1, n_2, \dots, n_r with $n_1 \leq n_2 \leq \dots \leq n_r$ (including its 1-cycles) then the integers n_1, n_2, \dots, n_r are called the **cycle type** of σ .
2. If $n \in \mathbb{Z}^+$, a partition of n is any nondecreasing sequence of positive integers whose sum is n .

Proposition 6.3.6. Two elements of S_n are conjugate in S_n if and only if they have the same cycle type. The number of conjugacy classes in S_n equals the number of partitions of S_n .

Proof. (Left to the reader - hint: D&F p.126) ■

§Right Group Actions

Definition 6.3.8. Let G be a group and define the **right group action** of G on a nonempty set A as a map from $A \times G$ to A , denoted by $a \cdot g$ for $a \in A$ and $g \in G$, that satisfies the axioms:

1. $(a \cdot g_1) \cdot g_2 = a \cdot (g_1 g_2)$ for all $a \in A$, and $g_1, g_2 \in G$, and
2. $a \cdot 1 = a$ for all $a \in A$.

Remark 6.3.2. Conjugation as a write group action is denoted by $a \cdot g = g^{-1}ag$, and it is sometimes notated with $a \cdot g = a^g$.

6.4.0 §P-Groups

Definition 6.4.1. If p is a prime, a group G is called a **p-group** if and only if the order of every element of G is a power of p .

Lemma 6.4.1. If G is a finite group and p is a prime, then $|G|$ is a power of p if and only if G is a p -group.

Proof. Firstly, suppose $|G|$ is a power of p . Then for all $g \in G$, $o(g) \mid |G|$ by 2, so in particular, $o(g) \mid p^n$ for some $n \in \mathbb{N}$, so $o(g) = p^k$ for $k \in \{0, 1, \dots, n\}$. Thus, G is a p -group.

Conversely, suppose for all $g \in G$ $o(g)$ is a power of p . Assume towards a contradiction that $|G|$ is not a power of p . Then by prime factorization there exists a prime $q \neq p$ such that $q \mid |G|$. But, by 10 it follows that G must have an element of order q , contrary to our assumption. Hence, $|G|$ is a power of p . ■

Theorem 6.4.2. Let $K \triangleleft G$ be groups and let p be a prime. Then G is a p -group if and only if both K and G/K are p -groups.

Proof. Firstly suppose G is a p -group. Then K is a p -group as every element of K is an element of G , and hence has order a power of p . Consider the canonical projection $\pi : G \rightarrow G/K$. It follows that for all $gK \in G/K$, $(gK)^{o(g)} = \pi(g)^{o(g)} = \pi(e) = K$, where $o(g)$ is a power of p by assumption. It follows that $o(gK) \mid o(g)$, which is a power of p , so $o(gK)$ is a power of p . Consequently we conclude that G/K is also a p -group.

Conversely, suppose K and G/K are both p -groups. Let $g \in G$ and consider $gK \in G/K$. Then $o(gK) = p^l$ for some nonnegative integer l . Thus $(gK)^{p^l} = g^{p^l}K = K$, which implies $g^{p^l} \in K$. But K is a p -group, so $o(g^{p^l}) = p^r$ for some nonnegative integer r . Consequently, we find

$$e = (g^{p^l})^{p^r} = g^{p^{l+r}}$$

Thus, we have that $o(g) \mid p^{l+r}$, so $o(g)$ is a power of p . As g was an arbitrary element of G , we find that G is indeed itself a p -group. ■

Theorem 6.4.3. If p is a prime and $G \neq \{1\}$ is a finite p -group, then $Z(G) \neq \{1\}$.

Proof. Let a_1, \dots, a_n be representatives of the nonsingleton conjugacy classes in G . Because $1 \notin N(a_i)$ for each i , $N(a_i) \neq G$, and since $|G : N(a_i)|$ divides $|G|$ by 2, it follows that p divides $|G : N(a_i)|$ for each i . But then p must divide $|Z(G)|$ by the class equation; in particular $Z(G) \neq \{1\}$. ■

Theorem 6.4.4. *If G is a group and $|G| = p^2$ where p is a prime, then G is abelian and either $G \cong \mathbb{Z}/p^2\mathbb{Z}$ or $G \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.*

Proof. By our previous result we know that $|Z(G)| \in \{p, p^2\}$. We aim to show that $|Z(G)| = p$ is impossible. Indeed, if it holds then $G/Z(G)$ is cyclic (being of order p), which implies G is abelian and hence $|Z(G)| = p^2$, a contradiction. Thus $|Z(G)| = p^2$, so $Z(G) = G$ and G is abelian. Now, if G is cyclic $G \cong \mathbb{Z}/p^2\mathbb{Z}$ and we're done, so suppose to the contrary. Then for every $g \in G$ we have $g^p = 1$. Choose $1 \neq a \in G$ and write $H = \langle a \rangle$. Then choose $b \notin H$ and write $K = \langle b \rangle$. Because $|K| = p = |H|$, we have $H \cap K = \{1\}$, and consequently $|HK| = |H||K| = p^2$. Thus $|HK| = p^2 = |G|$, so $G = HK \cong H \times K \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. ■

Theorem 6.4.5. *Let G be a finite p -group of order p^n . Then there exists a series*

$$G = G_0 \supset G_1 \supset \dots \supset G_n = \{1\}$$

of subgroups of G such that $G_i \triangleleft G$, $|G_i| = p^{n-i}$, and $|G_i/G_{i+1}| = p$ for all $i \in \{0, 1, \dots, n-1\}$.

Proof. If $n = 1$ then the statement is immediate, and we now proceed by induction on n . Suppose there exists $k \geq 1$ such that if $n = k$, then the hypothesis holds for all groups G of order $|G| = p^k$. Now, consider $|G| = p^{k+1}$. From our previous results $Z(G) \neq \{1\}$. Moreover, by 10, choose $a \in Z(G)$ such that $o(a) = p$, and write $G_k = \langle a \rangle$. Then $G_k \triangleleft G$ and G/G_k has order p^k so, by induction, let $(G/G_k) \supset X_1 \supset \dots \supset X_k = \{G_k\}$ be a series of subgroups of G/G_k such that $X_i \triangleleft G/G_k$ and $|X_i/X_{i+1}| = p$ for each i . By 6 we have that each X_i has the form $X_i = G_i/G_k$, where $G_i \triangleleft G$ and $|G_i/G_k| = p^{k-i}$. Furthermore, $X_i \supset X_{i+1}$ implies $G_i \supset G_{i+1}$, and $G_i/G_{i+1} \cong X_i/X_{i+1}$ by the third isomorphism theorem. Hence, $G \supset G_1 \supset \dots \supset G_k \supset \{1\}$ is the required series for G . Thus, by Mathematical Induction we conclude that the proposition holds for all $n \geq 1$. ■

6.5.0 §Sylow's Theorem

Definition 6.5.1. *Let G be a group and let p be a prime.*

1. *A group of order p^α for some $\alpha \geq 1$ is called a p -group. Subgroups of G which are p -groups are called p -subgroups.*
2. *If G is a group of order p^α , where $p \nmid m$, then a subgroup of order p^α is called a Sylow p -subgroup of G .*
3. *The set of Sylow p -subgroups of G will be denoted by $\text{Syl}_p(G)$ and the number of Sylow p -subgroups of G will be denoted by $n_p(G)$ (or just n_p when G is clear from context)*

Theorem 12 (Sylow's Theorem).

Let G be a group of order $p^\alpha m$, where p is a prime not dividing m .

1. *Sylow p -subgroups of G exist, i.e., $\text{Syl}_p(G) \neq \emptyset$*

2. If P is a Sylow p -subgroup of G and Q is any p -subgroup of G , then there exists $g \in G$ such that $Q \leq gPg^{-1}$, i.e., Q is contained in some conjugate of P . In particular, any two Sylow p -subgroups of G are conjugate in G .
3. The number of Sylow p -subgroups of G is of the form $1 + kp$, i.e.

$$n_p \equiv 1 \pmod{p}$$

Further, n_p is the index in G of the normalizer $N_G(P)$ for any Sylow p -subgroup P , hence n_p divides m .

To establish the proof of this important claim, we first state and prove some preliminary lemmas.

Lemma 6.5.1. *Let $P \in \text{Syl}_p(G)$. If Q is any p -subgroup of G , then $Q \cap N_G(P) = Q \cap P$.*

Proof. Let $H = N_G(P) \cap Q$. Since $P \leq N_G(P)$, it is clear that $P \cap Q \leq H$, so we only have the reverse inclusion to prove. Since by definition $H \leq Q$, this is equivalent to showing $H \leq P$. Consider the subset PH of G , containing both P and H . Since $H \leq N_G(P)$, for all $h \in H$ and $k \in P$ $hkh^{-1} \in P$, so $hk = (hkh^{-1})h \in PH$, so $HP \subseteq PH$. Similarly, $kh = h(h^{-1}kh) \in HP$, so $PH \subseteq HP$. Then $HKHK = KHHK = KHK = HKK = HK$ and $(HK)^{-1} = K^{-1}H^{-1} = KH = HK$, so HK is closed under the group operation and inversion so it is a subgroup as claimed. From another result we have that

$$|PH| = \frac{|P||H|}{|P \cap H|}$$

where each term on the right is a power of p , so PH is a p -group. Moreover, P is a subgroup of PH so the order of PH is divisible by p^α , the largest power of p which divides $|G|$. These two facts force $|PH| = p^\alpha = |P|$, so in turn $P = PH$, since $P \leq PH$, and $H \leq P$. This establishes that $N_G(P) \cap Q = H = P \cap Q$. ■

We can now prove the first point in Sylow's Theorem:

Sylow's Theorem 1. We proceed by induction on $|G|$. If $|G| = 1$, there is nothing to prove. Assume inductively the existence of Sylow p -subgroups for all groups of order less than $|G| \geq 2$.

If p divides $|Z(G)|$, then by 10 $Z(G)$ has a subgroup, N , of order p . Let $\overline{G} = G/N$, so that $|\overline{G}| = p^{\alpha-1}m$. By induction, \overline{G} has a subgroup \overline{P} of order $p^{\alpha-1}$. If we let P be the subgroup of G containing N such that $P/N = \overline{P}$, then $|P| = |P/N| \cdot |N| = p^\alpha$ and P is a Sylow p -subgroup of G . We now must handle the case when p does not divide $|Z(G)|$.

Let g_1, g_2, \dots, g_r be representatives of the distinct non-central conjugacy classes of G . The class equation for G is

$$|G| = |Z(G)| + \sum_{i=1}^r |G : C_G(g_i)|$$

If $p \nmid |G : C_G(g_i)|$ for all i , then since $p \nmid |G|$, we would also have $p \nmid |Z(G)|$, a contradiction. Thus, for some i , p does not divide $|G : C_G(g_i)|$. For this i let $H = C_G(g_i)$, so that

$$|H| = p^\alpha, \text{ where } p \nmid k$$

Since $g_i \notin Z(G)$, $|H| < |G|$. By induction H has a Sylow p -subgroup P , which of course is also a subgroup of G . Since $|P| = p^\alpha$, P is a Sylow p -subgroup of G . This completes the induction and establishes the first bullet of Sylow's Theorem. ■

Before proving 2. and 3. of the Sylow Theorems, we perform some calculations. Note that we now know there exists a Sylow p -subgroup, P , of G . Let

$$\{P_1, P_2, \dots, P_r\} = \mathcal{S}$$

be the set of all conjugates of P , and let Q be any p -subgroup of G . By definition of \mathcal{S} , G and hence Q acts by conjugation on \mathcal{S} . Write \mathcal{S} as a disjoint union of orbits under this action by Q :

$$\mathcal{S} = \mathcal{O}_1 \sqcup \mathcal{O}_2 \sqcup \dots \sqcup \mathcal{O}_s$$

so $r = \sum_{i=1}^s |\mathcal{O}_i|$. Renumber the elements of \mathcal{S} if necessary so that the first s elements of \mathcal{S} are representatives of the Q -orbits: $P_i \in \mathcal{O}_i$, $1 \leq i \leq s$. By the Orbit Stabilizer Theorem $|\mathcal{O}_i| = |Q : N_Q(P_i)|$. By definition, $N_Q(P_i) = N_G(P_i) \cap Q$, and by our previous Lemma $N_G(P_i) \cap Q = P_i \cap Q$. Combining these facts we obtain

$$|\mathcal{O}_i| = |Q : P_i \cap Q|, \quad 1 \leq i \leq s$$

We can now prove that $r \equiv 1 \pmod{p}$. Since Q was an arbitrary p -subgroup, we may take $Q = P_1$ above, so that $|\mathcal{O}_1| = 1$. Now, for all $i > 1$, $P_1 \neq P_i$, so $P_1 \cap P_i < P_1$. Then

$$|\mathcal{O}_i| = |P_1 : P_1 \cap P_i| > 1, \quad 2 \leq i \leq s$$

Since P_1 is a p -group, $|P_1 : P_1 \cap P_i|$ must be a power of p , so that $p \mid |\mathcal{O}_i|$ for all $2 \leq i \leq s$. Thus,

$$r = |\mathcal{O}_1| + \underbrace{\sum_{i=2}^s |\mathcal{O}_i|}_{1+p|\text{thing}} \equiv 1 \pmod{p}$$

We can now prove parts 2. and 3. of Sylow's Theorem:

Sylow's Theorem 1. and 2. Let Q be any p -subgroup of G . Suppose Q is not contained in P_i for any $i \in \{1, 2, \dots, r\}$ (i.e. $Q \not\leq gPg^{-1}$ for any $g \in G$). In this situation, $Q \cap P_i < Q$ for all i , so

$$|\mathcal{O}_i| = |Q : Q \cap P_i| > 1, \quad 1 \leq i \leq s$$

by our previous arguments. Thus $p \mid |\mathcal{O}_i|$ for all i , so p divides $|\mathcal{O}_1| + \dots + |\mathcal{O}_s| = r$. This contradicts the fact that $r \equiv 1 \pmod{p}$. This contradiction proves $Q \leq gPg^{-1}$ for some $g \in G$.

To see that all Sylow p -subgroups of G are conjugate, let Q be any Sylow p -subgroup of G . By the preceding argument, $Q \leq gPg^{-1}$ for some $g \in G$. Since $|gPg^{-1}| = |Q| = p^\alpha$, we must

have $gPg^{-1} = Q$. This establishes part 2. of the theorem. Moreover, this shows that $n_p = r \equiv 1 \pmod p$, which is the first part of 3.

Finally, since all Sylow p -subgroups are conjugate, we have that

$$n_p = |G : N_G(P)| \quad \text{for any } P \in \text{Syl}_p(G)$$

so $n_p | m$, completing the proof of Sylow's Theorem. ■

We now also offer an alternate proof of Sylow's Theorem:

Sylow's Theorem Alternate. Let G be a finite group such that $|G| = p^\alpha m$ for p a prime and $p \nmid m$.

(1) First, let G act by translation on the set \mathcal{J} of subset $J \subseteq G$, with $|J| = p^\alpha$. The number of such subsets is equal to $\binom{p^\alpha m}{p^\alpha}$, which is relatively prime to p . Thus, some orbit \mathcal{O}_J must have size prime to p . Thus, G_J (the stabilizer of J in G) has order divisible by p^α , since $|\mathcal{O}_J| = |G : G_J| = |G|/|G_J|$. But, J is equivalent to the union of the set of right cosets of G_J , so $|G_J| \leq |J| = p^\alpha$. Therefore, as $p^\alpha ||G_J|$, $p^\alpha \leq |G_J|$, so we obtain that $|G_J| = p^\alpha$. Thus G_J is a Sylow p -subgroup of G , proving existence.

(2) To be completed ■

Note that since conjugation is an automorphism on G , it is an isomorphism between subgroups which implies that every Sylow p -subgroup of G is isomorphic.

Corollary 6.5.2. *Let P be a Sylow p -subgroup of G . Then the following are equivalent:*

1. P is the unique Sylow p -subgroup of G , i.e., $n_p = 1$
2. P is normal in G
3. P is characteristic in G
4. All subgroups generated by elements of p -power order are p -groups, i.e., if X is any subset of G such that $|x|$ is a power of p for all $x \in X$, then $\langle X \rangle$ is a p -group.

Proof. If 1. holds, then $gPg^{-1} = P$ for all $g \in G$ since $gPg^{-1} \in \text{Syl}_p(G)$. Conversely, if $P \triangleleft G$ and $Q \in \text{Syl}_p(G)$, then by Sylow's Theorem there exists $g \in G$ such that $Q = gPg^{-1} = P$. Thus $\text{Syl}_p(G) = \{P\}$, so we have 1. \iff 2..

Since characteristic subgroups are normal, 3. implies 1.. Conversely, if P is the unique subgroup of G of order p^α , then p is characteristic in G since the image of P under any automorphism on G is a subgroup of order p^α . Thus we conclude 1. \iff 3..

Finally, assume 1. holds and suppose X is a subset of G such that $|x|$ is a power of p for all $x \in X$. By the conjugacy part of Sylow's Theorem, for each $x \in X$ there is some $g \in G$ such that $x \in gPg^{-1} = P$. Thus $X \subseteq P$, and so $\langle X \rangle \leq P$, and hence $\langle X \rangle$ is a p -group. Conversely,

if 4. holds, let X be the union of all Sylow p -subgroups of G . If P is any Sylow p -subgroup, P is any Sylow p -subgroup, P is a subgroup of the p -group $\langle X \rangle$. Since P is a p -subgroup of G of maximal order, we must have $P = \langle X \rangle$, so 1. holds. ■

Example 6.5.1. Let G be a finite group and let p be a prime.

1. If p does not divide the order of G , the Sylow p -subgroup of G is the trivial group, and all parts of Sylow's Theorem hold trivially. If $|G| = p^\alpha$, G is the unique Sylow p -subgroup of G .
2. A finite abelian group has a unique Sylow p -subgroup for each prime p . This subgroup consists of all elements x whose order is a power of p . This is sometimes called the p -primary component of the abelian group.
3. S_3 has three Sylow 2-subgroups: $\langle(1\ 2)\rangle, \langle(2\ 3)\rangle$ and $\langle(1\ 3)\rangle$. It has a unique Sylow 3-subgroup: $\langle(1\ 2\ 3)\rangle = A_3$. Note that $3 \equiv 1 \pmod{2}$.
4. A_4 has a unique Sylow 2-subgroup: $\langle(1\ 2)(3\ 4), (1\ 3)(2\ 4)\rangle \cong V_4$. It has four Sylow 3-subgroups: $\langle(1\ 2\ 3)\rangle, \langle(1\ 2\ 4)\rangle, \langle(1\ 3\ 4)\rangle$ and $\langle(2\ 3\ 4)\rangle$. Note that $4 \equiv 1 \pmod{3}$.

Applications of Sylow's Theorem

Example 6.5.2 (Groups of order pq , p and q primes with $p < q$). Suppose $|G| = pq$ for primes p and q with $p < q$. Let $P \in \text{Syl}_p(G)$ and let $Q \in \text{Syl}_q(G)$. We show that Q is normal in G and if P is also normal in G , then G is cyclic.

Now, the three conditions: $n_q = 1 + kq$ for some $k \geq 0$, n_q divides p and $p < q$, together force $k = 0$. Since $n_q = 1$, $Q \trianglelefteq G$.

Since n_p divides the prime q , the only possibility are $n_p = 1$ or q . In particular, if $p \nmid q - 1$, (that is $q \not\equiv 1 \pmod{p}$), then n_p cannot equal q , so $P \trianglelefteq G$.

Let $P = \langle x \rangle$ and $Q = \langle y \rangle$. If $P \trianglelefteq G$, then since $G/C_G(P)$ is isomorphic to a subgroup of $\text{Aut}(\mathbb{Z}/p\mathbb{Z})$ and the latter group has order $p - 1$, Lagrange's Theorem together with the observation that neither p nor q can divide $p - 1$ implies that $G = C_G(P)$. In this case $x \in P \leq Z(G)$, so x and y commute. This means $|xy| = pq$, hence in this case G is cyclic: $G \cong \mathbb{Z}/pq\mathbb{Z}$.

Example 6.5.3 (Groups of order 30). Let G be a group of order 30. We show that G has a normal subgroup isomorphic to $\mathbb{Z}/15\mathbb{Z}$. Note that any subgroup of order 15 is necessarily normal in G since it is of index 2, and cyclic by the preceding result, so it is only necessary to show there exists a subgroup of order 15.

Let $P \in \text{Syl}_5(G)$ and let $Q \in \text{Syl}_3(G)$. If either P or Q is normal in G , PQ is a subgroup of order 15. Note also that if either P or Q is normal, then both P and Q are characteristic subgroups of PQ , and since $PQ \trianglelefteq G$, both P and Q are normal. Assume therefore that neither Sylow subgroup is normal. The only possibilities are $n_5 = 6$ and $n_3 = 10$. Each element of order 5 lies in a Sylow 5-subgroup, each Sylow 5-subgroup contains 4 nonidentity elements, and by Lagrange's Theorem, distinct Sylow 5-subgroups intersect in the identity. Thus the

number of elements of order 5 in G is the number of nonidentity elements in one Sylow 5-subgroup times the number of Sylow 5-subgroups. This would be $4 \cdot 6 = 24$ elements of order 5. By similar reasoning, the number of elements of order 3 would be $2 \cdot 10 = 20$. This is absurd since a group of order 30 cannot contain $24 + 20 = 44$ distinct elements. One of P or Q (hence both) must be normal in G .

Example 6.5.4 (Groups of order 12). Let G be a group of order 12. We show that either G has a normal Sylow 3-subgroup, or $G \cong A_4$.

Suppose $n_3 \neq 1$ and let $P \in \text{Syl}_3(G)$. Since $n_3 | 4$ and $n_3 \equiv 1 \pmod{3}$, it follows that $n_3 = 4$. Since distinct Sylow 3-subgroups intersect in the identity and each contains two elements of order 3, G contains $2 \cdot 4 = 8$ elements of order 3. Since $|G : N_G(P)| = n_3 = 4$, $N_G(P) = P$. Now G acts by conjugation on its four Sylow 3-subgroups, so this action affords a permutation representation:

$$\varphi : G \rightarrow S_4$$

The kernel K of this action is the subgroup of G which normalizes all Sylow 3-subgroups of G . In particular, $K \leq N_G(P) = P$. Since P is not normal in G by assumption, $K = 1$, so φ is injective and

$$G \cong \varphi(G) \leq S_4$$

Since G contains 8 elements of order 3 and there are precisely 8 elements of order 3 in S_4 , all contained in A_4 , it follows that $\varphi(G)$ intersects A_4 in a subgroup of order at least 8. Since both groups have order 12 it follows that $\varphi(G) = A_4$, so that $G \cong A_4$.

Note that A_4 has 4 Sylow 3-subgroups, so such a group G does indeed exist. Also, letting V be a Sylow 2-subgroup of A_4 , $|V| = 4$ so it contains all the remaining elements of A_4 . In particular, there cannot be another Sylow 2-subgroup. Thus $n_2(A_4) = 1$, so $V \triangleleft A_4$.

Example 6.5.5 (Groups of order p^2q , p and q distinct primes). Let G be a group of order p^2q . We show that G has a normal Sylow subgroup (for either p or q). Let $P \in \text{Syl}_p(G)$ and let $Q \in \text{Syl}_q(G)$.

Consider first when $p > q$. Since $n_p | q$ and $n_p = 1 + kp$, we must have $n_p = 1$. Thus $P \trianglelefteq G$.

Consider now the case $p < q$. If $n_q = 1$, Q is normal in G . Assume therefore that $n_q > 1$, i.e., $n_q = 1 + tq$, for some $t > 0$. Now $n_q | p^2$ so $n_q = p$ or p^2 . Since $q > p$ we cannot have $n_q = p$, hence $n_q = p^2$. Thus

$$tq = p^2 - 1 = (p - 1)(p + 1)$$

Since q is prime, either $q | p - 1$ or $q | p + 1$. The former is impossible since $q > p$, so the latter holds. Since $q > p$ but $q | p + 1$, we must have $q = p + 1$. This forces $p = 2$, $q = 3$ and $|G| = 12$. This result now follows from the preceding example.

Groups of Order 60

Proposition 6.5.3. *If $|G| = 60$ and G has more than one Sylow 5-subgroup, then G is simple.*

Proof. Suppose by way of contradiction that $|G| = 60$ and $n_5 > 1$ but that there exists H a normal subgroup of G with $H \neq \{1\}$ or G . By Sylow's Theorem the only possibility for $n_5 = 6$. Let $P \in \text{Syl}_5(G)$, so that $|N_G(P)| = 10$ since its index is n_5 .

If $5 \nmid |H|$ then H contains a Sylow 5-subgroup of G and since H is normal, it contains all 6 conjugates of this subgroup. In particular, $|H| \geq 1 + 6 \cdot 4 = 25$, and the only possibility is $|H| = 30$. This leads to a contradiction since a previous example proved that any group of order 30 has a normal (hence unique) Sylow 5-subgroup. This argument shows 5 does not divide $|H|$ for any proper normal subgroup H of G .

If $|H| = 6$ or 12 , H has a normal, hence characteristic, Sylow subgroup, which is therefore also normal in G . Replacing H by this subgroup if necessary, we may assume $|H| = 2, 3$ or 4 . Let $\bar{G} = G/H$, so $|\bar{G}| = 30, 20$ or 15 . In each case, \bar{G} has a normal subgroup \bar{P} of order 5 by previous results. If we let H_1 be the complete preimage of \bar{P} in G , then $H_1 \trianglelefteq G$, $H_1 \neq G$ and $5 \nmid |H_1|$. This contradicts the preceding paragraph and so completes the proof. ■

Corollary 6.5.4. A_5 is simple.

Proof. The subgroups $\langle (1\ 2\ 3\ 4\ 5) \rangle$ and $\langle (1\ 3\ 2\ 4\ 5) \rangle$ are distinct Sylow 5-subgroups of A_5 , so the result follows from the proposition. ■

Proposition 6.5.5. If G is a simple group of order 60, then $G \cong A_5$.

Proof. Let G be a simple group of order 60, so $n_2 = 3, 5$ or 15 . Let $P \in \text{Syl}_2(G)$ and let $N = N_G(P)$, so $|G : N| = n_2$.

First observe that G has no proper subgroup H of index less than 5, as follows: if H were a subgroup of G of index 4, 3 or 2, then, G would have a normal subgroup K contained in H with G/K isomorphic to a subgroup of S_4, S_3 or S_2 . Since $K \neq G$, simplicity forces $K = \{1\}$. This is impossible since $|G| = 60$ does not divide $4!$. This argument shows, in particular, that $n_2 \neq 3$.

If $n_2 = 5$, then N has index 5 in G so the action of G by left multiplication on the set of left cosets of N gives a permutation representation of G into S_5 . Since the kernel of this representation is a proper normal subgroup and G is simple, the kernel is $\{1\}$ and G is isomorphic to a subgroup of S_5 . Identify G with this isomorphic copy so that we may assume $G \leq S_5$. If G is not contained in A_5 , then $S_5 = GA_5$ and, by the Second Isomorphism Theorem, $A_5 \cap G$ is of index 2 in G . Since G has no normal subgroup of index 2, this is a contradiction. This argument proves $G \leq A_5$. Since $|G| = |A_5|$, the isomorphic copy of G in S_5 coincides with A_5 , as desired.

Finally, assume $n_2 = 15$. If for every pair of distinct Sylow 2-subgroups P and Q of G , $P \cap Q = \{1\}$, then the number of nonidentity elements in Sylow 2-subgroups of G would be $(4 - 1) \cdot 15 = 45$. But $n_5 = 6$ so the number of elements of order 5 in G is $(5 - 1) \cdot 6 = 24$, accounting for 69 elements. This contradiction proves that there exist distinct Sylow 2-subgroups P and Q with $|P \cap Q| = 2$. Let $M = N_G(P \cap Q)$. Since P and Q are abelian (being groups of order 4), P and Q are subgroups of M and since G is simple, $M \neq G$. Thus $4 \nmid |M|$ and $|M| > 4$. The only possibility is $|M| = 12$, i.e., M has index 5 in G . But now the argument of the preceding paragraph applied to M in place of N gives $G \cong A_5$. This leads to a contradiction

in this case because $n_2(A_5) = 5$. The proof is complete. ■

Chapter 7

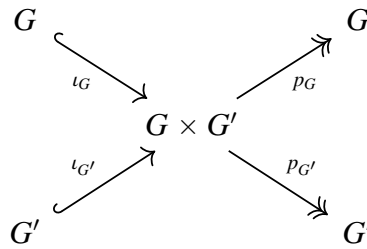
§§Product Groups

7.1.0 §Basic Definitions and Examples: Product Groups

Recall 7.1.1. For G, G' group, recall that $G \times G'$ with the binary operation $(a, a') \star (b, b') := (a \star_G b, a' \star_{G'} b')$ is a group with identity $(e_G, e_{G'})$ and inverse $(a, a')^{-1} = (a^{-1}, a'^{-1})$.

Definition 7.1.1 (Direct Product). The group $G \times G'$ is called the direct product of G and G' .

Remark 7.1.2. We have four homomorphisms which characterize the direct product of G and G' :



defined by

$$\iota_G(g) = (g, e_{G'}), \iota_{G'}(g') = (e_G, g') \quad (7.1.1)$$

$$p_G(g, g') = g, p_{G'}(g, g') = g' \quad (7.1.2)$$

for all $g \in G$ and $g' \in G'$. Moreover, this map satisfy the following properties:

1. ι_G and $\iota_{G'}$ are monomorphisms of homomorphic image $G \times \{e_{G'}\} \leq G \times G'$ and $\{e_G\} \times G' \leq G \times G'$, respectively.
2. p_G and $p_{G'}$ are epimorphisms called projections, with kernels $\ker(p_G) = \{e_G\} \times G'$ and $\ker(p_{G'}) = G \times \{e_{G'}\}$. Hence, we have that normal subgroups

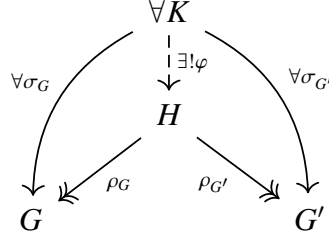
$$\{e_G\} \times G' \triangleleft G \times G' \triangleright G \times \{e_{G'}\} \quad (7.1.3)$$

Theorem 13 (Universal Property of Product Groups).

Let G and G' be groups. Then the direct product of G and G' is defined uniquely up to isomorphism by the triple

$$(H, \rho_G : H \rightarrow G, \rho_{G'} : H \rightarrow G') \quad (7.1.4)$$

that satisfies the universal property



Proof. (Left to the reader) ■

Remark 7.1.3. Given a group K , it is desirable to decompose K as a product $K \cong H \times H'$ for $H, H' \leq K$ proper subgroups. Indeed, H and H' are simpler groups, and it is easy to relate properties of K to properties of H and H' .

Note 7.1.4. A group cannot necessarily be written in this way for non-trivial H and H' .

Example 7.1.1. Observe that $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ as $([1]_2, [1]_3)$ is an element of order 6 and $|\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}| = 6$.

Proposition 7.1.1 (Cyclic Group Decomposition). *Let $m, n \in \mathbb{Z}$, for $m, n \geq 1$. Then we have that*

$$\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \quad (7.1.5)$$

if and only if m and n are relatively prime.

Proof. (Left to the reader) ■

Example 7.1.2 (Non-example). Observe that $\mathbb{Z}/4\mathbb{Z} \not\cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Indeed, every element of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ has order 1 or 2, whereas $\mathbb{Z}/4\mathbb{Z}$ is generated by an element of order 4.

Proposition 7.1.2. *Let $H, H' \leq K$ and let*

$$\begin{aligned} f : H \times H' &\rightarrow K \\ (h, h') &\mapsto hh' \end{aligned} \quad (7.1.6)$$

be the multiplication map (not a homomorphism in general). Then the image of f is

$$HH' = \{hh' : h \in H, h' \in H'\} \quad (7.1.7)$$

We then have that

1. (a) f is injective if and only if $H \cap H' = \{e_K\}$
- (b) f is surjective if and only if $K = HH'$

2. f is a group homomorphism from the direct product group $H \times H'$ to K if and only if $hh' = h'h$ for all $h \in H$ and all $h' \in H'$
3. f is a group isomorphism if and only if $H \cap H' = \{e_K\}$, $HH' = K$, and $H, H' \triangleleft K$.

1. a). First, let $x \in H \cap H'$ for $x \neq e_K$. Then we have that $f(x, e_K) = x = f(e_K, x)$, where $(x, e_K) \neq (e_K, x)$ by assumption, so f is not injective. Conversely, suppose $H \cap H' = \{e_K\}$ and $(a, b) \in \ker(f)$ so $f(a, b) = e_K$. Then we have that $H \ni a = b^{-1} \in H'$ since it is a subgroup, so $a, b^{-1} \in H \cap H'$. In particular, $a, b^{-1} = e_K$, so $b = e_K$ as well. Thus, $\ker(f) = \{(e_K, e_K)\}$, which implies f is injective.

[1. b)] Note that f is surjective if and only if $K = HH'$ by definition of f .

[2.] Let $h_1, h_2 \in H, h'_1, h'_2 \in H'$. Then f is a homomorphism if and only if

$$h_1 h'_1 h_2 h'_2 = f(h_1, h'_1) f(h_2, h'_2) = f(h_1 h_2, h'_1 h'_2) = h_1 h_2 h'_1 h'_2$$

which holds if and only if $h'_1 h_2 = h_2 h'_1$. But, this is true for all $h_2 \in H$ and all $h'_1 \in H'$, so the if and only if statement is true.

[3.] Note f is injective if and only if $H \cap H' = \{e_K\}$ by 1.a), and f is surjective if and only if $HH' = K$, by 1.b). First, suppose f is an isomorphism. Note that from the four fundamental makes of the group direct product we know that $H \times \{e_K\}, \{e_K\} \times H' \triangleleft H \times H'$. Thus, since f is assumed to be surjective and $f(H \times \{e_K\}) = H$, we have that $H \triangleleft K$, and similarly $H' \triangleleft K$. Conversely, suppose $H, H' \triangleleft K$. Then, let $h \in H$ and $h' \in H'$ and consider the commutator $[h, h'] = hh'h^{-1}h'^{-1}$. Since H and H' are normal we have that $H \ni h(h'h^{-1}h'^{-1}) = (hh'h^{-1})h'^{-1} \in H$. Hence, $[h, h'] \in H \cap H'$, but $H \cap H' = \{e_K\}$ so $[h, h'] = e_K$. Therefore, we have that $hh' = h'h$, so by 2. f is a homomorphism, and since f is shown to be injective and surjective, it is an isomorphism. ■

Proposition 7.1.3. Let $H, H' \leq K$. If H (or H') is a normal subgroup of K , then HH' is a subgroup of K .

Proof. (Left to the reader) ■

Remark 7.1.5. Note that the multiplication map can be bijective without being a homomorphism. For example, if we take $H = \langle x \rangle, H' = \langle y \rangle \in D_3$, and $H \cap H' = \{1\}$, $D_3 = HH'$, but $D_3 \not\cong \langle x \rangle \times \langle y \rangle$ because $\langle y \rangle$ is not a normal subgroup.

Corollary 7.1.4. Let G be a finite group with $H, H' \leq G$.

1. If $H \cap H' = \{e_G\}$, then $|H||H'| = |HH'|$
2. If $H \cap H' = \{e_G\}$, $H, H' \triangleleft G$, and $|G| = |H||H'|$, then

$$G \cong H \times H' \tag{7.1.8}$$

1. Suppose $H \cap H' = \{e_G\}$. Then by 1.a) of the previous proposition the multiplication map $f : H \times H' \rightarrow G$ is injective. Moreover, its image is precisely $HH' \subseteq G$. Thus, the corestriction

$f : H \times H' \rightarrow HH'$ is a bijection. Therefore

$$|H||H'| = |H \times H'| = |HH'| \quad (7.1.9)$$

[2.] Suppose $H \cap H' = \{e_G\}$, $H, H' \triangleleft G$, and $|G| = |H||H'|$. Since $H \cap H' = \{e_G\}$ we have by 1. that $|H||H'| = |HH'|$, so $|G| = |HH'|$ which implies $G = HH'$ as $HH' \subseteq G$. Thus, by 3. of the previous proposition $G \cong H \times H'$. ■

Remark 7.1.6 (Application). Suppose G is abelian and $|G| = p^2$ for a prime p . Then either G is cyclic or

$$G \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \quad (7.1.10)$$

Proof. Assume G is not cyclic. Then for all $g \in G$ with $g \neq e_G$ we have $o(g) = p$ by 2. Take $g, g' \in G$ such that $o(g) = o(g') = p$ and $g' \notin \langle g \rangle$, which is possible since there are p elements not in $\langle g \rangle$ of order p . Let $H = \langle g \rangle$ and $H' = \langle g' \rangle$. Since G is abelian H and H' are normal subgroups. Moreover, $H \cap H' = \{e_G\}$. Indeed, $H \cap H'$ is a subgroup of H and H' , so $|H \cap H'| \in \{1, p\}$. But, if $|H \cap H'| = p$ then $H = H \cap H' = H'$, which implies that $g' \in H$, contradicting our initial assumption. Thus $|H \cap H'| = 1$ so $H \cap H' = \{e_G\}$. Finally, $|G| = p^2 = |H||H'|$. Thus, by 2. of the previous corollary we conclude that

$$G \cong H \times H' \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \quad (7.1.11)$$

■

7.2.0 §Semi-Direct Products

Recall 7.2.1. Let G be a group and H, K subgroups such that $H \triangleleft G$. If additionally $H \cap K = \{1\}$, then HK is a subgroup of G and every element of HK can be written uniquely as a product hk for some $h \in H$ and $k \in K$.

Observe that if $H \triangleleft G$, $K \leq G$, then for any two elements $h_1k_1, h_2k_2 \in HK$,

$$\begin{aligned} (h_1k_1)(h_2k_2) &= h_1k_1h_2(k_1^{-1}k_1)k_2 \\ &= h_1(k_1h_2k_1^{-1})k_1k_2 \\ &= h_3k_3 \end{aligned}$$

where $h_3 = h_1(k_1h_2k_1^{-1})$ and $k_3 = k_1k_2$. Note since H is normal in G , the group K acts on H by conjugation:

$$k \cdot h = khk^{-1} \in H, \text{ for all } h \in H, k \in K$$

These observations inspire our following construction of a group given two groups H and K and a homomorphism $\phi : K \rightarrow \text{Aut}(H)$.

Theorem 7.2.1. Let H and K be groups and let ϕ be a group homomorphism from K into $\text{Aut}(H)$. Let \cdot denote the (left) action of K on H determined by ϕ . Let G be the set of ordered pairs (h, k) with $h \in H$ and $k \in K$, and define the following multiplication on G :

$$(h_1, k_1) \star (h_2, k_2) := (h_1k_1 \cdot h_2, k_1k_2)$$

for all $(h_1, k_1), (h_2, k_2) \in G$. Then

1. this multiplication makes G into a group of order $|G| = |H||K|$
2. the sets $\{(h, 1) | h \in H\}$ and $\{(1, k) | k \in K\}$ are subgroups of G and the maps $h \mapsto (h, 1)$ for $h \in H$ and $k \mapsto (1, k)$ for $k \in K$ are isomorphisms of these subgroups with the groups H and K respectively:

$$H \cong \{(h, 1) | h \in H\} \quad \text{and} \quad K \cong \{(1, k) | k \in K\}$$

Identifying H and K with their isomorphic copies in G we have

1. $H \triangleleft G$
2. $H \cap K = 1$
3. for all $h \in H$ and $k \in K$, $khk^{-1} = k \cdot h = \phi(k)(h)$

Proof. First, observe that that $(1_H, 1_K) \in G$ acts as the identity. Indeed given any $(h, k) \in G$,

$$(1_H, 1_K)(h, k) = (1_H 1_K \cdot h, 1_K k) = (h, k)$$

and

$$(h, k)(1_H, 1_K) = (hk \cdot 1_H, k 1_K) = (h, k)$$

because $k \cdot 1_H$ is the action of an automorphism of H , and hence must send the identity to itself. Let $(h_1, k_1) \in G$. Then I claim that $(h_1, k_1)^{-1} = (k_1^{-1} \cdot h_1^{-1}, k_1^{-1}) \in G$. Observe that

$$\begin{aligned} (h_1, k_1)(k_1^{-1} \cdot h_1^{-1}, k_1^{-1}) &= (h_1 k_1 \cdot (k_1^{-1} \cdot h_1^{-1}), k_1 k_1^{-1}) \\ &= (h_1 (k_1 k_1^{-1}) \cdot h_1^{-1}, 1_K) \\ &= (h_1 1_K \cdot h_1^{-1}, 1_K) \\ &= (h_1 h_1^{-1}, 1_K) \\ &= (1_H, 1_K) \end{aligned}$$

and

$$\begin{aligned} (k_1^{-1} \cdot h_1^{-1}, k_1^{-1})(h_1, k_1) &= (k_1^{-1} \cdot h_1^{-1} k_1^{-1} \cdot h_1, k_1^{-1} k_1) \\ &= (k_1^{-1} \cdot (h_1^{-1} h_1), 1_K) \quad (\text{since the action } k_1^{-1} \rightarrow \text{Aut}(H)) \\ &= (k_1^{-1} \cdot 1_H, 1_K) \\ &= (1_H, 1_K) \end{aligned}$$

as claimed, so G has inverses. Finally, for $(h_1, k_1), (h_2, k_2), (h_3, k_3) \in G$ we have

$$\begin{aligned} (h_1, k_1)[(h_2, k_2)(h_3, k_3)] &= (h_1, k_1)(h_2 k_2 \cdot h_3, k_2 k_3) \\ &= (h_1 k_1 \cdot (h_2 k_2 \cdot h_3), k_1 (k_2 k_3)) \\ &= (h_1 (k_1 \cdot h_2)(k_1 \cdot (k_2 \cdot h_3)), (k_1 k_2) k_3) \\ &= (h_1 (k_1 \cdot h_2)(k_1 k_2 \cdot h_3), (k_1 k_2) k_3) \\ &= (h_1 k_1 \cdot h_2, k_1 k_2)(h_3, k_3) \\ &= [(h_1, k_1)(h_2, k_2)](h_3, k_3) \end{aligned}$$

so multiplication in G is associative. Therefore G is indeed a group. For each $(h, k), (h', k') \in G$, $(h, k) = (h', k')$ if and only if $h = h'$ and $k = k'$, so $|G| = |H||K|$, finishing the proof of the first claim.

Let $H := \{(h, 1_K) | h \in H\}$ and $K := \{(1_H, k) | k \in K\}$. Then note $(a, 1_K)(b, 1_K) = (a1_K \cdot b, 1_K) = (ab, 1_K)$, and $(1_H, x)(1_H, y) = (1_H x \cdot 1_H, xy) = (1_H, xy)$, so H and K are indeed subgroups of G (as H and K are groups in their own right). Moreover, it follows that the maps defined in the second bullet connote isomorphisms $H \xrightarrow{\sim} H$ and $K \xrightarrow{\sim} K$. Now, observe that for all $(h, 1_K) \in H$ and all $(h_1, k_1) \in G$, we have

$$(h_1, k_1)(h, 1_K)(k_1^{-1} \cdot h_1^{-1}, k_1^{-1}) = (h_1 k_1 \cdot h, k_1)(k_1^{-1} \cdot h_1^{-1}, k_1^{-1}) = (h_1(k_1 \cdot h)(k_1 \cdot k_1^{-1} \cdot h_1^{-1}), k_1 k_1^{-1}) = (h_1(k_1 \cdot h)h_1^{-1}, 1_K)$$

so $H \triangleleft G$. Next, the fourth bullet follows immediately from the definitions of H and K , so $H \cap K = \{(1_H, 1_K)\}$.

Finally, let $(h, 1_K) \in H$ and $(1_H, k) \in K$. Then

$$(1_H, k)(h, 1_K)(1_H, k^{-1}) = (k \cdot h, k)(1_H, k^{-1}) = (k \cdot h, 1_K)$$

so identifying with H and K by the isomorphisms previously, we find $khk^{-1} = k \cdot h = \phi(k)(h)$. Moreover, under this identification $K \leq N_G(H)$ since the conjugation acts as an automorphism on H . Since $G = HK$ and $H \leq N_G(H)$, we have $N_G(H) = G$, i.e., which again proves $H \triangleleft G$ under our identification. ■

Definition 7.2.1. Let H and K be groups and let $\varphi : K \rightarrow \text{Aut}(H)$ be a group homomorphism. The group described in the previous theorem is called the **semidirect product** of H and K with respect to φ , and will be denoted by $H \rtimes_{\varphi} K$ (signifying that K is the group doing the action, and H is the normal “factor”).

We can now formalize direct products as special cases of semidirect products:

Proposition 7.2.2. Let H and K be groups and let $\varphi : K \rightarrow \text{Aut}(H)$ be a group homomorphism. Then the following are equivalent:

1. The identity (set) map between $H \rtimes_{\varphi} K$ and $H \times K$ is a group homomorphism (hence an isomorphism since the underlying set map is a bijection)
2. φ is the trivial homomorphism from K into $\text{Aut}(H)$
3. $K \triangleleft H \rtimes_{\varphi} K$

Proof. (1. \implies 2.) By definition of the group operation in $H \rtimes_{\varphi} K$

$$(h_1, k_1)(h_2, k_2) = (h_1 k_1 \cdot h_2, k_1 k_2)$$

for all $h_1, h_2 \in H$ and $k_1, k_2 \in K$. By assumption 1., we need $(h_1, k_1)(h_2, k_2) = (h_1 h_2, k_1 k_2)$, which is to say $\varphi(k_1)(h_2) = h_2$ for all $h_2 \in H$. In particular $\varphi(k_1) = \text{Id}_H$ for all $k_1 \in K$, so $\varphi(K) = \{\text{Id}_H\}$.

(2. \implies 3.) If φ is trivial, then the action of K on H is trivial, so that the elements of H commute with those of K by bullet 5. of our previous theorem. In particular, H normalizes K and K normalizes itself, so as $G = HK$, G normalizes K , proving 3..

(3. \implies 1.) If K is normal in $H \rtimes_{\varphi} K$ then for all $h \in H$ and $k \in K$, $[h, k] \in H \cap K = \{1\}$. Thus $hk = kh$ and the action of K on H is trivial. The multiplication in the semidirect products is then the same as that in the direct product:

$$(h_1, k_1)(h_2, k_2) = (h_1 h_2, k_1 k_2)$$

for all $h_1, h_2 \in H$ and $k_1, k_2 \in K$. This gives 1. and completes the proof. \blacksquare

Example 7.2.1. In all examples to follow let H and K be groups and φ a homomorphism from K into $\text{Aut}(H)$ with associated action of K on H denoted by \cdot . Let $G = H \rtimes_{\varphi} K$ and as in our previous work we identify H and K as subgroups of G .

1. Let H be any abelian group and let $K = \langle x \rangle \cong \mathbb{Z}/2\mathbb{Z}$ be the group of order 2. Define $\varphi : K \rightarrow \text{Aut}(H)$ by mapping x to the automorphism of inversion on H so that the associated action is $x \cdot h = h^{-1}$, for all $h \in H$. Then G contains the subgroup H of index 2 and

$$xhx^{-1} = h^{-1} \text{ for all } h \in H$$

When H is cyclic, we have the following special cases: if $H = \mathbb{Z}/n\mathbb{Z}$, one recognizes G as D_{2n} , and if $H = \mathbb{Z}$ we denote G by D_{∞} .

2. For H any group let $K = \text{Aut}(H)$ with φ the identity map from K to $\text{Aut}(H)$. The semidirect product $H \rtimes_{\varphi} \text{Aut}(H)$ is called the **holomorph** of H and will be denoted by $\text{Hol}(H)$. Some holomorphs are described below:

$$(a) \text{Hol}(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}) \cong S_4$$

- (b) If $|G| = n$ and $\pi : G \rightarrow S_n$ is the left regular representation, then $N_{S_n}(\pi(G)) \cong \text{Hol}(G)$. In particular, since the left regular representation of a generator of $\mathbb{Z}/n\mathbb{Z}$ is an n -cycle in S_n we obtain that for any n -cycle $(1 \ 2 \ \dots \ n)$:

$$N_{S_n}(\langle (1 \ 2 \ \dots \ n) \rangle) \cong \text{Hol}(\mathbb{Z}/n\mathbb{Z}) = \mathbb{Z}/n\mathbb{Z} \rtimes \text{Aut}(\mathbb{Z}/n\mathbb{Z})$$

with the latter group having order $n\varphi(n)$, for φ the Euler-toutient function.

Theorem 7.2.3. Suppose G is a group with subgroups H and K such that

1. $H \triangleleft G$, and
2. $H \cap K = \{1\}$

Let $\varphi : K \rightarrow \text{Aut}(H)$ be the homomorphism defined by mapping $k \in K$ to the automorphism of left conjugation by k on H . Then $HK \cong H \rtimes_{\varphi} K$. In particular, if $G = HK$ with H and K satisfying 1. and 2., then G is the semidirect product of H and K .

Proof. Note that since $H \triangleleft G$, HK is a subgroup of G . Every element of HK can be written uniquely in the form hk , for some $h \in H$ and $k \in K$ by properties 1. and 2.. Thus the map $hk \mapsto (h, k)$ is a set bijection from HK onto $H \rtimes_{\varphi} K$. The fact that this map is a homomorphism is given by

$$hkh'k' = (h(kh'k^{-1}))kk' \mapsto (hk \cdot h', kk') = (h, k)(h', k')$$

■

Definition 7.2.2. Let H be a subgroup of the group G . A subgroup K of G is called a **complement** for H in G if $G = HK$ and $H \cap K = \{1\}$.

Classifications of Certain Finite Groups

We shall apply our results on Semi-Direct product groups to classify certain finite groups. The argument shall follow the following structure:

1. show every group of order n has proper subgroups H and K satisfying $H \triangleleft G, K \leq G, H \cap K = \{1\}$ with $G = HK$.
2. find all possible isomorphism types for H and K
3. for each pair H, K found, find all possible homomorphisms $\varphi : K \rightarrow \text{Aut}(H)$
4. for each triple H, K, φ found form the semidirect product $H \rtimes_{\varphi} K$ and among all these semidirect products determine which pairs are isomorphic. This results in a list of the distinct isomorphism types of groups of order n .

Since H and K are proper subgroups of G one should think of the determination of H and K as being achieved inductively.

Example 7.2.2 (Groups of order pq , p and q primes with $p < q$). Let G be any group of order pq , let $P \in \text{Syl}_p(G)$ and let $Q \in \text{Syl}_q(G)$. Note that in the Sylow section we have shown $Q \triangleleft G$ and $P \leq G$ with $P \cap Q = \{1\}$, so $G \cong Q \rtimes_{\varphi} P$ for some $\varphi : P \rightarrow \text{Aut}(Q)$. Since P and Q are of prime order, they are cyclic. The group $\text{Aut}(Q)$ is cyclic of order $q - 1$. If p does not divide $q - 1$, the only homomorphism from P to $\text{Aut}(Q)$ is the trivial homomorphism, hence the only semidirect product in this case is the direct product, i.e., G is cyclic.

Consider now the case when $p | q - 1$, and let $P = \langle y \rangle$. Since $\text{Aut}(Q)$ is cyclic it contains a unique subgroup of order p , say $\langle \gamma \rangle$, and any homomorphism $\varphi : P \rightarrow \text{Aut}(Q)$ must map y to a power of γ . There are therefore p homomorphisms $\varphi_i : P \rightarrow \text{Aut}(Q)$ given by $\varphi_i(y) = \gamma^i, 0 \leq i \leq p - 1$. Since φ_0 is the trivial homomorphism, $Q \rtimes_{\varphi_0} P \cong Q \times P$ as before. Each φ_i for $i \neq 0$ gives rise to a non-abelian group, G_i , of order pq . These groups are all isomorphic because for each $\varphi_i, i > 0$, there is some generator y_i of P such that $\varphi_i(y_i) = \gamma$. Thus, up to a choice for the generator of P , these semidirect products are all the same.

Example 7.2.3 (Groups of order 30). From the examples following Sylow's Theorem, every group G of order 30 contains a subgroup H of order 15. By the preceding example H is cyclic

and H is normal in G (index 2). By Sylow's Theorem there is a subgroup K of G of order 2. Thus $G = HK$ and $H \cap K = \{1\}$ so $G \cong H \rtimes_{\varphi} K$, for some $\varphi : K \rightarrow \text{Aut}(H)$. Then

$$\text{Aut}(\mathbb{Z}/15\mathbb{Z}) \cong (\mathbb{Z}/15\mathbb{Z})^{\times} \cong \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$$

where the latter isomorphism follows from writing H as $\langle a \rangle \times \langle b \rangle \cong \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$, and since these subgroups are characteristic in H we have

$$\text{Aut}(H) \cong \text{Aut}(\mathbb{Z}/5\mathbb{Z}) \times \text{Aut}(\mathbb{Z}/3\mathbb{Z})$$

In particular, $\text{Aut}(H)$ contains precisely three elements of order 2, whose actions on the group $\langle a \rangle \times \langle b \rangle$ are the following:

$$\left\{ \begin{array}{l} a \mapsto a \\ b \mapsto b^{-1} \end{array} \right\} \quad \left\{ \begin{array}{l} a \mapsto a^{-1} \\ b \mapsto b \end{array} \right\} \quad \left\{ \begin{array}{l} a \mapsto a^{-1} \\ b \mapsto b^{-1} \end{array} \right\}$$

Thus there are three nontrivial homomorphisms from K into $\text{Aut}(H)$ given by sending the generator of K into one of these three elements of order 2 (as usual, the trivial homomorphism gives the direct product: $H \times K \cong \mathbb{Z}/30\mathbb{Z}$).

Let $K = \langle k \rangle$. If the homomorphism $\varphi_1 K \rightarrow \text{Aut}(H)$ is defined by mapping k to the first automorphism above, then $G_1 = H \rtimes_{\varphi_1} K$ is seen to be isomorphic to $\mathbb{Z}/5\mathbb{Z} \times D_6$ (in particular it is $\langle a \rangle \times \langle b, k \rangle$).

If φ_2 is defined by mapping k to the second automorphism above, then $G_2 = H \rtimes_{\varphi_2} K$ is seen to be isomorphic to $\mathbb{Z}/3\mathbb{Z} \times D_{10}$ (factorization: $\langle b \rangle \times \langle a, k \rangle$).

If φ_3 is defined by mapping k to the third automorphism above, then $G_3 = H \rtimes_{\varphi_3} K$ is isomorphic to D_{30} .

Chapter 8

§§ Nilpotent and Solvable Groups

8.1.0 § p -Groups

Definition 8.1.1. A maximal subgroup of a group G is a proper subgroup M of G such that there are no subgroups H of G with $M < H < G$.

Simply by order considerations we observe that any proper subgroup of a finite group is contained in a maximal subgroup. In contrast, infinite groups may or may not have maximal subgroups.

Theorem 8.1.1. Let p be a prime and let P be a group of order p^a for $a \geq 1$. Then

1. The center of P is nontrivial: $Z(P) \neq \{1\}$,
2. If H is a nontrivial normal subgroup of P then H intersects the center non-trivially: $H \cap Z(P) \neq \{1\}$. In particular, every normal subgroup of order p is contained in the center.
3. If H is a normal subgroup of P then H contains a subgroup of order p^b that is normal in P for each divisor p^b of $|H|$. In particular, P has a normal subgroup of order p^b for every $b \in \{0, 1, \dots, a\}$.
4. If $H < P$ then $H < N_P(H)$ (i.e.e, every proper subgroup of P is a proper subgroup of its normalizer in P).
5. Every maximal subgroup of P is of index p and is normal in P .

Proof. First note that 1. is a result previously proven using the class equation.

Now, let H be a nontrivial normal subgroup of P . Recall that for each conjugacy class C of P , either $C \subseteq H$ or $C \cap H = \emptyset$ because H is normal. Pick representatives of the conjugacy classes of P , a_1, \dots, a_r , with $a_1, \dots, a_k \in H$ and $a_{k+1}, \dots, a_r \notin H$. Let C_i be the conjugacy class of a_i in P , for all i . Thus

$$C_i \subseteq H, 1 \leq i \leq k$$

and

$$C_i \cap H = \emptyset, k+1 \leq i \leq r$$

By renumbering a_1, \dots, a_k if necessary, we may assume a_1, \dots, a_s represent classes of size 1 and a_{s+1}, \dots, a_k represent classes of size > 1 . Since H is the disjoint union of these we have

$$|H| = |H \cap Z(P)| + \sum_{i=s+1}^k \frac{|P|}{|C_p(a_i)|}$$

Now p divides $|H|$ and p divides each term in the sum $\sum_{i=s+1}^k |P : C_p(a_i)|$ so p divides their difference: $|H \cap Z(P)|$. This proves $H \cap Z(P) \neq \{1\}$. If $|H| = p$, since $H \cap Z(P) \neq \{1\}$ we must have $H \leq Z(P)$.

Next, we prove 3. by induction on a . If $a \leq 1$ or $H = \{1\}$, the result is immediate. Assume therefore that $a > 1$ and $H \neq \{1\}$. By part 2, $H \cap Z(P) \neq 1$ so by Cauchy's Theorem $H \cap Z(P)$ contains a normal subgroup of Z of order p . Then the quotient P/Z has order p^{a-1} and $H/Z \trianglelefteq P/Z$. By induction, for every nonnegative inter b such that p^b divides $|H/Z|$ there is a subgroup K/Z of H/Z of order p^b that is normal in P/Z . If K is the complete preimage of K/Z in P then $|K| = p^{b+1}$. The set of all subgroups of H obtained by this process together with the identity subgroup provides a subgroup of H that is normal in P for each divisor of $|H|$. This establishes part 3..

We prove 4. also by induction on $|P|$. If P is abelian then all subgroups of P are normal in P and the result is immediate. We may therefore assume $|P| > p$. Let H be a proper subgroup of P . Since all elements of $Z(P)$ commute with all elements of P , $Z(P)$ normalizes every subgroup of P . By part 1. we have that $Z(P) \neq \{1\}$. If $Z(P)$ is not contained in H , then H is properly contained in $\langle H, Z(P) \rangle$, and the latter subgroup is contained in $N_P(H)$ so 4. holds. We may therefore assume $Z(P) \leq H$. Since $P/Z(P)$ has smaller order than P , by induction $H/Z(P)$ is properly contained in $N_{P/Z(P)}(H/Z(P))$. It follows directly from the Lattice Isomorphism Theorem that $N_P(H)$ is the complete preimage in P of $N_{P/Z(P)}(H/Z(P))$, hence we obtain proper containment of H in its normalizer in this case as well. This completes the induction.

To prove 5. let M be a maximal subgroup of P . By definition, $M < P$ so by part 4., $M < N_P(M)$. By definition of maximality we must therefore have $N_P(M) = P$, so $M \triangleleft P$. The Lattice Isomorphism Theorem shows that P/M is a p -group with no proper nontrivial subgroups because M is a maximal subgroup. By part 3., however, P/M has subgroups of every order dividing $|P/M|$. The only possibility is $|P/M| = p$. This proves 5. and completes the proof. ■

8.2.0 §Nilpotent Groups

Definition 8.2.1. For any (finite or infinite) group G define the following subgroups inductively:

$$Z_0(G) = \{1\}, \quad Z_1(G) = Z(G)$$

and $Z_{i+1}(G)$ is the subgroup of G containing $Z_i(G)$ such that

$$Z_{i+1}(G)/Z_i(G) = Z(G/Z_i(G))$$

(i.e., $Z_{i+1}(G)$ is the complete preimage in G of the center of $G/Z_i(G)$ under the natural projection). The resulting chain of subgroups

$$Z_0(G) \leq Z_1(G) \leq Z_2(G) \leq \dots$$

is called the **upper central series** of G .

Definition 8.2.2. A group G is called **nilpotent** if $Z_c(G) = G$ for some $c \in \mathbb{Z}$. The smallest such c is called the **nilpotence class** of G .

Each $Z_i(G)$ in this series is in fact characteristic in G .

Remark 8.2.1.

1. If G is abelian then G is nilpotent of class 1 (provided $|G| > 1$), since in this case $G = Z(G) = Z_1(G)$. We can think of the heirarchy of structure as follows:

$$\text{cyclic} \subset \text{abelian} \subset \text{nilpotent} \subset \text{solvable} \subset \text{all groups}$$

2. For any finite group there must, by order considerations, be an integer n such that

$$Z_n(G) = Z_{n+1}(G) = Z_{n+2}(G) = \dots$$

For example, $Z_n(S_3) = \{1\}$ for all $n \in \mathbb{Z}^+$. Once two terms in the upper central series are the same, the chain stabilizes at that point. By definition, $Z_n(G)$ is a proper subgroup of G for all n for non-nilpotent groups.

3. For infinite groups G it may happen that all $Z_i(G)$ are proper subgroups of G (so G is not nilpotent) but

$$G = \bigcup_{i=1}^{\infty} Z_i(G)$$

Groups for which this hold are called **hypercentnilpotent**. Results that we shall derive which do not involve the notion of order, Sylow subgroups, etc. also hold for infinite groups.

Proposition 8.2.1. Let p be a prime and let P be a group of order p^a . Then P is nilpotent of nilpotence class at most $a - 1$.

Proof. For each $i \geq 0$, $P/Z_i(P)$ is a p -group so if $|P/Z_i(P)| > 1$, then $|Z(P/Z_i(P))| > 1$, which implies that $|Z_{i+1}(P)| > |Z_i(P)|$, and in particular if $P \neq Z_i(P)$, then $|Z_{i+1}(P)| \geq p|Z_i(P)|$. This implies inductively that $|Z_{i+1}(P)| \geq p^{i+1}$, so $|Z_a(P)| \geq p^a$. Hence, $P = Z_a(P)$ and P is nilpotent with nilpotence class $\leq a$. The only way for P to be of nilpotence class equal to a is if $|Z_i(P)| = p^i$ for all $1 \leq i \leq a$. In this case, however, $Z_{a-2}(P)$ would have index p^2 in P , so $P/Z_{a-2}(P)$ would be abelian. But then $P/Z_{a-2}(P)$ would be its own center and $Z_{a-1}(P) = P$, a contradiction. Thus, the nilpotence class of P is at most $a - 1$. ■

Example 8.2.1. $D_{2^{n-1}}$, the dihedral group of order 2^n , is nilpotent of nilpotence class $n - 1$. This can be proved inductively by showing $|Z(D_{2^{n-1}})| = 2$ and $D_{2^{n-1}}/Z(D_{2^{n-1}}) \cong D_{2^{n-2}}$ for $n \geq 3$. If n is not a power of 2, D_n is not nilpotent.

(To Be Continued)

8.3.0 §Composition Series and Solvable Groups

The following proposition and proof shows how one can use the information on a normal subgroup N and on the quotient G/N to determine information about G :

Proposition 8.3.1. *If G is a finite abelian group and p is a prime dividing $|G|$, then G contains an element of order p .*

Proof. The proof proceeds by induction on $|G|$, namely, we assume the result is valid for every group whose order is strictly smaller than the order of G and then prove the result valid for G . Since $|G| > 1$, there is an element $x \in G$ with $x \neq 1$. If $|G| = p$ then x has order p by 2 and we are done. We may therefore assume $|G| > p$.

Suppose p divides $|x|$ and write $|x| = pn$. Then $|x^n| = p$, and again we have an element of order p . We may therefore assume p does not divide $|x|$.

Let $N = \langle x \rangle$. Since G is abelian, $N \trianglelefteq G$. By 2, $|G/N| = |G|/|N|$ and since $N \neq 1$, $|G/N| < |G|$. Since p does not divide $|N|$, we must have $p \nmid |G/N|$. We can now apply the induction assumption to the smaller group G/N to conclude it contains an element yN of order p . Since $y \notin N$, but $y^p \in N$, we must have $\langle y^p \rangle \neq \langle y \rangle$, that is, $|y^p| < |y|$. Since $|y^p| = \frac{|y|}{\gcd(p, |y|)}$, we must have that $p \nmid |y|$. We are now in the situation described in the preceding paragraph, so that the argument again produces an element of order p . The induction is complete. ■

Note that simple groups, groups without any normal subgroups, are fundamental obstructions to this variety of proof. As simple groups cannot be “factored” into pieces like N and G/N , the role they play is analogous to that of primes in the arithmetic of \mathbb{Z} .

Definition 8.3.1. *In a group G , a sequence of subgroups*

$$1 = N_0 \leq N_1 \leq N_2 \leq \dots \leq N_{k-1} \leq N_k = G$$

*is called a **composition series** if $N_i \trianglelefteq N_{i+1}$ and N_{i+1}/N_i is a simple group, $0 \leq i \leq k-1$. If the above sequence is a composition series, the quotient groups N_{i+1}/N_i are called **composition factors** of G .*

As an example, two composition series of D_4 are

$$1 \trianglelefteq \langle s \rangle \trianglelefteq \langle s, r^2 \rangle \trianglelefteq D_4 \quad \text{and} \quad 1 \trianglelefteq \langle r^2 \rangle \trianglelefteq \langle r \rangle \trianglelefteq D_4$$

Theorem 14 (Jordan-Hölder).

Let G be a finite group with $G \neq \{1\}$. Then

1. G has a composition series
2. The composition factors in a composition series are unique, namely, if $1 = N_0 \leq N_1 \leq \dots \leq N_r = G$ and $1 = M_0 \leq M_1 \leq \dots \leq M_s = G$ are two composition series for G , then

$r = s$ and there is some permutation π of $\{1, 2, \dots, r\}$ such that

$$M_{\pi(i)}/M_{\pi(i)-1} \cong N_i/N_{i-1}, 1 \leq i \leq r$$

Proof. (To be completed) ■

Definition 8.3.2. A group G is **solvable** if there is a chain of subgroups

$$1 = G_0 \trianglelefteq G_1 \trianglelefteq G_2 \trianglelefteq \dots \trianglelefteq G_s = G$$

such that G_{i+1}/G_i is abelian for $i \in \{0, 1, \dots, s-1\}$.

A property of finite solvable groups is the following due to Philip Hall:

Theorem 8.3.2. The finite group G is solvable if and only if for every divisor n of $|G|$ such that $\gcd\left(n, \frac{|G|}{n}\right) = 1$, G has a subgroup of order n .

Another illustration of how using information on a normal subgroup N and a quotient group G/N is seen in the following result:

Proposition 8.3.3. If N and G/N are solvable, then G is solvable.

Proof. Let $1 = N_0 \trianglelefteq N_1 \trianglelefteq \dots \trianglelefteq N_n = N$ be a chain of subgroups of N such that N_{i+1}/N_i is abelian, $0 \leq i < n$, and $1_{G/N} = G_0/N \trianglelefteq G_1/N \trianglelefteq \dots \trianglelefteq G_m/N = G/N$ be a chain of subgroups of G/N such that $(G_{i+1}/N)/(G_i/N)$ is abelian, $0 \leq i < m$. Such G_i exist by the Lattice Isomorphism Theorem, with $N \leq G_i$ for each i . By the Third Isomorphism Theorem

$$(G_{i+1}/N)/(G_i/N) \cong G_{i+1}/G_i$$

Thus

$$1 = N_0 \trianglelefteq N_1 \trianglelefteq \dots \trianglelefteq N_n = N = G_0 \trianglelefteq G_1 \trianglelefteq \dots \trianglelefteq G_m = G$$

is a chain of subgroups of G all of whose successive quotient groups are abelian. This proves G is solvable. ■

The Hölder Program

The holder program has two goals:

1. Classify all finite simple groups
2. Find all ways of “putting simple groups together” to form other groups

The classification of finite simple groups was completed in 1980, resulting in a proof of the following result:

Theorem 8.3.4. *There is a list consisting of 18 (infinite) families of simple groups and 26 simple groups not belonging to these families (the sporadic simple groups) such that every finite simple group is isomorphic to one of the groups in this list.*

One such family is $\{\mathbb{Z}/p\mathbb{Z} \mid p \text{ a prime}\}$. The “extension problem” is one of a much higher difficulty, even for groups of relatively small order.

Part II

Ring Theory

Chapter 9

§§Basic Definitions and Examples: Rings

9.1.0 §Initial Definitions and Examples

Definition 9.1.1. A set R with two binary operations $+$ (addition) and \cdot (multiplication), is called a unital ring if the following are satisfied:

1. $(R, +)$ is an abelian group with identity 0
2. (R, \cdot) is a monoid with identity 1
3. Distributivity: for all $a, b, c \in R$

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

and

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

Remark 9.1.1.

1. The multiplicative identity 1 is unique for a given ring R and multiplication \cdot .
2. We often denote $a \cdot b$ by ab
3. For all $r \in R$, $0 \cdot r = r \cdot 0 = 0$. Indeed, $r \cdot 0 = r \cdot (0 + 0) = r \cdot 0 + r \cdot 0$, so $r \cdot 0 = 0$. $0 \cdot r = 0$ is similar.
4. For all $r \in R$, $(-1) \cdot r = -r$. Indeed, $r + (-1) \cdot r = 1 \cdot r + (-1) \cdot r = (1 + (-1)) \cdot r = 0 \cdot r = 0$, so $(-1) \cdot r = -r$.
5. Powers in the group $(R, +)$ are denoted

$$na = \begin{cases} 0, & \text{if } n = 0 \\ \underbrace{a + a + \dots + a}_{n\text{-fold times}}, & \text{if } n > 0 \\ \underbrace{(-a) + (-a) + \dots + (-a)}_{-n\text{-fold times}}, & \text{if } n < 0 \end{cases} \quad (9.1.1)$$

Moreover, for all $m, n \in \mathbb{Z}$, and for all $a \in R$:

- (a) $m(na) = (mn)a$
- (b) $(m + n)a = ma + na$

and for all $n' \cdot a = na$ for all $n \in \mathbb{Z}$ for $n' \in R$ defined by:

$$n = \begin{cases} 0, & \text{if } n = 0 \\ \underbrace{1_R + 1_R + \dots + 1_R}_{n\text{-fold times}}, & \text{if } n > 0 \\ \underbrace{(-1_R) + (-1_R) + \dots + (-1_R)}_{-n\text{-fold times}}, & \text{if } n < 0 \end{cases} \quad (9.1.2)$$

How small can a ring be?

Answer. The smallest ring is $R = \{0\}$, the zero ring, so $1 = 0$.

Construction 9.1.2 (Endomorphism Rings). *The best way to obtain rings (which are called **endomorphism rings**) is to start with an abelian group $(A, +, 0)$. Let $R = \text{End}(A) := \{f : A \rightarrow A : f \in \mathbf{Hom}_{\mathbf{Grp}}(A, A)\}$. We define addition on R to be*

$$(f + g)(x) := f(x) + g(x), \forall f, g \in R, \forall x \in A \quad (9.1.3)$$

Since the addition in the group A is commutative so is the addition on R . Moreover, we have zero element

$$0_R(x) = 0_A, \forall x \in A \quad (9.1.4)$$

so $0_R + f = f$. Additive inverses are defined such that $(-f)(a) = -(f(a))$. We define the multiplication law as

$$(f \times g)(a) = f(g(a)) \quad (9.1.5)$$

Then this operation is naturally associative, and the multiplicative identity is

$$1_R(a) = a, \forall a \in A \quad (9.1.6)$$

Note that from these definitions, we see that multiplication is not necessarily commutative, and does not necessarily have an inverse, as f has an inverse \iff it is an isomorphism of groups. That is, the group of units of R is $R^\times = \text{Aut}(A)$ equipped with the multiplication operation of function composition.

Example 9.1.1 (Constructing Rings from Endomorphisms on Cyclic Groups). I claim that the ring $(\mathbb{Z}, +, \cdot, 1, 0) = \text{End}(\mathbb{Z}, +, 0)$. Suppose we have a group homomorphism $f : \mathbb{Z} \rightarrow \mathbb{Z}$. Then $f(1) = n \in \mathbb{Z}$ determines everything, since $f(k) = f(1 + 1 + \dots + 1) = f(1) + \dots + f(1) = kf(1)$. We take f , and associate to it the integer $f(1)$, which then gives a multiplication on \mathbb{Z} . For example: Suppose $f(1) = n$, and $g(1) = m$, then $f \times g(k) = f(g(k)) = f(k \cdot m) = n(k \cdot m)$, which gives multiplication on \mathbb{Z} . Now, suppose f is associated to a negative integer, so $f(1) = n < 0$, then it switches the halves of the real line. Then, $f \times f(1) = f(f(1)) = f(n) = f(-1 - 1 - \dots - 1) = -f(1) - f(1) - \dots - f(1) = -n - n - \dots - n > 0$. Likewise, $\mathbb{Z}/n\mathbb{Z} = \text{End}(\mathbb{Z}/n\mathbb{Z}, +, 0)$, where we identify f by $f(1)$. This works to give a ring structure on cyclic groups.

Example 9.1.2 (Constructing Rings from Endomorphisms of other Abelian group). Take $A = (\mathbb{Z}/p\mathbb{Z})^2 = \{(a_1, a_2) : a_i \in \mathbb{Z}/p\mathbb{Z}\}$. Then $\text{End}(A) = M_2(\mathbb{Z}/p\mathbb{Z})$. If we have a matrix

$$B = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \quad (9.1.7)$$

matrix multiplication gives us the multiplication in our ring. This is an example of a non-commutative ring. In general, if $A = (\mathbb{Z}/p\mathbb{Z})^n$, then $\text{End}(A) = M_n(\mathbb{Z}/p\mathbb{Z})$.

Definition 9.1.3. The order of 1 in $(R, +)$ is called the characteristic of the ring R , and denoted $\text{char } R$, if $o(1_R) < +\infty$. If $o(1) = +\infty$ then $\text{char } R := 0$. In general, for a non-unital ring R' , $\text{char } R'$ is the smallest positive integer n such that $n \cdot r = 0_{R'}$ for all $r \in R'$. If no such n exists then $\text{char } R' := 0$.

Example 9.1.3.

1. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ with the usual addition and multiplication are all rings of characteristic 0.
2. For all $n > 0$, $\mathbb{Z}/n\mathbb{Z}$ is a ring with addition and multiplication modulo n of characteristic n .
3. Let $M_n(\mathbb{R})$ be the set of $n \times n$ real matrices. Then $M_n(\mathbb{R})$ is a ring for the usual addition and multiplication of matrices, 0 = the zero matrix and $1 = I_n$. Indeed, $M_n(R)$ is a ring for any arbitrary ring R .
4. Let X be a set and R a ring. Then, the set $F(X, R)$ of all functions $f : X \rightarrow R$ is a ring for pointwise addition and multiplication:

$$\begin{aligned} (f_1 + f_2)(x) &:= f_1(x) + f_2(x) \\ (f_1 \cdot f_2)(x) &:= f_1(x) \cdot_R f_2(x) \end{aligned} \quad \forall x \in X, \forall f_1, f_2 \in F(X, R) \quad (9.1.8)$$

Moreover, $0(x) = 0$ for all $x \in X$ and $1(x) = 1_R$ for all $x \in X$.

5. Let $X = \mathbb{R}$ and $R = \mathbb{R}$. A function $f \in F(\mathbb{R}, \mathbb{R})$ is a polynomial if it can be written in the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \quad (9.1.9)$$

for some $a_i \in \mathbb{R}$, $0 \leq i \leq n$ and $n \in \mathbb{Z}$, $n \geq 0$.

↳ Polynomial functions $\mathcal{P}(\mathbb{R}, \mathbb{R})$ form a ring for the addition and multiplication in $F(\mathbb{R}, \mathbb{R})$ - This is a subring

↳ (eg: $f(x) = 2x^2 + 1$, $g(x) = 6x$, $f \cdot g(x) = 12x^3 + 6x$ is a polynomial)

6. Let G be an abelian group and let $\text{End}(G)$ be the set of endomorphisms on G . Then, $\text{End}(G)$ is a ring with addition defined pointwise and multiplication defined by function composition:

$$\begin{aligned} (f_1 + f_2)(g) &:= f_1(g) + f_2(g) \\ (f_1 \circ f_2)(g) &:= f_1(f_2(g)) \end{aligned} \quad \forall g \in G, \forall f_1, f_2 \in \text{End}(G) \quad (9.1.10)$$

Moreover, $0(g) = e_G$ and $1(g) = \text{Id}(g) = g$ for all $g \in G$.

Note 9.1.2. The multiplication need not be commutative. For instance multiplication is not necessarily commutative in $M_n(\mathbb{R})$ for $n \geq 2$.

Definition 9.1.4. If the multiplication of a ring R is commutative ($ab = ba \forall a, b \in R$) then R is said to be a commutative ring

Definition 9.1.5. Let R be a ring, the set of invertible elements for the multiplication is called the group of units of R , and is denoted R^\times .

Recall 9.1.3. For all $a \in R$, a is invertible for \cdot if there exists $b \in R$ such that $ab = ba = 1_R$.

Exercise 9.1.4. (R^\times, \cdot) is a group.

Proof. (Left to the reader) ■

Example 9.1.5.

1. We've seen $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$, $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$, and $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$. But, in general, this is not enough. For instance, $(\mathbb{Z}/n\mathbb{Z})^\times \neq \mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}$ for all $n \geq 1$.
2. $\mathbb{Z}^\times = \{1, -1\} \cong \mathbb{Z}/2\mathbb{Z}$
3. $M_n(\mathbb{R})^\times = \mathbf{GL}_n(\mathbb{R})$
4. $F(X, \mathbb{R})^\times = \{f \in F(X, \mathbb{R}) : \forall x \in X, f(x) \neq 0\}$
5. In general, $F(X, R)^\times = \{f \in F(X, R) : \forall x \in X, f(x) \in R^\times\}$

Definition 9.1.6. If R is a ring such that $R^\times = R \setminus \{0_R\}$ (that is every nonzero element is invertible for \cdot), then R is called a division ring, or skew-field

↳ If R is a commutative division ring then R is called a field.

Example 9.1.6.

1. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields (\mathbb{Z} is not a field)
2. $\mathbb{Z}/p\mathbb{Z}$ for P a prime is a field, denoted \mathbb{F}_p , called the finite field of order p
3. $\mathbb{Z}/n\mathbb{Z}$ is not a field if n is not a prime
4. Division rings which are not commutative rings are rare. An example of them are the Quaternions.

§Integral Domains

Definition 9.1.7. For a ring R , $a \in R$ is called a zero divisor if there exists $b \in R$, $b \neq 0$, such that $ab = 0$ or $ba = 0$.

Remark 9.1.4.

1. If $1 \neq 0$, then 0 is a zero divisor.
2. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is a non-zero zero divisor of $M_2(\mathbb{R})$.
3. If r is a divisor of n and $r \neq 1$, then $[r]_n \in \mathbb{Z}/n\mathbb{Z}$ is a zero divisor. Indeed, $n = rq$ for $q \in \mathbb{Z}$. If $r = n$, $q = 1$, and $[r]_n[1]_n = [0]_n$, where $[1]_n \neq [0]_n$. Otherwise, if $1 < r < n$, $1 < q < n$, so $[q]_n \neq [0]_n$. But, $[r]_n[q]_n = [0]_n$ so $[r]_n$ is a zero divisor.

Example 9.1.7.

1. The set of zero divisors in a division ring is $K = \{0\}$. Indeed, if $a \in K$, $ab = 0$ for some $b \neq 0$, then $a = abb^{-1} = 0b^{-1} = 0$.
2. The set of zero divisors of \mathbb{Z} is $\{0\}$, even though \mathbb{Z} is not a division ring.

Proposition 9.1.1. *Let R be a ring. The following are equivalent:*

1. $0 \in R$ is the only zero divisor.
2. For all $a, b \in R$, $ab = 0$ implies $a = 0$ or $b = 0$
3. For all $a, b, c \in R$, $ab = ac$ and $a \neq 0$ implies $b = c$.
4. For all $a, b, c \in R$, $ba = ca$ and $a \neq 0$ implies $b = c$

\hookrightarrow ((3) and (4) are called cancellation laws)

Proof. (Left to the reader) ■

Definition 9.1.8. *If the equivalent conditions of the proposition are satisfied for a ring $R \neq \{0\}$, then R is called a domain. If R is also a commutative ring then R is said to be an integral domain.*

Remark 9.1.5. Every division ring is a domain and every field is an integral domain, but the converse is not true.

$\hookrightarrow \mathbb{Z}$ is an integral domain, but not a field. $\mathcal{P}(\mathbb{R}, \mathbb{R})$ is an integral domain.

§Subrings

Definition 9.1.9. *A subset S of a ring R that is closed under addition, subtraction, multiplication, and contains 1 is called a subring of R .*

$\hookrightarrow \forall a, b \in S, \{a + b, a - b, ab, 1\} \subseteq S$.

Remark 9.1.6. In other words, S is a subring of R if and only if $(S, +)$ is a subgroup of $(R, +)$ and (S, \cdot) is a monoid with identity $1 \in R$.

Note 9.1.7.

1. There is no standard notation for “ S is a subring of R ”
2. The definition directly implies that the intersection of an arbitrary number of subrings is again a subring:

Proof. (Left to the reader) ■

Definition 9.1.10. The subring generated by a subset $X \subseteq R$ is the intersection of all subrings of R containing X .

↳ (R is a subring of R , so there is always at least one subring of R containing X and the definition is well-defined)

Example 9.1.8.

1. $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are all subrings
2. $M_n(\mathbb{Z}) \subseteq M_n(\mathbb{Q}) \subseteq M_n(\mathbb{R}) \subseteq M_n(\mathbb{C})$ are all subrings
3. $\mathcal{P}(\mathbb{R}, \mathbb{R}) \subseteq F(\mathbb{R}, \mathbb{R})$ is a subring
4. The subring of \mathbb{C} generated by i is

$$\{a + bi : a, b \in \mathbb{Z}\} =: \mathbb{Z}[i] \subseteq \mathbb{C} \quad (9.1.11)$$

and is called the **Gaussian integers**.

Exercise 9.1.9. $\mathbb{Z}[i]$ is an integral domain.

Proof. (Left to the reader) ■

5. The subring of \mathbb{C} generated by $\frac{1}{2}$

$$\left\{ \frac{a}{2^n} : a \in \mathbb{Z}, n \geq 0 \right\} \subseteq \mathbb{Q} \subseteq \mathbb{C}$$

it is an integral domain as well

6. The set of all upper triangular matrices, $T_n(\mathbb{R})$, is a subring of $M_n(\mathbb{R})$. In general, $T_n(R)$ is a subring of $M_n(R)$ for an arbitrary ring R .
7. The **center** of a ring R is

$$Z(R) := \{r \in R : ra = ar \forall a \in R\} \quad (9.1.12)$$

It is a subring of R

↳ If $b \in Z(R)$, b is called a central element of R .

Example 9.1.10. i. $Z(\mathbb{R}) = \mathbb{R}$, and similarly $Z(R) = R$ for any commutative ring R

ii. $Z(M_n(\mathbb{R})) = \mathbb{R}I_n$

8. A subring of a field which is itself a field is called a subfield.

↳

Example 9.1.11. $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are all subfields

9.2.0 §Ring Homomorphisms

Definition 9.2.1. Let R, S be rings. A map

$$R \xrightarrow{f} S$$

is called a ring homomorphism if the following conditions are satisfied for all $r, r' \in R$:

1. $f(r + r') = f(r) + f(r')$
2. $f(rr') = f(r)f(r')$
3. $f(1_R) = 1_S$

A bijective ring homomorphism $A \xrightarrow{\phi} B$ is called a ring isomorphism, and A, B are said to be isomorphic rings.

↳ (In the case that f is surjective, $f(1_R) = 1_S$ follows from the multiplicative condition)

Remark 9.2.1. The image of a ring homomorphism is a subring of the codomain.

Example 9.2.1.

1. The identity $\text{Id} : R \rightarrow R$ is a ring isomorphism

2. $\begin{matrix} \mathbb{Z} \rightarrow R \\ n \mapsto n \cdot 1_R \end{matrix}$ is a ring homomorphism

↳

Example 9.2.2. $\begin{matrix} \mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z} \\ r \mapsto [r]_n \end{matrix}$

3. The inclusion $S \subseteq R$ of a subring is a ring homomorphism.

4. Let $a \in X$. Then

$$\begin{aligned} \mathbf{ev}_a : F(X, R) &\rightarrow R \\ f &\mapsto f(a) \end{aligned} \quad (9.2.1)$$

is a ring homomorphism called the evaluation at a

$$\begin{aligned} \hookrightarrow \text{Indeed, } \mathbf{ev}_a(f + g) &= (f + g)(a) = f(a) + g(a) = \mathbf{ev}_a(f) + \mathbf{ev}_a(g), \mathbf{ev}_a(f \cdot g) = \\ &= (f \cdot g)(a) = f(a) \cdot g(a) = \mathbf{ev}_a(f) \cdot \mathbf{ev}_a(g), \text{ and } \mathbf{ev}_a(1) = 1(a) = 1. \end{aligned}$$

5. If $|R| = p$, a prime number, then R is isomorphic to the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Proof. Suppose R is a ring of order p . Then, note that by definition $\text{char } R = o(1_R)$ in $(R, +)$. Then, by 2 $o(1_R) \mid p$. Hence, $o(1_R) \in \{1, p\}$. Note that if $o(1_R) = 1$ then $1_R = 0_R$ so R is the zero ring and $|R| = 1$. But, as 1 is not a prime integer this is impossible. Thus $o(1_R) = p$. Now, define the map

$$\begin{aligned} \mathbb{Z}/p\mathbb{Z} &\rightarrow R \\ [n]_p &\mapsto n \cdot 1_R \end{aligned}$$

. First, if $[n]_p = [m]_p$ then $n - m \mid p$. Hence, $(n - m) \cdot 1_R = 0_R$ since $\text{char } R = p$. Thus, by distributivity $n \cdot 1_R - m \cdot 1_R = 0_R$. By addition of $m \cdot 1_R$ on both sides we find $n \cdot 1_R = m \cdot 1_R$. Thus, the map is well-defined. Moreover, $\phi([1]_p) = 1 \cdot 1_R = 1_R$, and for all $[n]_p, [m]_p \in \mathbb{F}_p$, we have

$$\phi([n + m]_p) = (n + m) \cdot 1_R = n \cdot 1_R + m \cdot 1_R = \phi([n]_p) + \phi([m]_p)$$

and

$$\phi([nm]_p) = (nm) \cdot 1_R = n \cdot (m \cdot 1_R) = (n \cdot 1_R) \cdot (m \cdot 1_R) = \phi([n]_p) \cdot \phi([m]_p)$$

Hence, we find that ϕ is a ring homomorphism. Finally, if $[k]_p \in \ker(\phi)$, then $k \mid p$, which implies $[k]_p = [0]_p$ so $\ker(\phi) = \{[0]_p\}$, and since both sets are finite (and of the same order) we conclude that ϕ is a bijection. Therefore, ϕ is a ring isomorphism so $R \cong \mathbb{F}_p$, as claimed. ■

6. $\mathbb{C} \xrightarrow{\theta} M_2(\mathbb{R})$
 $a + bi \mapsto \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ is an injective ring homomorphism.

$$\hookrightarrow \text{Hence, } \mathbb{C} \xrightarrow{\sim} \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} : a, b \in \mathbb{R} \right\} \subseteq M_2(\mathbb{R})$$

Remark 9.2.2. The property of being a field or an integral domain is preserved under ring isomorphism.

Remark 9.2.3 (Canonical Map). If we have a commutative ring R , there is a natural ring homomorphism $f : \mathbb{Z} \rightarrow R$ which is completely characterized by $f(1) = 1_R$, so for $n \geq 1$, $f(n) = f(\underbrace{1 + 1 + \dots + 1}_{n\text{-times}}) = \underbrace{1_R + \dots + 1_R}_{n\text{-times}}$ and $f(-n) = -f(n)$. This is the canonical ring

homomorphism associated to any commutative ring. Moreover, we know that the kernel of f is an ideal of \mathbb{Z} , so it is of the form $n\mathbb{Z}$ for some $n \geq 0$. If $R = \{0\}$, then $\ker(f) = \mathbb{Z}$, and if $R = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, then $\ker(f) = 0\mathbb{Z} = \{0\}$. Moreover, if $R = \mathbb{Z}/n\mathbb{Z}$, then $\ker(f) = n\mathbb{Z}$.

Remark 9.2.4 (Think about this). If R is a field, and h is the natural homomorphism given above, then $\ker h = \{0_R\}$, or $\ker h = p\mathbb{Z}$, where p is prime.

Proposition 9.2.1. *If R is a field, then $\ker(f) = \{0\}$ or $\ker(f) = p\mathbb{Z}$ for p a prime.*

Proof. For the sake of contradiction suppose $\ker(f) = n\mathbb{Z}$ for $n = ab$ composite, so $1 < a, b < n$. Then $f(n) = 0$ in R . But, $f(n) = f(a)f(b) = a_R b_R = 0_R$, so since R is a field, a_R is zero or b_R is zero. However, this contradicts the fact that the $\ker(f)$ is a multiple of n , and $a, b \notin n\mathbb{Z}$. ■

9.3.0 §Domains and Fields of Fractions

Proposition 9.3.1. *The characteristic of a domain is zero ($o(1) = \infty$) or a prime number p .*

Proof. Suppose that R is a domain. If R has a zero characteristic we are done, so suppose $\text{char } R = n$ where $n \in \mathbb{Z}$ and $n \geq 1$. If $n = 1$ then $R = \{0\}$, which contradicts the fact that R is a domain. Thus, $n > 1$. We argue by contradiction and suppose n is not prime. Then there exist $r, s \in \mathbb{Z}$ with $1 < r, s < n$ such that $n = rs$. It follows that $(r \cdot 1_R)(s \cdot 1_R) = rs \cdot 1_R = n \cdot 1_R = 0_R$ by definition of the characteristic of a ring. However, since R is a domain it follows that $r \cdot 1_R = 0_R$ or $s \cdot 1_R = 0_R$. However, $r, s < n$, so either case would contradict the minimality of n in $o(1) = n$. Thus, n being composite leads to a contradiction so we conclude that n must be prime, as claimed. ■

Remark 9.3.1. Every subring of a field is an integral domain.

↳ ($R \subseteq F_{\text{field}}$, then for $a, b \in R$, if $ab = 0$ in F and $b \neq 0$, then $0 = abb^{-1} = a \in F$. Thus $a = 0 \in R$. Hence, R is an integral domain)

Remark 9.3.2. Actually, every subring of a division ring, or skew-field, is a domain.

Theorem 9.3.2. *Every integral domain is a subring of a field.*

Construction 9.3.1. Denote $R \setminus \{0\} = R^*$. We start with an integral domain R , and consider pairs $(a, b) \in R \times R^*$. We define a relation \sim on $R \times R^*$ by $(a, b) \sim (a', b')$ if and only if $ab' = a'b$.

Claim 9.3.3. \sim is an equivalence relation on $R \times R^*$.

Proof. Let $(a, b), (a', b'), (a'', b'') \in R \times R^*$. First, $(a, b) \sim (a, b)$ since by reflexivity of “=” $ab = ab$, so \sim is reflexive. Then, suppose $(a, b) \sim (a', b')$, so $ab' = a'b$. By the symmetry of “=” we have $a'b = ab'$, so $(a', b') \sim (a, b)$. Hence \sim is symmetric. Now, suppose $(a', b') \sim (a'', b'')$, so $a'b'' = a''b'$. Then observe that

$$\begin{aligned} ab''a' &= aa''b' \\ &= ab'a'' \end{aligned} \quad (\text{commutativity})$$

$$\begin{aligned}
 &= a'ba'' \\
 &= a''ba' \quad (\text{commutativity})
 \end{aligned}$$

Then, we have that $(ab'' - a''b)a' = 0_R$ by distributivity. Note if $a' = 0_R$, then $a''b' = a'b'' = 0_R$, so $a'' = 0_R$ since R is an integral domain, and similarly $ab' = a'b = 0_R$ so $a = 0$ and $ab'' = 0_R = a''b$. Now, suppose $a' \neq 0$. Then as R is an integral domain $ab'' = a''b$, so $(a, b) \sim (a'', b'')$ and the relation is transitive, as desired. Therefore, \sim is an equivalence relation of $R \times R^*$. ■

↳ We define addition and multiplication on the set

$$\text{Frac}(R) := \{[(a, b)]_{\sim} : (a, b) \in R \times R^*\} \quad (9.3.1)$$

of equivalence classes by

$$\begin{aligned}
 [(a, b)] + [(c, d)] &:= [(ad + cb, bd)] \\
 [(a, b)] \cdot [(c, d)] &:= [(ac, bd)]
 \end{aligned} \quad (9.3.2)$$

↳ Let us see that these operations are well-defined. Consider $[(ad + cb, bd)]$ and $[(a'd' + c'b', b'd')]$ for $[(a, b)] = [(a', b')]$ and $[(c, d)] = [(c', d')]$. Then, observe that

$$\begin{aligned}
 (ad + cb)(b'd') - (a'd' + c'b')(bd) &= adb'd' + cbb'd' - a'd'bd - c'b'bd \\
 &= a'bdd' + c'dbb' - a'bdd' - c'dbb' \\
 &= 0_R
 \end{aligned}$$

so $[(ad + cb, bd)] = [(a'd' + c'b', b'd')]$ and the addition is well defined. Similarly,

$$\begin{aligned}
 (ac)(b'd') - (a'c')(bd) &= acb'd' - a'c'bd \\
 &= a'bcd' - a'bcd' \\
 &= 0_R
 \end{aligned}$$

so $[(ac, bd)] = [(a'c', b'd')]$, and multiplication is also well-defined. Furthermore, we observe that $0_{\text{Frac}(R)} = [(0_R, b)]$ since $0_R \cdot b = 0_R = 0_R \cdot b'$ for all $b, b' \in R^*$. Additionally, $-[(a, b)] = [(-a, b)]$ and $1_{\text{Frac}(R)} = [(1, 1)]$. Note that it is a tedious but rudimentary check to see that $\text{Frac}(R)$ as defined is a commutative ring.

Claim 9.3.4. $\text{Frac}(R)$ is a field.

Proof. Since $\text{Frac}(R)$ is a commutative ring, all we must show is that all non-zero elements are invertible. Indeed, $[(a, b)] \neq 0_{\text{Frac}(R)}$ if and only if $a \neq 0$. Hence $a \in R^*$ so $[(b, a)] \in \text{Frac}(R)$ is well-defined. Then, $(ab, ba) \sim (1, 1)$ since $ab = ba$ by commutativity, so $[(a, b)]^{-1} = [(b, a)]$, and in particular $[(a, b)]$ is invertible. Thus, $\text{Frac}(R)$ is a field, as claimed. ■

Theorem 15 (Universal Property of Field of Fractions).

Let R be an integral domain.

1. $\text{Frac}(R)$ is a field containing R as a subring by the inclusion

$$\begin{aligned} R &\hookrightarrow \text{Frac}(R) \\ r &\mapsto [(r, 1)] \end{aligned} \quad (9.3.3)$$

2. If $R \xrightarrow{j} \mathbb{F}$ is an injective ring homomorphism of rings with \mathbb{F} a field, then there exists a unique injective homomorphism of rings $\text{Frac}(R) \xrightarrow{f} \mathbb{F}$ with $f \circ i = j$:

$$\begin{array}{ccc} R & \xrightarrow{j} & \mathbb{F} \\ & \searrow i & \nearrow \exists! f \\ & \text{Frac}(R) & \end{array}$$

Proof. 1. Define i as above. Then, observe that $i(1) = [(1, 1)] = 1_{\text{Frac}(R)}$, $i(a + b) = [(a + b, 1)] = [(a, 1)] + [(b, 1)] = i(a) + i(b)$, and $i(ab) = [(ab, 1)] = [(a, 1)][(b, 1)] = i(a)i(b)$. Moreover, $i(a) = 0_{\text{Frac}(R)}$ if and only if $a = 0_R$, so i is an injective ring homomorphism, as desired.

2. Suppose $j : R \hookrightarrow \mathbb{F}$ is an injective ring homomorphism and define $f : \text{Frac}(R) \rightarrow \mathbb{F}$ by $f([(a, b)]) := j(a)j(b)^{-1}$. Note that since $b \neq 0_R$ and j is injective, $j(b) \neq 0_{\mathbb{F}}$ so $j(b) \in \mathbb{F}^\times$. Now, suppose $[(a, b)] = [(a', b')]$, so $ab' = a'b$. Then $j(a)j(b') = j(a')j(b)$ by multiplicativity. It follows that $j(a)j(b)^{-1} = j(a')j(b')^{-1}$, so the map f is well-defined. Moreover, $f([(1, 1)]) = j(1)j(1)^{-1} = 1_{\mathbb{F}}$,

$$\begin{aligned} f([(aa', bb')]) &= j(aa')j(bb')^{-1} \\ &= (j(a)j(b)^{-1})(j(a')j(b')^{-1}) \\ &= f([(a, b)])f([(a', b')]) \end{aligned}$$

and

$$\begin{aligned} f([(ab' + a'b, bb')]) &= j(ab' + a'b)j(bb')^{-1} \\ &= (j(a)j(b')j(b)^{-1}j(b')^{-1} + j(a')j(b)j(b)^{-1}j(b')^{-1}) \\ &= j(a)j(b)^{-1} + j(a')j(b')^{-1} \\ &= f([(a, b)]) + f([(a', b')]) \end{aligned}$$

Hence, f is a ring homomorphism. Moreover, for all $a \in R$, $f \circ i(a) = f([(a, 1)]) = j(a)j(1)^{-1} = j(a)$, so $f \circ i = j$ as desired. Injectivity shall be shown by the Lemma to follow. Now, suppose $[(a, b)] \in \text{Frac}(R)$. Then, observe that

$$\begin{aligned} g([(a, b)]) &= g([(a, 1)])g([(1, b)]) && \text{(by multiplicativity)} \\ &= (g \circ i(a))g([(b, 1)])^{-1} && \text{(by multiplicativity and } b \neq 0) \\ &= j(a)(g \circ i(b))^{-1} \\ &= j(a)j(b)^{-1} \\ &= f([(a, b)]) && \text{(by definition)} \end{aligned}$$

Thus we have that $f = g$, so the map is unique. ■

Lemma 9.3.5. Let $K \xrightarrow{f} R$ be a ring homomorphism with K a field and $R \neq \{0\}$. Then f is injective.

Proof. Take $a \neq 0, a \in K$. Then $1_R = f(1_K) = f(aa^{-1}) = f(a)f(a^{-1})$. If $f(a) = 0$ then $1 = 0$, but by assumption $R \neq \{0\}$ so $f(a) \neq 0$. Then $a \notin \ker(f)$, and in particular $\ker(f) = \{0_K\}$. Thus, f is injective. ■

Example 9.3.1.

1. $\text{Frac}(\mathbb{Z}) \cong \mathbb{Q}$
2. $\text{Frac}(\mathbb{F}) \cong \mathbb{F}$ for any field \mathbb{F} .

Proof. Indeed, consider the identity $\text{Id}_{\mathbb{F}} : \mathbb{F} \rightarrow \mathbb{F}$, which is an injective ring homomorphism. Then, by the universal property of \mathbb{F} 's field of fractions there exists a unique injective ring homomorphism $j : \text{Frac}(\mathbb{F}) \rightarrow \mathbb{F}$ such that $j \circ i = \text{Id}_{\mathbb{F}}$. Then, let $f \in \mathbb{F}$, and observe that $j(i(f)) = \text{Id}_{\mathbb{F}}(f) = f$, so j is also a surjection. Thus, j is an isomorphism of rings, so $\text{Frac}(\mathbb{F}) \cong \mathbb{F}$ as claimed. ■

3. $\text{Frac}(\mathbb{Z}[1/2]) \cong \mathbb{Q}$
4. $\text{Frac}(\mathbb{Z}[i]) \cong \mathbb{Q}[i]$. Indeed, by the universal property we have an injection $\text{Frac}(\mathbb{Z}[i]) \hookrightarrow \mathbb{Q}[i]$, and $\mathbb{Q} \hookrightarrow \text{Frac}(\mathbb{Z}[i])$, which implies $\mathbb{Q}[i] \hookrightarrow \text{Frac}(\mathbb{Z}[i])$ since $i \in \text{Frac}(\mathbb{Z}[i])$.
5. For $R = \mathcal{P}(\mathbb{R}, \mathbb{R})$,

$$\text{Frac}(R) = \{f/g : f, g \in \mathcal{P}(\mathbb{R}, \mathbb{R}), g \neq 0\} \quad (9.3.4)$$

is called the field of rational polynomials.

9.4.0 §§Special Definitions and Facts

Definition 9.4.1. A ring R is said to be a local ring if the set of non-units in R is an ideal.

Proposition 9.4.1. If R is a local ring with ideal of non-units $J(R)$, then $R/J(R)$ is a division ring.

Proof. Let R be a local ring with ideal of non-units $J(R)$. Then, let $a + J(R) \in R/J(R)$ such that $a + J(R) \neq 0_{R/J(R)}$, so $a \notin J(R)$. It follows that a is a unit of R so there exists $b \in R$ such that either $ab = 1_R$ or $ba = 1_R$. Without loss of generality suppose $ab = 1_R$. Then it follows that $(a + J(R))(b + J(R)) = ab + J(R) = 1_R + J(R) = 1_{R/J(R)}$ in $R/J(R)$. Thus, every non-zero element in $R/J(R)$ has an inverse so $R/J(R)$ is a division ring. ■

Proposition 9.4.2. If R is a local ring with ideal of non-units $J(R)$, and $A \subseteq J(R)$ is an ideal of R , then R/A is local and $J(R/A) = \{r + A : r \in J(R)\}$.

Proof. Let R be a local ring with ideal of non-units $J(R)$, and let $A \subseteq J(R)$ be an ideal of R . Then, by the 19 for quotient rings we have that $J(R)/A = \{r + A : r \in J(R)\}$ is an ideal of R/A . I claim that $J(R)/A = J(R/A)$ in the proposition. Let $r + A \in R/A$ be a non-unit. For the sake of contradiction suppose that $r + A \notin J(R)/A$. Then $r \notin J(R)$, so r must be a unit of R . Then there exists $r' \in R$ such that $rr' = 1_R$ or $r'r = 1_R$. Without loss of generality suppose $rr' = 1_R$. Then $(r + A)(r' + A) = rr' + A = 1_R + A$, so $(r + A)$ is a unit of R/A , which contradicts the assumption that $r + A$ is a non-unit. Thus, $r + A \in J(R)/A$, so $J(R/A) \subseteq J(R)/A$. Next, let $r + A \in J(R)/A$, so $r \in J(R)$. Again towards a contradiction suppose $r + A \notin J(R/A)$. Then there exists $r' + A \in R/A$ such that $(r + A)(r' + A) = 1_R + A$ or $(r' + A)(r + A) = 1_R + A$. Without loss of generality suppose $(r + A)(r' + A) = 1_R + A$, so $rr' + A = 1_R + A$. Note that since $J(R)$ is an ideal, $rr' \in J(R)$ so $rr' + A \in J(R)/A$. Then we have that $rr' - 1_R \in A \subseteq J(R)$, so $1_R = rr' + (- (rr' - 1_R)) \in J(R)$. However, 1_R is a unit in R , and $J(R)$ is the set of non-units which is a contraction. Thus, we conclude that $r + A \in J(R/A)$, so $J(R)/A \subseteq J(R/A)$. Hence, $J(R/A) = J(R)/A$ is an ideal, so R/A is a local ring as claimed. ■

9.5.0 §The Gaussian Integers

Example 9.5.1. Consider $R = \mathbb{Z}[i]$. What if we want $2 + i = 0$? Let $I = (2 + i)$ and take $\bar{R} = R/I$. We wish to identify \bar{R} . First, let's identify the intersection $I \cap \mathbb{Z}$. Note that $0 = (2 + i)(2 - i) = 4 + 1 = 5$, so $5 \in I \cap \mathbb{Z}$. In particular, $5\mathbb{Z} \subset I \cap \mathbb{Z}$, where $5\mathbb{Z}$ is a maximal subgroup of \mathbb{Z} , so in fact it is a maximal ideal. Therefore, either $I \cap \mathbb{Z} = \mathbb{Z}$ or $I \cap \mathbb{Z} = 5\mathbb{Z}$. Secondly, observe that if $(2 + i)(a + bi) \in \mathbb{Z}$, then $(2a - b) + (2b + a)i \in \mathbb{Z}$. In particular, $2b + a = 0$, so $a = -2b$. It follows that $(2 + i)(a + bi) = 2(-2b) - b = -4b - b = 5(-b) \in 5\mathbb{Z}$. Therefore, $I \cap \mathbb{Z} = 5\mathbb{Z}$. Then, if we take the canonical homomorphism $\mathbb{Z} \rightarrow R/I = \bar{R}$, it has kernel $5\mathbb{Z}$, and image $\cong \mathbb{Z}/5\mathbb{Z}$. In fact, $\bar{R} \cong \mathbb{Z}/5\mathbb{Z}$ under this map, or in other words, the map is surjective. Note that since $2 + i \equiv 0 \pmod{I}$, $i \equiv -2 \pmod{I}$, and $a + bi \equiv a - 2b \pmod{I}$ in R/I , but $a - 2b \in \mathbb{Z}$. Thus, the map is surjective, so the image of the integers, $\mathbb{Z}/5\mathbb{Z}$, must be isomorphic to R/I .

Theorem 9.5.1. More generally, if p is a prime number with $p \equiv 1 \pmod{4}$, there is an ideal $I \subset \mathbb{Z}[i] = R$ with $R/I \cong \mathbb{Z}/p\mathbb{Z}$.

Proof. First, note that for the canonical homomorphism $f : R \rightarrow R/I$, if $R/I \cong \mathbb{Z}/p\mathbb{Z}$, then $f(i)$ must have order 4 multiplicatively since $i^4 = 1$ and $i^2 = -1$, so $f(i)^2 \cong -1 \pmod{p}$. If $f(i)$ has order 4 in $(\mathbb{Z}/p\mathbb{Z})^*$, then $p \equiv 1 \pmod{4}$. Then, recall by Wilson's Theorem that $(p - 1)! \equiv -1 \pmod{p}$. Now, consider the element $\left(\frac{p-1}{2}\right)!$, and complete it $1 * 2 * \dots * \frac{p-1}{2} * \frac{p+3}{2} * \dots * (p - 2) * (p - 1) \cong -1 \pmod{p}$. But, the terms in the first half are minus the terms in the second, so the product of the first half is equal to that of the second half times the number of minus signs. Note that the number of minus signs is $(-1)^{\frac{p-1}{2}}$, and since $p \equiv 1 \pmod{4}$, $\frac{p-1}{2}$ is even. Thus, the product of the first half is equal to the product of the second half. Hence, the square of $a \equiv \left(\frac{p-1}{2}\right)!$ is -1 , so it is our element of order 4. Then, let I be the ideal generated by p and $i - a$, so $I = (p, i - a)$. First, note that $I \cap \mathbb{Z} \supset p\mathbb{Z}$. Moreover, $(i - a)(b + ci) = (-ab - c) + (-ac + b)i$, where $-ac + b = 0$, so $-ab - c = -a^2c - c = -c(a^2 + 1)$. But, $a^2 \equiv -1 \pmod{p}$, so $a^2 + 1 \equiv 0 \pmod{p}$. Thus, $-c(a^2 + 1) \in p\mathbb{Z}$. Hence, $\mathbb{Z} \rightarrow R/I$ is surjective, as $i \equiv a \pmod{R/I}$, with kernel $p\mathbb{Z}$, so $R/I \cong \mathbb{Z}/p\mathbb{Z}$. ■

Theorem 9.5.2 (Gauss's Theorem). *For $R = \mathbb{Z}[i]$, every ideal $I \in R$ is principal.*

Corollary 9.5.3. *Since every $I \subset \mathbb{Z}[i]$ is principal, so is $(p, i - a)$, which implies $(p, i - a) = (a + bi)$ for some $a + bi \in \mathbb{Z}[i]$, and from above, $R/(a + bi) \cong \mathbb{Z}/p\mathbb{Z}$. This implies that $a^2 + b^2 = p$.*

Theorem 9.5.4 (Fermat's Theorem). *For any prime number p such that $p \equiv 1 \pmod{4}$, $p = a^2 + b^2$ for some $a, b \in \mathbb{Z}$.*

Remark 9.5.1. Gauss showed that if you can write all primes $p \equiv 1 \pmod{4}$ as $p = a^2 + b^2$, then every ideal $(a + bi) \subset \mathbb{Z}[i]$ must be principal.

Remark 9.5.2. The first step to prove Gauss's theorem is to show that for all prime $p \equiv 1 \pmod{4}$, there is an ideal $(a + bi)$ so that $R/(a + bi) \cong \mathbb{Z}/p\mathbb{Z}$, then the second step is to show that all ideals are principal, and then the third step is to show that if you have a quotient $R/(a + bi) \cong \mathbb{Z}/p\mathbb{Z}$, $a^2 + b^2 = p$.

Remark 9.5.3. More generally, the order of the finite ring $R/(a + bi)$ is $a^2 + b^2 = (a + bi)(a - bi)$ providing that $(a + bi) \neq (0)$.

Chapter 10

§§Ideals and Quotient Rings

10.1.0 §Basic Definitions and Examples: Ideals

Remark 10.1.1. Let $f : R \rightarrow R'$ be a ring homomorphism, then $\ker(f)$ is a subgroup of $(R, +)$. Moreover, for all $r_1, r_2 \in R$ and all $k \in \ker(f)$, $r_1kr_2 \in \ker(f)$. Indeed, $f(k) = 0_{R'}$ so $f(r_1kr_2) = f(r_1)f(k)f(r_2) = f(r_1)0_{R'}f(r_2) = 0_{R'}$.

Definition 10.1.1. Let R be a ring. A subgroup I of $(R, +)$ is called an ideal of R if for all $r_1, r_2 \in R$, and all $i \in I$,

$$r_1ir_2 \in I \quad (10.1.1)$$

↳ (This is actually the definition of a two-sided ideal, and the definitions of left and right sided ideals can be derived by relaxing the condition in this definition)

Example 10.1.1.

1. For an arbitrary ring R , $\{0\}$ and R are ideals of R .
2. If R is a commutative ring, then

$$aR := \{ar \in R : r \in R\} \quad (10.1.2)$$

is an ideal of R for all $a \in R$. I is called the principal ideal generated by a , and is denoted (a) .

3. $\ker(\mathbf{ev}_2 : \mathbb{R}[x] \rightarrow \mathbb{R})$ is an ideal of $\mathbb{R}[x]$. Indeed, $\ker(\mathbf{ev}_2) = (x - 2)$.
4. For a ring R , the ideal generated by a subset $X \subseteq R$ is given by

$$(X) := \left\{ \sum_{i=1}^n a_i x_i b_i : n \geq 1, a_i, b_i \in R, x_i \in X \right\} \quad (10.1.3)$$

5. Every ideal of \mathbb{Z} is principal. Indeed, every ideal is a subgroup of $(\mathbb{Z}, +)$, and every subgroup of $(\mathbb{Z}, +)$ is cyclic so every ideal is principal. Moreover, every subgroup is an

ideal. Indeed, we know that $n\mathbb{Z}$ is an ideal for all $n \in \mathbb{Z}$, but $n\mathbb{Z}$ is precisely the form for subgroups of \mathbb{Z} , so all subgroups are principal ideals of \mathbb{Z} and all ideals of \mathbb{Z} are principal.

6. The ideal $(2, X)$ of $\mathbb{Z}[X]$ is not a principal ideal.

Proof. For the sake of contradiction suppose $(2, X) = (p)$ for some $p \in \mathbb{Z}[X]$. Then $2 = pf$ for some $f \in \mathbb{Z}[X]$. Then $\deg(p) + \deg(f) \leq 0$, so $\deg(p) = 0$. Then $p = n \in \mathbb{Z}$ such that $n \mid 2$. Hence, $p = 2$ or $p = 1$. But, $1 \notin (2, X)$ since $2 \nmid 1$ and $X \nmid 1$ in \mathbb{Z} . Thus, $p = 2$. However, $2 \nmid X$ in $\mathbb{Z}[X]$, which contradicts the assumption that $(2, X) = (p)$. Thus, $(2, X)$ can not be principal in $\mathbb{Z}[X]$. ■

Definition 10.1.2. Let $I \subseteq R$ be an ideal of a ring R . The quotient group R/I , for the additive group $(R, +)$, is a ring for the multiplication

$$(a + I)(b + I) = ab + I \quad (\star)$$

Theorem 10.1.1. The addition on R/I and the multiplication given by (\star) makes R/I into a ring such that the canonical quotient map

$$\begin{aligned} \pi : R &\twoheadrightarrow R/I \\ r &\mapsto r + I \end{aligned} \quad (10.1.4)$$

is a surjective ring homomorphism of kernel I . We call R/I the quotient ring of R by I .

Proof. (Left to the reader) ■

Theorem 10.1.2. Any ideal I is the kernel of a natural ring homomorphism $R \rightarrow R/I$, where R/I is the quotient ring, taking $a \mapsto a + I$.

Remark 10.1.2. If R is a commutative ring then so is R/I for all ideals I of R .

Example 10.1.2. Consider $R = \mathbb{Z}/4\mathbb{Z}$. Then some ideals of R are, $I = (0), (1), (2) = \{0, 2\}$. In general, for $R = \mathbb{Z}/n\mathbb{Z}$ we have the ideal (d) for all $d \mid n$. In particular, if $R = \mathbb{Z}/p^k\mathbb{Z}$, then the distinct ideals are

$$(1) \supset (p) \supset (p^2) \supset \dots \supset (p^k) = (0)$$

Remark 10.1.3. In general the set of ideals form a lattice for the ring.

Example 10.1.3.

1. For any ring R , $R/(0) \cong R$ and $R/R \cong \{0\}$.
2. $\mathbb{Z}/n\mathbb{Z}$, where $n\mathbb{Z} = (n)$, is the ring of integers modulo n .
 - ↳ By the 19, we have that ideals of \mathbb{Z} containing $n\mathbb{Z}$ correspond to ideals of $\mathbb{Z}/n\mathbb{Z}$, which is to say, all ideals of the form $m\mathbb{Z}$ for $m \mid n$.

3. Let R be a commutative ring. Then $\mathbf{ev}_a : R[x] \rightarrow R$ for $a \in R$ is a surjective ring homomorphism, so by the First Isomorphism Theorem for rings and the fact that $\ker(\mathbf{ev}_a) = (x - a)$, we have that

$$\overline{\mathbf{ev}_a} : R[x]/(x - a) \xrightarrow{\sim} R \quad (10.1.5)$$

is a ring isomorphism. Note by the 19, ideals of $R[x]$ containing $(x - a)$ correspond to ideals of R . If $R = \mathbb{F}$ a field, then the only ideals of \mathbb{F} are $\{0\}$ and \mathbb{F} , which implies $(x - a)$ is maximal. Thus, the only ideals of $\mathbb{F}[x]$ containing $(x - a)$ are $(x - a)$ and $\mathbb{F}[x]$.

4. For X a set and $z \in X$,

$$\mathbf{ev}_z : F(X, R) \rightarrow R$$

is a surjective ring homomorphism of kernel

$$\ker(\mathbf{ev}_z) = \{f : X \rightarrow R \mid f(z) = 0_R\}$$

Definition 10.1.3. Let I be an ideal of R , $a, b \in R$. Then $a + I \in R/I$ is called the (congruence) class of a modulo I (sometimes denoted $a \bmod I$). If $a + I = b + I$, so $b - a \in I$, we say that a and b are congruent modulo I , written

$$a \equiv b \bmod I, a = b \bmod I, a = b(I) \quad (10.1.6)$$

§Simple Ideals

Definition 10.1.4. A ring R is simple if $R \neq \{0\}$ and if the only ideals of R are $\{0\}$ and R . That is, R has exactly two ideals.

Example 10.1.4.

1. Division rings are simple. If $a \neq 0$ in an ideal I of a division ring R , then $1 = a^{-1}a \in I$, so $I = (1) = R$. In particular, fields are simple rings.
2. $M_n(R)$ for R an arbitrary ring is simple if and only if R is simple.

Proof. Suppose R is simple and let \mathcal{A} be an ideal of $M_n(R)$. Then by the following lemma $\mathcal{A} = M_n(A)$ for some ideal A of R . Thus, $A = \{0\}$ or $A = R$, so $\mathcal{A} = M_n(0) = \{0\}$ or $\mathcal{A} = M_n(R)$. Consequently, $M_n(R)$ is simple. Now, suppose $M_n(R)$ is simple and let A be an ideal of R . Then $M_n(A)$ is an ideal of $M_n(R)$. Hence, $M_n(A) = \{0\}$ or $M_n(A) = M_n(R)$. Thus, $A = \{0\}$ or $A = R$, so R is a simple ring as claimed. ■

Lemma 10.1.3. For a ring R , every ideal of $M_n(R)$ has the form $M_n(A)$ for an ideal A of R .

Proposition 10.1.4. A commutative ring R is simple if and only if it is a field.

Proof. If R is a field then it is simple by the previous example. Now, suppose R is a commutative simple ring, and let $0 \neq r \in R$. Consider the ideal $(r) = rR$. Since R is simple, $(r) = R$ since $(r) \neq \{0\}$. Hence, $1 \in (r)$, so there exists $r' \in R$ such that $rr' = 1$. Hence, r is a unit of R , so in particular R is a field. ■

§Maximal and Prime Ideals

Definition 10.1.5. An ideal $I \subseteq R$ in a ring R is called maximal if

1. $I \neq R$
2. For all ideals $J \subseteq R$, if $I \subseteq J$, then either $I = J$ or $J = R$.

Proposition 10.1.5. An ideal $I \subseteq R$ is maximal if and only if the quotient ring R/I is simple.

Proof. Let R be a ring with ideal I .

\implies Suppose I is maximal. Then by the 19 the only ideals of R/I are $\{0_{R/I}\}$ and R/I corresponding to I and R respectively. Moreover, since $I \neq R$, $R/I \neq \{0\}$. Thus, we have that R/I is simple.

\impliedby Suppose R/I is a simple ring and let $I \subseteq J \subseteq R$ be an ideal containing I . Then by the 19 we have that $I/I \subseteq J/I \subseteq R/I$ is an ideal of R/I . However, as R/I is simple $J/I = I/I$ or $J/I = R/I$. By the bijectivity of the correspondence, $J = I$ or $J = R$. Hence, I is a maximal ideal in R as claimed. ■

Corollary 10.1.6. If R is commutative, then an ideal $I \subseteq R$ is maximal if and only if R/I is a field.

Example 10.1.5.

1. If F is a field, then $\{0\}$ is the only maximal ideal of F .
2. If p is a prime number, then $p\mathbb{Z} \subseteq \mathbb{Z}$ is a maximal ideal. (indeed $\mathbb{Z}/p\mathbb{Z}$ is a field)
 - ↳ Actually, for $n \geq 0$, $n\mathbb{Z} \subseteq \mathbb{Z}$ is a maximal ideal if and only if n is prime
3. In $F[x]$ for F a field, the ideal $(x-a)$ is maximal for all $a \in F$, because $F[x]/(x-a) \cong F$.

Definition 10.1.6. Let R be a commutative ring. Then an ideal I of R is prime if

1. $I \subsetneq R$
2. For all $r_1, r_2 \in R$, if $r_1 r_2 \in I$ then either $r_1 \in I$ or $r_2 \in I$.

Example 10.1.6. 1. A commutative ring R is an integral domain if and only if $\{0\}$ is a prime ideal.

↳ Indeed, R is a commutative integral domain $\iff R \neq \{0\}$ and whenever $ab = 0$, either $a = 0$ or $b = 0 \iff R \neq \{0\}$ and whenever $ab \in \{0\}$, $a \in \{0\}$ or $b \in \{0\} \iff \{0\}$ is a prime ideal of R .

2. $p\mathbb{Z} \subset \mathbb{Z}$ is a prime ideal for p a prime number. Indeed, $p\mathbb{Z} \neq \mathbb{Z}$ and $p \mid ab$ implies $p \mid a$ or $p \mid b$.

Proposition 10.1.7. *Let R be a commutative ring with ideal $I \subseteq R$. Then I is prime if and only if R/I is an integral domain.*

Proof. Indeed, I is prime $\iff I \subsetneq R$ and $ab \in I$ implies $a \in I$ or $b \in I \iff R/I \neq \{0\}$ and $ab + I = I$ implies $a + I = I$ or $b + I = I \iff R/I$ is an integral domain. ■

Corollary 10.1.8. *Every maximal ideal of a commutative ring R is a prime ideal.*

Example 10.1.7.

1. $(x) \subset \mathbb{Z}[x]$ is a prime ideal. Indeed, $\mathbb{Z}[x]/(x) \cong \mathbb{Z}$ is an integral domain (not a field, so (x) is not maximal)
2. $(p) \subset \mathbb{Z}[x]$ for a prime number p is a prime ideal which is not maximal. Consider

$$\begin{aligned} \mathbb{Z}[x] &\xrightarrow{f} \mathbb{Z}/p\mathbb{Z}[x] \\ \sum_i a_i x^i &\mapsto \sum_i [a_i]_p x^i \end{aligned} \quad (10.1.7)$$

Then f is a surjective ring homomorphism and $\ker(f) = (p)$. So, by the first isomorphism theorem f induces a ring isomorphism

$$\bar{f} : \mathbb{Z}[x]/(p) \xrightarrow{\sim} \mathbb{Z}/p\mathbb{Z}[x] \quad (10.1.8)$$

Since $\mathbb{Z}/p\mathbb{Z}[x]$ is an integral domain (p) is a prime ideal, but $\mathbb{Z}/p\mathbb{Z}[x]$ is not a field, so (p) is not maximal.

10.2.0 §Ideal Arithmetic and the Chinese Remainder Theorem

Definition 10.2.1. *Let R be a ring, and I, J ideals of R . We define their sum as*

$$I + J := \{i + j : i \in I, j \in J\} \quad (10.2.1)$$

Then $I + J$ is an ideal of R . Next, define their product

$$IJ := \langle ij : i \in I, j \in J \rangle = \left\{ \sum_{k=1}^n i_k j_k : n \geq 1, i_k \in I, j_k \in J, \forall 1 \leq k \leq n \right\} \quad (10.2.2)$$

Then IJ is an ideal of R as well. Recall $I \cap J$ is also an ideal of R

Example 10.2.1. For $R = \mathbb{Z}$ and $m, n > 0$, we have

$$m\mathbb{Z} + n\mathbb{Z} = \gcd(m, n)\mathbb{Z} \quad (10.2.3)$$

$$m\mathbb{Z} \cdot n\mathbb{Z} = mn\mathbb{Z} \quad (10.2.4)$$

$$m\mathbb{Z} \cap n\mathbb{Z} = \text{lcm}(m, n)\mathbb{Z} \quad (10.2.5)$$

So, if $\gcd(m, n) = 1$, then

$$m\mathbb{Z} \cdot n\mathbb{Z} = m\mathbb{Z} \cap n\mathbb{Z}$$

because $mn = \text{lcm}(m, n) \gcd(m, n)$.

Remark 10.2.1. Let I, J, K be ideals of a ring R . Then

1. $(I + J)K = IK + JK$
2. $K(I + J) = KI + KJ$
3. $IR = I = RI$

Proof. Let I, J, K be ideals of a ring R

1. Let $(i + j)k \in (I + J)K$. Then by distributivity $(i + j)k = ik + jk \in IK + JK$. Similarly, for all $ik + jk \in IK + JK$, $ik + jk = (i + j)k \in (I + J)K$. Thus, we have that $(I + J)K = IK + JK$.

2. This statement is identical to 1., replacing right distributivity with left distributivity.

3. Note that for all $i \in I$ and all $r \in R$, $ir, ri \in I$ since I is an ideal, and $i = 1 \cdot i \in RI$, $i = i \cdot 1 \in IR$ using the fact that R is unital. Hence, $RI = I = IR$, completing the proof. ■

Theorem 16 (Chinese Remainder Theorem (CRT)).

Let R be a ring, and I, J ideals of R . Assume that $I + J = R$ (I and J are said to be relatively prime). Then the ring homomorphism

$$\begin{aligned} R &\xrightarrow{\alpha} R/I \times R/J \\ r &\mapsto (r + I, r + J) \end{aligned} \quad (10.2.6)$$

is surjective of kernel $I \cap J$. Consequently, by the first isomorphism theorem for rings we have an isomorphism

$$\bar{\alpha} : R/I \cap J \xrightarrow{\sim} R/I \times R/J \quad (10.2.7)$$

Moreover, if R is commutative, then $I + J = R$ implies that $I \cap J = IJ$.

Proof. Suppose R is a ring with relatively prime ideals I, J . Define a map

$$\begin{aligned} R &\xrightarrow{\alpha} R/I \times R/J \\ r &\mapsto (r + I, r + J) \end{aligned}$$

Let $r, r' \in R$. Then, observe that

$$\begin{aligned} \alpha(r + r') &= (r + r' + I, r + r' + J) & \alpha(rr') &= (rr' + I, rr' + J) \\ &= (r + I + r' + I, r + J + r' + J) & &= ((r + I)(r' + I), (r + J)(r' + J)) \\ &= (r + I, r + J) + (r' + I, r' + J) & &= (r + I, r + J)(r' + I, r' + J) \\ &= \alpha(r) + \alpha(r') & &= \alpha(r)\alpha(r') \end{aligned}$$

and

$$\alpha(1_R) = (1_R + I, 1_R + J) = (1_{R/I}, 1_{R/J})$$

so α is a ring homomorphism. To show α is surjective, let $a, b \in R$. We want to find $r \in R$, $i \in I$, and $j \in J$ such that $r + i = a$ and $r + j = b$. But, we then have that $r = a - i = b - j$,

so $a - b = i - j$. Note that $I + J = R$ since they are relatively prime, so there exist $i' \in I$ and $j' \in J$ such that $a - b = i' + j'$. Set $i = i'$ and $j = -j'$. Then, observe that

$$\alpha(r) = (r + I, r + J) = (r + i + I, r + j + J) = (a + I, b + J)$$

Therefore α is a surjective ring homomorphism. Observe that $I \cap J \subseteq \ker(\alpha)$ since for all $k \in I \cap J$, $\alpha(k) = (k + I, k + J) = (I, J)$. Then, let $t \in \ker(\alpha)$. Observe that then $(I, J) = \alpha(t) = (t + I, t + J)$, so by definition $t \in I$ and $t \in J$. Thus, $t \in I \cap J$, so $\ker(\alpha) \subseteq I \cap J$. Consequently $\ker(\alpha) = I \cap J$. By the first isomorphism theorem for rings we have our desired result.

Now, suppose R is commutative. Then observe that

$$\begin{aligned} I \cap J &= (I \cap J)R \\ &= (I \cap J)(I + J) \\ &= (I \cap J)I + (I \cap J)J \\ &\subseteq JI + IJ \\ &= IJ + IJ \\ &= IJ \end{aligned}$$

Moreover, $IJ \subseteq I \cap J$ since IJ is generated by ij for $i \in I$ and $j \in J$. However, since I and J are ideals $ij \in I$ and $ij \in J$, so in particular $ij \in I \cap J$. Thus, we conclude that $IJ = I \cap J$. ■

Lemma 10.2.1. *If I_1, I_2, \dots, I_n are pairwise relatively prime ideals of R , with $n \geq 2$, then for all $i \in \{1, 2, \dots, n\}$ $\bigcap_{j \neq i} I_j$ and I_i are relatively prime. That is, $\bigcap_{j \neq i} I_j + I_i = R$.*

Proof. Let I_1, I_2, \dots, I_n be as in the statement, for $n \geq 2$. If $n = 2$ we are done by assumption. Then, if $n = 3$ there exist $i_1 \in I_{j_1}, i_2 \in I_{j_2}, i_3, i'_3 \in I_i$ such that $i_1 + i_3 = 1, i_2 + i'_3 = 1$, where $i \in \{1, 2, 3\}$. Then $1 = (i_1 + i_3)(i_2 + i'_3) = i_1 i_2 + i_3 i_2 + i_1 i'_3 + i_3 i'_3$, where $i_1 i_2 \in I_{j_1} \cap I_{j_2}$, and $i_3 i_2, i_1 i'_3, i_3 i'_3 \in I_i$ since they are ideals. Thus, $1 \in I_{j_1} \cap I_{j_2} + I_i$, so $I_{j_1} \cap I_{j_2} + I_i = R$, since it is an ideal. Hence, the base cases hold. Now, suppose there exists $k \geq 3$ such that if $n = k$, $I_{j_1} \cap I_{j_2} \cap \dots \cap I_{j_{k-1}} + I_i = R$ for $i \in \{1, 2, \dots, k\}$, and $\{j_1, j_2, \dots, j_{k-1}\} = \{1, 2, \dots, k\} \setminus \{i\}$. Then, choose $i \in \{1, 2, \dots, k+1\}$ and $\{j_1, j_2, \dots, j_k\} = \{1, 2, \dots, k+1\} \setminus \{i\}$. Let $I = I_{j_1} \cap I_{j_2} \cap \dots \cap I_{j_{k-1}}$. Then we have by the induction hypothesis and assumption that $I + I_i = R$, $I + I_{j_k} = R$ and $I_i + I_{j_k} = R$. Then, by our argument in the base case for $n = 3$ we have that $I \cap I_{j_k} + I_i = R$. In particular, $I_{j_1} \cap I_{j_2} \cap \dots \cap I_{j_{k-1}} \cap I_{j_k} + I_i = R$, as desired. Therefore, by mathematical induction we conclude that for all $n \geq 2$ and all $i \in \{1, 2, \dots, n\}$, $I_{j_1} \cap I_{j_2} \cap \dots \cap I_{j_{n-1}} + I_i = R$, completing the proof. ■

Corollary 10.2.2. *Let R be a ring with ideals I_1, I_2, \dots, I_n of R , for $n \geq 1$. Suppose that $I_i + I_j = R$ for all $i \neq j$. Then*

$$\begin{aligned} R &\xrightarrow{\alpha} R/I_1 \times R/I_2 \times \dots \times R/I_n = \bigotimes_{i=1}^n R/I_i \\ r &\mapsto (r + I_1, r + I_2, \dots, r + I_n) \end{aligned} \tag{10.2.8}$$

is a surjective ring homomorphism of kernel $\bigcap_{i=1}^n I_i = I_1 \cap I_2 \cap \dots \cap I_n$. Thus, we have an isomorphism

$$\bar{\alpha} : R / \bigcap_{i=1}^n I_i \xrightarrow{\sim} \bigotimes_{i=1}^n R/I_i \tag{10.2.9}$$

If R is commutative, we have $\bigcap_{i=1}^n I_i = \prod_{i=1}^n I_i$.

Proof. Let R be a ring with pairwise relatively prime ideals I_1, I_2, \dots, I_n of R , for $n \geq 1$. Define a map

$$\begin{aligned} R &\xrightarrow{\alpha} R/I_1 \times R/I_2 \times \dots \times R/I_n = \bigotimes_{i=1}^n R/I_i \\ r &\mapsto (r + I_1, r + I_2, \dots, r + I_n) \end{aligned}$$

From the proof of the Chinese Remainder Theorem α is a ring homomorphism. Moreover, $\ker(\alpha) = \bigcap_{i=1}^n I_i$. To show surjectivity let $a_1, a_2, \dots, a_n \in R$. Then, for each $i \in \{1, 2, \dots, n\}$ choose $c_i \in I_{i_1} \cap \dots \cap I_{i_{n-1}}$ and $i' \in I_i$ such that $c_i + i' = 1_R$. This choice is possible by the result of Lemma 10.2.1. We then define $a = a_1 c_1 + \dots + a_n c_n$. Note that for each $i \in \{1, 2, \dots, n\}$, $c_i \equiv 0 \pmod{I_j}$ if $j \neq i$, and $c_i \equiv 1 \pmod{I_i}$ since $c_i + i' = 1_R$. It follows that $a \equiv 0_R + \dots + a_i + \dots + 0_R \equiv a_i \pmod{I_i}$. Thus, we have that

$$\alpha(a) = (a + I_1, a + I_2, \dots, a + I_n) = (a_1 + I_1, a_2 + I_2, \dots, a_n + I_n)$$

Therefore, α is a surjective ring homomorphism with kernel $\ker(\alpha) = \bigcap_{i=1}^n I_i$, so by the first isomorphism theorem for rings

$$\bar{\alpha} : R / \bigcap_{i=1}^n I_i \xrightarrow{\sim} \bigotimes_{i=1}^n R/I_i$$

Then, by the existence of this isomorphism $b \equiv a_i \pmod{I_i}$ for each $i \in \{1, 2, \dots, n\}$ if and only if $b \equiv a \pmod{\bigcap_{i=1}^n I_i}$.

Finally, suppose R is a commutative ring. First, for each basic element of the form $i_1 i_2 \dots i_n \in \prod_{i=1}^n I_i$ we have $i_1 i_2 \dots i_n \in I_j$ for each $j \in \{1, 2, \dots, n\}$ since they are ideals. Hence, $\prod_{i=1}^n I_i \subseteq \bigcap_{i=1}^n I_i$. Next, write $I_1 \cap I_2 \cap \dots \cap I_{n-1} = I$. Then by our Lemma 10.2.1 $I + I_n = R$. It follows that

$$\begin{aligned} I \cap I_n &= (I \cap I_n)R && \text{(Lemma 10.2.1)} \\ &= (I \cap I_n)(I + I_n) \\ &= (I \cap I_n)I + (I \cap I_n)I_n && \text{(Lemma 10.2.1)} \\ &\subseteq I_n I + I I_n \\ &= I I_n + I I_n && \text{(Commutativity of } R) \\ &= I I_n && \text{(since } 0_R \in I_j, \forall j) \end{aligned}$$

Thus, we have that $\bigcap_{i=1}^n I_i = I \cap I_n \subseteq I I_n = \prod_{i=1}^n I_i$. Therefore, we conclude that if R is commutative, $\prod_{i=1}^n I_i = \bigcap_{i=1}^n I_i$. ■

Remark 10.2.2 (Solving Congruences). Suppose we have a system of congruences

$$\begin{cases} x \equiv a_1 \pmod{m_1} \\ x \equiv a_2 \pmod{m_2} \\ \vdots \\ x \equiv a_n \pmod{m_n} \end{cases} \quad (10.2.10)$$

we can find the c_i given in the above proof as follows. Write $M = m_1 \dots m_n$ and $M_i = m_1 \dots m_{i-1} m_{i+1} \dots m_n$. Then, note that $M_i = m_1 \dots m_{i-1} m_{i+1} \dots m_n \mid c_i$ by definition of c_i , and that since $c_i \equiv 1 \pmod{m_i}$, there exists $b_i \in \mathbb{Z}$ such that $c_i + b_i m_i = 1$. Then, write $c_i = y_i M_i$, where we can solve for y_i using the extended Euclidean Algorithm on (M_i, m_i) , or noting the inverse of M_i modulo m_i , as they are relatively prime.

10.3.0 §Adjunctions

Definition 10.3.1 (Ring Relations). *Creating Relations in a ring R : Suppose we have an element $a \in R$. If we want a ring \bar{R} which is an image of R , where $\bar{a} = 0$, then the largest such quotient is $\bar{R} = R/(a)$. If we want a ring where we have a number of relations $a_1 = a_2 = \dots = a_n = 0$, we can take $(R/(a_1))/(a_2)/\dots/(a_n) = \bar{R} = R/(a_1, a_2, \dots, a_n)$. This is valid because the ideal (a_1, \dots, a_n) contains (a_i) for all i , and then this is successive applications of the Isomorphism Theorem.*

Remark 10.3.1. If R is a ring and $a \in R$, if a is a unit then $R/(a) = \{0\}$ since $(a) = R$. I.e. modding out by a unit mods out all elements of the ring.

10.4.0 §Isomorphism Theorems and Correspondence

Theorem 17 (First Isomorphism Theorem (for rings)).

Let $f : R \rightarrow R'$ be a ring homomorphism. Then by the First Isomorphism Theorem for groups there exists a unique group isomorphism for $(R, +)$ such that

$$\begin{aligned} R/\ker(f) &\xrightarrow{\bar{f}} f(R) \\ r + \ker(f) &\mapsto f(r) \end{aligned} \tag{10.4.1}$$

and this is also a ring isomorphism. This theorem can be stated succinctly by the following diagram:

$$\begin{array}{ccc} R & \xrightarrow{\forall f} & \forall R' \\ \pi \downarrow & \nearrow \exists! \bar{f} & \\ R/\ker(f) & & \end{array}$$

Proof. (Left to the reader) ■

Theorem 18 (Third Isomorphism Theorem).

Let R be a ring and suppose $A \subseteq B \subseteq R$ are ideals of R . Then B/A is an ideal of R/A and

$$(R/A)/(B/A) \cong R/B \tag{10.4.2}$$

$$\begin{array}{ccc}
 R & \xrightarrow{\pi_B} & R/B \\
 \pi_A \downarrow & \exists! \overline{\pi_B} \nearrow & \uparrow \exists! \overline{\pi_B} \\
 R/A & \xrightarrow{\pi_{B/A}} & (R/A)/(B/A)
 \end{array}$$

Proof. Suppose R is a ring with ideals $A \subseteq B \subseteq R$. By the 19 we know that B/A is an ideal of R/A . Then, define a map

$$\begin{aligned}
 \phi : R/A &\rightarrow R/B \\
 a + A &\mapsto a + B
 \end{aligned}$$

First, suppose that $a + A = b + A$. Then we have that $a - b \in A \subseteq B$, so $a - b \in B$. Hence, $a + B = b + B$ by definition of coset equality for quotient groups, so ϕ is well defined. Moreover, observe that for all $a + A, c + A \in R/A$, we have

$$\phi(a + A + c + A) = \phi(a + c + A) = a + c + B = a + B + c + B = \phi(a + A) + \phi(c + A)$$

$$\phi((a + A)(c + A)) = \phi(ac + A) = ac + B = (a + B)(c + B) = \phi(a + A)\phi(c + A)$$

and

$$\phi(1_{R/A}) = \phi(1_R + A) = 1_R + B = 1_{R/B}$$

so ϕ is a ring homomorphism. Moreover, by construction we have that ϕ is surjective. Now, note that $B/A \subseteq \ker(\phi)$. Then, let $k + A \in \ker(\phi)$, so $\phi(k + A) = k + B = B$. Hence, $k \in B$ so $k + A \in B/A$. Therefore, we conclude that $\ker(\phi) = B/A$, so by the First Isomorphism Theorem for rings,

$$(R/A)/(B/A) \cong R/B \quad (10.4.3)$$

■

Theorem 19 (Correspondence Theorem (for rings)).

Let $\phi : R \rightarrow R'$ be a surjective ring homomorphism of kernel $K \subseteq R$. Then, the correspondence between subgroups of $(R', +')$ and subgroups of $(R, +)$ containing K induces a bijection between ideals of R' and ideals of R containing K :

$$\begin{aligned}
 \{I : I \subseteq R' \text{ an ideal}\} &\leftrightarrow \{J : K \subseteq J \subseteq R \text{ an ideal}\} \\
 I &\mapsto \phi^{-1}(I) \\
 \phi(J) &\leftarrow J
 \end{aligned} \quad (10.4.4)$$

are inverse bijections.

Proof. (Left to the reader)

■

Remark 10.4.1 (Warning about image of ideals). If $f : R \rightarrow R'$ is a ring homomorphism that is not surjective, and $J \subset R$ is an ideal, then $f(J)$ is not necessarily an ideal as well. For example, $i : \mathbb{Z} \hookrightarrow \mathbb{Q}$ is a ring homomorphism and \mathbb{Z} is an ideal in \mathbb{Z} , but $i(\mathbb{Z})$ is not an ideal in \mathbb{Q} since \mathbb{Q} is a field with only trivial ideals.

Chapter 11

§§Adjunction of Elements

Chapter 12

§§Unique Factorization Domains

12.1.0 §Basic Definitions and Examples: UFDs

12.2.0 §Unique Factorization in $F[x]$

Theorem 20 (Unique Factorization Theorem (for $F[x]$)).

)] Take F a field and $f \in F[x]$ of degree greater than or equal to 1. Then

1. $f = aP_1P_2\dots P_m$ where $a \in F$ is the leading coefficient of f and P_i is monic irreducible in $F[x]$ for all i .
2. The factorization in 1. is unique up to the order of the factors.

Proof. First Attempt Let F be a field and let $f \in F[x]$ of degree $n \geq 1$. It is sufficient to consider f monic since we can replace f with $a^{-1}f$, where a is the leading coefficient of f . Then, we proceed by induction on the degree of f . If $\deg(f) = 1$ then f is already a monic irreducible polynomial in $F[x]$, so we're done. Inductively suppose there exists $k \geq 1$ such that for all $j \leq k$, if $\deg(f) = j$ then $f = q_1q_2\dots q_m$ for monic irreducible polynomials $q_i \in F[x]$. Then, suppose $\deg(f) = k + 1$. If f is irreducible then we're done. On the other hand, if f is not irreducible there exist $g, h \in F[x]$ such that $f = gh$ and $\deg(g), \deg(h) \geq 1$. Consequently, $\deg(f) = \deg(g) + \deg(h) > \deg(g), \deg(h)$, so $\deg(g), \deg(h) \leq k$. Thus, by the induction hypothesis $g = g_1g_2\dots g_m$ and $h = h_1h_2\dots h_l$ for monic irreducible polynomials $g_i, h_j \in F[x]$, $1 \leq i \leq m$, $1 \leq j \leq l$. Hence, $f = g_1g_2\dots g_mh_1h_2\dots h_l$ is the product of monic irreducible polynomials in $F[x]$ as desired. Therefore, by mathematical induction we have that for all $f \in F[x]$, $\deg(f) \geq 1$, f can be factored as the product of monic irreducible polynomials and its leading coefficient. Now, suppose the factorization is not necessarily unique. Let f be a polynomial of lowest degree with two such factorizations $f = aP_1P_2\dots P_m$ and $f = bQ_1Q_2\dots Q_n$. Since the P_i and Q_j are monic we must have that $a = b \neq 0$, so multiplying by a^{-1} on both sides we obtain $P_1P_2\dots P_m = Q_1Q_2\dots Q_n$. Note that (P_i) is a prime ideal for each P_i , so in particular each P_i is a prime element. Hence, P_1 divides $Q_1Q_2\dots Q_n$ which implies that P_1 divides Q_j for some $1 \leq j \leq n$. Reorder the Q_i if need be so that P_1 divides Q_1 . Then, there exists $f \in F[x]$

such that $P_1 f = Q_1$. But, Q_1 is irreducible so as $\deg(P_1) \geq 1$, we must have that $f \in F[x]^\times$, so $f = c \in F$ for some c . But, P_1 and Q_1 are monic, so $c = 1$. Thus $P_1 = Q_1$. Since $F[x]$ is an integral domain we can cancel elements, so $P_2 \dots P_m = Q_2 \dots Q_n$. But, this is a polynomial of strictly lower degree than f with a distinct factorization, contradicting the minimality of f . Therefore, the factorization of f must be unique up to reordering. ■

Example 12.2.1. Consider $f = 5(x^3 - 1) \in F[x]$ with $5 \neq 0 \in F$. We always have $f = 5(x - 1)(x^2 + x + 1)$ (for any field)

1. If $F = \mathbb{R}$, this is the factorization from the UFT because $x^2 + x + 1$ is monic irreducible in $\mathbb{R}[x]$, as it has no roots over \mathbb{R} .
2. if $F = \mathbb{Z}/2\mathbb{Z}$, $f = 5(x - 1)(x^2 + x + 1) = (x + 1)(x^2 + x + 1)$ is the UFT factorization as $x^2 + x + 1$ has no roots over $\mathbb{Z}/2\mathbb{Z}$
3. If $F = \mathbb{Z}/3\mathbb{Z}$, $f = 5(x - 1)(x^2 + x + 1) = 2(x - 1)(x - 1)^2 = 2(x + 2)^3$ is the UFT factorization
4. If $F = \mathbb{C}$ then we have $f = 5(x - 1)(x^2 + x + 1) = 5(x - 1)(x - u)(x - \bar{u})$ for $u = \frac{-1 + \sqrt{3}i}{2}$.

Chapter 13

§§Principal Ideal Domains

13.1.0 §Basic Definitions and Examples: PIDs

Definition 13.1.1. *An integral domain of which every ideal is principal is called a principal ideal domain (PID).*

Remark 13.1.1. \mathbb{Z} is a prototypical example of a PID. Additionally, for a field \mathbb{F} , $\mathbb{F}[X]$ is also a PID.

Theorem 13.1.1. *Let \mathbb{F} be a field. Then $\mathbb{F}[X]$ is a principal ideal domain. Moreover, every non-zero ideal I of $\mathbb{F}[X]$ is generated by the monic polynomial of lowest degree in I .*

Proof. Note that since \mathbb{F} is an integral domain so is $\mathbb{F}[X]$. Consider an ideal $I \subseteq \mathbb{F}[X]$. If $I = \{0\} = (0)$ we are done, so assume $I \neq \{0\}$ and that $0 \neq g \in I$ is of minimal degree. We can assume that g is monic; if a is the leading coefficient of g , then $a^{-1} \in \mathbb{F}$ so $a^{-1}g$ is monic, and we replace g with $a^{-1}g \in I$ since I is an ideal. We claim that $I = (g)$. Since $g \in I$ and I is an ideal we have $g\mathbb{F}[X] = (g) \subseteq I$. Let $P \in I$. By the division algorithm there exist unique $q, r \in \mathbb{F}[X]$ such that $P = qg + r$ where $r = 0$ or $\deg(r) < \deg(g)$ since the leading coefficient of g is a unit. Then $r = P - qg \in I$ since $qg \in I$ as it is an ideal. Then $r = 0$, since otherwise $\deg(r) < \deg(g)$, contradicting the minimality of g 's degree. Thus, $P = qg \in (g)$, so $I \subseteq (g)$. Therefore $I = (g)$ and $\mathbb{F}[X]$ is a PID as claimed. ■

Chapter 14

§§Euclidean Domains

14.1.0 §Basic Definitions and Examples: Euclidean Domains

Definition 14.1.1 (Norm). For an integral domain R , any function $N : R \rightarrow \mathbb{Z}^+ \cup \{0\}$ with $N(0_R) = 0$ is called a norm on the integral domain R . If $N(a) > 0$ for $a \neq 0$ define N to be a positive norm.

Definition 14.1.2 (Euclidean Domain). The integral domain R is said to be a Euclidean Domain (or posses a Division Algorithm) if there is a norm N on R such that for any two elements $a, b \in R$, with $b \neq 0_R$, there exist elements $q, r \in R$ with

$$a = qb + r$$

where $r = 0_R$ or $N(r) < N(b)$. The element q is called the quotient and the element r the remainder of the division.

Example 14.1.1.

1. Fields are trivial examples of Euclidean Domains where any norm will satisfy the defining condition. (e.g., $N(a) = 0, \forall a$) This is because for all $a, b, b \neq 0$, we have $a = qb + 0$, where $q = ab^{-1}$.
2. The integers are a Euclidean Domain with norm $N(a) = |a|$, the usual absolute value.
3. If F is a field, the polynomial ring $F[x]$ is a Euclidean Domain with norm $N(p(x)) = \deg(p(x))$
4. The Gaussian integers, $\mathbb{Z}[i]$, is a Euclidean domain with norm $N(a + bi) = a^2 + b^2$

Proposition 14.1.1. Every ideal in a Euclidean Domain is principal. In particular, if I is any nonzero ideal in the Euclidean Domain R , then $I = (d)$ for d any nonzero element of I of minimal norm.

Proof. If I is the zero ideal there is nothing to prove. Otherwise, let d be a nonzero element of I of minimum norm. Such a d exists since the set $\{N(a) : a \in I\}$ is a nonempty subset of \mathbb{Z} , which is bounded below, and hence has a minimum element by the well-ordering of \mathbb{Z} . Clearly $(d) \subseteq I$ since $d \in I$. To show the reverse inclusion let a be any element of I , and use the division algorithm to write $a = qd + r$ for $q, r \in R$, with $N(r) < N(d)$. Then, since I is an ideal $-qd \in I$, so $r = a - qd \in I$. Thus, as d is an element of minimal norm in I , so we must have that $r = 0$. Thus $a = qd \in (d)$, showing $I = (d)$. ■

Remark 14.1.1. This proposition can be used to prove that some integral domains R are not Euclidean Domains if they have non-principal ideals.

Definition 14.1.3. Let R be a commutative ring and let $a, b \in R$ with $b \neq 0$.

1. a is said to be a **multiple** of b if there exists an element $x \in R$ with $a = bx$. In this case b is said to **divide** a or be a **divisor** of a , written $b \mid a$
2. A **greatest common divisor** of a and b is a nonzero element $d \in R$ such that
 - (a) $d \mid a$ and $d \mid b$, and
 - (b) if $d' \mid a$ and $d' \mid b$ then $d' \mid d$.

A greatest common divisor of a and b is denoted $\gcd(a, b)$.

Remark 14.1.2. Translating into the language of ideals, if $I = (a, b) \subseteq R$, then d is the greatest common divisor of a and b if

1. $I \subseteq (d)$, and
2. if $I \subseteq (d')$, then $(d) \subseteq (d')$

Hence, a greatest common divisor for a and b (if one exists) is a generator for the unique smallest principal ideal containing a and b .

Proposition 14.1.2. If a, b are nonzero elements in the commutative ring R such that the ideal generated by a and b is a principal ideal (d) , then d is the greatest common divisor of a and b .

Definition 14.1.4. An integral domain in which every ideal (a, b) generated by two elements is principal is called a **Bezout Domain**. Note that Bezout Domain's can have non-principal ideals.

Proposition 14.1.3. Let R be an integral domain. If two elements d and d' of R generate the same principal ideal, i.e. $(d) = (d')$, then there exists $u \in R^\times$ such that $d = ud'$. d and d' are called **associates** in this case.

Theorem 14.1.4. Let R be a Euclidean Domain and let $a, b \in R$ be nonzero elements of R . Let $d = r_n$ be the last nonzero remainder in the Euclidean Algorithm for a and b . Then

1. $d = \gcd(a, b)$, and

2. $(d) = (a, b)$, so there exist $x, y \in R$ such that

$$d = ax + by \tag{14.1.1}$$

Proof. (Left to the reader) ■

Definition 14.1.5. Let $\tilde{R} := R^\times \cup \{0\}$. An element $u \in R - \tilde{R}$ is called a **universal side divisor** if for every $x \in R$ there is some $z \in \tilde{R}$ such that u divides $x - z \in R$.

Proposition 14.1.5. Let R be an integral domain that is not a field. If R is a Euclidean Domain then there are universal side divisors in R .

Proof. (Left to the reader) ■

Chapter 15

§§Polynomial Rings

15.1.0 §Basic Definitions and Examples: Polynomial Rings

Definition 15.1.1. Let R be a ring and let x be a formal symbol (not related to R). We want to define a ring $R[x]$ such that

1. $x \in Z(R[x])$
2. $R \subseteq R[x]$ is a subring
3. $R[x]$ is generated by $\{x\} \cup R$

Then, for $P \in R[x]$ we have that $P = \sum_{j=0}^n b_j x^j$ where $b_j \in R$ for all j , and $n \in \mathbb{Z}$, with $n \geq 0$ (note we define $x^0 = 1_R \in R$). We also require that

$$\sum_{j=0}^n b_j x^j = \sum_{i=0}^m a_i x^i, \iff b_j = a_j \forall j \geq 0 \quad (15.1.1)$$

where if $j > n$, $b_j = 0$, and if $j > m$, $a_j = 0$. Next, we define addition and multiplication as

Addition:

$$\sum_{j=0}^n b_j x^j + \sum_{i=0}^m a_i x^i := \sum_{i=0}^{\max(m,n)} (a_i + b_i) x^i$$

Multiplication:

$$\left[\sum_{j=0}^n b_j x^j \right] \left[\sum_{i=0}^m a_i x^i \right] := \sum_{i=0}^{m+n} \left(\sum_{k=0}^i a_k b_{i-k} \right) x^i$$

Remark 15.1.1. Formally, the polynomial is determined by a sequence of coefficients a_i

$$\mathbf{a} = (a_0, a_1, a_2, \dots) \quad (15.1.2)$$

where $a_i \in R$ and only a finite number of a_i are not zero. The sequence with 1 in the i th position and zero everywhere else corresponds to the indeterminate monomial x^i , and the monomials form a basis of the space of polynomials.

Claim 15.1.1. For all rings R , $(R[x], +, \cdot)$ is a ring with

$$0_{R[x]} = 0_R + 0_R \cdot x + \dots \quad (15.1.3)$$

and

$$1_{R[x]} = 1_R + 0_R \cdot x + \dots \quad (15.1.4)$$

Consequently R can naturally be embedded in $R[x]$ by $\begin{matrix} R \hookrightarrow R[x] \\ r \mapsto r \cdot x^0 \end{matrix}$. Also, $x \in Z(R[x])$ and $R[x]$ is generated by $R \cup \{x\}$.

Proof. (Left to the reader) ■

Proposition 15.1.2. There is a unique commutative ring structure on the set of polynomials $R[x]$ having these properties:

1. Addition of polynomials is done coefficient wise for equal degree monomials (like vector addition)
2. Multiplication of monomials is given by the rule above
3. The ring R is a subring of $R[x]$, when the elements of R are identified with the constant polynomials

Remark 15.1.2.

1. $(a + bx + cx^2)(\alpha + \beta x) = a\alpha + (a\beta + b\alpha)x + (b\beta + c\alpha)x^2 + c\beta x^3$
2. $R = M_2(\mathbb{R})$, $I_2 \cdot x^0 = 1 \in R[x]$, and $P = A_0 + A_1x + A_2x^2 + \dots + A_nx^n$, $A_i \in M_2(\mathbb{R})$.
3. $R = \mathbb{Z}/2\mathbb{Z} = \mathbb{F}_2$. In $\mathbb{F}_2[x]$, $(x + 1)^2 = x^2 + 2x + 1 = x^2 + 1 \pmod{2}$.

Definition 15.1.2 (Polynomials in Multiple Variables). Let x_1, \dots, x_n be variables (indeterminates). A monomial is a formal product of these variables of the form

$$x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \quad (15.1.5)$$

where the exponents i_v are nonnegative numbers. The n -tuple (i_1, \dots, i_n) of exponents determines the monomial. Such an n -tuple is called a multi-index, and vector notation $\mathbf{i} = (i_1, \dots, i_n)$ for multi-indices is convenient. Using it, we may write the monomial symbolically as

$$x^{\mathbf{i}} = x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \quad (15.1.6)$$

The monomial x^0 is denoted by 1.

A polynomial with coefficients in a ring R is a finite linear combination of monomials with coefficients in R . Using the shorthand, any polynomial $f(x) = f(x_1, \dots, x_n)$ can be written uniquely in the form

$$f(x) = \sum_i a_i x^i \quad (15.1.7)$$

And only finitely many of the coefficients $a_i \in R$ are different from zero.

A polynomial which is the product of a nonzero element $r \in R$ with a monomial is also called a monomial

$$m = r x^i \quad (15.1.8)$$

Using multi-index notation, the addition and multiplication for polynomials in multiple variables is analogous to the case for one variable using the formulas defined above, and the above proposition also holds analogously for polynomials in multiple variables.

A ring of polynomials in several variables with coefficients in the ring R is denoted by

$$R[x_1, \dots, x_n] \text{ or } R[x], \quad x = (x_1, \dots, x_n) \quad (15.1.9)$$

Remark 15.1.3. For a general ring R , $R[x]$ is not isomorphic to $\mathcal{P}(R, R) \subseteq F(R, R)$ (polynomial functions over R).

Example 15.1.1. Consider $R = \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ for p a prime. Consider now a polynomial function $f \in F(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z})$ defined by

$$f = \sum_{i=0}^n b_i (\hat{x})^i$$

where $(\hat{x})^i$ is the product in $F(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z})$. This gives us the subring \mathcal{P}

Claim 15.1.3. $\mathcal{P}(\mathbb{F}_p, \mathbb{F}_p) \cong \mathbb{F}_p[x]$

Proof. First, observe that $\mathcal{P}(\mathbb{F}_p, \mathbb{F}_p)$ is actually finite because $|F(\mathbb{F}_p, \mathbb{F}_p)| = p^p < +\infty$ and $\mathcal{P}(\mathbb{F}_p, \mathbb{F}_p) \subseteq F(\mathbb{F}_p, \mathbb{F}_p)$. However, $|\mathbb{F}_p[x]|$ is infinite. Indeed, for all $i \neq j \geq 0$, $x^i \neq x^j$, so $\{1, x, x^2, \dots\} = \{x^n : n \geq 0\} \subseteq \mathbb{F}_p[x]$ is a subset of infinite order. Moreover, note that $(\hat{x})^p - \hat{x} = 0 \in F(\mathbb{F}_p, \mathbb{F}_p)$ by Fermat's theorem because

$$((\hat{x})^p - \hat{x})(a) = a^p - a = a - a = 0 \pmod{p}$$

But, $x^p - x \neq 0 \in \mathbb{F}_p[x]$. ■

Thus, it is not sufficient to consider polynomial functions.

Claim 15.1.4. However, $\mathbb{R}[x] \cong \mathcal{P}(\mathbb{R}, \mathbb{R})$

Proof. Define a function

$$\begin{aligned} \Phi : \mathbb{R}[x] &\rightarrow \mathcal{P}(\mathbb{R}, \mathbb{R}) \\ p &\mapsto \begin{pmatrix} \mathbf{ev}_p : \mathbb{R} \rightarrow \mathbb{R} \\ r \mapsto \mathbf{ev}_r(p) \end{pmatrix} \end{aligned}$$

First, note that for all $p = \sum_i a_i x^i \in \mathbb{R}[x]$ and all $r \in \mathbb{R}$, $\mathbf{ev}_r(p) = \sum_i a_i r^i$, so we have that $\mathbf{ev}_p = \sum_i a_i (\hat{x})^i \in \mathcal{P}(\mathbb{R}, \mathbb{R})$, so the function is well-defined. Recall that $\mathbf{ev}_r(p)$ is a ring homomorphism for all $r \in \mathbb{R}$. Now, let $p = \sum_i a_i x^i, q = \sum_i b_i x^i \in \mathbb{R}[x]$. Then, observe that

$$\begin{aligned} \Phi(p+q)(r) &= \mathbf{ev}_{p+q}(r) & \Phi(p \cdot q)(r) &= \mathbf{ev}_{p \cdot q}(r) \\ &= \mathbf{ev}_r(p+q) & &= \mathbf{ev}_r(p \cdot q) \\ &= \mathbf{ev}_r(p) + \mathbf{ev}_r(q) & &= \mathbf{ev}_r(p) \cdot \mathbf{ev}_r(q) \\ &= \mathbf{ev}_p(r) + \mathbf{ev}_q(r) & &= \mathbf{ev}_p(r) \cdot \mathbf{ev}_q(r) \\ &= \Phi(p)(r) + \Phi(q)(r) & &= \Phi(p)(r) \cdot \Phi(q)(r) \\ &= (\Phi(p) + \Phi(q))(r) & &= (\Phi(p) \cdot \Phi(q))(r) \end{aligned}$$

and

$$\Phi(1)(r) = \mathbf{ev}_1(r) = \mathbf{ev}_r(1) = 1$$

for all $r \in \mathbb{R}$. Thus, we have that Φ is a homomorphism of rings. Then, let $p \in \ker(\Phi)$ and for the sake of contradiction suppose $\deg(p) = n$ for some $n \geq 0$. Then, we have that p has at most n roots, $\{r_1, r_2, \dots, r_n\}$. Let $r \in \mathbb{R}$ with $r \notin \{r_1, r_2, \dots, r_n\}$. Then, by assumption we have

$$0 = \Phi(p)(r) = \mathbf{ev}_p(r) = \mathbf{ev}_r(p)$$

But, $\mathbf{ev}_r(p)$ is the remainder for the division of p by $x - r$, so this implies $x - r$ divides p . However, we would then have r as a root of p , but by assumption r is not one of the n roots of p , so this is a contradiction. Therefore, we must have that $p = 0$, so $\ker(\Phi)$ is trivial. Now, let $f = \sum_i a_i (\hat{x})^i \in \mathcal{P}(\mathbb{R}, \mathbb{R})$. Observe that for $q = \sum_i a_i x^i$, we have for all $t \in \mathbb{R}$

$$\Phi(q)(t) = \mathbf{ev}_q(t) = \mathbf{ev}_t(q) = \sum_i a_i t^i = \left(\sum_i a_i (\hat{x})^i \right)(t)$$

so $\Phi(q) = f$, and in particular Φ is a surjection. Therefore, we conclude that Φ is an isomorphism of rings, so

$$\mathbb{R}[x] \cong \mathcal{P}(\mathbb{R}, \mathbb{R}) \quad (15.1.10)$$

■

Definition 15.1.3. Let R be an arbitrary unital ring. $P \in R[x]$ is called a *polynomial with coefficients in R* . For $P = a_0 + a_1x + \dots + a_nx^n$, a_0 is called the **constant coefficient**. Now, assume $P \neq 0$. Then $\max\{i \geq 0 : a_i \neq 0\}$ is the **degree of P** , denoted $\deg(P)$. $a_{\deg(P)}$ is called the **leading coefficient of P** .

Definition 15.1.4. Let R be a ring and let $r \in Z(R)$. Then, the map

$$\begin{aligned} \mathbf{ev}_r : R[x] &\rightarrow R \\ P = \sum_i a_i x^i &\mapsto P(r) = \sum_i a_i r^i \end{aligned} \quad (15.1.11)$$

is a surjective ring homomorphism, called the *evaluation at r* . Denote $\mathbf{ev}_r(P)$ by $P(r)$.

Proof. (Left to the reader)

■

15.2.0 §Division Algorithm

Definition 15.2.1. A non-zero polynomial is called monic if its leading coefficient is 1.

Example 15.2.1.

$$1. \deg(\underbrace{3x^2 + x + 1}_{\in \mathbb{Z}[x]}) = 2, \text{ leading coefficient} = 3.$$

$$2. x^2 + 7x_1 \in \mathbb{Z}[x] \text{ is monic of degree } 2$$

$$3. 2 \text{ has degree } 0.$$

Remark 15.2.1. $P, Q \in R[x]$, then if $PQ \neq 0$, then $\deg(PQ) \leq \deg(P) + \deg(Q)$. If R is a domain and $P \neq 0, Q \neq 0$, then $\deg(PQ) = \deg(P) + \deg(Q)$.

Proof. (Left to the reader) ■

Example 15.2.2. $\mathbb{Z}/6\mathbb{Z}[x]$, $(2x)(3x + 1) = 6x^2 + 2x = 2x$ of degree $1 < \deg(P) + \deg(Q) = 1 + 1 = 2$.

Corollary 15.2.1.

1. If R is a domain, then $R[x]$ is a domain
2. The units of $R[x]$ are the units of R .

Proof. (Left to the reader) ■

Theorem 15.2.2 (Division Algorithm). Let R be a ring. Let $P, Q \in R[x]$. Assume that $P \neq 0$ and that the leading coefficient of P is a unit ($\in R$). Then there are unique polynomials f and $g \in R[x]$ such that:

1. $Q = fP + g$
2. $g = 0$ or $\deg(g) < \deg(P)$

(g is called the remainder of the division of Q by P)

Proof. Write $\deg(Q) = m$ and $\deg(P) = n$. If $Q = 0$ or $m < n$ then $Q = 0P + Q$ does it. Hence, suppose $m \geq n$, and proceed by induction on m . Write $P = ux^n + a_{n-1}x^{n-1} + \dots$ and $Q = b_mx^m + b_{m-1}x^{m-1} + \dots$, where $u \in R^\times$ by hypothesis. Consider the new polynomial

$$\begin{aligned} g_1 &= Q - b_mu^{-1}x^{m-n}P \\ &= (b_{m-1} - b_mu^{-1}a_{n-1})x^{m-1} + \dots \end{aligned}$$

where we use the fact that x is central in $R[x]$. Hence, either $g_1 = 0$ or $\deg(g_1) < m$ so, by induction, polynomials q_1 and r exist such that $g_1 = Pq_1 + r$ and either $r = 0$ or $\deg(r) < \deg(P)$. But then

$$Q = g_1 + b_m u^{-1} x^{m-n} P = (q_1 + b_m u^{-1} x^{m-n}) P + r$$

Hence, the induction is satisfied, so f and g exist satisfying the claim.

To prove uniqueness, suppose that also $Q = f_1 P + g_1$, where either $g_1 = 0$ or $\deg(g_1) < \deg(P)$. Then $(g - g_1) = (f_1 - f)P$. If $(f_1 - f) \neq 0$, then since the leading coefficient of P is a unit $(f_1 - f)P \neq 0$, and that

$$\deg(g - g_1) = \deg[(f_1 - f)P] = \deg(f_1 - f) + \deg(P)$$

But, this implies $\deg(g - g_1) \geq \deg(P)$, but $\deg(g - g_1) \leq \max\{\deg(g), \deg(g_1)\} < \deg(P)$, a contradiction. Thus, we must have that $(f_1 - f) = 0$, and whence $(g - g_1) = (f_1 - f)P = 0$, proving uniqueness. ■

Corollary 15.2.3. *A non-zero polynomial of degree n over any field has at most n roots.*

Proof. (Left to the reader) ■

Example 15.2.3.

1. For $\mathbb{Z}[x]$, $P = -x + 1$, $Q = x^2$, $Q = P(-x - 1) + 1$, where $\deg(1) = 0 < 1 = \deg(P)$
2. Conversely, $P = 0 \cdot Q + (-x + 1)$, where $\deg(-x + 1) = 1 < 2 = \deg(x^2)$.

Corollary 15.2.4. *Let R be a commutative ring with $a \in R$, and $P \in R[x]$.*

1. $\mathbf{ev}_a(P) = 0$ if and only if $P = (x - a)Q$ for some $Q \in R[x]$.
2. The remainder of the division of P by $x - a$ is $P(a)$.

Proof. First, observe that if $P = (x - a)Q$ for some $Q \in R[x]$. Then $\mathbf{ev}_a(P) = 0 \mathbf{ev}_a(Q) = 0$, satisfying the implication. On the other hand, since the leading coefficient of $x - a$ is a unit, we have by the division algorithm that there exists $f, g \in R[x]$ such that $P = f(x - a) + g$, where $g = 0$ or $\deg(g) < \deg(x - a) = 1$. Thus, $g = r$ for some $r \in R$. Then, observe that $\mathbf{ev}_a(P) = \mathbf{ev}_a(f(x - a) + r) = \mathbf{ev}_a(f)0 + r = r$, since \mathbf{ev}_a is a ring homomorphism. Note that this proves the second claim. Now, by assumption $\mathbf{ev}_P = 0$, $r = 0$. Thus, $P = (x - a)f$, completing the proof. ■

Definition 15.2.2. *Let R be a commutative ring. We say that a polynomial $f \in R[X]$ **divides** a polynomial $g \in R[X]$ denoted $f \mid g$, and that f is a **divisor** of g if there exists $P \in R[X]$ such that $g = fP$.*

Corollary 15.2.5. *Let \mathbb{F} be a field. Let $f, g \in \mathbb{F}[X]$ not both zero. Then*

$$(f, g) := \{fP + gQ : P, Q \in \mathbb{F}[X]\} = (d) \quad (15.2.1)$$

where d is the monic generator of minimal degree, which satisfies

1. d divides f and g
2. If a polynomial P divides f and g , then P divides d
3. There exist $Q_1, Q_2 \in \mathbb{F}[X]$ such that $d = fQ_1 + gQ_2$ (one says that d is a linear combination of f and g)

Thus, every ideal I in the ring $R = F[x]$ is principal, $I = (f)$, generated by the monic polynomial f in I of least degree.

A. Let I be an ideal. If $I \neq (0)$, take $f \in I$ of minimal degree, n . Scale f by $c = a_n^{-1}$ to make f monic. Note that since $c \in F[x]$, $c \cdot f \in I$. Let h be another polynomial in I , and write $h(x) = q(x)f(x) + r(x)$, with the degree of $r(x)$ less than $f(x)$. Note that $q(x) \in F[x]$, so $q(x)f(x) \in I$, and $h(x) - q(x)f(x) \in I$. Thus, $r(x) \in I$. But, $r(x)$ has a smaller degree than f , so $r(x) = 0$. Thus, $h(x) = q(x)f(x)$, so $f(x)$ divides $h(x)$, and $I = (f)$. ■

B. (Left to the reader) ■

Remark 15.2.2. Thus, the set of ideals is in a one-to-one correspondence with the set of monic polynomials, and the ideal associated to f , I_f , contains the ideal generated by the monic polynomial g , that is $I_f \supset I_g$, if and only if f divides the polynomial g . Namely, $g(x) = f(x)q(x)$ for some $q(x) \in F[x]$.

Example 15.2.4 (Non-principal Ideals). Take the ring $R = F[x, y] = \left\{ \sum_{i=1, j=1}^{n, m} a_{ij}x^i y^j : a_{ij} \in F \right\}$.

Consider the map $h : R \rightarrow F$ by $f(x, y) \mapsto f(0, 0)$. The kernel of h is not generated by one element (so it's not principal). In fact, $\ker h = (x, y) = \{rx + sy : r, s \in R\}$.

Definition 15.2.3. Let $f, g \in \mathbb{F}[x]$ not both zero, then d in the above corollary is called the **greatest common divisor** of f and g , denoted $\gcd(f, g)$.

↳ If $\gcd(f, g) = 1$, then f and g are said to be **relatively prime**.

How do we compute the gcd?

Example 15.2.5. In $\mathbb{Q}[x]$

1. $(x^2 + 1, x) = (1)$ because $1 = (x^2 + 1) + (-x)(x) \in (x^2 + 1, x)$
2. $\gcd(x^2 - 1, 2x + 2) = \gcd((x - 1)(x + 1), 2(x + 1)) = (x + 1)$

Remark 15.2.3. In general, there is the **euclidean algorithm** for polynomials over a field.

Lemma 15.2.6. Let $f, g \in \mathbb{F}[x]$, \mathbb{F} a field, not both zero. If $g = Pf + Q$ for some $P, Q \in \mathbb{F}[x]$ then $\gcd(f, g) = \gcd(f, Q)$.

Proof. Let $d = \gcd(f, g)$, $d' = \gcd(f, Q)$. Then I claim $(f, Q) = (f, g)$. Note $f \in (f, Q)$ and $g = Pf + Q \in (f, Q)$ so $(f, g) \subseteq (f, Q)$. Similarly, $f \in (f, g)$ and $Q = g - Pf \in (f, g)$, so $(f, Q) \subseteq (f, g)$. Thus, $(f, g) = (f, Q)$ so $(d) = (f, g) = (f, Q) = (d')$. Hence, $d = d'$ and the proof is complete. ■

Theorem 21 (Euclidean Algorithm).

Let $f, g \in \mathbb{F}[x]$, not both zero. Assume $f \neq 0$, so the leading coefficient of f is $a \neq 0$. Then a is a unit in \mathbb{F} since \mathbb{F} is a field. By the division algorithm there exist unique $P, Q \in \mathbb{F}[x]$ such that

$$g = Pf + Q \quad (15.2.2)$$

with $Q = 0$ or $\deg(Q) < \deg(f)$. By the Lemma $\gcd(f, g) = \gcd(f, Q)$. If $Q = 0$, $\gcd(f, g) = \gcd(f, 0) = a^{-1}f$, where a is the leading coefficient of f . Otherwise, if $Q \neq 0$ we apply the division algorithm again to obtain $P', Q' \in \mathbb{F}[x]$ satisfying $f = P'Q + Q'$ for $Q' = 0$ or $\deg(Q') < \deg(Q)$. We then have $\gcd(f, g) = \gcd(f, Q) = \gcd(Q, Q')$. This process must terminate eventually because the degree of the remainder strictly decreases at every step and must be non-negative (or the remainder is 0).

Example 15.2.6. The greatest common divisor of $f = x^3 + x + 2$, $g = x^3 + x^2 + x + 1$ in $\mathbb{Q}[x]$: First $f = g \cdot 1 + (1 - x^2)$. Then $g = (1 - x^2)(-x - 1) + 2x + 2$, and finally $(-x^2 + 1) = (-x/2 + 1/2)(2x + 2) + 0$. Thus, $\gcd(f, g) = \gcd(g, 1 - x^2) = \gcd(1 - x^2, 2x + 2) = \gcd(2x + 2, 0) = x + 1$, by making the remainder monic. Working backwards we have

$$(x + 1) = g/2 - (x + 1)(x^2 - 1)/2 = g/2 - (x + 1)(g - f)/2 = -gx/2 + f(x + 1)/2$$

§Solutions to Polynomials

Proposition 15.2.7. If $p \in \mathbb{Z}[x]$ and $\mathbf{ev}_c(p) = 0$ for some $c \in \mathbb{Z}$, then $\mathbf{ev}_c(p) = 0$ in $\mathbb{Z}/m\mathbb{Z}$ for all $m \in \mathbb{Z}^+$. The contrapositive is important so it shall be stated: If there exists $m \in \mathbb{Z}^+$ such that $\mathbf{ev}_c(p) \neq 0$ in $\mathbb{Z}/m\mathbb{Z}$ for a polynomial $p \in \mathbb{Z}[x]$, then $\mathbf{ev}_c(p) \neq 0$ in \mathbb{Z} , and in particular c is not a root of p in \mathbb{Z} .

Proof. Let $p \in \mathbb{Z}[x]$ and suppose $\mathbf{ev}_c(p) = 0$. Then, let $m \in \mathbb{Z}^+$. Write $p = \sum_i a_i x^i$. Then observe that $\pi(p) = \sum_i [a_i]_m x^i = [\sum_i a_i x^i]$. Then, since \mathbf{ev}_c is a ring homomorphism we have that

$$\mathbf{ev}_c(\pi(p)) = \sum_i [a_i]_m c^i = \left[\sum_i a_i c^i \right] = [\mathbf{ev}_c(p)] = [0]$$

in $\mathbb{Z}/m\mathbb{Z}$, completing the proof. ■

15.3.0 §Substitution Principle

Theorem 22 (Substitution Principle).

Let $\phi : R \rightarrow R'$ be a ring homomorphism.

1. Given an element $\alpha \in R'$, there is a unique homomorphism $\Phi : R[x] \rightarrow R'$ which agrees with the map ϕ on constant polynomials, and which sends $x \mapsto \alpha$
2. More generally, given $\alpha_1, \dots, \alpha_n \in R'$, there is a unique homomorphism $\Phi : R[x_1, \dots, x_n] \rightarrow R'$ from the polynomial ring in n variables to R' , which agrees with ϕ on constant polynomials and which sends $x_i \mapsto \alpha_i$, for $i = 1, 2, \dots, n$

Proof. With vector notation for indices, the proof of (2) is identical to that of (1). Let us denote the image of an element $r \in R$ in R' by r' . Using the fact that Φ is a homomorphism which restricts to ϕ on R , and sends $x_v \mapsto \alpha_v$, we find that it acts on a polynomial $f(x) = \sum r_i x^i$ by sending

$$\sum r_i x^i \mapsto \sum \phi(r_i) \alpha^i = \sum r'_i \alpha^i \quad (15.3.1)$$

In other words, Φ acts on the coefficients of a polynomial as ϕ , and it substitutes α for x . Since this formula describes Φ completely for us, we have proved the uniqueness of the substitution homomorphism. To prove its existence, we take this formula as the definition of Φ , and we show that the map is a ring homomorphism $R[x] \rightarrow R'$. Since ϕ is a ring homomorphism, Ψ sends 1 to 1, and by the above formula, it is compatible with addition of polynomials. Using the formula we also find that it is compatible with multiplication as

$$\begin{aligned} \Psi(fg) &= \Psi\left(\sum a_i b_j x^{i+j}\right) \\ &= \sum \Psi(a_i b_j x^{i+j}) \\ &= \sum_{i,j} a'_i b'_j \alpha^{i+j} \\ &= \left(\sum_i a'_i \alpha^i\right) \left(\sum_j b'_j \alpha^j\right) \\ &= \Psi(f) \Psi(g) \end{aligned}$$

■

Example 15.3.1. We consider the case of a homomorphism $\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}$. This map extends to a homomorphism

$$\mathbb{Z}[x] \rightarrow \mathbb{Z}/p\mathbb{Z}[x], f(x) = a_n x^n + \dots + a_0 \mapsto \overline{a_n} x^n + \dots + \overline{a_0} = \overline{f}(x) \quad (15.3.2)$$

where $\overline{a_i}$ denotes the residue class of a_i modulo p . We call the polynomial $\overline{f}(x)$ the residue of $f(x)$ modulo p .

Corollary 15.3.1. Let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_m)$ denote set of variables. There is a unique isomorphism $R[x, y] \xrightarrow{\sim} R[x][y]$ which is the identity on R and which sends the variables to themselves.

Proof. Note that R is a subring of $R[x]$, and that $R[x]$ is a subring of $R[x][y]$. So R is also a subring of $R[x][y]$. Consider the inclusion map $\phi : R \hookrightarrow R[x][y]$. The Substitution Principle tells us that there is a unique homomorphism $\Phi : R[x, y] \rightarrow R[x][y]$ which extends the map and sends variables x_μ, y_ν wherever we wish. Thus, we can send the variables to themselves. The map Φ constructed is thus the desired isomorphism. Using the Substitution Principle

once more, we note that $R[x]$ is a subring of $R[x, y]$, so we can extend the inclusion map $\psi : R[x] \rightarrow R[x, y]$ to a map $\Psi : R[x][y] \rightarrow R[x, y]$ by sending y_j to itself. The composed homomorphism $\Psi\Phi : R[x, y] \rightarrow R[x, y]$ is the identity on R and on $\{x_\mu, y_\nu\}$. By uniqueness of the Substitution Principle, $\Psi\Phi$ is the identity map. Similarly, $\Phi\Psi$ is the identity on $R[x][y]$. Thus, Φ is a bijective homomorphism, so it is an isomorphism. ■

Proposition 15.3.2. *Let \mathcal{R} denote the ring of continuous real-valued functions on \mathbb{R}^n . The map $\phi : \mathbb{R}[x_1, \dots, x_n] \rightarrow \mathcal{R}$ sending a polynomial to its associated polynomial function is an injective homomorphism.*

Proof. The existence of this homomorphism follows from the Substitution Principle. To prove injectivity, it is enough to show that if the function associated to a polynomial $f(x)$ is the zero function, then $f(x)$ is the zero polynomial. Let the associated function be $\tilde{f}(x)$. If $\tilde{f}(x)$ is identically zero, then all its derivatives are zero too. On the other hand we can differentiate a formal polynomial by using the power rule and the linearity of the derivative. If some coefficients of $f(x)$ are nonzero, then the constant term of a suitable derivative will be nonzero too. Hence, that derivative will not vanish at the origin. Therefore, $\tilde{f}(x)$ can't be the zero function. ■

Proposition 15.3.3. *There is exactly one ring homomorphism*

$$\phi : \mathbb{Z} \rightarrow R \quad (15.3.3)$$

from the ring of integers to an arbitrary ring R . It is the map defined by $\phi(n) = 1_R + \dots + 1_R$ n -times if $n > 0$, and $\phi(-n) = -\phi(n)$.

Remark 15.3.1. This allows us to identify the images of the integers in an arbitrary ring R . We can hence interpret the symbol 3 as $1 + 1 + 1$ in R .

15.4.0 §Roots and Factorization

For this section let R be a commutative ring and $P \in R[x]$.

Definition 15.4.1. $r \in R$ is a root of P if $\text{ev}_r(P) = P(r) = 0$. Note that this happens if and only if $P = (x - r)Q$ for some $Q \in R[x]$.

What if $Q(r) = 0$ as well?

Definition 15.4.2. Let $\alpha \in R$ be a root of P . We say that α is a root of P of multiplicity $n \geq 1$ if $P = (x - \alpha)^n Q'$ for some $Q' \in R[x]$ and $Q'(\alpha) \neq 0$.

Proposition 15.4.1. *Let R be an integral domain, and let $P \neq 0$ be a polynomial of degree n . Then P has at most n roots counted with multiplicities.*

Proof. We argue by induction on the degree of P . If $n = 0$ then $P = c \neq 0$, so P has no roots. Thus, the base case holds. Then, suppose that there exists $k \geq 0$ such that for all $j \leq k$, if $n = k$

P has at most k roots counting multiplicities. Then, consider $n = k + 1$. If P has no roots then we are done. Otherwise, let α be a root of P . Then $P = (x - \alpha)Q$ for some $Q \in R[x]$ such that $\deg(Q) = k + 1 - 1 = k$. Thus, by our induction hypothesis Q has at most k roots counting with multiplicities. Thus, P has at most $k + 1$ roots counting with multiplicities, as desired. ■

Remark 15.4.1. Note that if R is not an integral domain this is not true.

↳ If $R = \mathbb{Z}/6\mathbb{Z}$, then observe

$$P = (x - 2)(x - 3) = x^2 - 5x + 6 = x^2 - 5x = x(x - 5)$$

so $\{0, 2, 3, 5\}$ are roots of P and P is of degree 2.

Moreover, even if R is an integral domain, $P \in R$ with $\deg(P) > 1$ may have no roots. For example, $x^2 + 1 \in \mathbb{R}[x]$ has no roots in \mathbb{R} .

Corollary 15.4.2. If R is an integral domain and $0 \neq P \in R[x]$ of degree n , then P has at most n distinct roots.

Definition 15.4.3 (Irreducible). Let R be an integral domain.

1. An element $a \in R$ is irreducible if

(a) $a \notin R^\times$ and $a \neq 0$

(b) If $a = bc$ for some $b, c \in R$, then $b \in R^\times$ or $c \in R^\times$

2. If $P \in R[x]$ is irreducible, we say that P is irreducible over R .

Example 15.4.1.

1. $x^2 + 1$ is irreducible over \mathbb{R}

2. $x^2 + 1 = (x + i)(x - i)$ is not irreducible over \mathbb{C}

3. $2x + 2$ is irreducible over \mathbb{Q} , but not over \mathbb{Z} as $2 \notin \mathbb{Z}^\times = \{1, -1\}$

4. A linear polynomial (i.e.) of degree 1) over a field is irreducible.

Proof. Let F be a field and let $P \in F[x]$ of degree 1. Then $P \neq 0$ and $P \notin F[x]^\times$. Moreover, if $P = fg$ for some $f, g \in F[x]$ then $\deg(fg) = \deg(f) + \deg(g) = \deg(P) = 1$, so either $\deg(f) = 0$ or $\deg(g) = 0$. Thus, either $g \in F$ or $f \in F$ with $g, f \neq 0$, so in particular either $g \in F[x]^\times$ or $f \in F[x]^\times$. Therefore, P is irreducible over F . ■

Proposition 15.4.3. Let F be a field and $P \in F[x]$ with $\deg(P) \geq 2$.

1. If P is irreducible over F , then P has no roots in F

2. If $\deg(P) \in \{2, 3\}$, then P is irreducible over F if and only if P has no roots in F .

Proof. If P has a root $a \in F$, then $P = (x - a)Q$ for some $Q \in F[x]$. But, $\deg(P) \geq 2$ so $\deg(Q) \geq 1$, and hence $Q \notin F[x]^\times$. Thus, P is not irreducible over F . On the other hand, suppose $\deg(P) \in \{2, 3\}$. Suppose P is not irreducible over F so $P = p_1 p_2$ for some $p_1, p_2 \in F[x]$ such that $\deg(p_1), \deg(p_2) \geq 1$. But, $\deg(p_1) + \deg(p_2) = \deg(P) \in \{2, 3\}$ so either $\deg(p_1) = 1$ or $\deg(p_2) = 1$. Without loss of generality suppose $\deg(p_1) = 1$. Then $p_1 = ax + b$ for $a, b \in F$ and $a \neq 0$. Since F is a field $a^{-1} \in F$ and $\text{ev}_{a^{-1}(-b)}(p_1) = 0$. Hence, as $P = p_1 p_2$ P has a root in F , completing the proof. ■

Remark 15.4.2.

1. In general no roots $\not\Rightarrow$ irreducible

↳ For example, $(x^2 + 1)^2 \in \mathbb{R}[x]$ is not irreducible but it has no roots in \mathbb{R} .

2. $x^2 + x + 1$ is irreducible over $\mathbb{Z}/2\mathbb{Z}$ (no roots)

3. $x^2 - 2$ is irreducible over \mathbb{Q} (no roots), but it has roots over \mathbb{R} .

Definition 15.4.4 (Algebraically Closed). A field F is algebraically closed if every non-constant polynomial $P \in F[x]$ has a root.

Theorem 23 (Fundamental Theorem of Algebra).

The field of complex numbers \mathbb{C} is algebraically closed. So, $P = a(x - r_1)(x - r_2)\dots(x - r_n)$ for $P \in \mathbb{C}[x]$, $\deg(P) = n$, $a \in \mathbb{C} \setminus \{0\}$ the leading coefficient of P , and $\{r_1, r_2, \dots, r_n\}$ the roots of P (not necessarily distinct).

§Polynomials over \mathbb{Q} and \mathbb{Z}

Theorem 24 (Rational Roots Theorem).

Let $P \in \mathbb{Z}[x]$, $P = a_0 + a_1x + \dots + a_nx^n$, $a_n \neq 0 \neq a_0$. Then every root of P in \mathbb{Q} is of the form $\frac{c}{d}$ such that $c \mid a_0$ and $d \mid a_n$. In particular, if P is monic, i.e. $a_n = 1$, then every rational root of P is in \mathbb{Z} .

Proof. Suppose $\frac{c}{d}$ is a root of P in \mathbb{Q} , and assume $\gcd(c, d) = 1$. Then,

$$P\left(\frac{c}{d}\right) = a_0 + a_1\frac{c}{d} + \dots + a_n\frac{c^n}{d^n} = 0$$

Multiply by d^n to obtain

$$a_0d^n + a_1cd^{n-1} + \dots + a_{n-1}c^{n-1}d + a_nc^n = 0$$

Then, since $a_1cd^{n-1}, \dots, a_nc^n$ are divisible by c , we must have that c divides a_0d^n . But, $\gcd(c, d) = 1$, so as can be shown by induction, $\gcd(c, d^n) = 1$. Hence, c divides a_0 . Indeed, $cx + d^ny = 1$ for some $x, y \in \mathbb{Z}$, so $a_0 = c(a_0x + ky)$, where $a_0d^n = ck$. Similarly, a_nc^n is divisible by d as $a_0d^n, \dots, a_{n-1}c^{n-1}d$ are divisible by d . Thus, again $\gcd(d, c^n) = 1$ so d divides a_n , completing the proof. ■

Example 15.4.2. $P = x^3 + 2x^2 + \frac{3}{5}x + 2$ is irreducible over \mathbb{Q} . Indeed, since $\deg(P) = 3$, P is irreducible over \mathbb{Q} if it has no roots in \mathbb{Q} . To put P in the form of Rational Roots Theorem we multiply by 5:

$$5P = 5x^3 + 10x^2 + 3x + 10 \in \mathbb{Z}[x]$$

By the Rational Roots Theorem, if $5P$ has a root $\frac{c}{d} \in \mathbb{Q}$ then $c \mid 10$ and $d \mid 5$. Thus, $d \in \{\pm 1, \pm 5\}$ and $c \in \{\pm 1, \pm 2, \pm 5, \pm 10\}$. In particular

$$\frac{c}{d} = \{\pm 1, \pm 2, \pm 5, \pm 10, \pm \frac{1}{5}, \pm \frac{2}{5}\}$$

Upon direct computation none of these values are roots of $5P$, so in particular $5P$ has no roots in \mathbb{Q} . Thus, P has no roots over \mathbb{Q} and is hence irreducible over \mathbb{Q} .

Theorem 25 (Gauss' Lemma).

Let $f, g, h \in \mathbb{Z}[x]$ such that $f = gh$. If a prime p divides every coefficient of f , then p divides every coefficient of g or p divides every coefficient of h .

Proof. Consider the surjective ring homomorphism

$$\begin{aligned} \mathbb{Z}[x] &\xrightarrow{\alpha} \mathbb{Z}/p\mathbb{Z}[x] \\ a_0 + a_1x + \dots + a_nx^n &\mapsto [a_0]_p + [a_1]_px + \dots + [a_n]_px^n \end{aligned}$$

Then $[0]_p = \alpha(f) = \alpha(g)\alpha(h)$. But, since $\mathbb{Z}/p\mathbb{Z}$ is a field, it is an integral domain and we must have that $\alpha(g) = 0$ or $\alpha(h) = 0$. That is, $g \in (p)$ or $h \in (p)$, completing the proof. ■

Definition 15.4.5. For all $f \in \mathbb{Z}[x]$, $\alpha(f) \in \mathbb{Z}/p\mathbb{Z}[x]$ is called reduction modulo p of f .

Corollary 15.4.4. Let f be a non-constant polynomial in $\mathbb{Z}[x]$.

1. If $f = gh$ with $g, h \in \mathbb{Q}[x]$, then there exist $g_0, h_0 \in \mathbb{Z}[x]$ so that $\deg(g_0) = \deg(g)$, $\deg(h_0) = \deg(h)$ and $f = g_0h_0$
2. f is irreducible over \mathbb{Q} if and only if f cannot be written $f = g_0h_0 \in \mathbb{Z}[x]$, where g_0, h_0 are non-constant.

Proof. Let $f \in \mathbb{Z}[x]$, with $\deg(f) \geq 1$.

1) Suppose $f = gh$ with $g, h \in \mathbb{Q}[x]$. Let a and b be least common multiples of the denominators of the coefficients of g and h , respectively. Then $g' = ag$ and $h' = bh$ are in $\mathbb{Z}[x]$. Moreover, we have the equation $abf = g'h'$ in $\mathbb{Z}[x]$. If $ab = 1$ then we're done, so suppose $ab > 1$ and let p be a prime dividing ab . Then by Gauss' Lemma p divides g' or h' . Hence, p can be cancelled to give

$$\frac{ab}{p}f = g_2h_2 \tag{15.4.1}$$

in $\mathbb{Z}[x]$. Repeat for all prime factors of ab to obtain $f = g_0h_0$ in $\mathbb{Z}[x]$ with $\deg(g_0) = \deg(g)$ and $\deg(h_0) = \deg(h)$.

2) If $f = g_0h_0 \in \mathbb{Z}[x]$ for g_0, h_0 non-constant, then $0 \neq g_0, h_0 \notin \mathbb{Q}[x]^\times$, so f is not irreducible. Conversely, if f is not irreducible over \mathbb{Q} then $f = gh \in \mathbb{Q}[x]$ and the result follows from 1), completing the proof. ■

Definition 15.4.6. Let $f \in \mathbb{Z}[x]$. A proper factorization of f is $f = g_0 h_0$ where $\deg(g_0) \geq 1$ and $\deg(h_0) \geq 1$.

Theorem 26 (Modular Irreducibility Test).

Let $0 \neq f \in \mathbb{Z}[x]$ such that there exists a prime number p with:

1. p does not divide the leading coefficient of f
2. The reduction $\alpha(f)$ of f modulo p is irreducible over $\mathbb{Z}/p\mathbb{Z}$

Then f is irreducible over \mathbb{Q} .

Proof. Suppose $f \in \mathbb{Z}[x]$ such that f satisfies the conditions of the theorem. Then, for the sake of contradiction suppose f is not irreducible over \mathbb{Q} . Then $f = gh$ for some non-constant $g, h \in \mathbb{Q}[x]$. By the corollary to Gauss' Lemma we have that $f = g_0 h_0$ for $g_0, h_0 \in \mathbb{Z}[x]$ non-constant polynomials. Then, we have that $\alpha(f) = \alpha(g_0)\alpha(h_0)$. Note that if a and b are the leading coefficients of g_0 and h_0 respectively, then ab is the leading coefficient of f , which by assumption is not divisible by p . Thus, a and b are not divisible by p , so $\deg(\alpha(g_0)) \geq 1$ and $\deg(\alpha(h_0)) \geq 1$. But, this implies that $\alpha(g_0), \alpha(h_0)$ are not units in $\mathbb{Z}/p\mathbb{Z}[x]$, so $\alpha(f) = \alpha(g_0)\alpha(h_0)$ is not irreducible over $\mathbb{Z}/p\mathbb{Z}$, contradicting the initial assumptions. Therefore, f must be irreducible over \mathbb{Q} . ■

Example 15.4.3.

1. $f = x^3 + 4x^2 + 6x + 2 \in \mathbb{Z}[x]$ is irreducible over \mathbb{Q} . Indeed, take $p = 3$, then $f \bmod 3 = x^3 + x^2 + 2 \in \mathbb{Z}/3\mathbb{Z}[x]$, which has no roots in $\mathbb{Z}/3\mathbb{Z}$, so $f \bmod 3$ is irreducible over $\mathbb{Z}/3\mathbb{Z}[x]$. Applying the Modular Irreducibility Test, f is irreducible over \mathbb{Q} .

Remark 15.4.3. Irreducible over $\mathbb{Q} \not\Rightarrow$ irreducible over $\mathbb{Z}/p\mathbb{Z}$.

↳ Eg: $x^2 - 2$ is irreducible over \mathbb{Q} , but has a root in $\mathbb{Z}/2\mathbb{Z}$.

↳ For $p = 2$, $x^4 + 1$ is irreducible over \mathbb{Q} , but not over $\mathbb{Z}/p\mathbb{Z}$.

Theorem 27 (Eisenstein's Criterion).

Let $f = a_0 + a_1x + \dots + a_nx^n \in \mathbb{Z}[x]$ with $n \geq 1$ and $a_n \neq 0$. Suppose there exists a prime p such that

1. $p \mid a_i$ for all $0 \leq i < n$,
2. $p \nmid a_n$
3. $p^2 \nmid a_0$

Then f is irreducible in $\mathbb{Q}[x]$

Proof. Suppose $f \in \mathbb{Z}[x]$, $\deg(f) \geq 1$ and let $p \in \mathbb{Z}$ a prime satisfying the conditions of the theorem. If f is not irreducible in $\mathbb{Q}[x]$, then there exists a proper factorization $f = gh$, $g, h \in \mathbb{Z}[x]$ by the corollary to Gauss' Lemma. Write $g = b_0 + b_1x + \dots + b_kx^k$ and $h = c_0 + c_1x + \dots + c_lx^l$, so we have $a_0 = b_0c_0$, $a_n = b_m c_l$, and $n = m + l$. Note, since p does not divide a_n , p does not divide b_k nor c_l . Let $\alpha : \mathbb{Z}[x] \rightarrow \mathbb{Z}/p\mathbb{Z}[x]$ be the reduction modulo p . Then $\alpha(f) = [a_n]_p x^n = \alpha(g)\alpha(h)$. Since p does not divide the leading coefficients of g nor h , this is a proper factorization. Note that since p^2 does not divide a_0 , p divides b_0 or c_0 but not both. Without loss of generality suppose p divides b_0 . Then, let b_m be the first element of b_0, b_1, \dots, b_k for which p does not divide (this is possible as p does not divide b_k). Then, note that

$$a_m = b_m c_0 + b_{m-1} c_1 + \dots + b_1 c_{m-1} + b_0 c_m$$

Then, since $m \leq k < n$, p divides a_m . Moreover, by construction p divides $b_{m-i} c_i$ for all $i \geq 1$. Thus, it follows that p must divide $b_m c_0$, so p divides b_m or p divides c_0 . But, by assumption p does not divide b_m and p does not divide c_0 , leading to a contradiction. Therefore, f must be irreducible over \mathbb{Q} . ■

Theorem 28 (General Eisenstein Criterion).

Let P be a prime ideal of an integral domain R . Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in R[x]$, where $n \geq 1$. Also, suppose $a_{n-1}, a_{n-2}, \dots, a_0 \in P$, but $a_n \notin P$ and $a_0 \notin P^2 = \{\sum_{i=1}^n p_i q_i : n \geq 1, p_i, q_i \in P\}$.

Proof. Towards a contradiction suppose $f(x) = b(x)c(x)$ for some $b(x), c(x) \in R[x]$ such that $\deg(b(x)), \deg(c(x)) \geq 1$. Then consider the reduction map $\varphi : R[x] \rightarrow (R/P)[x]$. Since P is a prime ideal, R/P is an integral domain. Denote a reduced element with a bar. Then we have that $\overline{f(x)} = \overline{a(x)b(x)}$ and $\overline{f(x)} = a_n x^n$. Let F denote the field of fractions for (R/P) , and extend the natural injection $\iota : R/P \hookrightarrow F$ to an injection $\iota' : (R/P)[x] \hookrightarrow F[x]$. Then, since F is a field $F[x]$ is a unique factorization domain. Hence, the factorization $\overline{a_n x^n} = \overline{\alpha x^j \beta x^i}$ is unique up to associates, i.e. units. Thus, we must have that $\overline{a(x)} = \overline{\alpha_j x^j}$ for some $j \geq 1$, $\alpha_j \in R$ and $\overline{b(x)} = \overline{\beta_i x^i}$ for $i = n - j$, $\beta_i \in R$. Thus, $\overline{a_0}, \overline{\beta_0} = \overline{0}$. Hence, $a_0 = \alpha_0 \beta_0 \in P^2$, which contradicts the assumption that $a_0 \notin P^2$. Therefore, we conclude that $f(x)$ must be irreducible in $R[x]$. ■

Note 15.4.4. The reduction modulo p of f , $[a_n]_p x^n$, is not irreducible over $\mathbb{Z}/p\mathbb{Z}$ if $n > 1$.

Example 15.4.4.

1. $x^{1000} + 3x + 6$ is irreducible over \mathbb{Q} . Apply Eisenstein's Criterion for $p = 3$, where $p \mid 3, 6$, $p^2 \nmid 6$ and $p \nmid 1$. Thus, it is irreducible over \mathbb{Q} .
2. If p is a prime, the p th cyclotomic polynomial

$$\Phi_p = x^{p-1} + x^{p-2} + \dots + x + 1 \quad (15.4.2)$$

is irreducible over \mathbb{Q} .

Proof. Replacing x by $x + 1$, it suffices to show that $\Phi_p(x + 1)$ is irreducible. Observe that

$$(x - 1)\Phi_p = (x - 1)(x^{p-1} + x^{p-2} + \dots + x + 1) = x^p - 1$$

Replacing x by $x + 1$, $x\Phi_p(x + 1) = (x + 1)^p - 1$, so by the binomial theorem

$$\Phi_p(x + 1) = x^{p-1} + \binom{p}{1} x^{p-2} + \dots + \binom{p}{p-2} x + p$$

But, p divides $\binom{p}{k}$ for all $1 \leq k \leq p - 1$, and $p^2 \nmid p$, so by Eisenstein's Criterion, $\Phi_p(x + 1)$ is irreducible over \mathbb{Q} . ■

§Parallels between the Integers and Polynomials over a Field

Remark 15.4.5 (Parallels).

1. \mathbb{Z} Integral domain that is not a field
 $F[x]$ Same
2. \mathbb{Z} Principal ideal domain ($I = n\mathbb{Z}$)
 $F[x]$ Same ($I = (d)$ for d monic)
3. \mathbb{Z} For $n \in \mathbb{Z}$, $n \neq 0, \pm 1$, $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if $n = \pm p$, p a prime
 if and only if $\mathbb{Z}/n\mathbb{Z}$ is a field
 $F[x]$ For $P \in F[x]$, $\deg(P) \geq 0$ (i.e. $P \neq 0$ and $P \notin F[x]^\times$) $F[x]/(P)$ is an integral domain if and only if P is irreducible over F (i.e. $P = aQ$ for Q monic irreducible and $a \in F^\times$) if and only if $F[x]/(P)$ is a field
4. \mathbb{Z} Unique Factorization Domain (in terms of unique prime numbers)
 $F[x]$ Same (in terms of unique monic irreducible polynomials)
5. \mathbb{Z} Expression of gcd as largest common factor in the UFD factorization
 $F[x]$ Same

15.5.0 §Polynomials over a Field

Theorem 15.5.1. *Let F be a field, $P \in F[x]$, $\deg(P) > 0$. Then the following are equivalent:*

1. $F[x]/(P)$ is an integral domain
2. P is irreducible in $F[x]$
3. $F[x]/(P)$ is a field

Proof. Suppose F is a field and $P \in F[x]$, with $\deg(P) > 0$.

[1 \implies 2] Suppose $F[x]/(P)$ is an integral domain. Then, (P) is a prime ideal in $F[x]$, so in particular P is a prime element of $F[x]$. Then, suppose $P = fg$ for some $f, g \in F[x]$. Since

P is a prime element P divides f or g . Without loss of generality suppose P divides f and write $f = Pq$ for some $q \in F[x]$. Then, we have that $P = Pqg$, so $P(1 - qg) = 0$. But, $P \neq 0$ and $F[x]$ is an integral domain, so $1 - qg = 0$. Hence, $1 = qg$, so $g \in F[x]^\times$. Therefore, P is an irreducible element in $F[x]/(P)$.

[2 \implies 3] Suppose P is irreducible in $F[x]$. Let $I \subset F[x]$ be an ideal containing (P) . Then, since F is a field, $F[x]$ is a PID so $I = (f)$ for some $f \in F[x]$. It follows that $(P) \subset (f)$, so $P \in (f)$. Thus, $P = fq$ for some $q \in F[x]$. Then either f or q is a unit. If f is a unit then $(f) = F[x]$. On the other hand, if q is a unit then $f = q^{-1}P \in (P)$, so $(f) = (P)$. Therefore, (P) is a maximal ideal in $F[x]$, so $F[x]/(P)$ is a field, as claimed.

[3 \implies 1] Suppose $F[x]/(P)$ is a field. Then, in particular it is an integral domain.

Thus, all implications hold so the statements are equivalent. \blacksquare

Note 15.5.1. The element $x + (P) \in \mathbb{F}[x]/(P) = K$ is a root of P treated with coefficients in K . Indeed, we have an injective homomorphism

$$\begin{aligned} F &\hookrightarrow F[x]/(P) \\ a &\mapsto a + (P) \end{aligned} \tag{15.5.1}$$

Then we can consider $P \in F[x] \hookrightarrow K[x]$. We claim that P as a polynomial in $K[x]$ has a root $\alpha = x + (P) \in K$. In particular, denote elements of K by $\bar{p} \in K$ where $p \in F[x]$. Then, observe that if $P = a_n x^n + \dots + a_1 x + a_0$ in $F[x]$, then $P = \bar{a}_n x^n + \dots + \bar{a}_1 x + \bar{a}_0 \in K[x]$. It follows that

$$\begin{aligned} \text{ev}_\alpha(P) &= P(\alpha) \\ &= \bar{a}_n \alpha^n + \dots + \bar{a}_1 \alpha + \bar{a}_0 \\ &= \overline{a_n x^n + \dots + a_1 x + a_0} \\ &= \bar{P} = \bar{0} \in K = \mathbb{F}[x]/(P) \end{aligned}$$

Example 15.5.1. $\mathbb{R}[x]/(x^2 + 1)$ is a field such that

$$(x + (x^2 + 1))^2 = x^2 + (x^2 + 1) = -1 + (x^2 + 1) \tag{15.5.2}$$

because $x^2 + 1 \in (x^2 + 1)$, so letting $\alpha = x + (x^2 + 1)$ we have that

$$\alpha^2 + 1 = 0 \in K = \mathbb{R}[x]/(x^2 + 1) \tag{15.5.3}$$

Corollary 15.5.2. Let $P \in F[x]$ be irreducible over F , and let $f_1, \dots, f_n \in F[x]$. If $P \mid f_1 f_2 \dots f_n$, then there is i such that $P \mid f_i$.

Proof. By the previous theorem $F[x]/(P)$ is a field since $P \in F[x]$ is irreducible over F . Hence, (P) is a maximal ideal so in particular (P) is a prime ideal. Then if $fg \in (P)$, $f \in (P)$ or $g \in (P)$. We shall proceed by induction on n . For $n = 1$ and $n = 2$ the base case holds trivially by definition of P . Hence, suppose there exists $k \geq 2$ such that if $n = k$, $f_1 f_2 \dots f_k \in (P)$ implies $f_i \in (P)$ for some $i \in \{1, 2, \dots, k\}$. Then, consider $n = k + 1$, so $f_1 f_2 \dots f_k f_{k+1} \in (P)$. Since (P) is a prime ideal either $f_1 f_2 \dots f_k \in (P)$ or $f_{k+1} \in (P)$. If $f_{k+1} \in (P)$ we're done. Hence, suppose $f_1 f_2 \dots f_k \in (P)$. But then, by the induction hypothesis there exists $i \in \{1, 2, \dots, k\}$ such that $f_i \in (P)$. Therefore, by mathematical induction we conclude that if $f_1 f_2 \dots f_n \in (P)$, then there exists $i \in \{1, 2, \dots, n\}$ such that $f_i \in (P)$ for all $n \geq 1$. \blacksquare

Remark 15.5.2. If $P \in F[x]$, $\deg(P) = n \geq 1$, (not necessarily irreducible) then

$$\begin{aligned} \prod_{i=1}^n F &\xrightarrow{\phi} F[x]/(P) \\ (a_0, a_1, \dots, a_{n-1}) &\mapsto a_0 + a_1x + \dots + a_{n-1}x^{n-1} + (P) \end{aligned} \quad (15.5.4)$$

is a group isomorphism for $(F[x]/(P), +)$.

Proof. By definition of addition in $F[x]/(P)$ we note that ϕ is a group homomorphism. First, let $(a_0, a_1, \dots, a_{n-1}) \in \ker(\phi)$. Then in particular $a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in (P)$ so there exists $g \in F[x]$ such that

$$a_0 + a_1x + \dots + a_{n-1}x^{n-1} = gP$$

But, since F is a field we have that $\deg(gP) = \deg(g) + \deg(P) \geq n$ or $gP = 0$, and $a_0 + a_1x + \dots + a_{n-1}x^{n-1} = 0$ or $\deg(a_0 + a_1x + \dots + a_{n-1}x^{n-1}) = n - 1$. Thus, we must have that $a_0 + a_1x + \dots + a_{n-1}x^{n-1} = gP = 0$. Therefore, $a_0 = a_1 = \dots = a_{n-1} = 0$, so $\ker(\phi)$ is trivial. Hence, ϕ is injective. Write $P = b_0 + b_1x + \dots + b_nx^n$. Then, for any $c_0 + c_1x + \dots + c_kx^k + (P) \in F[x]/(P)$, for all $m \geq n$ we can replace x^m by $x^{m-n}b_n^{-1}(-b_0 - b_1x - \dots - b_{n-1}x^{n-1})$. Repeat this step until all powers of x are less than or equal to $n - 1$, so $c_0 + c_1x + \dots + c_kx^k + (P) = c'_0 + c'_1x + \dots + c'_{n-1}x^{n-1} + (P)$. Then, we have that $\phi(c'_0, c'_1, \dots, c'_{n-1}) = c'_0 + c'_1x + \dots + c'_{n-1}x^{n-1} + (P)$, so ϕ is indeed surjective. Hence, we have that ϕ is a group isomorphism.

[Alternative Surjectivity] Let $Q \in F[x]$. If $\deg(Q) \leq n - 1$ then $Q = b_0 + b_1x + \dots + b_{n-1}x^{n-1} + (P) \in F[x]/(P)$ is the image of $(b_0, b_1, \dots, b_{n-1})$. If $\deg(Q) \geq n$, then $Q = aP + q$ by the division algorithm, with $q = 0$ or $\deg(q) < \deg(P) = n$. So, $Q + (P) = aP + q + (P) = q + (P)$ which is in the image by the first case. ■

Corollary 15.5.3. If F is a finite field of order q and P is an irreducible polynomial over F of degree n , then $F[x]/(P)$ is a finite field of order q^n .

Example 15.5.2. $\mathbb{Z}/2\mathbb{Z}[x]/(x^2 + x + 1)$ is a field of order $2^2 = 4$. Indeed, $x^2 + x + 1$ has no roots in $\mathbb{Z}/2\mathbb{Z}$, and is consequently irreducible.

§Field Extensions

Remark 15.5.3. Let $R \xrightarrow{f} S$ be a ring homomorphism for a commutative rings R, S and let

$$P = a_0 + a_1x + \dots + a_nx^n \in R[x] \quad (15.5.5)$$

then for all $\alpha \in R$

$$f(P(\alpha)) = P'(f(\alpha)) \quad (15.5.6)$$

where $P' = f(a_0) + f(a_1)x + \dots + f(a_n)x^n \in S[x]$. Indeed,

$$\begin{aligned} f(P(\alpha)) &= f(a_0 + a_1\alpha + \dots + a_n\alpha^n) \\ &= f(a_0) + f(a_1)f(\alpha) + \dots + f(a_n)f(\alpha)^n \\ &= P'(f(\alpha)) \end{aligned}$$

Definition 15.5.1. A ring homomorphism $F \xrightarrow{\iota} F'$ for F, F' fields is called a **field extension** and F' is called an **extension field** of F . Note that ι must be injective since F is a field and it is assumed to be a ring homomorphism, so one often identifies F with the isomorphic subfield $\iota(F)$ to F in F' .

Theorem 15.5.4 (Kronecker's Theorem). If F is a field and $P \in F[x]$, $\deg(P) > 0$, then there is an extension field of F in which P has a root.

Proof. By the Unique Factorization Theorem we have that $P = aP_1 \dots P_n$ for a constant $a \neq 0$ in F and monic irreducible polynomials P_i , for all i . Then we know that $F' = F[x]/(P_1)$ is a field since P_1 is irreducible. Moreover, P has a root ($\alpha = x + (P_1) \in F'$) in F' , where we see P as a polynomial with coefficients in F' via the embedding

$$\begin{aligned} F &\xrightarrow{\iota} F' \\ a &\mapsto a + (P_1) \end{aligned} \quad (15.5.7)$$

which is a field extension. Then, because $P_1(\alpha) = 0$ in $F'[x]$ and $P = P_1 Q$ for $Q = aP_2 \dots P_n$, we have that $P(\alpha) = 0$ in $F'[x]$. Hence, α is a root of P in the extension field F' . ■

15.6.0 §GCD of Polynomials

Definition 15.6.1 (GCD). We have seen that for $f, g \in F[x]$, F a field, if $f \neq 0$ or $g \neq 0$ and d is the monic generator of

$$(f, g) = \{Pf + Qg \mid P, Q \in F[x]\} \quad (15.6.1)$$

Then

1. d is monic
2. $d \mid f$ and $d \mid g$
3. If $P \mid f$ and $P \mid g$, then $P \mid d$ ($\forall P \in F[x]$)

Remark 15.6.1. If $d' \in F[x]$ satisfies 1.-2.-3. above, then d' is the monic generator of (f, g)

Proof. Let $d = \gcd(f, g)$ be the monic generator of (f, g) . By 1. $(d') \supseteq (f, g) = (d)$ since $d' \mid f$ and $d' \mid g$. Thus, $d' \mid d$. By 2., since $d \mid f$ and $d \mid g$, $d \mid d'$. By the lemma below, $d = d'$ ■

Lemma 15.6.1. For F a field, $f, g \in F[x]$, and f, g monic, if $f \mid g$ and $g \mid f$, then $f = g$.

Proof. If $f \mid g$ and $g \mid f$ then $g = Qf$ and $f = Pg$ for some $P, Q \in F[x]$. Thus $f = Pg = PQf$, so $(1 - PQ)f = 0$, where $f \neq 0$ and $F[x]$ is an integral domain, so $1 = PQ$. Hence, $P, Q \in F[x]^\times = F^\times = F \setminus \{0\}$. Then $g = af$ for $a \in F \setminus \{0\}$. Since f is monic, the leading coefficient of g is a . But, g is monic as well, so $a = 1$ and $f = g$. ■

Claim 15.6.2 (Greatest Common Factor). *Let $f, g \in F[x]$, $\deg(f), \deg(g) \geq 1$. Let $f = aP_1 \dots P_n$, $g = bQ_1 \dots Q_m$ be their unique factorization into a constant times a product of monic irreducible polynomials. Let $h = P_{j_1} P_{j_2} \dots P_{j_l}$ be the greatest common factor (set $h = 1$ if they don't have a common monic irreducible factor). Then $\gcd(f, g) = h$, as h satisfies 1.-2.-3. from the definition.*

Proof. By definition h is monic and divides both f and g . Now, let $P \in F[x]$ such that $P \mid f$ and $P \mid g$. Then there exists $Q, H \in F[x]$ such that $f = PQ$, $g = PH$. If P is constant then $P \mid h$ automatically. If $\deg(P) \geq 1$ then by the unique factorization theorem $P = cU_1 \dots U_t$ for some constant c and irreducible monic polynomials U_i , for all $i \in \{1, \dots, t\}$. Similarly **To be continued** ■

Example 15.6.1. Let $F = \mathbb{Q}$, $f = 10(x-1)(x-2)^2(x-3)^2$, and $g = \frac{1}{11}(x-1)^3(x-2)^2(x-3)$. Then

$$\gcd(f, g) = (x-1)(x-2)^2(x-3)$$

Part III

Field Theory

Chapter 16

§§Basic Definitions and Examples: Fields

Definition 16.0.1. A field is a commutative division ring.

Definition 16.0.2. The characteristic of a field F , denoted $\text{char}(F)$, is the smallest $p \in \mathbb{N} = \{1, 2, \dots\}$ such that $p \cdot 1_F = 0_F$ if it exists, and $\text{char}(F) = 0$, otherwise.

From here on out we will denote the multiplicative identity by 1 and the additive identity by 0 for any field.

Remark 16.0.1. If F is a field and $\text{char}(F) = p$ for $p \neq 0$, then p is a prime.

Proof. Suppose that F . If $\text{char}(F) = 0$ then we are done, so suppose $\text{char}(F) = n$ for $n \geq 1$. If $n = 1$ then we have that $1 = 1 \cdot 1 = 0$, so $F = \{0\}$. But, since F is a field $1 \neq 0$, so we have a contradiction. Thus, $n > 1$. Towards a contradiction suppose that $n = a \cdot b$ is composite, where $1 < a, b < n$. Then $0 = n \cdot 1 = (a \cdot 1) \cdot (b \cdot 1)$. Thus, since F is a field either $a \cdot 1 = 0$ or $b \cdot 1 = 0$. But, $a, b < n$ and by assumption $\text{char}(F) = n$ so n is the minimal positive integer for which $n \cdot 1 = 0$, so we obtain a contradiction. Therefore, n must be prime, as claimed. Moreover, suppose $m \cdot 1 = 0$ for some $m \neq 0$. Then let $m = nq + r$ be the division of m by n with remainder $0 \leq r < n$. Then we have that $0 = m \cdot 1 = nq \cdot 1 + r \cdot 1 = r \cdot 1$, but since $r < n$ we must have that $r = 0$ by minimality of n . Hence, $n|m$. ■

Example 16.0.1.

1. $\mathbb{R}, \mathbb{Q}, \mathbb{C}$, fields of characteristic zero.
2. $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ the finite field of p elements.
3. For $\mathbb{F}_p[x]$ polynomials, we have $\mathbb{F}_p(x)$ the field of rational functions with coefficients from \mathbb{F}_p . This is isomorphic to the quotient or fraction field of the integral domain $\mathbb{F}_p[x]$.

Remark 16.0.2. For a field F , we have a unique ring homomorphism $\varphi : \mathbb{Z} \rightarrow F$ defined by $\varphi(n) := n \cdot 1 = \underbrace{1 + 1 + \dots + 1}_{n \text{ times}}$ for $n \in \mathbb{N}$, and $\varphi(-n) = -\varphi(n)$. Note that $\ker(\varphi) \subseteq \mathbb{Z}$ is an ideal, so $\ker(\varphi) = n\mathbb{Z}$ for $n \in \mathbb{N} \cup \{0\}$. Moreover, $\mathbb{Z}/\ker(\varphi) \cong \varphi(\mathbb{Z})$, which is a subring of F .

since φ is a ring homomorphism. Moreover, $\varphi(\mathbb{Z})$ is isomorphic to \mathbb{Z} if $\ker(\varphi) = (0)$, and $\varphi(\mathbb{Z})$ is isomorphic to $\mathbb{Z}/p\mathbb{Z}$ for p a prime if $\ker(\varphi) = p\mathbb{Z}$.

Thus, each field F has a subring which is isomorphic to either \mathbb{Z} or $\mathbb{Z}/p\mathbb{Z}$. By the Field of Fractions technique we have that F has a subfield isomorphic to either \mathbb{Q} or \mathbb{F}_p .

Proof. First, suppose $\text{char}(F) = 0$. Then since $\varphi(n) = n \cdot \varphi(1) = n \cdot 1$, we have that $n \cdot 1 \neq 0$ for all $n \neq 0$. Hence, $\ker(\varphi) = (0)$, and \mathbb{Z} is isomorphic to a subring of F . Now, suppose $\text{char}(F) = p$ for some prime p . Then, we have that $\varphi(pq) = pq\varphi(1) = pq \cdot 1 = 0$ for all $pq \in p\mathbb{Z}$, so $p\mathbb{Z} \subseteq \ker(\varphi)$. Next, let $m \in \ker(\varphi)$. Then $m \cdot 1 = 0$, which implies $p|m$. Hence, $m \in p\mathbb{Z}$ so $\ker(\varphi) = p\mathbb{Z}$. Thus, we conclude that F has a subring isomorphic to $\mathbb{Z}/p\mathbb{Z}$. ■

Recall 16.0.3. Reminder that the only ideals in a field F are either (0) or F .

Definition 16.0.3. The prime subfield of a field F is generated by 1_F and is isomorphic to either \mathbb{Q} or \mathbb{F}_p for some prime p .

Chapter 17

§§Field Extensions

17.1.0 §Initial Definitions and Examples

Definition 17.1.1. If K is a field containing the subfield F , then K is said to be an extension field of F denoted K/F .

$$\begin{array}{c} K \\ | \\ F \end{array}$$

We call F , the field being extended, the base field.

Remark 17.1.1. Suppose K/F . Let $x, y \in K$ and $c_1, c_2 \in F$. Then we have

1. $c_1(x + y) = c_1x + c_1y$
2. $c_1(c_2x) = (c_1c_2)x$
3. $(c_1 + c_2)x = c_1x + c_2x$
4. $1 \cdot x = x$

along with the fact that K is an abelian group over $+$ as it is a field, so K is an F -vector space.

Note that what the “base field” or “extension field” is dependent on the context. For instance \mathbb{R} is the base field in the context of:

$$\begin{array}{c} \mathbb{C} \\ | \\ \mathbb{R} \end{array}$$

while \mathbb{R} is the extension field in the context of:

$$\begin{array}{c} \mathbb{R} \\ | \\ \mathbb{Q} \end{array}$$

Definition 17.1.2. The degree of K over F is $[K : F] = \dim_F(K)$. If $\dim_F(K)$ is finite then K is said to be a **finite extension** of F . If $\dim_F(K)$ is infinite then K is said to not be a finite extension of F (or an **infinite extension**).

Theorem 17.1.1. Let F be a field and $p(x) \in F[x]$ an **irreducible polynomial**. Then there exists a field K containing $F' \cong F$, in which $p(x)$ has a root; that is $p(\alpha) = 0$ for some $\alpha \in K$, where p is now envisioned as the corresponding polynomial in F' .

Proof. Let F be a field and $p(x) \in F[x]$ an irreducible polynomial. Then the ideal generated by $p(x)$, $(p(x))$, is maximal in $F[x]$ so $F[x]/(p(x)) = K$ is a field. Consider $p \in K[X]$ such that for $p(x) = a_n x^n + \dots + a_0$ in $F[x]$, we have

$$p = (a_n + (p(x)))X^n + \dots + (a_0 + (p(x)))$$

Moreover, for $\alpha = x + (p(x))$ in K , we have that

$$\begin{aligned} p(\alpha) &= (a_n + (p(x)))\alpha^n + \dots + (a_0 + (p(x))) \\ &= (a_n x^n + (p(x))) + \dots + (a_0 + (p(x))) \\ &= (a_n x^n + \dots + a_0) + (p(x)) \\ &= p(x) + (p(x)) \\ &= 0 + (p(x)) \end{aligned}$$

so α is a root of p in $K[x]$. Finally, note that the map $\varphi : F \rightarrow K$ sending $\varphi(a) = a + (p(x))$ is a ring monomorphism since F is a field, and thus restricting the codomain to $\varphi(F) = F'$, we have that $F \cong F'$, a subfield of K , as desired. ■

Theorem 17.1.2. Let $p(x) \in F[x]$ for F a field, irreducible of degree n over F , and $K = F[x]/p(x)$. Let $\theta = x \bmod p(x)$. Then $1, \theta, \theta^2, \dots, \theta^{n-1}$ is a basis for K over F . That is,

$$K = \text{span}(1, \theta, \theta^2, \dots, \theta^{n-1}) = \{a_0 + a_1\theta + \dots + a_{n-1}\theta^{n-1} : a_0, \dots, a_{n-1} \in F\} \quad (17.1.1)$$

so $\dim_F(K) = n = [K : F]$.

Proof. Let $f(x) \in F[x]$. Then since F is a field we have by the Division Algorithm that there exists $P, Q \in F[x]$ such that $f = pQ + P$, where either $P = 0$ or $\deg(P) < \deg(p)$. Then $f = P \bmod p(x)$, where $P = \sum_{i=0}^{n-1} a_i x^i$ since $\deg(p) = n$. Then

$$f = P \bmod p(x) = \sum_{i=0}^{n-1} a_i \theta^i \in \text{span}(1, \theta, \theta^2, \dots, \theta^{n-1})$$

so we conclude that

$$K = \text{span}(1, \theta, \theta^2, \dots, \theta^{n-1})$$

Now, suppose there existed a linear dependence $\sum_{i=0}^{n-1} c_i \theta^i = 0$ in K . Then by definition $\sum_{i=0}^{n-1} c_i x^i \in (p(x))$, so $p(x) \mid \sum_{i=0}^{n-1} c_i x^i$. But, $\deg(p(x)) = n$ while $\deg\left(\sum_{i=0}^{n-1} c_i x^i\right) \leq n-1$ or $\sum_{i=0}^{n-1} c_i x^i = 0$, so we must have that $c_i = 0$ for all i . Thus, $\{1, \theta, \theta^2, \dots, \theta^{n-1}\}$ is indeed a basis for K . ■

Example 17.1.1. Consider $\mathbb{R}[x]/(x^2 + 1) \cong \{a + b\theta : a, b \in \mathbb{R}\}$ by the previous Theorem. Moreover, in $\mathbb{R}[x]/(x^2 + 1) = \mathbb{R}(\theta)$, for $p = x^2 + 1$ $p(\theta) = 0$, so $\theta^2 = -1$.

Example 17.1.2. In general, if $\theta \in K$ is a root of the irreducible polynomial

$$p(x) = \sum_{i=0}^n p_i x^i$$

we can compute $\theta^{-1} \in K$ from

$$\theta(p_n \theta^{n-1} + p_{n-1} \theta^{n-2} + \dots + p_1) = -p_0$$

namely

$$\theta^{-1} = -p_0^{-1}(p_n \theta^{n-1} + p_{n-1} \theta^{n-2} + \dots + p_1)$$

where $p_0 \neq 0$ since $p(x)$ is irreducible.

Definition 17.1.3. If K is an extension field of F containing α, β, \dots . Then the smallest field containing both α, β, \dots and F is denoted $F(\alpha, \beta, \dots)$. When we just adjoin α , then $F(\alpha)$ is said to be a simple extension of F with primitive element α .

Theorem 17.1.3. Let F be a field and $p(x) \in F[x]$ and irreducible polynomial. Suppose K is an extension field of F containing the root α of $p(x)$. That is $p(\alpha) = 0$ in K . Let $F(\alpha)$ denote the subfield of K generated by F and α . Then $F(\alpha) \cong F[x]/(p(x))$.

Proof. Let $\varphi : F[x] \rightarrow F(\alpha) \subseteq K$, defined by $\varphi(f(x)) = f(\alpha)$. Indeed, if $f(x) = c_0 + c_1 x + \dots + c_n x^n$, then $\varphi(f(x)) = c_0 + c_1 \alpha + \dots + c_n \alpha^n \in F(\alpha)$. Moreover, φ is ring homomorphism, since the evaluation map is a ring homomorphism. Then we have that

$$\ker(\varphi) = \{g(x) \in F[x] : g(\alpha) = 0\}$$

It follows that for all $h(x)p(x) \in (p(x))$, $\varphi(h(x)p(x)) = h(\alpha)p(\alpha) = 0$ in $F(\alpha)$, so $(p(x)) \subseteq \ker(\varphi)$. Since $(p(x)) \subseteq \ker(\varphi)$ we have by the Factor Theorem a ring homomorphism

$$\bar{\varphi} : F[x]/(p(x)) \rightarrow F(\alpha)$$

But, since $p(x)$ is irreducible in $F[x]$, $F[x]/(p(x))$ so $\bar{\varphi}$ must be injective. Then, we have that $F[x]/(p(x))$ is isomorphic to a subfield of $F(\alpha)$ containing F and α . But then by definition $\bar{\varphi}(F[x]/(p(x)))$ contains $F(\alpha)$, so φ must be surjective and hence an isomorphism since $F(\alpha)$ is the smallest such field. Thus $F[x]/(p(x)) \cong F(\alpha)$. ■

Example 17.1.3. Consider $\mathbb{Q}[x]/(x^2 - 2)$. Then by our previous theorem $\mathbb{Q}[x]/(x^2 - 2) \cong \mathbb{Q}(\sqrt{2})$, as $(\sqrt{2})^2 - 2 = 0$, so $p(x) = x^2 - 2$ has $\alpha = \sqrt{2}$ as a root. Note $\beta = -\sqrt{2}$ is also a root, so $\mathbb{Q}(\sqrt{2}) \cong \mathbb{Q}(-\sqrt{2})$.

Example 17.1.4. Consider $\mathbb{Q}[x]/(x^3 - 2)$, so $\mathbb{Q}[x]/(x^3 - 2) \cong \mathbb{Q}(\sqrt[3]{2})$ has $p(x) = x^3 - 2$ has $p(\alpha) = 0$ for $\alpha = \sqrt[3]{2}$. Then, for $\omega = e^{2\pi i/3}$, we have roots $\omega\alpha$ and $\omega^2\alpha$ of $p(x)$. Then we have by our previous theorem $\mathbb{Q}(\sqrt[3]{2}) \cong \mathbb{Q}(\omega\sqrt[3]{2}) \cong \mathbb{Q}(\omega^2\sqrt[3]{2})$. By 17.1.3, these fields are algebraically indistinguishable.

Remark 17.1.2. Suppose $\phi : F \xrightarrow{\sim} F'$ is an isomorphism between fields F and F' . Then we can extend ϕ to the isomorphism

$$\begin{aligned} \phi' : F[x] &\xrightarrow{\sim} F'[x] \\ \phi'(c_0 + c_1x + \dots + c_nx^n) &= \phi(c_0) + \phi(c_1)x + \dots + \phi(c_n)x^n \end{aligned} \quad (17.1.2)$$

If $p(x)$ is irreducible over F , then $(p(x))$ is maximal in $F[x]$, so $\phi'((p(x))) = (p'(x))$ is maximal, which implies $p'(x)$ is irreducible in $F'[x]$. Then the theorem follows:

Theorem 17.1.4. Let α be a root of $p(x)$ and β a root of $p'(x) = \phi'(p(x))$ in the extension fields $F(\alpha)$ and $F'(\beta)$. Then there is an isomorphism

$$\sigma : F(\alpha) \xrightarrow{\sim} F'(\beta) \quad (17.1.3)$$

where $\sigma|_F = \phi$ and $\alpha \mapsto \beta$. Then the diagram

$$\begin{array}{ccc} F(\alpha) & \xrightarrow{\sigma} & F(\beta) \\ \downarrow & & \downarrow \\ F & \xrightarrow{\phi} & F' \end{array}$$

commutes.

Proof. Let $p(x) \in F[x]$ be irreducible. Then as ϕ is an isomorphism, so is ϕ' and hence we have that $\phi'((p(x))) = (\phi'(p(x)))$ is a maximal ideal. Indeed, let $(\phi'(p(x))) \subseteq J$ for J an ideal in $F'[x]$. Then $\phi'^{-1}(J)$ is an ideal in $F[x]$ containing $(p(x))$ by the correspondence theorem. Thus, $(p(x)) = \phi'^{-1}(J)$ or $F[x] = \phi'^{-1}(J)$, which implies that $\phi'((p(x))) = J$ or $F'[x] = J$ by bijectivity of the correspondence. Then, I claim that $\phi' : F[x] \xrightarrow{\sim} F'[x]$ induces an isomorphism $\Phi : F[x]/(p(x)) \xrightarrow{\sim} F'[x]/\phi'((p(x)))$. Indeed, $\Phi(f(x) + (p(x))) := \phi'(f(x)) + (\phi'(p(x)))$ is a well defined isomorphism (since ϕ' is an isomorphism). Then since $F[x]/(p(x)) \cong F(\alpha)$ and $F'[x]/\phi'((p(x))) \cong F'(\beta)$, by 17.1.3, we have that $F(\alpha) \cong F'(\beta)$. ■

17.2.0 §Algebraic Extensions

Definition 17.2.1. $\alpha \in K$ over F is **algebraic** over F if α is a root of some nonzero polynomial $f(x) \in F[x]$. If α is not algebraic over F , α is said to be **transcendental** over F . The extension K/F is **algebraic** if and only if every element of K is algebraic.

Proposition 17.2.1. Let α be algebraic over F . Then there exists a unique monic irreducible polynomial $m_{\alpha,F}(x) \in F[x]$ which has α as a root. Furthermore, $m_{\alpha,F}(x)$ divides any $f(x) \in F[x]$ such that $f(\alpha) = 0$.

Proof. First, since α is algebraic over F , there exist $f(x) \in F[x]$ such that $f(\alpha) = 0$. Then, by well-ordering of \mathbb{N} , let $g(x) \in F[x]$ be a polynomial of minimal degree such that $g(\alpha) = 0$. Multiplying by a constant we may assume that $g(x)$ is monic. Suppose that $g(x)$ were reducible over F , then $g(x) = a(x)b(x)$ in F for $\deg(a(x)), \deg(b(x)) \geq 1$. But, then $g(\alpha) = a(\alpha)b(\alpha)$, since the evaluation map is a homomorphism, so as F is a field either $a(\alpha) = 0$ or $b(\alpha) = 0$, contradicting the minimality of g . Hence, g is irreducible. Now suppose $f(x) \in F[x]$ such that $f(\alpha) = 0$, and by the division algorithm obtain $P, Q \in F[x]$ such that $f = gP + Q$, with $\deg(Q) < \deg(g)$. But then $0 = f(\alpha) = g(\alpha)P(\alpha) + Q(\alpha) = Q(\alpha)$, so by minimality of the degree of g we must have that $Q(x) = 0$. Hence, $g(x) | f(x)$ as desired. Finally, by this result $g(x)$ would divide any other monic irreducible polynomial having α as a root, so $g(x)$ must be unique. ■

Definition 17.2.2. $m_{\alpha,F}(x)$ in the Proposition 17.2.1 is the minimal polynomial of α over F .

Example 17.2.1. $m_{\sqrt{2},\mathbb{Q}} = x^2 - 2$, but $m_{\sqrt{2},\mathbb{R}} = x - \sqrt{2}$.

Corollary 17.2.2. $\alpha \in F$ if and only if $m_{\alpha,F} = x - \alpha$.

Corollary 17.2.3. For a field extension L/F and α is algebraic over both F and L , then $m_{\alpha,L}(x)$ divides $m_{\alpha,F}(x)$ in $L[x]$.

Proof. Note that since $F \subseteq L$, we have by Proposition 17.2.1 that $m_{\alpha,L}(x) | m_{\alpha,F}(x)$ since α is a root of $m_{\alpha,F}(x)$, and $m_{\alpha,L}(x)$ is the minimal such polynomial in L . ■

Proposition 17.2.4. Let α be algebraic over F , $F(\alpha)$ generated by α and F . Then $F(\alpha) \cong F[x]/(m_{\alpha,F}(x))$, and $[F(\alpha) : F] = \deg(m_{\alpha,F}(x)) = \deg(\alpha)$.

Proof. Since $m_{\alpha,F}(x)$ is an irreducible polynomial over F with root α , we have by 17.1.3 that $F(\alpha) \cong F[x]/(m_{\alpha,F}(x))$. Then, by Theorem 17.1.2 we have that $[F(\alpha) : F] = [F[x]/(m_{\alpha,F}(x)) : F] = \deg(m_{\alpha,F}(x))$, as claimed. ■

Proposition 17.2.5. The element α is algebraic over F if and only if the simple extension $F(\alpha)/F$ is a finite extension.

\implies . By Proposition 17.2.4. [\Leftarrow] Suppose $F(\alpha)/F$ is finite, so $[F(\alpha) : F] = n$ for some $n \in \mathbb{N}$. Thus $\{1, \alpha, \alpha^2, \dots, \alpha^n\}$ must be linearly dependent over F . Thus, there exist $c_i \in F$, $0 \leq i \leq n$, not all zero such that

$$c_0 + c_1\alpha + \dots + c_n\alpha^n = 0$$

Thus, we have that $f(x) = c_0 + c_1x + \dots + c_nx^n$ is a non-zero polynomial in $F[x]$ which takes α as a root, so α is algebraic over F by definition. ■

Corollary 17.2.6. If the extension K/F is finite, then its algebraic.

Proof. Let $\alpha \in K$ so $F(\alpha)$ is a subfield of K , and $[F(\alpha) : F] \leq [K : F] = n$ for some $n \in \mathbb{N}$. Thus, by the previous proposition α is algebraic. ■

Example 17.2.2 (Quadratic Extensions over Fields of Characteristic not equal to 2). Let F be a field of characteristic $\neq 2$, and let K be an extension of F of degree 2, $[K : F] = 2$. Let α be any element of K not contained in F . By the previous corollary α satisfies an equation of degree at most 2 over F . Moreover, since $\alpha \notin F$, it cannot be of degree 1, so the minimal polynomial of α is a monic quadratic polynomial

$$m_{\alpha,F}(x) = x^2 + bx + c, b, c \in F$$

Since $F \subset F(\alpha) \subseteq K$ and $F(\alpha)$ is already a vector space over F of dimension 2, we have $K = F(\alpha)$. By the quadratic formula, which is valid for any field of characteristic $\neq 2$, the roots of this quadratic formula are

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4c}}{2}$$

Here $b^2 - 4c$ is not a square in F since α is not an element of F , and the symbol $\sqrt{b^2 - 4c}$ denotes a root of the equation $x^2 - (b^2 - 4c) = 0$ in K . Note that here there is no natural choice of one of the roots analogous to choosing the positive square root of 2 in \mathbb{R} - the roots are algebraically indistinguishable.

Now $F(\alpha) = F(\sqrt{b^2 - 4c})$ as follows: by the formula above, α is an element of the field on the right, hence $F(\alpha) \subseteq F(\sqrt{b^2 - 4c})$. Conversely, $\sqrt{b^2 - 4c} = \mp(b + 2\alpha)$ shows that $\sqrt{b^2 - 4c}$ is an element of $F(\alpha)$, which gives the reverse inclusion $F(\sqrt{b^2 - 4c}) \subseteq F(\alpha)$.

Corollary 17.2.7. *It follows that any extension K of F of degree 2 is of the form $F(\sqrt{D})$ where D is an element of F which is not a square in F , and conversely, every such extension is an extension of degree 2 of F . For this reason extions of degree 2 of a field F are called quadratic extensions of F .*

Theorem 17.2.8. *Let $F \subseteq K \subseteq L$ be fields. Then*

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

$$\begin{array}{c} L \\ \left| \begin{array}{c} [L:K] \end{array} \right. \\ [L:F] \quad K \\ \left| \begin{array}{c} [K:F] \end{array} \right. \\ F \end{array}$$

Proof. First, suppose that $[L : K] = m$ and $[K : F] = n$ are finite. Let $\alpha_1, \alpha_2, \dots, \alpha_m$ be a basis for L over K , and let $\beta_1, \beta_2, \dots, \beta_n$ be a basis for K over F . Then every element of $z \in L$ can be written as a linear combination

$$z = \sum_{i=1}^m a_i \alpha_i$$

where $a_i \in K$ for each i . Thus, there exist $b_{ij} \in F$ such that

$$a_i = \sum_{j=1}^n b_{ij} \beta_j$$

for all $1 \leq i \leq m$. Substituting we have that

$$z = \sum_{i=1}^m a_i \alpha_i = \sum_{i=1}^m \sum_{j=1}^n b_{ij} \beta_j \alpha_i$$

Hence,

$$z \in \text{span}(\beta_1 \alpha_1, \dots, \beta_1 \alpha_m, \beta_2 \alpha_1, \dots, \beta_2 \alpha_m, \dots, \beta_n \alpha_1, \dots, \beta_n \alpha_m)$$

so the $\beta_j \alpha_i$ are a spanning set for K . Now, suppose $\sum c_{ij} \alpha_i \beta_j = 0$ is a linear dependence over F . In particular, we have that

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

where $\sum_{j=1}^n c_{ij} \beta_j \in K$ for each i . Hence, by the linear independence of the α_i , we have for each i that

$$\sum_{j=1}^n c_{ij} \beta_j = 0$$

But then $c_{ij} \in F$ for all i, j , so by the independence of the β_j over F , we find that $c_{ij} = 0$ for all i and j . Hence, the set is linearly independent and consequently a basis for L . Therefore, we conclude that $[L : F] = mn = [L : K][K : F]$.

Now, if $[L : F]$ is infinite, then either $[L : K]$ is infinite or $[K : F]$ is infinite since otherwise by the previous argument $[L : F]$ would be finite, a contradiction.

If $[L : K]$ is infinite, there are infinitely many elements of L linearly independent over K , so as $F \subseteq K$ there are infinitely many elements of L linearly independent over F , so $[L : F]$ is infinite. If $[K : F]$ is infinite, there are infinitely many elements of K linearly independent over F . But as $L \supseteq K$, there are also infinitely many elements of L that are linearly independent over F . ■

Corollary 17.2.9. Suppose L/F is a finite extension and let K be a subfield of L containing F , $F \subseteq K \subseteq L$. Then $[K : F]$ divides $[L : F]$.

Proof. By Theorem 17.2.8 $[L : F] = [L : K][K : F]$, so indeed $[K : F] \mid [L : F]$. ■

Example 17.2.3. Let α be a real root of $x^3 - 3x - 1 = 0$, which exists between $0 < x < 2$ (exists by IVT). Then I claim $\sqrt{2} \notin \mathbb{Q}(\alpha)$. Recall that $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ as it is a quadratic extension. Now, note that by the Rational Roots Theorem $x^3 - 3x - 1$ is indeed irreducible over \mathbb{Q} , so $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3$. Hence, since $2 \nmid 3$, by the previous corollary we conclude that $\mathbb{Q}(\sqrt{2}) \not\subseteq \mathbb{Q}(\alpha)$.

Example 17.2.4. Consider $\alpha = \sqrt[6]{2}$ adjoined to \mathbb{Q} . Then by Eisenstein $m_{\alpha, \mathbb{Q}}(x) = x^6 - 2$. In particular, $[\mathbb{Q}(\sqrt[6]{2}), \mathbb{Q}] = 6$. Note that $\sqrt[6]{2}^3 = \sqrt{2}$, so $\sqrt{2} \in \mathbb{Q}(\sqrt[6]{2})$, and $\mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\sqrt[6]{2})$. Then, by the corollary $[\mathbb{Q}(\sqrt[6]{2}), \mathbb{Q}(\sqrt{2})] = 6/2 = 3$, so $\sqrt[6]{2}$ is of degree 3 over $\mathbb{Q}(\sqrt{2})$. Indeed, $m_{\sqrt[6]{2}, \mathbb{Q}(\sqrt{2})}(x) = x^3 - \sqrt{2}$ is a minimal polynomial by the corollary, so it is also irreducible over $\mathbb{Q}(\sqrt{2})$.

Definition 17.2.3. An extension K/F is finitely generated if there are elements $\alpha_1, \alpha_2, \dots, \alpha_k \in K$ such that $K = F(\alpha_1, \alpha_2, \dots, \alpha_k)$.

Lemma 17.2.10. $F(\alpha, \beta) = (F(\alpha))(\beta)$, that is, the field generated over F by α and β is the field generated by β over the field $F(\alpha)$ generated by α .

Proof. First, we note that $F(\alpha, \beta)$ contains F , α , and β . Thus, since $F(\alpha)$ is the smallest field containing F and α it contains $F(\alpha)$ and β . Thus, as $(F(\alpha))(\beta)$ is the smallest such field, $(F(\alpha))(\beta) \subseteq F(\alpha, \beta)$. Conversely, since $(F(\alpha))(\beta)$ contains F , α , and β , by minimality of $F(\alpha, \beta)$ we find that $F(\alpha, \beta) \subseteq (F(\alpha))(\beta)$. ■

By the lemma we have that

$$K = F(\alpha_1, \dots, \alpha_k) = (F(\alpha_1, \dots, \alpha_{k-1}))(\alpha_k)$$

and so by iterating, we see that K is obtained by taking the field F_1 generated over F by α_1 , then the field F_2 generated over F_1 by α_2 , and so on, with $F_k = K$. This gives a sequence of fields:

$$F = F_0 \subseteq F_1 \subseteq F_2 \subseteq \dots \subseteq F_k = K$$

where $F_{i+1} = F_i(\alpha_{i+1})$ for $i = 0, 1, \dots, k-1$. By the multiplicativity of extension degrees we see that

$$[K : F] = [F_k : F_{k-1}] \dots [F_2 : F_1][F_1 : F_0]$$

Theorem 17.2.11. The extension K/F is finite if and only if K is generated by a finite number of algebraic elements over F .

Proof. If K/F is finite of degree n , let $\alpha_1, \alpha_2, \dots, \alpha_n$ be a basis for K as a vector space over F . By Corollary 17.2.9 we have that $[F(\alpha_i) : F]$ divides $[K : F] = n$ for all $i = 1, 2, \dots, n$, so they are finite and hence each α_i is algebraic over F . Thus, K is generated by a finite number of algebraic elements over F . Conversely, if K was generated by a finite number of algebraic elements over F , say β_1, \dots, β_n with $[F(\beta_i) : F] = m_i$, we have that

$$[K : F] = [F_n : F_{n-1}] \dots [F_2 : F_1][F_1 : F_0] \leq [F(\beta_n) : F] \dots [F(\beta_2) : F][F(\beta_1) : F] = m_1 m_2 \dots m_n$$

so K is a finite extension of F . ■

Corollary 17.2.12. Suppose α and β are algebraic over F . Then $\alpha \pm \beta, \alpha\beta, \alpha/\beta$ (for $\beta \neq 0$) are all algebraic.

Proof. By the theorem $F(\alpha, \beta)$ is a finite extension over F , and it contains all of these elements (if they exist). Moreover, since $F(\alpha, \beta)$ is finite all of its elements are algebraic by Corollary 17.2.6, so these elements are algebraic. ■

Corollary 17.2.13. Let L/F be an arbitrary extension. Then the collection of elements of L that are algebraic over F form a subfield of K .

Proof. Let \mathcal{A} be such a subset. Then by the previous corollary \mathcal{A} is closed under addition, subtraction, multiplication, and inverses, so it is indeed a subfield of L . ■

Example 17.2.5 (Algebraic Numbers). Consider the extension \mathbb{C}/\mathbb{Q} , and let $\overline{\mathbb{Q}}$ denote the subfield of all elements of \mathbb{C} that are algebraic over \mathbb{Q} . In particular, the elements $\sqrt[n]{2}$ (positive n th roots of 2 in \mathbb{R}) are all elements of $\overline{\mathbb{Q}}$, so that $[\overline{\mathbb{Q}}, \mathbb{Q}] \geq n$ for all integers $n > 1$. Hence $\overline{\mathbb{Q}}$ is an infinite algebraic extension of \mathbb{Q} , called the algebraic numbers.

Consider $\overline{\mathbb{Q}} \cap \mathbb{R}$, the subfield of \mathbb{R} consisting of elements algebraic over $\overline{\mathbb{Q}}$. The field $\overline{\mathbb{Q}}$ is countable. The number of polynomials in $\mathbb{Q}[x]$ of any given degree n is therefore also countable, being the union of a finite number of countable sets. Since these polynomials have at most n roots in \mathbb{R} , the number of algebraic elements of \mathbb{R} of degree n is countable, being contained in a finite union of countable sets. Finally, the collection of all algebraic elements in \mathbb{R} is the countable union (indexed by n) of countable sets, hence is countable. Since \mathbb{R} is uncountable, it follows that there exist (in fact many) elements of \mathbb{R} which are not algebraic, i.e. are transcendental, over \mathbb{Q} .

Theorem 17.2.14. *If K is algebraic over F and L is algebraic over K , then L is algebraic over F .*

Proof. Let $\alpha \in L$. Then there exists $f(x) = \sum_{i=0}^n a_i x^i \in K[x]$ such that $f(\alpha) = 0$ in L . Consider the field $F(\alpha, a_0, a_1, \dots, a_n)$ generated over F by α and the coefficients of this polynomial. Since K/F is algebraic, the elements a_0, a_1, \dots, a_n are algebraic over F , so the extension $F(a_0, a_1, \dots, a_n)/F$ is finite by Theorem 17.2.11. By the equation above we know that $\sum_{i=0}^n a_i \alpha^i = 0$, so α generates an extension of this field of degree at most n , since its minimal polynomial over this field is a divisor of the polynomial above. Therefore

$$[F(\alpha, a_0, a_1, \dots, a_n) : F] = [F(\alpha, a_0, a_1, \dots, a_n) : F(a_0, a_1, \dots, a_n)][F(a_0, a_1, \dots, a_n) : F]$$

is also finite and by Theorem 17.2.11 $F(\alpha, a_0, a_1, \dots, a_n)/F$ is an algebraic extension. In particular, the element α is algebraic over F , which proves that L is algebraic over F . ■

Definition 17.2.4. Let K_1 and K_2 be two subfields of a field K . Then the **composite field** of K_1 and K_2 , denoted $K_1 K_2$, is the smallest subfield of K containing both K_1 and K_2 . Similarly, the composite of any collection of subfields of K is the smallest subfield containing all the subfields.

Proposition 17.2.15. Let K_1 and K_2 be two finite extensions of a field F contained in K . Then

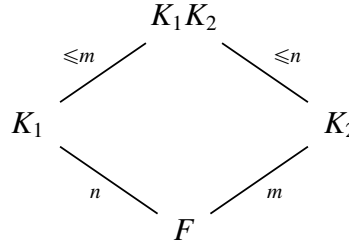
$$[K_1 K_2 : F] \leq [K_1 : F][K_2 : F]$$

with equality if and only if an F -basis for one of the fields remains linearly independent over the other field. If $\alpha_1, \alpha_2, \dots, \alpha_n$ and $\beta_1, \beta_2, \dots, \beta_m$ are bases for K_1 and K_2 over F , respectively, then the elements $\alpha_i \beta_j$ for $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$ span $K_1 K_2$ over F .

Proof. Note that $K_1 = F(\alpha_1, \alpha_2, \dots, \alpha_n)$ and $K_2 = F(\beta_1, \beta_2, \dots, \beta_m)$. Then observe that $K_1 K_2$ contains $F, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m$ so $F(\alpha_i, \beta_j | 1 \leq i \leq n, 1 \leq j \leq m) \subseteq K_1 K_2$, and similarly $F(\alpha_i, \beta_j | 1 \leq i \leq n, 1 \leq j \leq m)$ contains both K_1 and K_2 by their minimality, so $K_1 K_2 \subseteq F(\alpha_i, \beta_j | 1 \leq i \leq n, 1 \leq j \leq m)$, proving equality. Now, this is equal to $K_1(\beta_1, \dots, \beta_m)$,

so $[K_1K_2 : K_1] \leq m = [K_2 : F]$, with equality if and only if these elements are linearly independent over K_1 . Since $[K_1K_2 : F] = [K_1K_2 : K_1][K_1 : F] \leq [K_2 : F][K_1 : F]$, the proposition is proven. ■

By this proposition, and its proof, we have the following diagram:



Corollary 17.2.16. Suppose that $[K_1 : F] = n$, $[K_2 : F] = m$ in the previous proposition, where n and m are relatively prime: $\gcd(n, m) = 1$. Then $[K_1K_2 : F] = [K_1 : F][K_2 : F] = nm$.

Proof. First observe that $[K_1K_2 : F] = [K_1K_2 : K_1][K_1 : F]$ and $[K_1K_2 : F] = [K_1K_2 : K_2][K_2 : F]$, so n and m divide $[K_1K_2 : F]$. Let $l = \text{lcm}(m, n)$. Then note that l must divide $[K_1K_2 : F]$. Now, since $\gcd(n, m) = 1$, we have that nm divides $[K_1K_2 : F]$, so $nm \leq [K_1K_2 : F] \leq nm$. Thus, $[K_1K_2 : F] = nm$, as claimed. ■

17.3.0 §Classical Straightedge and Compass Constructions

The following three geometric problems posed by the Greeks can now be shown to not be possible:

- I. (Doubling the Cube) Is it possible using only straightedge and compass to construct a cube with precisely twice the volume of a given cube?
- II. (Trisecting the Angle) Is it possible using only straightedge and compass to trisect any given angle θ ?
- III. (Squaring the Circle) Is it possible using only straightedge and compass to construct a square whose area is precisely the area of a given circle?

Let 1 denote some fixed unit distance. Then any distance is determined by its length $a \in \mathbb{R}$. We construct the usual cartesian plane \mathbb{R}^2 , and view all of our constructions as occurring in \mathbb{R}^2 .

Definition 17.3.1. A point $(x, y) \in \mathbb{R}^2$ is **constructible** starting with the given distance 1 if and only if its coordinates x and y are constructible elements of \mathbb{R} . Elements of \mathbb{R} are called **constructible** starting with the given distance 1 if and only if they can be obtained by straightedge and compass constructions from 1.

Each straightedge and compass construction consists of a series of operations of the following four types:

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

Theorem 17.3.1. *Given lengths a and b , you can construct $a + b, a - b, ab, a/b$, given $b \neq 0$.*

$a + b$ and $a - b$ can be attained by parallel lines. Then $ab, a/b$ can be attained by similar triangles. Moreover, using an upper half circle, with diameter $a + 1$, we can attain \sqrt{a} . These facts imply that the set of constructible numbers is a subfield of \mathbb{R} . Moreover, this field is an extension of the rationals.

It can be shown that for any field F , the straightedge constructions are closed in F , while the compass constructions at most bring F to a quadratic extension of itself. Since quadratic extensions have degree 2, and extension degrees are multiplicative, it follows that if $\alpha \in \mathbb{R}$ is obtained from elements in F by a finite series of straightedge and compass constructions, then α is an element of an extension K of F of degree a power of 2: $[K : F]2^m$ for some m . Since $[F(\alpha) : F]$ divides this extension, it must also be a power of 2.

Proposition 17.3.2. *If the element $\alpha \in \mathbb{R}$ is obtained from a field $F \subset \mathbb{R}$ by a series of compass and straightedge constructions, then $[F(\alpha) : F] = 2^k$ for some integer $k \geq 0$.*

Theorem 17.3.3. *None of the classical Greek problems: (I) Doubling the Cube, (II) Trisecting an Angle, (III) Squaring the Circle, is possible.*

Sketch. (I) Doubling the cube amounts to constructing $\sqrt[3]{2}$ in the reals starting with the unit 1. Since $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ is not a power of 2, this is impossible.

(II) Trisecting the angle is possible for some angles, but not all. If an angle θ can be constructed, then so can $\cos \theta$ and $\sin \theta$. In attempt to trisect 60 degrees, we see that an extension of \mathbb{Q} by $\cos 20^\circ$ has degree 3, and is therefore not constructible, and consequently 20 degrees is not constructible.

(III) Squaring the circle is equivalent to determining if π is constructible. In fact, π is transcendental, so $[\mathbb{Q}(\pi) : \mathbb{Q}]$ is infinite and hence π is not constructible and we cannot square the circle. ■

17.4.0 §Splitting Fields

Let F be a field, and $f(x) \in F[x]$ with $\deg(f(x)) = n$. Then we have shown there exists E_1/F such that there is $\alpha_1 \in E_1$ for which α_1 is a root of $f(x)$ and $f(x) = (x - \alpha_1)g_1(x)$ for some $g_1(x) \in E_1[x]$, where $[E_1 : F] \leq n$.

Example 17.4.1. $f(x) = x^2 - 3x + 2 \in \mathbb{Q}[x]$. Then $f(x) = (x - 1)(x - 2)$, so \mathbb{Q} is its own extension field.

Next, we can consider $g_1(x) \in E_1[x]$, and find an extension E_2/E_1 in which there exists $\alpha_2 \in E_2$ such that $g_1(x) = (x - \alpha_1)g_2(x)$ for some $g_2(x) \in E_2[x]$, and $[E_2 : E_1] \leq n - 1$, as $\deg(g_1(x)) = n - 1$. It follows that $f(x) = (x - \alpha_1)(x - \alpha_2)g_2(x)$ in $E_2[x]$. In principle we can proceed inductively and obtain E/F such that $f(x) = A(x - \alpha_1)(x - \alpha_2)\dots(x - \alpha_n)$, where $\alpha_1, \alpha_2, \dots, \alpha_n \in E$ and

$$[E : F] = [E_n : E_{n-1}][E_{n-1} : E_{n-2}]\dots[E_2 : E_1][E_1 : F] \leq n!$$

Definition 17.4.1. The extension field K of F is called a **splitting field** for the polynomial $f(x) \in F[x]$ if $f(x)$ factors completely into linear factors (or **splits completely**) in $K[x]$ and $f(x)$ does not factor completely into linear factors over any proper subfield of K containing F .

Note that if $f(x)$ is of degree n , then $f(x)$ has at most n roots in F and has precisely n roots in F if and only if $f(x)$ splits completely in $F[x]$.

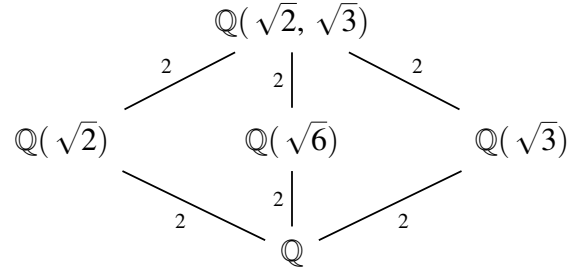
Theorem 17.4.1. For any field F , if $f(x) \in F[x]$ then there exists an extension K of F which is a splitting field for $f(x)$.

Proof. We first show that there exists an extension E of F over which $f(x)$ splits completely into linear factors by induction on the degree n of $f(x)$. If $n = 1$, then take $E = F$. Suppose now that $n > 1$. If the irreducible factors of $f(x)$ over F are all of degree 1, then F is the splitting field for $f(x)$ and we may take $E = F$. Otherwise, at least one of the irreducible factors, say $p(x)$ of $f(x)$ in $F[x]$ is of degree at least 2. Then by Theorem 17.1.1 there exists an extension E_1 of F containing a root α of $p(x)$. Over $E_1[x]$ $f(x)$ has the linear factor $x - \alpha$. The degree of the remaining factor $f_1(x)$ is $n - 1$, so by the induction hypothesis there is an extension E of E_1 containing all of the roots of $f_1(x)$. Since $\alpha \in E$, E is an extension of F containing all of the roots of $f(x)$. Now, let K be the intersection of all the subfields of E containing F which also contain all roots of $f(x)$. Then K is a field which is a splitting field for $f(x)$. ■

Definition 17.4.2. If K is an algebraic extension of F which is the splitting field over F for a collection of polynomials $f(x) \in F[x]$ then K is called a **normal extension** of F .

Example 17.4.2. $x^2 - 2 \in \mathbb{Q}[x]$ splits in $\mathbb{Q}(\sqrt{2})$, with $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ and $\pm\sqrt{2} \in \mathbb{Q}(\sqrt{2})$.

Example 17.4.3. Consider $(x^2 - 2)(x^2 - 3) \in \mathbb{Q}$. Then $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is the splitting field, and we have the field diagram:

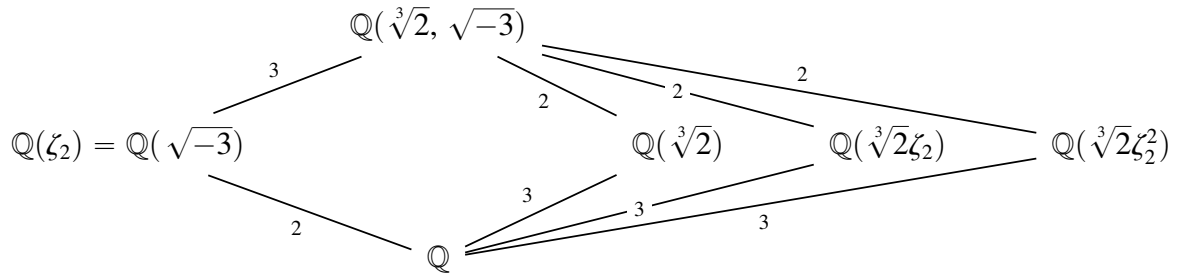


Moreover, $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$.

Example 17.4.4. Consider $x^3 - 2$ over \mathbb{Q} . Then for $\alpha = \sqrt[3]{2} \in \mathbb{R}$, $\alpha^3 - 2 = 0$. Complex roots are $\alpha, \alpha\zeta, \alpha\zeta^2$ for $\zeta = e^{\frac{2\pi i}{3}}$ the third root of unity. Then

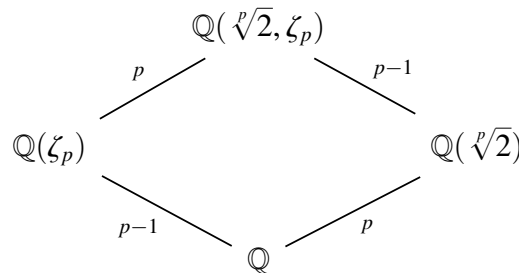
$$x^3 - 2 = (x - \alpha)(x - \alpha\zeta)(x - \alpha\zeta^2)$$

So the splitting field is $K = \mathbb{Q}(\alpha, \alpha\zeta, \alpha\zeta^2)$. Note $\alpha\zeta/\alpha = \zeta \in K$, so as $\zeta = \frac{-1+i\sqrt{3}}{2}$, $i\sqrt{3} = 2\zeta + 1 \in K$. Then we can also write $K = \mathbb{Q}(\sqrt[3]{2}, \sqrt{-3})$, so



Adjoining $\sqrt{-3}$ to $\mathbb{Q}(\sqrt[3]{2})$, $[K : \mathbb{Q}(\sqrt[3]{2})] = 2$, with minimal polynomial $x^2 + 3 \in \mathbb{Q}(\sqrt[3]{2})[x]$.

Example 17.4.5. Consider $x^p - 2 \in \mathbb{Q}[x]$. Then for $\zeta_p = e^{\frac{2\pi i}{p}}$, the p th root of unity, we have



Example 17.4.6. Consider $(x^4 - 1) = (x^2 + 1)(x^2 - 1) = (x - i)(x + i)(x - 1)(x + 1)$, for $x^4 - 1 \in \mathbb{Q}[x]$. Then $\mathbb{Q}(i)$ is a splitting field for $x^4 - 1$.

Conversely, $(x^4 + 1) = 0$ has roots $e^{\frac{\pi i}{4}}\zeta_4^k$, for $\zeta_4 = e^{\frac{2\pi i}{4}}$ the principal root of unity. Then we have that

$$(-1)^{1/4} = \{e^{\frac{\pi i}{4}}, e^{\frac{\pi i}{4}}\zeta_4, e^{\frac{\pi i}{4}}\zeta_4^2, e^{\frac{\pi i}{4}}\zeta_4^3\}$$

Then the splitting field is $\mathbb{Q}(i, \sqrt{2})$, and we have $[\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}] = 4$.

Proposition 17.4.2. *A splitting field of a polynomial of degree n over F is of degree at most $n!$ over F .*

Proof. (By induction on degree) ■

Cyclotomic Fields

Consider the splitting field of $x^n - 1$ over \mathbb{Q} . The roots of this are the roots of unity

$$z \in (1)^{1/n} = \{1, \zeta_n, \zeta_n^2, \dots, \zeta_n^{n-1}\}$$

for $\zeta_n = e^{\frac{2\pi i}{n}}$, a primitive generator. Note ζ_n^k also generates $(1)^{1/n}$ if and only if $\gcd(k, n) = 1$, and $o(\zeta_n^k) = \frac{o(\zeta_n)}{\gcd(o(\zeta_n), k)} = \frac{n}{\gcd(n, k)}$, treating $(1)^{1/n}$ as a group.

Definition 17.4.3. *A generator of the cyclic group of all n th roots of unity is called a primitive n th root of unity.*

Note there are $\varphi(n)$ primitive roots of unity, for φ the Euler-totient function.

Definition 17.4.4. *The splitting field of $x^n - 1$ is $\mathbb{Q}(\zeta_n)$, which is called the cyclotomic field of n th roots of unity.*

Theorem 17.4.3. *Let $\varphi : F \xrightarrow{\sim} F'$ be an isomorphism of fields. Let $f(x) \in F[x]$ be a polynomial and let $f'(x) \in F'[x]$ be the polynomial obtained by applying φ to the coefficients of $f(x)$. Let E be a splitting field for $f(x)$ over F and let E' be a splitting field for $f'(x)$ over F' . Then the isomorphism φ extends to an isomorphism $\sigma : E \xrightarrow{\sim} E'$, i.e., restricting σ to F is the isomorphism φ :*

$$\begin{array}{ccc} E & \xrightarrow{\sigma} & E' \\ \downarrow & & \downarrow \\ F & \xrightarrow{\varphi} & F' \end{array}$$

Proof. We proceed by induction on the degree of $f(x)$. Let Φ denote the corresponding isomorphism between $F[x]$ and $F'[x]$ induced by φ . Then if $f(x)$ corresponds to $f'(x)$ under this isomorphism, then so do the irreducible factors of $f(x)$ and $f'(x)$.

If $f(x)$ splits completely in F , $f'(x)$ will split completely in F' , and vice-versa. Hence, $E = F \cong F' = E'$, taking $\sigma = \varphi$. This shows the result holds for $n = 1$, and when all irreducible factors are linear.

Assume now by induction that the result holds for any field F , isomorphism φ , and polynomial $f(x) \in F[x]$ of degree less than n . Let $p(x)$ be an irreducible factor of $f(x)$ of degree at least 2, and $p'(x)$ the corresponding irreducible factor of $f'(x)$. If $\alpha \in E$ is a root of $p(x)$ and $\beta \in E'$ is a root of $p'(x)$, then by Theorem 17.1.3 we can extend φ to an isomorphism

$\sigma' : F(\alpha) \xrightarrow{\sim} F'(\beta)$. Then $f(x) = (x - \alpha)f_1(x)$ over $F(\alpha)$ and $f'(x) = (x - \beta)f'_1(x)$ over $F'(\beta)$. The field E is a splitting field for $f_1(x)$ over $F(\alpha)$. Similarly, E' is a splitting field for $f'_1(x)$ over $F'(\beta)$. Since the degrees are less than n , we have by the induction hypothesis that there exists an isomorphism $\sigma : E \xrightarrow{\sim} E'$ which restricts to σ' . Thus, σ restricted to F is φ , as desired. ■

Corollary 17.4.4 (Uniqueness of Splitting Fields). *Any two splitting fields for a polynomial $f(x) \in F[x]$ over a field F are isomorphic.*

Definition 17.4.5. *The field \overline{F} is called an algebraic closure of F if \overline{F} is algebraic over F and if every polynomial $f(x) \in F[x]$ splits completely over \overline{F} (so that \overline{F} can be said to contain all the elements algebraic over F).*

Definition 17.4.6. *A field K is said to be algebraically closed if every polynomial with coefficients in K has a root in K .*

Example 17.4.7. \mathbb{R} has algebraic closure $\overline{\mathbb{R}} = \mathbb{C}$, which is also the splitting field $\mathbb{R}(i)$ of $x^2 + 1 \in \mathbb{R}[x]$.

Example 17.4.8. \mathbb{Q} has $\overline{\mathbb{Q}}$ algebraic closure. But, $\mathbb{Q}(i)$ is the splitting field for $x^2 + 1 \in \mathbb{Q}[x]$ is not equal to this.

Proposition 17.4.5. *Let \overline{F} be an algebraic closure of F . Then \overline{F} is algebraically closed.*

Proposition 17.4.6. *For any field F , there exists an algebraically closed field K containing F .*

Proposition 17.4.7. *Let K be an algebraically closed field and let F be a subfield of K . Then the collection of elements of \overline{F} of K that are algebraic over F is an algebraic closure of F . An algebraic closure of F is unique up to isomorphism.*

Theorem 17.4.8 (Fundamental Theorem of Algebra). *The field of \mathbb{C} is algebraically closed.*

Corollary 17.4.9. *The field \mathbb{C} contains an algebraic closure for any of its subfields. In particular, $\overline{\mathbb{Q}}$, the collection of complex numbers algebraic over \mathbb{Q} , is an algebraic closure of \mathbb{Q} .*

17.5.0 §Separable and Inseparable Extensions

Let F be a field and let $f(x) \in F[x]$. Over a splitting field for $f(x)$ we have a factorization

$$f(x)(x - \alpha_1)^{n_1} \dots (x - \alpha_k)^{n_k}$$

where $\alpha_1, \dots, \alpha_k$ are distinct elements of the splitting field, and $n_i \geq 1$ for all i . The integer n_i is called the multiplicity of α_i . α_i is called a multiple root if $n_i > 1$, and a simple root if $n_i = 1$.

Definition 17.5.1. *A polynomial over F is separable if it has no multiple roots (i.e. all roots are distinct). A polynomial which is not separable is called inseparable.*

Example 17.5.1. $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ is separable over \mathbb{Q} .

Example 17.5.2. $x^2 - t \in \mathbb{F}_2(t)[x]$, then there exists $\sqrt{t} \in K/\mathbb{F}_2(t)$ for some extension K . Then

$$x^2 - t = (x - \sqrt{t})(x + \sqrt{t}) = (x + \sqrt{t})(x + \sqrt{t}) = (x + \sqrt{t})^2$$

so $x^2 - t$ is inseparable over $\mathbb{F}_2(t)$.

Definition 17.5.2. The derivative of the polynomial

$$f(x) = a_n x^n + \dots + a_1 x + a_0 \in F[x]$$

is defined to be the polynomial

$$D_x f(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \dots + a_1$$

Moreover, this satisfies the identities

$$D_x(f(x)g(x)) = D_x(f(x))g(x) + f(x)D_x(g(x))$$

and

$$D_x(f(x) + g(x)) = D_x f(x) + D_x g(x)$$

Proposition 17.5.1. A polynomial $f(x)$ has a multiple root α if and only if α is also a root of $D_x f(x)$, i.e., $f(x)$ and $D_x f(x)$ are both divisible by the minimal polynomial for α . In particular, $f(x)$ is separable if and only if it is relatively prime to its derivative: $\gcd(f(x), D_x f(x)) = 1$, i.e. $(f(x)) + (D_x f(x)) = F[x]$.

Proof. Suppose first that α is a multiple root of $f(x)$. Then over a splitting field $f(x) = (x - \alpha)^n g(x)$ for some integer $n \geq 2$ and some polynomial $g(x)$ in the splitting field. Taking derivatives we obtain

$$D_x f(x) = n(x - \alpha)^{n-1} g(x) + (x - \alpha)^n D_x g(x)$$

which shows that $D_x f(x)$ has α as a root since $n \geq 2$.

Conversely, suppose α is a root of $D_x f(x)$ and $f(x)$. Write $f(x) = (x - \alpha)h(x)$, and observe that since $D_x f(x) = h(x) + (x - \alpha)D_x h(x)$ and $D_x f(\alpha) = 0$, we have that $h(\alpha) = 0$. Hence, $h(x) = (x - \alpha)h'(x)$ for some polynomial $h'(x)$, so $f(x) = (x - \alpha)^2 h'(x)$ and α is a multiple root of $f(x)$ as claimed. ■

Example 17.5.3. Consider $x^{p^n} - x \in \mathbb{F}_p$, then its derivative is $p^n x^{p^n-1} - 1 = -1$ since the field is of characteristic p . Then since the derivative has no roots, $x^{p^n} - x$ has no multiple roots and is therefore separable.

Corollary 17.5.2. Every irreducible polynomial over a field of characteristic 0 is separable (i.e. has no multiple roots). A polynomial over such a field is separable if and only if it is the product of distinct irreducible polynomials.

Proof. Suppose F is a field of characteristic 0 and $p(x) \in F[x]$ is irreducible of degree n . Then the derivative $D_x p(x)$ is a polynomial of degree $n - 1$. Up to constant factors the only factors of $p(x)$ in $F[x]$ are 1 and $p(x)$, so $D_x p(x)$ must be relatively prime to $p(x)$. Then $p(x)$ is separable by Proposition 17.5.1. Then, by Proposition 17.2.1 all distinct irreducible polynomials have distinct roots, which gives the second result. ■

In characteristic p , the derivative of any polynomial in x^p is 0, unlike in the characteristic 0 case. However, if the derivative is nonzero we can again conclude that the polynomial is separable by the same argument as above.

Proposition 17.5.3. *Let F be a field of characteristic p . Then for any $a, b \in F$,*

$$(a + b)^p = a^p + b^p, \text{ and } (ab)^p = a^p b^p$$

Put another way, the p th power map defined by $\varphi(a) = a^p$ is an injective field homomorphism from F to F .

Proof. The binomial Theorem for expanding $(a + b)^n$ for any positive integer n holds over any commutative ring:

$$(a + b)^n = \sum_{i=0}^n \binom{n}{i} a^{n-i} b^i$$

If p is a prime, then for $i = 1, 2, \dots, p - 1$, $\binom{p}{i}$ is divisible by p since for these values of i the numbers $i!$ and $(p - i)!$ only involve factors smaller than p , hence are relatively prime to p and so cannot cancel the factor of p in the numerator of the expression $\frac{p!}{i!(p-i)!}$. It follows that over a field of characteristic p all the intermediate terms in the expansion of $(a + b)^p$ are 0, which gives the first equation of the proposition. The second equation is from commutivity of multiplication in the field, and φ is injective since it is a field homomorphism. ■

Definition 17.5.3. *The field homomorphism $\varphi : F \rightarrow F$ defined by $\varphi(a) = a^p$, where F is a field of characteristic p , is called the Frobenius endomorphism of F .*

Corollary 17.5.4. *Suppose that F is a finite field of characteristic p . Then every element of F is a p th power in F (Notationally, $F = F^p$).*

Let F be a finite field of characteristic p , and $f(x) \in F[x]$ an irreducible polynomial. If $f(x)$ were inseparable, then we would need that $f(x) = q(x^p)$ for some $q(x) \in F[x]$, as otherwise by our previous result $f(x)$ would be separable. Let $q(x) = \sum_{i=0}^n a_i x^i$, and note that by our previous result we have $a_i = b_i^p$ for each i , so we may write

$$f(x) = \sum_{i=0}^n b_i^p x^{ip} = \left(\sum_{i=0}^n b_i x^i \right)^p$$

where the last equality follows from our endomorphism. But, $f(x)$ is assumed irreducible and we've written it as a product of p non-constant polynomials, which is a contradiction.

Proposition 17.5.5. *Every irreducible polynomial over a finite field F is separable. A polynomial in $F[x]$ is separable if and only if it is the product of distinct irreducible polynomials in $F[x]$.*

Definition 17.5.4. *A field K is of characteristic p is called perfect if every element of K is a p th power in K . Any field of characteristic 0 is also called perfect.*

From this we see that we've shown that every irreducible polynomial over a perfect field is separable.

Example 17.5.4 (Existence and Uniqueness of Finite Fields). Let $n > 0$ be any positive integer and consider the splitting field of the polynomial $x^{p^n} - x$ over \mathbb{F}_p . We know that this polynomial is separable, hence has precisely p^n distinct roots. Let α and β be any two roots of this polynomial, so $\alpha^{p^n} = \alpha$ and $\beta^{p^n} = \beta$. Then $(\alpha\beta)^{p^n} = \alpha\beta$, $(\alpha^{-1})^{p^n} = \alpha^{-1}$, and by induction on our endomorphism:

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

Hence, the set, F , of roots of the polynomial is a field, and consequently is the splitting field of the polynomial. Since the number of elements is p^n , $[F : \mathbb{F}_p] = n$, which gives the existence of finite fields of degree n over \mathbb{F}_p , for all $n > 0$.

Letting \mathbb{F} be another such finite field of characteristic p , if $[\mathbb{F} : \mathbb{F}_p] = n$, then \mathbb{F} has p^n elements. Since the multiplicative group \mathbb{F}^\times is of order $p^n - 1$, we have $\alpha^{p^n-1} = 1$ for all $\alpha \in \mathbb{F}^\times$, so that $\alpha^{p^n} = \alpha$ for every $\alpha \in \mathbb{F}$. Hence, \mathbb{F} is contained in a splitting field for $x^{p^n} - x$, but since we know the splitting field for such a polynomial is of order p^n and $|\mathbb{F}| = p^n$, we have that \mathbb{F} is a splitting field for the polynomial. Since splitting fields are unique up to isomorphism, this shows the existence and uniqueness up to isomorphism of finite fields of order p^n .

Proposition 17.5.6. Let $p(x)$ be an irreducible polynomial over a field \mathbb{F} of characteristic p . Then there is a unique integer $k \geq 0$ and a unique irreducible separable polynomial $p_{sep}(x) \in \mathbb{F}[x]$ such that

$$p(x) = p_{sep}(x^{p^k})$$

Definition 17.5.5. Let $p(x)$ be an irreducible polynomial over a field of characteristic p . The degree of $p_{sep}(x)$ in the last proposition is called the separable degree of $p(x)$, denoted $\deg_s p(x)$. The integer p^k in the proposition is called the inseparable degree of $p(x)$, denoted $\deg_i p(x)$.

From this we observe that $p(x)$ is separable if and only if its inseparable degree is 1 if and only if its separable degree is equal to its actual degree. Computing degrees in the equation $p(x) = p_{sep}(x^{p^k})$ gives

$$\deg p(x) = \deg_s p(x) \deg_i p(x)$$

Definition 17.5.6. The field K is said to be separable (or separably algebraic) over F if every element of K is the root of a separable polynomial over F (equivalently, the minimal polynomial over F of every element of K is separable). A field which is not separable is inseparable.

The separability of finite extensions of perfect fields is immediate since for these fields the minimal polynomial of an algebraic element is irreducible and hence separable.

Corollary 17.5.7. Every finite extension of a perfect field is separable. In particular, every finite extension of either \mathbb{Q} or a finite field is separable.

17.6.0 Cyclotomic Polynomials and Extensions

The purpose of this section is to prove the cyclotomic extension $\mathbb{Q}(\zeta_n) = \mathbb{Q}$ generated by the n th root of unity over \mathbb{Q} is of degree $\varphi(n) = |\{1 \leq a < n : \gcd(a, n) = 1\}|$.

Definition 17.6.1. Let μ_n denote the group of n th roots of unity over \mathbb{Q} .

Recall that $\mathbb{Z}/n\mathbb{Z} \cong \mu_n$, with the map $a \mapsto \zeta_n^a$ for a fixed primitive n th root of unity. Also recall that the primitive n th roots of unity are precisely the residue classes prime to n , so there are $\varphi(n)$ primitive n th roots of unity.

If d is a divisor of n , and ζ_d is a d th root of unity, then ζ_d is also a n th root of unity as $\zeta_d^n = (\zeta_d^d)^{n/d}$. Thus, we have that

$$\mu_d \subseteq \mu_n, \quad \text{for all } d|n$$

Definition 17.6.2. Define the n th cyclotomic polynomial $\Psi_n(x)$ to be the polynomial whose roots are the primitive n th roots of unity:

$$\Phi_n(x) = \prod_{\zeta \text{ primitive } \in \mu_n} (x - \zeta) = \prod_{\substack{1 \leq a < n \\ \gcd(a, n) = 1}} (x - \zeta_n^a)$$

(which is of degree $\varphi(n)$)

The roots of $x^n - 1$ are precisely the n th roots of unity, so we have the factorization

$$x^n - 1 = \prod_{\zeta \in \mu_n} (x - \zeta)$$

If we group the factors $(x - \zeta)$ where ζ is an element of order d in μ_n , then we obtain

$$x^n - 1 = \prod_{d|n} \prod_{\zeta \text{ primitive } \in \mu_d} (x - \zeta)$$

Equivalently, we can write

$$x^n - 1 = \prod_{d|n} \Phi_d(x)$$

Incidentally, comparing degrees we have that

$$n = \sum_{d|n} \varphi(d)$$

Lemma 17.6.1. The cyclotomic polynomial $\Phi_n(x)$ is a monic polynomial in $\mathbb{Z}[x]$ of degree $\varphi(n)$.

Proof. $\Phi_n(x)$ being a product of monic polynomials is consequently monic. Moreover, by construction the degree of $\Phi_n(x)$ is $\varphi(n)$. To show coefficients lie in \mathbb{Z} we proceed by induction on n . The result is immediate for $n = 1$, with $\Phi_1(x) = (x - 1)$. Assume by induction that $\Phi_d(x) \in \mathbb{Z}[x]$ for all $1 \leq d < n$. Then $x^n - 1 = f(x)\Phi_n(x)$ where $f(x) = \prod_{d|n, d < n} \Phi_d(x)$ is monic and has coefficients in \mathbb{Z} by the induction hypothesis. Since $f(x)$ divides $x^n - 1$ in $F[x]$ where $F = \mathbb{Q}(\zeta_n)$ is the field of n th roots of unity and both $f(x)$ and $x^n - 1$ have coefficients in \mathbb{Q} , $f(x)$ divides $x^n - 1$ in $\mathbb{Q}[x]$ by the division algorithm. By Gauss' Lemma, $f(x)$ divides $x^n - 1$ in $\mathbb{Z}[x]$, hence $\Phi_n(x) \in \mathbb{Z}[x]$. ■

Theorem 17.6.2. *The cyclotomic polynomial $\Phi_n(x)$ is an irreducible monic polynomial in $\mathbb{Z}[x]$ of degree $\varphi(n)$.*

Proof. Suppose towards a contradiction that $\Phi_n(x) = f(x)g(x)$, with $f(x), g(x) \in \mathbb{Z}[x]$ monic, where we take $f(x)$ to be an irreducible factor of $\Phi_n(x)$. Let ζ be a primitive n th root of unity which is a root of $f(x)$ (so then $f(x)$ is the minimal polynomial for ζ over \mathbb{Q}) and let p denote any prime not dividing n . Then ζ^p is again a primitive n th root of unity, hence is a root of either $f(x)$ or $g(x)$.

Suppose $g(\zeta^p) = 0$. Then ζ is a root of $g(x^p)$ and since $f(x)$ is the minimal polynomial for ζ , $f(x)$ must divide $g(x^p)$ in $\mathbb{Z}[x]$, say $g(x^p) = f(x)h(x)$ for some $h(x) \in \mathbb{Z}[x]$. If we reduce this equation mod p , we obtain $\bar{g}(x^p) = \bar{f}(x)\bar{h}(x)$ in $\mathbb{F}_p[x]$. By the Frobenius endomorphism for finite fields we have that $\bar{g}(x^p) = (\bar{g}(x))^p$, so we have the equation $(\bar{g}(x))^p = \bar{f}(x)\bar{h}(x)$ in the U.F.D $\mathbb{F}_p[x]$. It follows that $\bar{f}(x)$ and $\bar{g}(x)$ have a factor in common in $\mathbb{F}_p[x]$.

Now, from $\Phi_n(x) = f(x)g(x)$ we see by reducing mod p that $\bar{\Phi}_n(x) = \bar{f}(x)\bar{g}(x)$, and so by the above it follows that $\bar{\Phi}_n(x) \in \mathbb{F}_p[x]$ has a multiple root. But then also $x^n - 1$ would have a multiple root over \mathbb{F}_p since it has $\bar{\Phi}_n(x)$ as a factor. This is a contradiction since we have seen in the last section that there are n distinct roots of $x^n - 1$ over any field of characteristic not dividing n .

Hence ζ^p must be a root of $f(x)$. Since this applies to every root ζ of $f(x)$, it follows that ζ^a is a root of $f(x)$ for every integer a relatively prime to n : write $a = p_1 p_2 \dots p_k$ as a product of not necessarily distinct primes not dividing n so that ζ^{p_1} is a root of $f(x)$, so also $(\zeta^{p_1})^{p_2}$ is a root of $f(x)$, etc. But this implies that every primitive n th root of unity is a root of $f(x)$, which is to say $f(x) = \Phi_n(x)$, showing $\Phi_n(x)$ is irreducible. ■

Corollary 17.6.3. *The degree over \mathbb{Q} of the cyclotomic field of n th roots of unity is $\varphi(n)$:*

$$[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \varphi(n)$$

Chapter 18

§§Galois Theory

18.1.0 §Basics Definitions and Examples: Galois Theory

Definition 18.1.1. An isomorphism of $K \xrightarrow{\sim} K$ is called an automorphism of K . The collection of all automorphisms is denoted $\text{Aut}(K)$. We say a $\sigma \in \text{Aut}(K)$ fixes $a \in K$ if $\sigma(a) = a$. If F is a subset of K , then an automorphism $\sigma \in \text{Aut}(K)$ is said to fix F if it fixes all the elements of F , i.e., $\sigma(a) = a$ for all $a \in F$.

Note that all fields have at least one automorphism, namely the identity automorphism, or trivial automorphism. The prime subfield of K is generated by $1 \in K$, and since any automorphism σ takes 1 to 1 (and 0 to 0), i.e., $\sigma(1) = 1$, it follows that $\sigma(a) = a$ for all a in the prime subfield. In particular, we see that \mathbb{Q} and \mathbb{F}_p only have the trivial automorphism.

Definition 18.1.2. Let K/F be an extension of fields. Let $\text{Aut}(K/F)$ denote the collection of automorphisms of K which fix F ; $\text{Aut}(K/F) = \{\sigma \in \text{Aut}(K) : \sigma(x) = x, \forall x \in F\}$.

Example 18.1.1. For K a field of characteristic zero, $\text{Aut}(K/\mathbb{Q}) = \text{Aut}(K)$, and for K a field of prime characteristic p , $\text{Aut}(K/\mathbb{F}_p) = \text{Aut}(K)$.

Proposition 18.1.1. $\text{Aut}(K)$ is a group under composition. Moreover, $\text{Aut}(K/F)$ is a subgroup of that group.

Proof. Note that the composition of bijections is a bijection and the composition of algebraic structure preserving maps still preserves that structure so $\text{Aut}(K)$ is indeed closed under the binary operation of composition. Moreover, composition is associative, and the inverse of a field isomorphism is again a field isomorphism. Finally, every field has the identity isomorphism which acts as an identity under composition, so $\text{Aut}(K)$ is indeed a group. First, we note that $1_K \in \text{Aut}(K/F)$ as it fixes F . If $\sigma, \tau \in \text{Aut}(K/F)$ and $x \in F$, then $\sigma \circ \tau(x) = \sigma(\tau(x)) = \sigma(x) = x$, so $\sigma \circ \tau \in \text{Aut}(K/F)$. Moreover, if $\sigma \in \text{Aut}(K/F)$, then $\sigma^{-1}(x) = \sigma^{-1}(\sigma(x)) = x$, so $\sigma^{-1} \in \text{Aut}(K/F)$. Thus, $\text{Aut}(K/F)$ is a subgroup of $\text{Aut}(K)$ as desired. ■

Proposition 18.1.2. *Let K/F be a field extension and $\alpha \in K$ algebraic over F . Then for any $\sigma \in \text{Aut}(K/F)$, $\sigma(\alpha)$ is a root of the minimal polynomial for α over F i.e., $\text{Aut}(K/F)$ permutes roots of irreducible polynomials in F . Equivalently, any polynomial with coefficients in F having α as a root also has $\sigma(\alpha)$ as a root.*

Proof. Suppose that $\alpha \in K$ satisfies the equation

$$\alpha^n + a_{n-1}\alpha^{n-1} + \dots + a_1\alpha + a_0 = 0$$

which is minimal, for $a_i \in F$. Let $\sigma \in \text{Aut}(K/F)$, and we act on the above:

$$\sigma(\alpha)^n + a_{n-1}\sigma(\alpha)^{n-1} + \dots + a_1\sigma(\alpha) + a_0 = 0$$

since elements of F are fixed, and field homomorphisms are additive and multiplicative. But, this precisely says that $\sigma(\alpha)$ is a root of the same polynomial over F in K as α . ■

Example 18.1.2. Consider $K = \mathbb{Q}(\sqrt{2})$, and let $\tau \in \text{Aut}(K/\mathbb{Q}) = \text{Aut}(\mathbb{Q}(\sqrt{2}))$. Then τ is either given by $\tau(\sqrt{2}) = \sqrt{2}$, which gives $\tau = 1_K$, or $\tau(\sqrt{2}) = -\sqrt{2}$, extended algebraically. Thus, $\text{Aut}(K/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z}$.

Example 18.1.3. Consider $K = \mathbb{Q}(\sqrt[3]{2})$. Then, for $\tau \in \text{Aut}(K/\mathbb{Q})$, we have

$$\tau(a + b\sqrt[3]{2} + c\sqrt[3]{2}^2) = a + b\tau(\sqrt[3]{2}) + c\tau(\sqrt[3]{2})^2$$

so τ is determined by where it sends $\sqrt[3]{2}$. This must again be a root of $x^3 - 2$. But, the other roots are in \mathbb{C} , and hence not in K , so τ can only send $\sqrt[3]{2}$ to itself and $\text{Aut}(K/\mathbb{Q}) = \{1_K\}$.

If K is generated over F by some collection of elements, then any automorphism $\sigma \in \text{Aut}(K/F)$ is completely determined by what it does to the generators. If K/F is finite then K is finitely generated over F by algebraic elements so by the proposition the number of automorphisms of K fixing F is finite. In particular, the automorphisms of a finite extension can be considered as permutations of the roots of a finite number of equations.

We have associated to each field extension K/F a group, $\text{Aut}(K/F)$, the group of automorphisms of K which fix F . One can also reverse this process:

Proposition 18.1.3. *Let $H \leq \text{Aut}(K)$ be a subgroup of the group automorphisms of K . Then the collection F of elements of K fixed by all elements of H is a subfield of K .*

Proof. Let F be the fixed collection under H . Then since all automorphisms fix 0 and 1, we have that $0, 1 \in F$. Let $a, b \in F$, $b \neq 0$. Then observe that for any $\sigma \in H$, $\sigma(a - b) = \sigma(a) - \sigma(b) = a - b$ by additivity of automorphisms, so $a - b \in F$. Moreover, $\sigma(ab^{-1}) = \sigma(a)\sigma(b)^{-1} = ab^{-1}$, so $ab^{-1} \in F$. Therefore, F is closed under subtraction and division so by the subfield test F is a subfield of K . ■

Note that it is not important in this proposition that H be a subgroup. Indeed, the subcollection of K fixed by any subset of $\text{Aut}(K)$ is again a subfield of K .

Definition 18.1.3. *If H is a subgroup of the group of automorphisms of K , the subfield of K fixed by all elements of H is called the fixed field of H .*

The Galois Correspondence

Let L/K be a field extension, and consider $\sigma : L \rightarrow L$ such that $\sigma(k) = k$ for all $k \in K$, so that $\sigma \in \text{Aut}(L/K)$.

Example 18.1.4. Consider $\sigma : \mathbb{C} \rightarrow \mathbb{C}$ by $\sigma(z) = \bar{z}$. Then $\sigma \in \text{Aut}(\mathbb{C})$, and moreover, $\sigma(x + i0) = x - i0 = x$, so $\sigma \in \text{Aut}(\mathbb{C}/\mathbb{R})$.

Definition 18.1.4. For any subgroup H of $\text{Aut}(L/K)$, then the fixed field of H is denoted by

$$L^H = \{\alpha \in L : \forall \sigma \in H, \sigma(\alpha) = \alpha\}$$

We're now studying the following correspondence for a given L/K :

$$\{K \subseteq F \subseteq L : F \rightsquigarrow \text{Aut}(L/F)\} \quad \text{and} \quad \{H \leq \text{Aut}(L) : H \rightsquigarrow L^H\}$$

In particular, we have such diagrams as

$$\begin{array}{ccc} L & & \{id\} \\ | & & | \\ F & & \text{Aut}(L/F) \\ | & & | \\ F' & & \text{Aut}(L/F') \\ | & & | \\ K & & \text{Aut}(L/K) \end{array}$$

Proposition 18.1.4. The association of groups to fields and fields to groups defined above is inclusion reversing, namely:

- if $F_1 \subseteq F_2 \subseteq L$ are two subfields of K then $\text{Aut}(L/F_2) \leq \text{Aut}(L/F_1)$, and
- if $H_1 \leq H_2 \leq \text{Aut}(L)$ are two subgroups of automorphisms with associated fixed fields L^{H_1} and L^{H_2} , $L^{H_2} \subseteq L^{H_1}$.

Proof. First suppose $F_1 \subseteq F_2 \subseteq L$ are subfields. Take $\sigma \in \text{Aut}(L/F_2)$. Then let $x \in F_1$. It follows that $x \in F_2$, so $\sigma(x) = x$ and hence $\sigma \in \text{Aut}(L/F_1)$. Thus, $\text{Aut}(L/F_2) \leq \text{Aut}(L/F_1)$.

Next, let $H_1 \leq H_2 \leq \text{Aut}(L)$ be subgroups. Let $x \in L^{H_2}$. Then let $\sigma \in H_1$, so $\sigma \in H_2$. Then $\sigma(x) = x$, and consequently we find that $x \in L^{H_1}$ by definition. Therefore, $L^{H_2} \subseteq L^{H_1}$. ■

We observe that $L^{\text{Aut}(L/K)} \supseteq K$. Equality does not always hold. Indeed:

Example 18.1.5. $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ has $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, so there are no proper intermediate fields. Moreover, $\text{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \{\text{Id}\}$, so $\text{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \text{Aut}(\mathbb{Q}(\sqrt[3]{2}))$, so $\mathbb{Q}(\sqrt[3]{2})^{\text{Aut}(\mathbb{Q}(\sqrt[3]{2}))} = \mathbb{Q}(\sqrt[3]{2})$.

Definition 18.1.5. The roots of a common irreducible polynomial in $K[x]$ are called **K-conjugates**.

Example 18.1.6. $\pm\sqrt{2}$ are \mathbb{Q} -conjugates as they share the minimal polynomial $x^2 - 2$ in $\mathbb{Q}[x]$. But, they are not \mathbb{R} -conjugates since $x^2 - 2$ is not irreducible over \mathbb{R} since $x - \sqrt{2}$ is irreducible with root only $\sqrt{2}$, not $-\sqrt{2}$.

Thus, for α a root of $f(x) \in K[x]$, we have that for any $\sigma \in \text{Aut}(L/K)$, $\sigma(\alpha)$ and α are K -conjugates.

The goal of this section is to bound the size of L/K :

Theorem 18.1.5. For any finite extension L/K , the group $\text{Aut}(L/K)$ is finite.

Proof. Let $L = K(\alpha_1, \alpha_2, \dots, \alpha_n)$. Then for any $\sigma \in \text{Aut}(L/K)$ is given by $\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)$, with each root associated to a polynomial of finite degree. Then for each α_i , there are a finite number of choices for where it can be sent since it must be sent to another root of its minimal polynomial in $K[x]$. Hence, each σ is determined by a finite choice, and hence $\text{Aut}(L/K)$ itself is finite. ■

Theorem 18.1.6. Let $\sigma : K \rightarrow K'$ be an isomorphism of fields, $f(X) \in K[X]$, L a splitting field of $f(X)$ over K and L' be a splitting field of $(\sigma f)(X)$ over K' . Then $[L : K] = [L' : K']$, and σ extends to an isomorphism $L \rightarrow L'$, and the number of such extensions is at most $[L : K]$.

Corollary 18.1.7. If L is a splitting field over K of a polynomial $f(X) \in K[X]$, then $|\text{Aut}(L/K)| \leq [L : K]$.

Proof. Apply Theorem 18.1.6 with $K' = K$, $L' = L$, and $\sigma = \text{Id} : K \rightarrow K$. Extensions of the identity isomorphism on K to automorphisms of L are precisely the elements of $\text{Aut}(L/K)$. ■

Theorem 18.1.8. Let $\sigma : K \rightarrow K'$ be an isomorphism of fields, $f(X) \in K[X]$, L be a splitting field of $f(X)$ over K and L' be a splitting field of $(\sigma f)(x)$ over K' . If $f(X)$ is separable then there are $[L : K]$ extensions of σ to an isomorphism $L \rightarrow L'$.

Proof. We may assume $[L : K] > 1$. Let α be a root of $f(X)$ that is not in K , and let $\pi(X)$ be its minimal polynomial over K , and $d = \deg(\pi(X))$. Because $f(X)$ is separable, $(\sigma f)(X)$ is separable too. One way to show this is with the characterization of separability in terms of relative primality to the derivative: we can write

$$f(X)u(X) + f'(X)v(X) = 1$$

for some $u(X)$ and $v(X)$ in $K[X]$. Applying σ to coefficients commutes with forming derivatives (i.e. $\sigma(f') = (\sigma f)'$), so if we apply σ to coefficients in our expression to get

$$(\sigma f)(X)(\sigma u)(X) + (\sigma f)'(X)(\sigma v)(X) = 1$$

so $(\sigma f)(X)$ and its derivative are relatively prime in $K'[X]$. This last polynomial identity proves $(\sigma f)(X)$ is separable. Every factor of a separable polynomial is separable, so $(\sigma\pi)(X)$ and therefore has d roots in L' since it splits completely over L' .

(To be continued-relies on proof of 18.1.6) ■

Corollary 18.1.9. *If L/K is a splitting field of a separable polynomial then $|\operatorname{Aut}(L/K)| = [L : K]$.*

Theorem 18.1.10. *For a finite extension L/K , the following are equivalent characterizations of Galois extensions:*

- $|\operatorname{Aut}(L/K)| = [L : K]$
- $L^{\operatorname{Aut}(L/K)} = K$
- L/K is separable and normal
- L is a splitting field over K of a separable polynomial in $K[x]$.

Definition 18.1.6. *Let K/F be a finite extension. Then K is said to be Galois over F and K/F is a **Galois extension** if $|\operatorname{Aut}(K/F)| = [K : F]$. If K/F is Galois the group of automorphisms $\operatorname{Aut}(K/F)$, is called the Galois group of K/F , denoted $\operatorname{Aut}(K/F) = \operatorname{Gal}(K/F)$.*

In this case $\operatorname{Aut}(K/F)$ has the maximal number of possible automorphisms.

Definition 18.1.7. *If $f(x)$ is a separable polynomial over F , then the Galois group of $f(x)$ over F is the Galois group of the splitting field of $f(x)$ over F .*

Example 18.1.7. Consider $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$, the splitting field of $x^2 - 2 \in \mathbb{Q}[x]$. This is separable. Then $\operatorname{Aut}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$ is Galois with $|\operatorname{Aut}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})| = 2$, with $\operatorname{Aut}(\mathbb{Q}(\sqrt{2})/\mathbb{Q}) = \{\operatorname{Id}, \sigma\}$.

Example 18.1.8. Any quadratic extension of any field F of characteristic different from 2 is Galois, $K = F(\sqrt{D})$. This holds as it is the splitting field of $x^2 - D$ in F , for $\sqrt{D} \neq -\sqrt{D} \in K$.

Example 18.1.9. For $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over \mathbb{Q} , it is the splitting field of $(x^2 - 2)(x^2 - 3)$, which is separable. Thus, $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is Galois. It's elements are

$$\sigma: \begin{matrix} \sqrt{2} \mapsto -\sqrt{2} \\ \sqrt{3} \mapsto \sqrt{3} \end{matrix}, \quad \tau: \begin{matrix} \sqrt{2} \mapsto \sqrt{2} \\ \sqrt{3} \mapsto -\sqrt{3} \end{matrix}$$

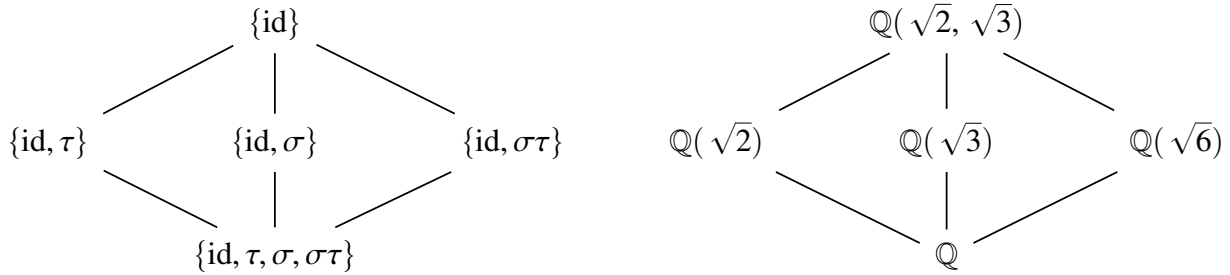
and

$$\sigma\tau: \begin{matrix} \sqrt{2} \mapsto -\sqrt{2} \\ \sqrt{3} \mapsto -\sqrt{3} \end{matrix}, \quad \operatorname{Id}: \begin{matrix} \sqrt{2} \mapsto \sqrt{2} \\ \sqrt{3} \mapsto \sqrt{3} \end{matrix}$$

This is in fact the Klien-4 group, as indeed since $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is Galois and $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$, so $\operatorname{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) = \{\operatorname{Id}, \sigma, \tau, \sigma\tau\} \cong K_4 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Subgroup	Fixed
$\{\operatorname{Id}\}$	$\mathbb{Q}(\sqrt{2}, \sqrt{3})$
$\{\operatorname{Id}, \sigma\}$	$\mathbb{Q}(\sqrt{3})$
$\{\operatorname{Id}, \tau\}$	$\mathbb{Q}(\sqrt{2})$
$\{\operatorname{Id}, \sigma\tau\}$	$\mathbb{Q}(\sqrt{6})$
$\{1, \sigma, \tau, \sigma\tau\}$	\mathbb{Q}

So we can draw:



Example 18.1.10. The field $\mathbb{Q}(\sqrt[4]{2})$ is not Galois over \mathbb{Q} since any automorphism is determined by where it sends $\sqrt[4]{2}$ by where it sends $\sqrt[4]{2}$ and of the four possibilities, only two elements are real and hence in the field. Note

$$\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\sqrt[4]{2})$$

So $[\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = 4$, but $\text{Aut}(\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}) = \{\text{Id}, \tau\}$. However, both $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ and $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}(\sqrt{2})$ are in fact Galois, so the composition of Galois extensions is not necessarily Galois.

Example 18.1.11. Consider $\mathbb{F}_{p^n}/\mathbb{F}_p$ for prime p . Recall $x^{p^n} - x$ is separable, and \mathbb{F}_{p^n} is the splitting field for $x^{p^n} - x$, so it is a Galois extension. Consider the Frobenius automorphism $\sigma : \mathbb{F}_{p^n} \rightarrow \mathbb{F}_{p^n}$, with $\sigma(\alpha) = \alpha^p$. Then $\sigma^2(\alpha) = \alpha^{p^2}$, and in general $\sigma^j(\alpha) = \alpha^{p^j}$ for $j = 1, 2, \dots, n-1$, and $\sigma^n(\alpha) = \alpha^{p^n} = \alpha$, so $\sigma^i \sigma^j = \sigma^{i+j}$ modulo n . Then the Frobenius automorphism is of order n , so as $[\mathbb{F}_{p^n} : \mathbb{F}_p] = n$ the Frobenius automorphism is a generator for the Galois group $\text{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)$.

18.2.0 §The Fundamental Theorem of Galois Theory

Definition 18.2.1. A **character** χ of a group G with values in a field L is a homomorphism $\chi : G \rightarrow L^\times$. In particular, for $g_1, g_2 \in G$,

$$\chi(g_1 g_2) = \chi(g_1) \chi(g_2)$$

and $\chi(g) \neq 0$ for all $g \in G$.

Definition 18.2.2. If $\chi_1, \chi_2, \dots, \chi_n$ are characters of G with values in L , then they're linearly independent over L provided there does not exist $a_1, a_2, \dots, a_n \in L$, not all zero, such that

$$a_1 \chi_1 + a_2 \chi_2 + \dots + a_n \chi_n = 0$$

(as a function, $a_1 \chi_1(g) + \dots + a_n \chi_n(g) = 0$ for all $g \in G$)

Theorem 18.2.1. If $\chi_1, \chi_2, \dots, \chi_n$ are distinct characters of G with values in L , then they are linearly independent over L .

Proof. Towards a contradiction suppose the characters were linearly dependent. Among all of the linear dependence relations, choose one with minimal number m of non zero coefficients a_i . With possible renumbering we may write $a_1\chi_1 + \dots + a_m\chi_m = 0$. In particular, for any $g \in G$, $a_1\chi_1(g) + \dots + a_m\chi_m(g) = 0$. Let $g_0 \in G$ such that $\chi_1(g_0) \neq \chi_m(g_0)$, which is possible as the characters are distinct. Then we have $a_1\chi_1(g_0g) + \dots + a_m\chi_m(g_0g) = 0$. Multiplying our previous equation by $\chi_m(g_0)$ and subtracting it from this one we obtain

$$[\chi_1(g_0) - \chi_m(g_0)]a_1\chi_1(g) + [\chi_2(g_0) - \chi_m(g_0)]a_2\chi_2(g) + \dots + [\chi_{m-1}(g_0) - \chi_m(g_0)]a_{m-1}\chi_{m-1}(g) = 0$$

But, since $\chi_1(g_0) - \chi_m(g_0) \neq 0$, this is a linear dependence of length less than or equal to $m - 1 < m$, contradicting the minimality of m . ■

Consider an injective homomorphism σ of a field K into a field L , called an **embedding** of K into L . Then in particular σ restricts to a homomorphism of the multiplicative group K^\times into L^\times . Note also that this character contains all of the useful information about the values of σ viewed simply as a function on K .

Corollary 18.2.2. *If $\sigma_1, \sigma_2, \dots, \sigma_n$ are distinct embeddings of a field K into a field L , then they are linearly independent as functions on K . In particular, distinct automorphisms of a field K are linearly independent as functions on K .*

This follows from our last theorem, and the discussion above about restricting embeddings to group characters.

Theorem 18.2.3. *Let $G = \{\sigma_1 = 1, \sigma_2, \dots, \sigma_n\}$ be a subgroup of automorphisms of a field K and let F be the associated fixed field. Then*

$$[K : F] = n = |G|$$

Proof. Suppose first that $n > [K : F]$ and let $\omega_1, \dots, \omega_m$ be a basis for K over F ($m < n$). Then the system

$$\sigma_1(\omega_1)x_1 + \sigma_2(\omega_1)x_2 + \dots + \sigma_n(\omega_1)x_n = 0$$

$$\vdots$$

$$\sigma_1(\omega_m)x_1 + \sigma_2(\omega_m)x_2 + \dots + \sigma_n(\omega_m)x_n = 0$$

of m equations in n unknowns has a nontrivial kernel, and hence solution β_1, \dots, β_n in K since $m < n$.

Let a_1, \dots, a_m be m arbitrary elements of F . The field is by definition fixed by the σ_i 's, so each of these elements is fixed by every σ_i , i.e. $\sigma_i(a_j) = a_j$. Multiplying the i th equation by a_i we have

$$\sigma_1(a_1\omega_1)x_1 + \sigma_2(a_2\omega_1)x_2 + \dots + \sigma_n(a_1\omega_1)x_n = 0$$

$$\vdots$$

$$\sigma_1(a_m\omega_m)x_1 + \sigma_2(a_m\omega_m)x_2 + \dots + \sigma_n(a_m\omega_m)x_n = 0$$

Adding these equations and substituting β_i for x_i we obtain

$$\sigma_1(a_1\omega_1 + \dots + a_m\omega_m)\beta_1 + \dots + \sigma_n(a_1\omega_1 + \dots + a_m\omega_m)\beta_n = 0$$

for all choices of $a_i \in F$. But since the ω_i form a basis for K over F , this consequently holds for all elements of K . However, this implies that this is a linear dependence on the set $\sigma_1, \dots, \sigma_n$ of automorphisms, contradicting the linear independence of the previous corollary. Thus, $n \leq [K : F]$.

Now, suppose $n < [K : F]$. Then there are more than n F -linearly independent elements of K , say $\alpha_1, \dots, \alpha_{n+1}$. Then the system

$$\begin{aligned} \sigma_1(\alpha_1)x_1 + \sigma_1(\alpha_2)x_2 + \dots + \sigma_1(\alpha_{n+1})x_{n+1} &= 0 \\ \vdots \\ \sigma_n(\alpha_1)x_1 + \sigma_n(\alpha_2)x_2 + \dots + \sigma_n(\alpha_{n+1})x_{n+1} &= 0 \end{aligned}$$

of n equations in $n + 1$ unknowns has a non trivial solution $\beta_1, \dots, \beta_{n+1}$ in K . If all the β_i were elements of F , then the first equation would contradict the linear independence of the α_j , recalling that $\sigma_1 = 1$. Hence, at least 1 β_i is not in F .

Choose a solution with a minimal number of nonzero β_i , r . By renumbering if necessary write β_1, \dots, β_r nonzero. Dividing by β_r we may also assume that $\beta_r = 1$. We have already seen that at least one of $\beta_1, \dots, \beta_{r-1}, 1$ is not in F . With possible reordering, suppose $\beta_1 \notin F$. Then our system reads

$$\begin{aligned} \sigma_1(\alpha_1)\beta_1 + \sigma_1(\alpha_2)\beta_2 + \dots + \sigma_1(\alpha_{r-1})\beta_{r-1} + \sigma_1(\alpha_r) &= 0 \\ \vdots \\ \sigma_n(\alpha_1)\beta_1 + \sigma_n(\alpha_2)\beta_2 + \dots + \sigma_n(\alpha_{r-1})\beta_{r-1} + \sigma_n(\alpha_r) &= 0 \end{aligned}$$

so $\sigma_i(\alpha_1)\beta_1 + \sigma_i(\alpha_2)\beta_2 + \dots + \sigma_i(\alpha_{r-1})\beta_{r-1} + \sigma_i(\alpha_r) = 0$ for each i . Since $\beta_1 \notin F$, and F is the fixed field of the subgroup of automorphisms, there exists σ_{k_0} such that $\sigma_{k_0}\beta_1 \neq \beta_1$ for some $k_0 \in \{1, 2, \dots, n\}$. If we apply the automorphism to the previous equation we obtain

$$\sigma_{k_0}\sigma_j(\alpha_1)\sigma_{k_0}\beta_1 + \dots + \sigma_{k_0}\sigma_j(\alpha_{r-1})\sigma_{k_0}\beta_{r-1} + \sigma_{k_0}\sigma_j(\alpha_r) = 0$$

But the action of left multiplication of a group element on the group itself simply permutes the elements, so $\sigma_{k_0}\sigma_1, \dots, \sigma_{k_0}\sigma_n$ equals $\sigma_1, \dots, \sigma_n$ with some reordering. Let $\sigma_{k_0}\sigma_j = \sigma_i$. Then we have

$$\sigma_i(\alpha_1)\sigma_{k_0}\beta_1 + \dots + \sigma_i(\alpha_{r-1})\sigma_{k_0}\beta_{r-1} + \sigma_i(\alpha_r) = 0$$

Subtracting our original linear dependence from this equation we have

$$\sigma_i(\alpha_1)[\sigma_{k_0}\beta_1 - \beta_1] + \dots + \sigma_i(\alpha_{r-1})[\sigma_{k_0}\beta_{r-1} - \beta_{r-1}] = 0$$

But $[\sigma_{k_0}\beta_1 - \beta_1] \neq 0$, so this is a linear dependence with fewer than r nonzero elements, contradicting the minimality of r . ■

Corollary 18.2.4. *Let K/F be any finite extension. Then*

$$|\text{Aut}(K/F)| \leq [K : F]$$

with equality if and only if F is the fixed field of $\text{Aut}(K/F)$. Put another way, K/F is Galois if and only if F is the fixed field of $\text{Aut}(K/F)$.

Proof. Let F_1 be the fixed field of $\text{Aut}(K/F)$, so that $F \subseteq F_1 \subseteq K$. By our previous theorem $[K : F_1] = |\text{Aut}(K/F)|$. Hence, $[K : F] = [K : F_1][F_1 : F]$, which proves the corollary. ■

Corollary 18.2.5. *Let G be a finite subgroup of automorphisms of a field K and let F be the fixed field. Then every automorphism of K fixing F is contained in G , i.e., $\text{Aut}(K/F) = G$, so that K/F is Galois, with Galois group G .*

Proof. By definition F is fixed by all elements of G , so we have $G \leq \text{Aut}(K/F)$. Hence $|G| \leq |\text{Aut}(K/F)|$. By the theorem we have $|G| = [K : F]$ and by the previous corollary we have $|\text{Aut}(K/F)| \leq [K : F]$. This gives

$$|G| \leq |\text{Aut}(K/F)| \leq |G|$$

so $|\text{Aut}(K/F)| = |G|$, and since $G \leq \text{Aut}(K/F)$, $G = \text{Aut}(K/F)$. ■

Corollary 18.2.6. *If $G_1 \neq G_2$ are distinct finite subgroups of automorphisms of a field K , then their fixed fields are distinct.*

Proof. Suppose F_1 is the fixed field of G_1 and F_2 is the fixed field of G_2 . Then suppose $F_1 = F_2$, so by definition F_1 is fixed by G_2 . By the previous corollary, any automorphism fixing F_1 is in G_1 , so $G_2 \leq G_1$. Similarly, F_2 is fixed by G_1 so $G_1 \leq G_2$ and we have equality, $G_1 = G_2$, which is a contradiction. ■

Theorem 18.2.7. *The extension K/F is Galois if and only if K is the splitting field of some separable polynomial over F . Furthermore, if this is the case then every irreducible polynomial with coefficients in F which has a root in K is separable and has all of its roots in K (so in particular K/F is a separable extension).*

Proof. By Corollary 18.1.9 we have that the splitting field of some separable polynomial over F is Galois.

We now suppose that K/F is Galois. Set $G = \text{Gal}(K/F)$. Let $p(x) \in F[x]$ be an irreducible polynomial having a root $\alpha \in K$, and consider the elements $\alpha, \sigma_2(\alpha), \dots, \sigma_n(\alpha) \in K$ for $\{1, \sigma_2, \dots, \sigma_n\}$ are the elements of G . Let $\alpha, \alpha_2, \dots, \alpha_r$ denote the distinct roots in the list. If $\tau \in G$, then since G is a group, $\{\tau, \tau\sigma_2, \dots, \tau\sigma_n\} = \{1, \sigma_1, \dots, \sigma_n\}$ in some order. It follows that applying τ to $\alpha, \alpha_2, \dots, \alpha_r$ also permutes these elements. The polynomial

$$f(x) = (x - \alpha)(x - \alpha_2) \dots (x - \alpha_r)$$

has coefficients which are all fixed by all elements of G since the elements of G simply permute the factors. Hence the coefficients lie in the fixed field of G , so by a preceding corollary, this field is F . Hence $f(x) \in F[x]$.

Since $p(x)$ is irreducible and has α as a root, $p(x)$ is the minimal polynomial for α over F , hence $p(x)$ divides $f(x)$ in $F[x]$. Since $f(x)$ is a product of linear factors, each of which is a root of $p(x)$ (since roots are permuted under fixed field automorphisms), $f(x)$ divides $p(x)$ as well, so $p(x) = f(x)$ since they are both monic. Thus, $p(x)$ is separable and all of its roots lie in K .

To complete the proof let $\omega_1, \dots, \omega_n$ be a basis for K/F . Let $p_i(x)$ be the minimal polynomial for ω_i for each i . Then by what we've just shown $p_i(x)$ is separable and has all of its roots in K . Let $g(x)$ be the polynomial obtained by removing any multiple factors from $p_1(x) \dots p_n(x)$. Then the splitting field of the two polynomials is the same, and this field is K . Indeed, all of the roots lie in K so K contains the splitting field, and the generators of K are among the roots so K is also contained in the splitting field. Hence K is the splitting field of the separable polynomial $g(x)$. ■

Definition 18.2.3. Let K/F be a Galois extension. If $\alpha \in K$ the elements $\sigma\alpha$ for σ in $\text{Gal}(K/F)$ are called the **conjugates** (or **Galois conjugates**) of α over F . If E is a subfield of K containing F , the field $\sigma(E)$ is called the **conjugate field** of E over F .

The proof of the above theorem shows that in a Galois extension K/F , the other roots of a minimal polynomial over F which has $\alpha \in K$ are precisely the distinct conjugates of α under the Galois group of K/F .

The second statement of the theorem also shows that K is not Galois over F if we can find even one irreducible polynomial over F having a root in K but not having all of its roots in K .

Remark 18.2.1. We now have the 4 characterizations of Galois extensions K/F :

1. splitting fields of separable polynomials over F
2. fields where F is precisely the set of elements fixed by $\text{Aut}(K/F)$
3. fields with $[K : F] = |\text{Aut}(K/F)|$
4. finite, normal and separable extensions.

Theorem 29 (Fundamental Theorem of Galois Theory).

Let K/F be a Galois extension and set $G = \text{Gal}(K/F)$. Then there is a bijection

$$\left\{ \begin{array}{c} \text{subfields } E \\ \text{of } K \\ \text{containing } F \end{array} \begin{array}{c} K \\ | \\ E \\ | \\ F \end{array} \right\} \leftrightarrow \left\{ \begin{array}{c} \text{subgroups } H \\ \text{of } G \end{array} \begin{array}{c} 1 \\ | \\ H \\ | \\ G \end{array} \right\}$$

given by the correspondences

$$\begin{array}{ccc} E & \rightarrow & \left\{ \begin{array}{c} \text{the elements of } G \\ \text{fixing } E \end{array} \right\} \\ \left\{ \begin{array}{c} \text{the fixed field} \\ \text{of } H \end{array} \right\} & \leftarrow & H \end{array}$$

which are inverse under this correspondence ($E \mapsto \text{Aut}(K/E)$ and $H \mapsto K^H$).

1. (Inclusion reversing) If E_1, E_2 correspond to H_1, H_2 , respectively, then $E_1 \subseteq E_2$ if and only if $H_2 \leq H_1$

2. $[K : E] = |H|$ and $[E : F] = |G : H|$, the index of H in G , then

$$\left. \begin{array}{c} K \\ | \\ E \end{array} \right\} |H|$$

$$\parallel$$

$$\left. \begin{array}{c} E \\ | \\ F \end{array} \right\} |G : H|$$

3. K/E is always Galois, with Galois group $\text{Gal}(K/E) = H$:

$$\left. \begin{array}{c} K \\ | \\ E \end{array} \right\} |H|$$

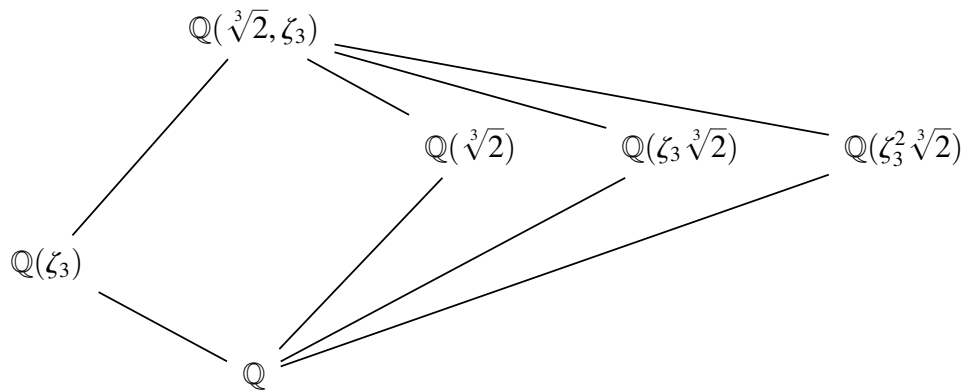
4. E is Galois over F if and only if H is a normal subgroup in G . If this is the case, then the Galois group is isomorphic to the quotient group

$$\text{Gal}(E/F) \cong G/H$$

More generally, if H is not necessarily normal in G , the isomorphisms of E into a fixed algebraic closure of F containing K which fix F are in one to one correspondence with the cosets σH of H in G .

5. If E_1, E_2 correspond to H_1, H_2 , respectively, then the intersection $E_1 \cap E_2$ corresponds to the group $\langle H_1, H_2 \rangle$ generated by H_1 and H_2 and the composite $E_1 E_2$ corresponds to the intersection $H_1 \cap H_2$. Hence the lattice of subfields of K containing F and the lattice of subgroups of G are “dual” (the lattice diagram for one is the lattice diagram for the other turned upside down)

Example 18.2.1. Consider $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$, for $\zeta_3 = e^{2\pi i/3}$. Then $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ is the splitting field of $x^3 - 2$, which is irreducible and separable over \mathbb{Q} so the extension is Galois. Then we can write:



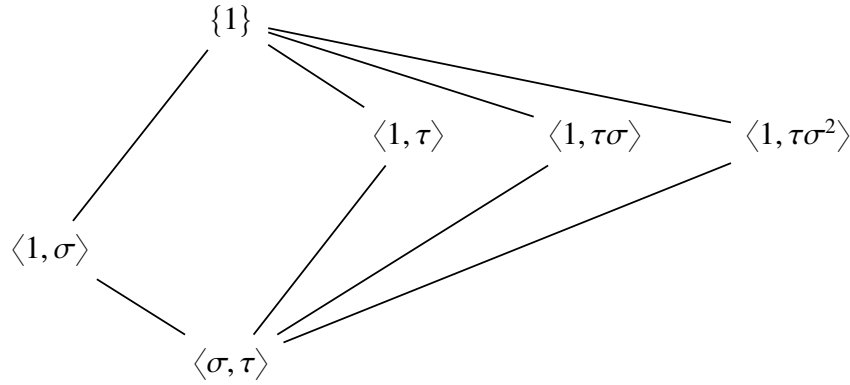
Then $\text{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) = \langle \sigma, \tau \rangle$ for σ, τ defined by

$$\sigma : \begin{cases} \sqrt[3]{2} \mapsto \zeta_3 \sqrt[3]{2} \\ \zeta_3 \mapsto \zeta_3 \end{cases}$$

and

$$\tau : \begin{cases} \sqrt[3]{2} \mapsto \sqrt[3]{2} \\ \zeta_3 \mapsto \zeta_3^2 = -1 - \zeta_3 \end{cases}$$

and we have the relations $\sigma^3 = \tau^2 = 1$, and $\sigma\tau = \tau\sigma^2$, so $\sigma\tau\sigma = \tau$. (Dihedral group of order 3). We also obtain the diagram



Example 18.2.2. The extension $\mathbb{Q}(\sqrt[4]{2}, i)/\mathbb{Q}$ is Galois since it is the splitting field of the separable polynomial $x^4 - 2 \in \mathbb{Q}[x]$. We note that $[\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}] = 8$, so as the extension is Galois there are 8 automorphisms of $\mathbb{Q}(\sqrt[4]{2}, i)$ fixing \mathbb{Q} . But, $x^4 - 2$ has 4 roots, which can be permuted in 24 ways, so not all permutations of the roots result in automorphisms of the field. For example, $\sqrt[4]{2}$ and $-\sqrt[4]{2}$ add to zero, so they must be mapped under field automorphisms to two roots which are negatives of each other.

To think about the Galois group concretely, we think of an automorphism σ and what it does to $\sqrt[4]{2}$ and i , rather than what it does to all fourth roots of 2. Since $\sigma(\sqrt[4]{2})$ must be a root of $X^4 - 2$, there are four possible values for which we can send it to. On the other hand $\sigma(i)$ must be a root of $x^2 + 1$, and hence there are two values for which we can send it to, so there are at most $4 \cdot 2 = 8$ automorphisms of $\mathbb{Q}(\sqrt[4]{2}, i)/\mathbb{Q}$. But, we have shown that there are exactly 8 automorphisms in the Galois group, so all choices of where to send these two roots result in automorphisms of the extension field. Let r be the automorphism determined by $r(\sqrt[4]{2}) = i\sqrt[4]{2}$ and $r(i) = i$, and s be the automorphism determined by $s(\sqrt[4]{2}) = \sqrt[4]{2}$ and $s(i) = -i$. Taking compositions we obtain the following eight automorphisms:

Table 18.1: Galois Group for $\mathbb{Q}(\sqrt[4]{2}, i)/\mathbb{Q}$

σ	id	r	r^2	r^3	s	rs	r^2s	r^3s
$\sigma(\sqrt[4]{2})$	$\sqrt[4]{2}$	$i\sqrt[4]{2}$	$-\sqrt[4]{2}$	$-i\sqrt[4]{2}$	$\sqrt[4]{2}$	$i\sqrt[4]{2}$	$-\sqrt[4]{2}$	$-i\sqrt[4]{2}$
$\sigma(i)$	i	i	i	i	$-i$	$-i$	$-i$	$-i$

A calculation shows $r^4 = s^2 = id$, and $rs = sr^{-1}$, so $\text{Gal}(\mathbb{Q}(\sqrt[4]{2}, i)/\mathbb{Q}) \cong D_4$, where D_4 can be viewed as the 8 symmetries of the square whose vertices are the four complex roots of $X^4 - 2$; in particular, r is a counterclockwise rotation by 90° , and s is a reflection in the real axis (complex conjugation), which is a diagonal of the square. Note r is not multiplication by i everywhere, so it is not generally rotating all elements of $\mathbb{Q}(\sqrt[4]{2}, i)$ by 90° .

18.3.0 § Finite Fields

Recall a finite field \mathbb{F} has characteristic p for some prime, and hence is a finite dimensional vector space over $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. Moreover, if the dimension is n , i.e. $[\mathbb{F} : \mathbb{F}_p] = n$, then \mathbb{F} has precisely p^n elements. Recall that we have seen previously that \mathbb{F} is then isomorphic to the splitting field of the separable polynomial $x^{p^n} - x$ is $\mathbb{F}_p[x]$, and hence is a Galois extension. Let \mathbb{F}_{p^n} denote this field of order p^n .

Further, the Galois group of \mathbb{F}_{p^n} is of cyclic of order n , generated by the Frobenius automorphism σ_p :

$$\text{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) = \langle \sigma_p \rangle \cong \mathbb{Z}/n\mathbb{Z}$$

where $\sigma_p : \mathbb{F}_{p^n} \rightarrow \mathbb{F}_{p^n}$ is given by $\sigma_p(a) = a^p$. By the Fundamental Theorem of Galois Theory we have a bijective correspondence between subfields of \mathbb{F}_{p^n} containing \mathbb{F}_p and subgroups of $\mathbb{Z}/n\mathbb{Z}$. In particular, for every divisor d of n we have a subfield of degree d over \mathbb{F}_p , namely the fixed field generated by σ_p^d of order n/d . Moreover, from a previous discussion we know this field is isomorphic to \mathbb{F}_{p^d} , the unique finite field of order p^d .

Proposition 18.3.1. *Any finite field is isomorphic to \mathbb{F}_{p^n} for some prime p and some integer $n \geq 1$. The field \mathbb{F}_{p^n} is the splitting field over \mathbb{F}_p of the polynomial $x^{p^n} - x$, with cyclic Galois group of order n generated by the Frobenius automorphism σ_p . The subfields of \mathbb{F}_{p^n} are all Galois over \mathbb{F}_p and are in one to one correspondence with divisors d of n . They are the fields \mathbb{F}_{p^d} , the fixed fields of σ_p^d .*

We observe that the multiplicative group $\mathbb{F}_{p^n}^\times$ is finite group composed of field elements, and is hence cyclic. Then, if θ is any generator of this group we obtain $\mathbb{F}_{p^n} = \mathbb{F}_p(\theta)$. Consequently:

Proposition 18.3.2. *The finite field \mathbb{F}_{p^n} is simple. In particular, there exists an irreducible polynomial of degree n over \mathbb{F}_p for each $n \geq 1$.*

Recall the elements of \mathbb{F}_{p^n} are precisely the roots of $x^{p^n} - x \in \mathbb{F}_p[x]$. Grouping together the factors $x - \alpha$ of this polynomial based on the degree d of their minimal polynomials over \mathbb{F}_p , we obtain

Proposition 18.3.3. *The polynomial $x^{p^n} - x$ is precisely the product of all the distinct irreducible polynomials in $\mathbb{F}_p[x]$ of degree d where d runs through all divisors of n .*

This result gives a method for determining the product of all the irreducible polynomials over \mathbb{F}_p of a given degree. Note also that since the finite field \mathbb{F}_{p^n} is unique up to isomorphism, the quotients of $\mathbb{F}_p[x]$ by any irreducible polynomial of degree n are all isomorphic.

Define the Möbius μ -functions by

$$\mu(n) = \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{if } n \text{ has a square factor} \\ (-1)^r & \text{if } n \text{ has } r \text{ distinct prime factors} \end{cases}$$

If $f(n)$ is a function defined for all nonnegative integers and $F(n)$ is defined by

$$F(n) = \sum_{d|n} f(d)$$

Then the Möbius inversion formula states that one can recover the function $f(n)$ from $F(n)$:

$$f(n) = \sum_{d|n} \mu(d) F(n/d)$$

Then define $\psi(n)$ to be the number of irreducible polynomials of degree n in $\mathbb{F}_p[x]$. Counting degrees in our previous proposition

$$p^n = \sum_{d|n} d\psi(d)$$

Applying the inversion formula for $f(n) = n\psi(n)$, we obtain

$$n\psi(n) = \sum_{d|n} \mu(d) p^{n/d}$$

which gives us a formula for the number of irreducible polynomials of degree n over $\mathbb{F}_p[x]$:

$$\psi(n) = \frac{1}{n} \sum_{d|n} \mu(d) p^{n/d}$$

Next, recall that $\mathbb{F}_{p^d} \subseteq \mathbb{F}_{p^n}$ if and only if d divides n . In particular, for finite fields $\mathbb{F}_{p^{n_1}}$ and $\mathbb{F}_{p^{n_2}}$, we have a third field containing isomorphic copies of them, $\mathbb{F}_{p^{n_1 n_2}}$. This gives us a partial ordering on finite fields of order p , and allows us to take their union. Since these give all the finite extensions of \mathbb{F}_p , we see that the union of \mathbb{F}_{p^n} for $n \geq 1$ is an algebraic closure of \mathbb{F}_p , unique up to isomorphism:

$$\overline{\mathbb{F}_p} = \bigcup_{n \geq 1} \mathbb{F}_{p^n}$$

18.4.0 § Composite Extensions and Simple Extensions

Proposition 18.4.1. *Suppose K/F is a Galois extension and F'/F is any extension. Then KF'/F' is a Galois extension, with Galois group*

$$\text{Gal}(KF'/F') \cong \text{Gal}(K/K \cap F')$$

isomorphic to a subgroup of $\text{Gal}(K/F)$.

Proof. If K/F is a Galois extension, then K is a splitting field of a separable polynomial $f(x)$ over $F[x]$. Then observe that KF'/F' is a splitting field for $f(x)$ viewed as a polynomial in $F'[x]$, since F'/F . Hence, KF'/F' is Galois. Since K/F is Galois, every embedding of K fixing F is an automorphism of K , so the map

$$\begin{aligned} \varphi : \text{Gal}(KF'/F') &\rightarrow \text{Gal}(K/F) \\ \sigma &\mapsto \sigma|_K \end{aligned}$$

defined by restricting an automorphism σ to the subfield K is well defined. Further it is a homomorphism with kernel

$$\ker \varphi = \{\sigma \in \text{Gal}(KF'/F') \mid \sigma|_K = 1\}$$

Since an element of $\text{Gal}(KF'/F')$ is the identity on F' , and elements of the kernel are identities on both K and F' , they are also the identity on their composite. Hence, the kernel is trivial and φ is injective.

Let H denote the image of φ , and let K_H denote the corresponding fixed subfield. Since every element in H fixes F' , K_H contains $K \cap F'$. On the other hand, the composite $K_H F'$ is fixed by $\text{Gal}(KF'/F')$ (any $\sigma \in \text{Gal}(KF'/F')$ fixes F' , and fixes $K_H \subseteq K$ when restricted to $\sigma|_K$). By bijectivity in the fundamental theorem of Galois theory, $K_H F' = F'$, so that $K_H \subseteq F'$, which gives the reverse inclusion $K_H \subseteq K \cap F'$. Hence $K_H = K \cap F'$, so again by the Fundamental Theorem of Galois Theory, $H = \text{Gal}(K/K \cap F')$, completing the proof. ■

(To be finished)

18.5.0 §Cyclotomic Extensions and Abelian Extensions

Recall that the cyclotomic field $\mathbb{Q}(\zeta_n)$ of n th roots of unity is a Galois extension of \mathbb{Q} , being the splitting field of the separable polynomial $x^n - 1 \in \mathbb{Q}[x]$, and has degree $\varphi(n)$. Moreover, the roots of $x^n - 1$ form a cyclic group of order n , and hence any automorphism of $\mathbb{Q}(\zeta_n)$ is fully determined by its action on ζ_n .

Theorem 18.5.1. *The Galois group of the cyclotomic field $\mathbb{Q}(\zeta_n)$ of n th roots of unity is isomorphic to the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^\times$. The isomorphism is given explicitly by the map*

$$\begin{aligned} (\mathbb{Z}/n\mathbb{Z})^\times &\xrightarrow{\sim} \text{Gal}(\mathbb{Q}(\zeta_n)) \\ a \pmod n &\mapsto \sigma_a \end{aligned}$$

where σ_a is the automorphism defined by $\sigma_a(\zeta_n) = \zeta_n^a$.

Example 18.5.1. $\mathbb{Q}(\zeta_5)$ is Galois over \mathbb{Q} with $(\mathbb{Z}/5\mathbb{Z})^\times \cong \mathbb{Z}/4\mathbb{Z}$. Indeed, $x^5 - 1$ is separable as $\{1, \zeta_5, \zeta_5^2, \zeta_5^3, \zeta_5^4\}$ are its five distinct roots, and $\mathbb{Q}(\zeta_5)$ is consequently a splitting field for this polynomial. The elements of the Galois group $\text{Gal}(\mathbb{Q}(\zeta_5)/\mathbb{Q})$ are generated by $\sigma(\zeta_5) = \zeta_5^2$, so $\sigma^2(\zeta_5) = \zeta_5^4$, $\sigma^3(\zeta_5) = \zeta_5^3$, $\sigma^4(\zeta_5) = \zeta_5$, so $\sigma^4 = 1$, and $\text{Gal}(\mathbb{Q}(\zeta_5)/\mathbb{Q}) \cong \mathbb{Z}/4\mathbb{Z}$. Then, we can consider the subgroup $H = \{1, \sigma^2\}$ of order 2. Then, the fixed field of H is $\mathbb{Q}(\zeta_5 + \zeta_5^{-1})$.

(To be finished)

18.6.0 §Galois Group of Polynomials

Recall that the Galois group of a separable polynomial $f(x) \in F[x]$ is the Galois group of the splitting field of $f(x)$ over F .

Further, if K/F is Galois, then K is the splitting field for some separable polynomial $f(x) \in F[x]$, and $[K : F] = |\text{Aut}(K/F)|$. Moreover, any automorphism $\sigma \in \text{Gal}(K/F)$ maps a root of an irreducible factor of $f(x)$ to another root of the irreducible factor, and σ is uniquely determined by its action on these roots (since they generate K over F). Fix a labeling $\alpha_1, \dots, \alpha_n$ the roots of $f(x)$, and observe that $\sigma \in \text{Gal}(K/F)$ defines a unique permutation of $\alpha_1, \dots, \alpha_n$. This gives an injection $\text{Gal}(K/F) \hookrightarrow S_n$.

Since the degree of the splitting field is the same as the order of the Galois group, this explains from the group-theoretic side why the splitting field for a polynomial of degree n over F is of degree at most $n!$ over F .

In general, if the factorization of $f(x) = f_1(x) \dots f_k(x)$ into (distinct) irreducibles where $f_i(x)$ has degree n_i , then since the Galois group permutes the roots of the irreducible factors among themselves we have $\text{Gal}(K/F) \hookrightarrow S_{n_1} \times \dots \times S_{n_k}$.

If $f(x)$ is itself irreducible, then given any two roots of $f(x)$ there is an automorphism in the Galois group which maps the first root to the second. Such a group is said to be transitive on the roots. Thus, the Galois group must be transitive on blocks of roots.

Example 18.6.1. Consider the biquadratic extension $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over \mathbb{Q} , which is the splitting field of $(x^2 - 2)(x^2 - 3)$. Label the roots $\alpha_1 = \sqrt{2}, \alpha_2 = -\sqrt{2}, \alpha_3 = \sqrt{3}, \alpha_4 = -\sqrt{3}$. From our previous discussion $\text{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) \hookrightarrow S_2 \times S_2 = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} = V_4$, the Klein-4 group. But, $|\text{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q})| = 4$, so $\text{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) \cong V_4$. Viewing it as a subgroup of S_4 , we have from a previous discussion $\text{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) = \{1, \sigma, \tau, \sigma\tau\}$ so

$$\sigma = (12), \tau = (34)$$

as viewed in S_4 with the above labelling. Further, $\sigma\tau = (12)(34)$. Hence

$$\text{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) \cong \{1, (12), (34), (12)(34)\} \subseteq S_4$$

Example 18.6.2. Consider the Galois group of $x^3 - 2$, with extension $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$, with roots $\sqrt[3]{2}, \zeta_3 \sqrt[3]{2}, \zeta_3^2 \sqrt[3]{2}$. With this ordering the generators σ and τ of the Galois group give the following permutations

$$\sigma = (123), \tau = (23)$$

which gives

$$\text{Gal}(\sqrt[3]{2}, \zeta_3) \cong \{1, \sigma, \sigma^2, \tau, \tau\sigma, \tau\sigma^2\} = S_3$$

in this case the full symmetric group on 3 letters.

Recall that every finite group is isomorphic to a subgroup of S_n , for some n (by Cayley's Theorem). It is currently an open problem as to whether every finite group appears as the Galois group for some polynomial over \mathbb{Q} .

Definition 18.6.1. Let x_1, x_2, \dots, x_n be indeterminates. The elementary symmetric functions s_1, s_2, \dots, s_n are defined by

$$\begin{aligned} s_1 &= x_1 + x_2 + \dots + x_n \\ s_2 &= \sum_{i < j} x_i x_j = x_1 x_2 + x_1 x_3 + \dots + x_2 x_3 + x_2 x_4 + \dots + x_{n-1} x_n \\ &\vdots \\ s_n &= x_1 x_2 \dots x_n \end{aligned}$$

i.e., the i th symmetric function s_i of x_1, \dots, x_n is the sum of all products of the x_j 's, taken i at a time.

Definition 18.6.2. The general polynomial of degree n is the polynomial

$$(x - x_1)(x - x_2) \dots (x - x_n) \in (\mathbb{F}[x_1, x_2, \dots, x_n])[x]$$

whose roots are the indeterminates x_1, x_2, \dots, x_n .

By induction, it can be shown that the coefficients of the general polynomial of degree n are given by the elementary symmetric functions in the roots:

$$(x - x_1)(x - x_2) \dots (x - x_n) = x^n - s_1 x^{n-1} + s_2 x^{n-2} - \dots + (-1)^n s_n$$

For any field F , the extension $F(x_1, x_2, \dots, x_n)$ is then a Galois extension of the field $F(s_1, s_2, \dots, s_n)$ since it is a splitting field of the general polynomial of degree n .

If $\sigma \in S_n$ is any permutation of $\{1, 2, \dots, n\}$, then σ acts on the rational functions in $F(x_1, x_2, \dots, x_n)$ by permuting the subscripts of x_1, x_2, \dots, x_n , and this gives an automorphism of $F(x_1, \dots, x_n)$. Identifying $\sigma \in S_n$ with this automorphism, identifies S_n as a subgroup of $\text{Aut}(F(x_1, \dots, x_n))$. The elementary symmetric functions are fixed under any permutation of their subscripts (this is the reason they are called symmetric), which shows that the subfield $F(s_1, \dots, s_n)$ is contained in the fixed field of S_n as a subgroup of $\text{Aut}(F(x_1, \dots, x_n))$. By the Fundamental Theorem of Galois Theory, the fixed subfield of S_n has index $n!$ in $F(x_1, \dots, x_n)$. Since $F(x_1, \dots, x_n)$ is the splitting field over $F(s_1, \dots, s_n)$ of the polynomial of degree n above, we have

$$[F(x_1, \dots, x_n) : F(s_1, \dots, s_n)] \leq n!$$

Since the subfield must contain $F(s_1, \dots, s_n)$, it follows that we have equality and that $F(s_1, \dots, s_n)$ is precisely the fixed field of S_n .

Proposition 18.6.1. The fixed field of the symmetric group S_n acting on the field of rational functions in n variables $F(x_1, x_2, \dots, x_n)$ is the field of the rational functions in the elementary symmetric functions $F(s_1, \dots, s_n)$.

Definition 18.6.3. A rational function $f(x_1, \dots, x_n)$ is called symmetric if it is not changed by any permutation of the variables x_1, x_2, \dots, x_n .

Corollary 18.6.2 (Fundamental Theorem of Symmetric Functions). Any symmetric function in the variables x_1, x_2, \dots, x_n is a rational function in the elementary symmetric functions s_1, s_2, \dots, s_n .

Proof. A symmetric function lies in the fixed field of the subgroup isomorphic to S_n in $\text{Gal}(F(x_1, \dots, x_n)/F(s_1, \dots, s_n))$ and hence lies in $F(s_1, \dots, s_n)$ by the previous proposition. ■

Remark 18.6.1. If $f(x_1, \dots, x_n)$ is a polynomial in x_1, \dots, x_n which is symmetric then it can be seen that f is actually a polynomial in s_1, \dots, s_n . It is in fact true that a symmetric polynomial whose coefficients lie in R , where R is any commutative ring with identity, is a polynomial in the elementary symmetric functions with coefficients in R .

Example 18.6.3. Consider $(x_1 - x_2)^2$, which is symmetric in x_1 and x_2 , so we can write $(x_1 - x_2)^2 = (x_1 + x_2)^2 - 4x_1x_2 = s_1^2 - 4s_2$ a polynomial in terms of symmetric functions.

Example 18.6.4. Consider the three hyperbolic numbers $\mathcal{H}_3 = \langle j \rangle$, $j^3 = 1$, $\zeta = x + jy + j^2z$, which in matrix form is

$$M[\zeta] = \begin{bmatrix} x & z & y \\ y & x & z \\ z & y & x \end{bmatrix}$$

which has determinant $\det(M[\zeta]) = x^3 + y^3 + z^3 - 3xyz$, which is symmetric in x, y, z . This can then be written as $(x + y + z)^3 - 2(xy + xz + yz) - 3xyz = s_1(s_1^2 - 3s_2)$.

Theorem 18.6.3. *The general polynomial*

$$x^n - s_1x^{n-1} + s_2x^{n-2} + \dots + (-1)^n s_n$$

over the field $F(s_1, \dots, s_n)$ is separable with Galois group S_n .

This result says that if there are no relations among the coefficients of a polynomial of degree n , then the Galois group of this polynomial over the field generated by its coefficients is the full symmetric group S_n . To say the s_i are *indeterminants* in x_1, \dots, x_n this means, informally, that there does not exist nonzero polynomial relations over s_1, \dots, s_n .

For $n \geq 5$, there is only one normal subgroup of S_n , namely A_n , hence there is only one normal subfield of $F(x_1, \dots, x_n)$ containing $F(s_1, \dots, s_n)$ and it is an extension of degree 2.

Definition 18.6.4. Define the discriminant D of x_1, x_2, \dots, x_n by the formula

$$D = \prod_{i < j} (x_i - x_j)^2$$

Define the discriminant of a polynomial to be the discriminant of the roots of the polynomial.

The discriminant D is a symmetric function in x_1, \dots, x_n , hence is an element of $K = F(s_1, \dots, s_n)$.

Recall that $\sigma \in S_n$ is an element of A_n if and only if σ fixes the product

$$\sqrt{D} = \prod_{i < j} (x_i - x_j) \in \mathbb{Z}[x_1, x_2, \dots, x_n]$$

It follows by the fundamental theorem that if F has characteristic different from 2, then \sqrt{D} generates the fixed field of A_n and generates a quadratic extension of $K = F(s_1, \dots, s_n)$.

Proposition 18.6.4. *If $\text{char}(F) \neq 2$, then the permutation $\sigma \in S_n$ is an element of A_n if and only if it fixes the square root of the discriminant D .*

If the roots of the polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ are $\alpha_1, \alpha_2, \dots, \alpha_n$, then the discriminant of $f(x)$ is

$$D = \prod_{i < j} (\alpha_i - \alpha_j)^2$$

Note that $D = 0$ if and only if $f(x)$ is inseparable, so we have nondistinct roots, or multiples. In any perfect field this implies $f(x)$ is reducible. The discriminant D is symmetric in the roots of $f(x)$, and is hence fixed by automorphisms in the Galois group of $f(x)$. Moreover, we have \sqrt{D} is in the splitting field of $f(x)$, being a product of its roots. Then, if the roots of $f(x)$ are distinct, fix some ordering on the roots and view the Galois group of $f(x)$ as a subgroup of S_n , as before.

Proposition 18.6.5. *The Galois group of $f(x) \in F[x]$ is a subgroup of A_n if and only if the discriminant $D \in F$ is the square of an element in F .*

Proof. The Galois group is contained in A_n if and only if every element of the Galois group fixes

$$\sqrt{D} = \prod_{i < j} (\alpha_i - \alpha_j)$$

which is true if and only if $\sqrt{D} \in F$, since Galois groups and their associated fixed fields are unique for a Galois extension by the Fundamental Theorem of Galois. ■

Polynomials of Degree 2

Consider the polynomial $x^2 + ax + b$ with roots α, β . The discriminant D for this polynomial is $(\alpha - \beta)^2$, which can be written as a polynomial in the elementary symmetric functions of the roots, which gives

$$D = s_1^2 - 4s_2 = (-a)^2 - 4(b) = a^2 - 4b$$

which is the usual discriminant for a quadratic. The polynomial is separable if and only if $a^2 - 4b \neq 0$. The Galois group is a subgroup of $S_2 \cong \mathbb{Z}/2\mathbb{Z}$, the cyclic group of order 2, and is trivial (A_2 in this case) if and only if $a^2 - 4b$ is a rational square, which completely determines the possible Galois groups. Moreover, as we found before the splitting field for a non trivial Galois group is the quadratic extension $F(\sqrt{D})$.

Polynomials of Degree 3

Suppose the cubic polynomial is $f(x) = x^3 + ax^2 + bx + c$. If we make the substitution $x = y - a/3$ the polynomial becomes

$$g(y) = y^3 + py + q$$

which is called the ***depressed*** polynomial for $f(x)$, where

$$p = \frac{1}{3}(3b - a^2), \quad q = \frac{1}{27}(2a^3 - 9ab + 27c)$$

The splitting fields for these polynomials are the same since their roots differ by the constant $\frac{a}{3} \in F$ and since the formula for the discriminant involves the differences of roots, we see that these two polynomials also have the same discriminant.

Let the roots of $g(y)$ be α, β , and γ . We first compute the discriminant of this polynomial in terms of p and q . Note that

$$g(y) = (y - \alpha)(y - \beta)(y - \gamma)$$

so that if we differentiate we have

$$D_y g(y) = (y - \alpha)(y - \beta) + (y - \beta)(y - \gamma) + (y - \alpha)(y - \gamma)$$

Then

$$D_y g(\alpha) = (\alpha - \beta)(\alpha - \gamma)$$

$$D_y g(\beta) = (\beta - \alpha)(\beta - \gamma)$$

$$D_y g(\gamma) = (\gamma - \alpha)(\gamma - \beta)$$

Taking the product we see that

$$D = [(\alpha - \beta)(\alpha - \gamma)(\beta - \gamma)]^2 = -D_y g(\alpha)D_y g(\beta)D_y g(\gamma)$$

(Galois Group of the Cubic)

- If the cubic $f(x)$ is reducible, then it splits either into three linear factors or into a linear factor and a irreducible quadratic. In the first case the Galois group is trivial, and in the second case the Galois group is isomorphic to $S_2 \cong \mathbb{Z}/2\mathbb{Z}$.
- If the cubic polynomial $f(x)$ is irreducible, then a root of $f(x)$ generates an extension of degree 3 over F , so the degree of the splitting field over F is divisible by 3. Since the Galois group is a subgroup of S_3 , there are only two possibilities, namely A_3 or S_3 . The Galois group is A_3 (i.e. cyclic of order 3) if and only if the discriminant D is a square ($\sqrt{D} \in F$).

Explicitly, if D is the square of an element of F , then the splitting field of the irreducible cubic $f(x)$ is obtained by adjoining any single root of $f(x)$ to F . If D is not a square of an element of F then the splitting field of $f(x)$ is of degree 6 over F , hence is the field $F(\theta, \sqrt{D})$ for any of the roots θ of $f(x)$. This extension is Galois over F with Galois group S_3 .

Polynomials of Degree 4

Let $f(x) = x^4 + ax^3 + bx^2 + cx + d$ which under the substitution $x = y - a/4$ becomes the depressed quartic

$$g(y) = y^4 + py^2 + qy + r$$

with

$$\begin{aligned} p &= \frac{1}{8}(-3a^2 + 8b) \\ q &= \frac{1}{8}(a^3 - 4ab + 8c) \\ r &= \frac{1}{256}(-3a^4 + 16a^2b - 64ac + 256d) \end{aligned}$$

Suppose first that $g(y)$ is reducible. If it splits into a linear and a cubic, then the Galois group of the cubic is the Galois group of $g(y)$. Suppose then that $g(y)$ splits into two irreducible quadratics. Then the splitting field is the extension $F(\sqrt{D_1}, \sqrt{D_2})$ where D_1 and D_2 are the discriminants of the two quadratics. If D_1 and D_2 do not differ by a square factor then this extension is a biquadratic extension and the Galois group is isomorphic to the Klein 4 subgroup of S_4 , i.e. $S_2 \times S_2$. If D_1 is a square times D_2 , then this extension is a quadratic extension and the Galois group is isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

Now suppose that $g(y)$ is irreducible. Then we have that the Galois group is transitive on the four roots (recall this is to say that for any pair of roots, there exists an automorphism in the Galois group which sends one to the other). Also recall the Galois group must be a subgroup of S_4 . The only transitive subgroups of S_4 (as considered an action on the set of 4 letters), hence the only possibilities are $S_4, A_4, D_4 = \langle (1324), (13)(24) \rangle$ and its conjugates, V_4 , and $C = \langle (1234) \rangle$ and its conjugates. Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be the roots of g .

(To be continued)

18.7.0 § Solvable and Radical Extensions

Definition 18.7.1. The extension K/F is said to be cyclic if it is Galois with a cyclic Galois group.

Proposition 18.7.1. Let F be a field of characteristic not dividing n which contains the n th roots of unity. The extension $F(\sqrt[n]{a})$ for $a \in F$ is cyclic over F of degree dividing n .

Proof. First, the extension $F(\sqrt[n]{a}) = K$ is Galois over F if F contains the n th roots of unity since K would then be the splitting field of the separable polynomial $x^n - a$. For $\sigma \in \text{Gal}(K/F)$, $\sigma(\sqrt[n]{a})$ is another root of this polynomial, hence $\sigma(\sqrt[n]{a}) = \zeta_\sigma \sqrt[n]{a}$ for some n th root of unity ζ_σ . This gives a map

$$\begin{aligned} \text{Gal}(K/F) &\rightarrow \mu_n \\ \sigma &\mapsto \zeta_\sigma \end{aligned}$$

where μ_n denotes the group of n th roots of unity. Since F contains μ_n , every n th root of unity is fixed by every element of $\text{Gal}(K/F)$. Hence

$$\begin{aligned} \sigma\tau(\sqrt[n]{a}) &= \sigma(\zeta_\tau \sqrt[n]{a}) \\ &= \zeta_\tau \sigma(\sqrt[n]{a}) \\ &= \zeta_\tau \zeta_\sigma \sqrt[n]{a} = \zeta_{\sigma\tau} \sqrt[n]{a} \end{aligned}$$

which shows that $\zeta_{\sigma\tau} = \zeta_\sigma \zeta_\tau$, so the map is a homomorphism. The kernel consists precisely of the automorphisms which fix $\sqrt[n]{a}$, namely the identity. This gives an injection of $\text{Gal}(K/F)$ into the cyclic group μ_n of order n , which proves the proposition. ■

Let now K be any cyclic extension of degree n over a field F of characteristic not dividing n which contains the n th roots of unity. Let σ be a generator for the cyclic Galois group $\text{Gal}(K/F)$:

Definition 18.7.2. For $\alpha \in K$ and any n th root of unity ζ , define the Lagrange resolvent $(\alpha, \zeta) \in K$ by

$$(\alpha, \zeta) = \alpha + \zeta\sigma(\alpha) + \zeta^2\sigma^2(\alpha) + \dots + \zeta^{n-1}\sigma^{n-1}(\alpha)$$

If we apply the automorphism σ to (α, ζ) we obtain

$$\sigma(\alpha, \zeta) = \sigma(\alpha) + \zeta\sigma^2(\alpha) + \dots + \zeta^{n-1}\sigma^n(\alpha)$$

since ζ is an element of the base field F , and consequently is fixed by σ . Note we have $\zeta^n = 1$ in μ_n and $\sigma^n = 1$ in $\text{Gal}(K/F)$, so this can be written as

$$\begin{aligned} \sigma(\alpha, \zeta) &= \sigma(\alpha) + \zeta\sigma^2(\alpha) + \dots + \zeta^{-1}\alpha \\ &= \zeta^{-1}(\alpha + \zeta\sigma(\alpha) + \dots + \zeta^{n-1}\sigma^{n-1}(\alpha)) \\ &= \zeta^{-1}(\alpha, \zeta) \end{aligned}$$

It follows that

$$\sigma(\alpha, \zeta)^n = (\zeta^{-1})^n(\alpha, \zeta)^n = (\alpha, \zeta)^n$$

so that $(\alpha, \zeta)^n$ is fixed by $\text{Gal}(K/F)$, and hence is an element of F for any $\alpha \in K$.

Let ζ be a primitive n th root of unity. By the linear independence of the automorphisms $1, \sigma, \dots, \sigma^{n-1}$, there is an element $\alpha \in K$ such that $(\alpha, \zeta) \neq 0$. Iterating our previous result we have $\sigma^i(\alpha, \zeta) = \zeta^{-i}(\alpha, \zeta)$, and it follows that σ^i does not fix (α, ζ) for any $i < n$, since ζ is primitive. Hence, this element cannot lie in any proper subfield of K by the Fundamental Theorem of Galois Theory, so $K = F((\alpha, \zeta))$. Since we proved $(\alpha, \zeta)^n = a \in F$ above, we have $F(\sqrt[n]{a}) = F((\alpha, \zeta)) = K$. This proves the following converse to the previous proposition.

Proposition 18.7.2. Any cyclic extension of degree n over a field F of characteristic not dividing n which contains the n th roots of unity is of the form $F(\sqrt[n]{a})$ for some $a \in F$.

An extension of the form $F(\sqrt[n]{a})$ is called a simple radical extension.

Remark 18.7.1. The proofs above are part of something known as Kummer Theory. A group G is said to have exponent n if $g^n = 1$ for all $g \in G$. Let F be a field of characteristic not dividing n , which contains the n th roots of unity. If we take $a_1, \dots, a_k \in F^\times$ then we can see that the extension

$$F(\sqrt[n]{a_1}, \dots, \sqrt[n]{a_k})$$

is an abelian extension of F whose Galois group is of exponent n .

For simplicity we now consider F as a field of characteristic 0, but the results hold equally well for finite fields of characteristic not dividing the degree of any of the roots to be taken.

Definition 18.7.3. An element α which is algebraic over F can be expressed by radicals or solved for in terms of radicals if α is an element of a field K which can be obtained by a succession of simple radical extensions:

$$F = K_0 \subset K_1 \subset \dots \subset K_s = K$$

where $K_{i+1} = K_i(\sqrt[n_i]{a_i})$ for some $a_i \in K_i$, $i = 0, 1, \dots, s-1$. Here $\sqrt[n_i]{a_i}$ denotes some root of the polynomial $x^{n_i} - a_i$. Such a field K will be called a root extension of F .

Definition 18.7.4. A polynomial $f(x) \in F[x]$ can be solved by radicals if all of its roots can be solved for in terms of radicals.

Example 18.7.1. Consider $K_0 = \mathbb{Q}$, $K_1 = \mathbb{Q}(\sqrt{17})$, $K_2 = K_1(\sqrt{2(17 - \sqrt{17})})$, $K_3 = K_2(\sqrt{2(17 + \sqrt{17})})$, and finally

$$K_4 = K_3 \left(\sqrt{17 + 3\sqrt{17} - \sqrt{2(17 - \sqrt{17})} - 2\sqrt{2(17 + \sqrt{17})}} \right)$$

Each of these extensions is a radical extension. This shows that the element

$$-1 + \sqrt{17} + \sqrt{2(17 - \sqrt{17})} + 2\sqrt{17 + 3\sqrt{17} - \sqrt{2(17 - \sqrt{17})} - 2\sqrt{2(17 + \sqrt{17})}}$$

which is used to construct the 17-gon is constructible starting only with the unit length using a straight edge and compass.

Note that in considering radical extensions we can always adjoin roots of unity as they are radicals. Thus, cyclic extensions become radical extensions and conversely.

Lemma 18.7.3. If α is contained in a root extension K ($F = K_0 \subset K_1 \subset \dots \subset K_s = K$), then α is contained in a root extension which is Galois over F and where each extension K_{i+1}/K_i is cyclic.

Proof. Let L be the Galois closure of K over F . For any $\sigma \in \text{Gal}(L/F)$ we have a chain of subfields

$$F = \sigma K_0 \subset \sigma K_1 \subset \dots \subset \sigma K_s = \sigma K$$

where $\sigma K_{i+1}/\sigma K_i$ is again a simple radical extension (since it is generated by the element $\sigma(\sqrt[n_i]{a_i})$, which is a root of $x^{n_i} - \sigma(a_i)$ over $\sigma(K_i)$). It can be shown that the composite of any two root extensions is again a root extension. Then the composite of all the conjugate field $\sigma(K)$ for $\sigma \in \text{Gal}(L/F)$ is again a root extension. Since this field is precisely L , we see that α is contained in a Galois root extension.

Now adjoin to F all of the n_i th roots of unity for all the roots $\sqrt[n_i]{a_i}$ of the simple radical extensions in the Galois root extension K/F , obtaining the field F' , say, and then form the composite of F' with the root extensions:

$$F \subseteq F' = F'K_0 \subseteq F'K_1 \subseteq \dots \subseteq F'K_s = F'K$$

The field $F'K$ is a Galois extension of F since it is the composite of two Galois extensions. The extension from F to $F' = F'K_0$ can be given as a chain of subfields with each individual extension cyclic (this is true for any abelian extension). Each extension $F'K_{i+1}/F'K_i$ is a simple radical extension and since we now have the appropriate roots of unity in the base fields, each of these individual extensions from F' to $F'K$ is a cyclic extension by our first proposition in this section. Hence $F'K/F$ is a root extension which is Galois over F with cyclic intermediate extensions, completing the proof. ■

Theorem 18.7.4. *The polynomial $f(x)$ can be solved by radicals if and only if its Galois group is a solvable group.*

Proof. Suppose first that $f(x)$ can be solved by radicals. Then each root of $f(x)$ is contained in a Galois root extension over F . The composite L of such extensions is again Galois, and by the work in the proof, a root extension. Let G_i be the subgroups corresponding to the subfields $K_i, i = 0, 1, \dots, s - 1$. Since $\text{Gal}(K_{i+1}/K_i) = G_i/G_{i+1}$ it follows that the Galois group $G = \text{Gal}(L/F)$ (To be continued) ■

Part IV

Modules

Chapter 19

§§General Definitions and Examples

19.1.0 §Basic Definitions and Examples: Modules

Fix a (unital) ring R for this section.

Definition 19.1.1. An (left) R -module is an abelian group M with an additional structure of a map

$$\text{act}_M : R \times M \rightarrow M, \quad (r, m) \mapsto r \cdot m \quad (19.1.1)$$

such that the following properties hold:

1. For every $m \in M$, we have $1 \cdot m = m$.
2. For every $r_1, r_2 \in R$ and $m \in M$, we have

$$r_1 \cdot (r_2 \cdot m) = (r_1 \cdot r_2) \cdot m$$

3. For every $r_1, r_2 \in R$ and $m \in M$, we have

$$(r_1 + r_2) \cdot m = r_1 \cdot m + r_2 \cdot m$$

For every $r \in R$ and $m_1, m_2 \in M$, we have

$$r \cdot (m_1 + m_2) = r \cdot m_1 + r \cdot m_2$$

Note that the last condition is equivalent to saying that for any fixed $r \in R$, the map

$$\begin{aligned} M &\rightarrow M \\ m &\mapsto r \cdot m \end{aligned}$$

is a group endomorphism.

Definition 19.1.2 ((General)). Let R be a ring (not necessarily commutative nor unital). A left R -module or a left module over R is a set M together with

1. A binary operation $+$ on M under which M is an abelian group, and
2. An action of R on M (that is, a map $R \times M \rightarrow M$) denoted by $r.m$, for all $r \in R$ and for all $m \in M$ which satisfies
 - (a) $(r + s).m = r.m + s.m$, for all $r, s \in R$ and all $m \in M$
 - (b) $r.(m + n) = r.m + r.n$, for all $r \in R$ and all $m, n \in M$
 - (c) $r.(s.m) = (rs).m$, for all $r, s \in R$ and all $m \in M$

If the ring R has a 1 we impose the additional axiom:

 - (d) $1.m = m$ for all $m \in M$

Note that when R is a field k , our definition induces the definition of a k -vector space.

Definition 19.1.3. Let R be a ring and let M be an R -module. An **R -submodule** of M is a subgroup N of M which is closed under the action of ring elements, i.e., $r.n \in N$ for all $n \in N$ and all $r \in R$.

We note that if R is a field, then R -submodules correspond to subspaces. Moreover, a submodule of a module M is precisely a subset of M which is itself an R -module under the restricted action by ring elements.

Every R -module M has the submodules M and $\{0\}$, the second being the **trivial submodule**.

Lemma 19.1.1. For any $r \in R$ we have $r \cdot 0_M = 0_M$. For any $m \in M$ we have $0_R \cdot m = 0_M$ and $(-1) \cdot m = -m$.

Example 19.1.1.

1. The 0 module is an R -module.
2. Take $M = R$, with the structure of an abelian group the same as that on R . We define $act_M := mult_R$. The module axioms follow from the ring axioms on R . Moreover, it follows that every field can be considered as a 1-dimensional vector space over itself. Additionally, when R is considered as a left module over itself in this fashion, its submodules are precisely its left ideals. If R is not commutative it has a left and right module structure over itself, and these may be different.
3. Take $M = R^{1 \times 2} := R \times R$. The abelian group structure is defined component wise, and so is the action of R .
4. Generalizing the previous example we can take $M = R^{1 \times n}$ for any positive integer n . In particular, for $n \in \mathbb{Z}^+$ we define

$$R^n \cong R^{1 \times n} = \{[a_1, a_2, \dots, a_n] : a_i \in R, \forall i\}$$

The module R^n is called the **free module of rank n over R** . A clear submodule of R^n is the i th component, in which arbitrary ring elements can exist in the i th component while zeros are in the j th component for all $j \neq i$.

5. If we replace the ring in the previous example with a field F , we obtain the **affine n -space over F** , F^n .

We note that if M is an R -module and S is a subring of R with $1_S = 1_R$ (if identity exists), then M is automatically an S -module.

Definition 19.1.4. If M is an R -module, and I is a two-sided ideal such that $a.m = 0$ for all $a \in I$ and all $m \in M$, then we say M is **annihilated** by I . In this case we can make M into an (R/I) -module by defining an action of the quotient ring R/I on M as follows: for each $m \in M$ and each coset $r + I \in R/I$, let

$$(r + I).m := r.m$$

Since $a.m = 0$ for all $a \in I$ and $m \in M$, this is well-defined. In particular, if I is a maximal ideal and R is a commutative ring, then M is a vector space over the field R/I .

Example 19.1.2 ((\mathbb{Z} -Modules)). Let $R = \mathbb{Z}$, let A be an abelian group and write the operation of A as $+$. We can make A into a \mathbb{Z} -module as follows: for any $n \in \mathbb{Z}$ and $a \in A$ define

$$n.a := \begin{cases} \underbrace{a + a + \dots + a}_{n\text{-times}} & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ \underbrace{(-a) + (-a) + \dots + (-a)}_{-n\text{-times}} & \text{if } n < 0 \end{cases}$$

where 0 is the identity of the additive group A . This definition makes A into a \mathbb{Z} -module, and by the module actions this is the only definition which makes A into a (unital) \mathbb{Z} -module. Thus, every abelian group is a \mathbb{Z} -module.

Conversely, if M is any \mathbb{Z} -module, a fortiori M is an abelian group, so

\mathbb{Z} – modules are the same as abelian groups

Furthermore, from the definition it is clear that

\mathbb{Z} – submodules are the same as subgroups

Example 19.1.3 (($F[x]$ -modules)). Let F be a field, let x be an indeterminate and let R be the polynomial ring $F[x]$. Let V be a vector space over F and let T be a linear transformation from V to V . We already know that V is an F -module; the linear map G will enable us to make V into an $F[x]$ -module.

First, for the nonnegative integer n , define

$$\begin{aligned} T^0 &:= I, \\ &\vdots \\ T^n &:= \underbrace{T \circ T \circ \dots \circ T}_{n\text{-times}} \end{aligned}$$

where I is the identity map from V to V and \circ denotes function composition. Also, for any two linear transformations A, B from V to V and elements $\alpha, \beta \in F$, let $\alpha A + \beta B$ be defined by

$$(\alpha A + \beta B)(v) := \alpha(A(v)) + \beta(B(v))$$

for all $v \in V$. Note that this is again a linear transformation from V to V .

Now let us define the action of any polynomial in x on V . Let $p(x) \in F[x]$, $p(x) = a_n x^n + \dots + a_1 x + a_0$, where $a_0, \dots, a_n \in F$. For each $v \in V$ we define an action for the ring element $p(x)$ on the module element v by

$$\begin{aligned} p(x).v &= (a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0)(v) \\ &= a_n T^n(v) + a_{n-1} T^{n-1}(v) + \dots + a_1 T(v) + a_0 v \end{aligned}$$

Put another way, x acts on V as the linear transformation T , and we extend this to an action of all of $F[x]$ on V , satisfying all the module axioms.

Note that the action of $F[x]$ on V is consistent with the original action of F on the vector space V when restricted to constant polynomials. This construction in fact describes all $F[x]$ -modules.

Moreover, there is a bijection between the collections of $F[x]$ -modules and the collection of pairs V, T :

$$\{V \text{ an } F[x]\text{-module}\} \leftrightarrow \left\{ \begin{array}{c} V \text{ a vector space over } F \\ \text{and} \\ T : V \rightarrow V \text{ a linear transformation} \end{array} \right\}$$

given by

the element x acts on V as the linear transformation T

Next, the $F[x]$ -submodules U of V are precisely the T -stable (or invariant) subspaces of V as seen with V as a vector space over F . We obtain a similar bijection as before:

$$\{W \text{ an } F[x]\text{-submodule}\} \leftrightarrow \left\{ \begin{array}{c} W \text{ a vector subspace of } V \\ \text{and} \\ W \text{ is } T\text{-stable} \end{array} \right\}$$

Proposition 19.1.2 ((The Submodule Criterion)). *Let R be a ring and let M be an R -module. A subset N of M is a submodule of M if and only if*

1. $N \neq \emptyset$, and
2. $x + r.y \in N$ for all $r \in R$ and for all $x, y \in N$

Proof. If N is a submodule, then $0 \in N$ so $N \neq \emptyset$. Also N is closed under addition and is stable under the action of elements of R , so $x + r.y \in N$ for all $r \in R$ and $x, y \in N$.

Conversely, suppose the two points hold. Let $m \in N$ since N is non-empty. Then $m + (-1).m = (1 + (-1)).m = 0.m = 0$, so $0 \in N$. Moreover, for all $m, n \in N$ we have that $m + (-n) = m + (-1).n \in N$ by hypothesis, so N is a subgroup of M . Now, we can take $x = 0$ and observe that for all $y \in N$ and all $r \in R$, $r.y = 0 + r.y \in R$, so N is stable under the action. Thus N is indeed a submodule of M . ■

Definition 19.1.5. Let R be a commutative ring with identity. An **R -algebra** is a ring A with identity together with a ring homomorphism $f : R \rightarrow A$ mapping 1_R to 1_A such that the subring $f(R)$ of A is contained in the center of A .

Observe that if A is an R -algebra, then A has a natural left and right (unital) R -module structure defined by $r \cdot a = a \cdot r = f(r)a$ where $f(r)a$ is just the multiplication in the ring A . Other R -module structures are possible on A , but this is the standard one.

Definition 19.1.6. If A and B are two R -algebras, an **R -algebra homomorphism (or isomorphism)** is a ring homomorphism (isomorphism, respectively) $\phi : A \rightarrow B$ mapping 1_A to 1_B such that $\phi(r \cdot a) = r \cdot \phi(a)$ for all $r \in R$ and $a \in A$.

Example 19.1.4. Let R be a commutative ring with 1.

1. Any ring with identity is a \mathbb{Z} -algebra.
2. For any ring A with identity, if R is a subring of the center of A containing the identity of A then A is an R -algebra.
3. If A is an R -algebra then the R -module structure of A depends only on the subring $f(R)$ contained in the center of A as in the previous example. If we identify R by its image $f(R)$ we see that “up to a ring homomorphism” every algebra A arises from a subring of the center of A that contains 1_A .

If A is an R -algebra, then A is a ring with identity that is a (unital) left R -module satisfying $r \cdot (ab) = (r \cdot a)b = a(r \cdot b)$ for all $r \in R$ and $a, b \in A$.

19.2.0 §Module Homomorphisms

Definition 19.2.1. Let M_1 and M_2 be R -modules. An **R -module homomorphism** from M_1 to M_2 is a map of sets $\phi : M_1 \rightarrow M_2$ such that

1. ϕ is a group homomorphism
2. For every $r \in R$ and $m_1 \in M_1$, we have $\phi(r \cdot m_1) = r \cdot \phi(m_1)$.

The last condition can be rewritten in terms of the following commutative diagram:

$$\begin{array}{ccc} R \times M_1 & \xrightarrow{\text{Id}_R \times \phi} & R \times M_2 \\ \text{act}_{M_1} \downarrow & & \downarrow \text{act}_{M_2} \\ M_1 & \xrightarrow{\phi} & M_2 \end{array}$$

We denote the set of R -module homomorphisms $M_1 \rightarrow M_2$ by $\mathbf{Hom}_{R\text{-Mod}}(M_1, M_2)$.

Definition 19.2.2 (D&F). Let R be a ring and let M and N be R -modules.

1. A map $\varphi : M \rightarrow N$ is an **R -module homomorphism** if it respects the R -module structures of M and N , i.e.,
 - (a) $\varphi(x + y) = \varphi(x) + \varphi(y)$, for all $x, y \in M$, and
 - (b) $\varphi(r \cdot_M x) = r \cdot_N \varphi(x)$, for all $r \in R, x \in M$.
2. An R -module homomorphism is an **isomorphism of R -modules** if it is both injective and surjective. The modules M and N are said to be **isomorphic**, denoted $M \cong N$, if there is some R -module isomorphism $\varphi : M \rightarrow N$.
3. If $\varphi : M \rightarrow N$ is an R -module homomorphism, let $\ker \varphi = \{m \in M \mid \varphi(m) = 0_N\}$ and let $\varphi(M) = \{n \in N \mid \exists m \in M; n = \varphi(m)\}$.
4. Let M and N be R -modules, and define $\mathbf{Hom}_R(M, N)$ to be the set of all R -module homomorphisms from M into N .

Remark 19.2.1. Give R -modules M_1, M_2, M_3 , and R -module homomorphisms $\phi : M_1 \rightarrow M_2$, $\psi : M_2 \rightarrow M_3$, the composed map

$$\psi \circ \phi : M_1 \rightarrow M_3 \quad (19.2.1)$$

is an R -module homomorphism. We can regard the operation of composition as a map of sets

$$\mathbf{Hom}_{R\text{-Mod}}(M_2, M_3) \times \mathbf{Hom}_{R\text{-Mod}}(M_1, M_2) \rightarrow \mathbf{Hom}_{R\text{-Mod}}(M_1, M_3) \quad (19.2.2)$$

Example 19.2.1.

1. If R is a ring and $M = R$ is a module over itself, then R -module homomorphisms need not be ring homomorphisms and ring homomorphisms need not be R module homomorphisms. For example, take $R = \mathbb{Z}$ and the \mathbb{Z} -module homomorphism $x \mapsto 2x$ (doesn't send 1 to 1). When $R = F[x]$, the ring homomorphism $\varphi : f(x) \mapsto f(x^2)$ is not an $F[x]$ -module homomorphism.
2. Let R be a ring, let $n \in \mathbb{Z}^+$ and let $M = R^n$. It is a straightforward exercise to show that for each $i \in \{1, 2, \dots, n\}$, the canonical projection map

$$\pi_i : R^n \rightarrow R \text{ by } \pi_i(x_1, \dots, x_n) = x_i$$

is a surjective R -module homomorphism with kernel equal to the submodule of n -tuples which have a zero in position i .

3. If R is a field, the R -module homomorphisms are called **linear transformations**.
4. For a ring $R = \mathbb{Z}$ the action of ring elements on any \mathbb{Z} -module amounts to adding and subtracting within the abelian group structure of the module, so \mathbb{Z} -module homomorphisms are the same as abelian group homomorphisms.

Proposition 19.2.1. Let M, N , and L be R -modules.

1. A map $\varphi : M \rightarrow N$ is an R -module homomorphism if and only if $\varphi(rx+y) = r\varphi(x) + \varphi(y)$ for all $x, y \in M$ and all $r \in R$.
2. Let $\varphi, \psi \in \mathbf{Hom}_R(M, N)$. Define $\varphi + \psi$ by

$$(\varphi + \psi)(m) = \varphi(m) + \psi(m), \forall m \in M$$

Then $\varphi + \psi \in \mathbf{Hom}_R(M, N)$ and with this operation $\mathbf{Hom}_R(M, N)$ is an abelian group. If R is a commutative ring, then for $r \in R$ define $r\varphi$ by

$$(r\varphi)(m) = r(\varphi(m)), \forall m \in M$$

Then $r\varphi \in \mathbf{Hom}_R(M, N)$ and with this action of the commutative ring R the abelian group $\mathbf{Hom}_R(M, N)$ is an R -module.

3. If $\varphi \in \mathbf{Hom}_R(L, M)$ and $\psi \in \mathbf{Hom}_R(M, N)$ then $\psi \circ \varphi \in \mathbf{Hom}_R(L, N)$.
4. With addition as above and multiplication defined as function composition, $\mathbf{Hom}_R(M, M)$ is a ring with 1. When R is commutative $\mathbf{Hom}_R(M, M)$ is an R -algebra.

Proof. (1): Let $\varphi : M \rightarrow N$ be a map. If φ is an R -module homomorphism then for all $x, y \in M$ and all $r \in R$,

$$\varphi(rx + y) = \varphi(rx) + \varphi(y) = r\varphi(x) + \varphi(y)$$

Conversely, suppose this holds for all $r \in R$ and $x, y \in M$. First, take $r = 1$, so

$$\varphi(x + y) = \varphi(1 \cdot x + y) = 1 \cdot \varphi(x) + \varphi(y) = \varphi(x) + \varphi(y)$$

Additionally, if we take $y = 0_M$, then

$$\varphi(rx) = \varphi(rx + 0_M) = r\varphi(x) + \varphi(0_M) = r\varphi(x) + 0_N = r\varphi(x)$$

Thus φ is indeed an R -module homomorphism.

(2): Let $\varphi, \psi \in \mathbf{Hom}_R(M, N)$. Then observe that for all $r \in R, x, y \in M$, we have

$$\begin{aligned} (\varphi + \psi)(rx + y) &= \varphi(rx + y) + \psi(rx + y) \\ &= (r\varphi(x) + \varphi(y)) + (r\psi(x) + \psi(y)) \\ &= r(\varphi(x) + \psi(x)) + (\varphi(y) + \psi(y)) \\ &= r(\varphi + \psi)(x) + (\varphi + \psi)(y) \end{aligned}$$

so by result (1) $\varphi + \psi \in \mathbf{Hom}_R(M, N)$. Since N is an abelian group $\varphi + \psi = \psi + \varphi$, and as -1 is in the center of R , $-\varphi \in \mathbf{Hom}_R(M, N)$ where $\varphi + (-\varphi) = \mathbf{0}$, where $\mathbf{0}$ is the trivial map $m \mapsto 0_N$ for all $m \in M$. If R is commutative, then for all $r_1, r_2 \in R$, and all $x, y \in M$,

$$\begin{aligned} (r_1\varphi)(r_2x + y) &= r_1\varphi(r_2x + y) \\ &= r_1(r_2\varphi(x) + \varphi(y)) \\ &= r_1r_2\varphi(x) + r_1\varphi(y) \\ &= r_2(r_1\varphi)(x) + (r_1\varphi)(y) \end{aligned}$$

so $r_1\varphi \in \mathbf{Hom}_R(M, N)$, and by the R -module structure of N we conclude that $\mathbf{Hom}_R(M, N)$ is an R -module under this action.

(3): Let $\varphi \in \mathbf{Hom}_R(L, M)$ and $\psi \in \mathbf{Hom}_R(M, N)$, and let $r \in R, x, y \in L$. Then it follows that

$$\begin{aligned} (\psi \circ \varphi)(rx + y) &= \psi(\varphi(rx + y)) \\ &= \psi(r\varphi(x) + \varphi(y)) \\ &= r\psi(\varphi(x)) + \psi(\varphi(y)) \\ &= r(\psi \circ \varphi)(x) + (\psi \circ \varphi)(y) \end{aligned}$$

so $\psi \circ \varphi \in \mathbf{Hom}_R(L, N)$.

(4): Since the domain and codomain of $\mathbf{Hom}_R(M, M)$ are the same, composition is well defined, and by (3) it is a binary operation on $\mathbf{Hom}_R(M, M)$. Moreover, function composition is associative and Id_M the identity map on M acts as a 1 in $\mathbf{Hom}_R(M, M)$ under composition. Additionally by (2) $\mathbf{Hom}_R(M, M)$ is an abelian group under addition. Then, to prove distributivity let $r, r' \in R$ and $\varphi, \psi \in \mathbf{Hom}_R(M, M)$. Then for all $x \in M$ we have that

$$(r \cdot (\varphi + \psi))(x) = r(\varphi(x) + \psi(x)) = r\varphi(x) + r\psi(x) = (r \cdot \varphi + r \cdot \psi)(x)$$

and

$$((r + r') \cdot \varphi)(x) = r\varphi(x) + r'\varphi(x) = (r \cdot \varphi + r' \cdot \varphi)(x)$$

so indeed $r \cdot (\varphi + \psi) = r \cdot \varphi + r \cdot \psi$ and $(r + r') \cdot \varphi = r \cdot \varphi + r' \cdot \varphi$.

Finally, if R is commutative then $\mathbf{Hom}_R(M, M)$ has an R -module structure. Moreover, if we define $r\varphi = \varphi r$ for all $\varphi \in \mathbf{Hom}_R(M, M)$ and $r \in R$, then we observe that for all $\varphi, \psi \in \mathbf{Hom}_R(M, M)$ and $x \in M$

$$\begin{aligned} (r \cdot (\varphi \circ \psi))(x) &= r(\varphi \circ \psi)(x) &= r(\varphi(\psi(x))) &= r\varphi(\psi(x)) \\ &= \varphi(r\psi(x)) &= \varphi((r \cdot \psi)(x)) &= (\varphi \circ (r \cdot \psi))(x) \end{aligned}$$

so $\mathbf{Hom}_R(M, M)$ becomes an R -algebra. ■

Definition 19.2.3. The ring $\mathbf{Hom}_R(M, M)$ is called the **endomorphism ring** of M and will often be denoted by $\mathbf{End}_R(M)$, or just $\mathbf{End}(M)$ when the ring R is clear from context. Elements of $\mathbf{End}(M)$ are called **endomorphisms**.

We now wish to show that we can assign an R -module structure to quotients of R -modules with submodules:

Proposition 19.2.2. Let R be a ring, let M be an R -module and let N be a submodule of M . The (additive, abelian) quotient group M/N can be made into an R -module by defining an action of elements of R by:

$$r \cdot (x + N) = (r \cdot x) + N, \forall r \in R, x + N \in M/N$$

The natural projection map $\pi : M \rightarrow M/N$ defined by $\pi(x) = x + N$ is an R -module homomorphism with kernel N .

Proof. Let R be a ring, M an R -module, and N a submodule of M . Since M is an abelian group N is a normal subgroup of M so the quotient group M/N is well defined. Now, define an action of elements of R on M as above. Let $r, r' \in R$ and $x + N, y + N \in M/N$. Then it follows that

$$r \cdot (r' \cdot (x + N)) = r \cdot ((r' \cdot x) + N) = r \cdot (r' \cdot x) + N = (rr') \cdot x + N = (rr') \cdot (x + N)$$

$$\begin{aligned} (r + r') \cdot (x + N) &= (r + r') \cdot x + N \\ &= (r \cdot x + r' \cdot x) + N \\ &= (r \cdot x + N) + (r' \cdot x + N) \\ &= r \cdot (x + N) + r' \cdot (x + N) \end{aligned}$$

and

$$\begin{aligned} r \cdot ((x + N) + (y + N)) &= r \cdot ((x + y) + N) \\ &= r \cdot (x + y) + N \\ &= (r \cdot x + r \cdot y) + N \\ &= (r \cdot x + N) + (r \cdot y + N) \\ &= r \cdot (x + N) + r \cdot (y + N) \end{aligned}$$

Therefore, the action is a well defined module action on M/N , so M/N is an R -module under this action.

Consider the natural projection map $\pi : M \rightarrow M/N$. Recall that π is an abelian group homomorphism between M and M/N . To show that it is indeed an R -module homomorphism, let $r \in R$ and $m \in M$. Then

$$\pi(r \cdot m) = r \cdot m + N = r \cdot (m + N) = r \cdot \pi(m)$$

proving that π is an R -module homomorphism as desired. Finally, $\ker \pi = \{x \in M : x + N = n\} = \{x \in M : x \in N\} = N$. ■

Definition 19.2.4. Let A, B be submodules of the R -module M . The sum of A and B is the set

$$A + B := \{a + b \mid a \in A, b \in B\}$$

This is indeed a submodule, and in fact the small submodule containing both A and B .

§Isomorphism Theorems for Modules

Theorem 30 (The First Isomorphism Theorem for Modules).

Let M and N be R -modules and let $\varphi : M \rightarrow N$ be an R -module homomorphism. Then $\ker \varphi$ is a submodule of M and $M/\ker \varphi \cong \varphi(M)$.

Proof. First, observe that $0_M \in \ker \varphi$ as $\varphi(0_M) = 0_N$ since φ is an abelian group homomorphism over M . Now, let $m, n \in \ker \varphi$. Then

$$\varphi(m + (-n)) = \varphi(m) + \varphi(-n) = 0_N + (-\varphi(n)) = -0_N = 0_N$$

so $m + (-n) \in \ker \varphi$. Hence, $\ker \varphi$ is indeed a subgroup of M . Now, let $r \in R$. Then

$$\varphi(r \cdot m) = r \cdot \varphi(m) = r \cdot 0_N = 0_N$$

so $r \cdot m \in \ker \varphi$, and we conclude that $\ker \varphi$ is indeed a submodule of M .

Hence, $M/\ker \varphi$ is an R -module, as by our previous results. Now, define a map $\bar{\varphi} : M/\ker \varphi \rightarrow N$ by $\bar{\varphi}(x + \ker \varphi) = \varphi(x)$. Then, if $x + \ker \varphi = y + \ker \varphi$, there exists $m \in \ker \varphi$ such that $x = y + m$. In particular,

$$\bar{\varphi}(x + \ker \varphi) = \varphi(x) = \varphi(y + m) = \varphi(y) + \varphi(m) = \bar{\varphi}(y + \ker \varphi)$$

so the map is indeed well-defined. Furthermore, $\ker(\bar{\varphi}) = \{x + \ker \varphi \in M/\ker \varphi : x \in \ker \varphi\} = \{0_{M/\ker \varphi}\}$, so $\bar{\varphi}$ is injective. Moreover, restricting the codomain to the image $\varphi(M)$, the map is also surjective. Finally, since φ is an R -module homomorphism and $\bar{\varphi}$ is defined in terms of φ , it is also an R -module homomorphism. Hence, $\bar{\varphi}$ is an R -module isomorphism and $M/\ker \varphi \cong \varphi(M)$. ■

Theorem 31 (The Second Isomorphism Theorem for Modules).

Let A and B be submodules of the R -module M . Then $(A + B)/B \cong A/(A \cap B)$.

Proof. Suppose A and B are submodules of the R -module M . Then I claim $A + B$ and $A \cap B$ are submodules of M . Indeed, $A + B$ and $A \cap B$ are nonempty, as they contain 0_M , and for all $a + b, a' + b' \in A + B$, $k, k' \in A \cap B$, and $r \in R$,

$$r(a + b) + (a' + b') = (ra + rb) + (a' + b') = (ra + a') + (rb + b') \in A + B$$

and $kr + k' \in A \cap B$ since $kr + k' \in A$ and $kr + k' \in B$ as they are submodules. Now, define a map $\varphi : A \rightarrow (A + B)/B$ by $\varphi(a) = a + B$. This map is indeed well defined as $a = a + 0_M \in A + B$, and for all $x, y \in A$ and $r \in R$,

$$\varphi(rx + y) = (rx + y) + B = (rx + B) + (y + B) = r(x + B) + (y + B) = r\varphi(x) + \varphi(y)$$

so φ is an R -module homomorphism. Furthermore, $\ker \varphi = \{a \in A : a \in B\} = A \cap B$. Thus, by 30 we conclude that $A/(A \cap B) \cong (A + B)/B$, as desired. ■

Theorem 32 (The Third Isomorphism Theorem for Modules).

Let M be an R -module, and let A and B be submodules of M with $A \subseteq B$. Then $(M/A)/(B/A) \cong M/B$.

Proof. Define a map $f : M/A \rightarrow M/B$ by $f(x + A) = x + B$. Then, if $x + A = y + A$ we have that $x = y + a$ for some $a \in A$, so as $A \subseteq B$

$$f(x + A) = x + B = (y + a) + B = y + B$$

since $a \in B$. Hence f is well defined. Moreover, for all $x + A, y + A \in M/A$ and $r \in R$,
 $f(r(x + A) + (y + A)) = f(rx + y + A) = rx + y + B = r(x + B) + (y + B) = rf(x + A) + f(y + A)$
 so f is an R -module homomorphism. Now, observe that

$$\ker f = \{x + A \in M/A : x \in B\} = B/A$$

so by 30 we conclude that

$$(M/A)/(B/A) \cong M/B$$

■

Theorem 33 (The Fourth of Lattice Isomorphism Theorem).

Let N be a submodule of the R -module M . There is a bijection between the submodules of M which contain N and the submodules of M/N . The correspondence is given by $A \leftrightarrow A/N$, for all $A \supseteq N$. This correspondence commutes with the processes of taking sums and intersections.

Proof. (To be finished)

■

§Evaluation Bijections

Definition 19.2.5. Let M be an arbitrary R -module. Consider the set $\mathbf{Hom}_{R\text{-Mod}}(R, M)$, where R is considered as an R -module. We define the map of sets

$$\begin{aligned} \mathbf{ev} : \mathbf{Hom}_{R\text{-Mod}}(R, M) &\rightarrow M \\ \phi &\mapsto \phi(1) \in M \end{aligned} \quad (19.2.3)$$

\mathbf{ev} as defined is a bijection of sets. That is, to give a map of modules $R \rightarrow M$ is the same as to give an element of M .

Definition 19.2.6. Generalizing the previous definition we obtain the bijection

$$\begin{aligned} \mathbf{ev} : \mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n}, M) &\rightarrow M^{1 \times n} \\ \phi &\mapsto (\phi(1, 0, \dots, 0), \phi(0, 1, \dots, 0), \dots, \phi(0, 0, \dots, 1)) \in M^{1 \times n} \end{aligned} \quad (19.2.4)$$

Remark 19.2.2. In particular, taking $M = R^{1 \times m}$, we obtain a bijection

$$\mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n}, R^{1 \times m}) \xrightarrow{\mathbf{ev}} (R^{1 \times m})^{1 \times n} \cong \text{Mat}_{m \times n}(R) \quad (19.2.5)$$

In particular, for the composition map

$$\mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n_2}, R^{1 \times n_3}) \times \mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n_1}, R^{1 \times n_2}) \xrightarrow{\text{comp}} \mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n_1}, R^{1 \times n_3}) \quad (19.2.6)$$

we obtain the commutative diagram

$$\begin{array}{ccc} \mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n_2}, R^{1 \times n_3}) \times \mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n_1}, R^{1 \times n_2}) & \xrightarrow{\text{comp}} & \mathbf{Hom}_{R\text{-Mod}}(R^{1 \times n_1}, R^{1 \times n_3}) \\ \mathbf{ev} \times \mathbf{ev} \downarrow & & \downarrow \mathbf{ev} \\ \text{Mat}_{n_3 \times n_2}(R) \times \text{Mat}_{n_2 \times n_1}(R) & \xrightarrow{\text{mult}_{\text{mat}}} & \text{Mat}_{n_3 \times n_1}(R) \end{array}$$

19.3.0 §Submodules

19.4.0 §Free Modules and Generators

In section we assume R is a ring with 1.

Definition 19.4.1. Let M be an R -module and let N_1, \dots, N_n be submodules of M .

1. The **sum** of N_1, \dots, N_n is the set of all finite sums of elements from the sets N_i : $\{\sum_{i=1}^n a_i \mid a_i \in N_i, \forall i\}$. Denote this sum by $N_1 + \dots + N_n$.
2. For any subset A of M let

$$RA := \left\{ \sum_{i=1}^m r_i a_i \mid r_1, \dots, r_m \in R, a_1, \dots, a_m \in A, m \in \mathbb{Z}^+ \right\}$$

(where by convention $RA = \{0\}$ if $A = \emptyset$). If A is the finite set $\{a_1, a_2, \dots, a_n\}$ we shall write $Ra_1 + Ra_2 + \dots + Ra_n$ for RA . Call RA the **submodule of M generated by A** . If N is a submodule of M and $N = RA$ for some subset A of M , we call A a **set of generators** or **generating set** for N , and we say N is **generated** by A .

3. A submodule N of M is **finitely generated** if there is some finite subset A of M such that $N = RA$, that is, if N is generated by some finite subset.
4. A submodule N of M is **cyclic** if there exists an element $a \in M$ such that $N = Ra$, that is, if N is generated by one element:

$$N = Ra = \{ra \mid r \in R\}$$

Note that these definitions do not require that the ring R contain a 1, however this condition ensures that A is contained in RA .

Remark 19.4.1. Let N be a submodule of an R -module M which is finitely generated. Then there is a smallest nonnegative integer d such that N is generated by d elements. We then call any generating set consisting of d elements a **minimal set of generators for N** .

Definition 19.4.2. Let M_1, \dots, M_k be a collection of R -modules. The collection of k -tuples (m_1, m_2, \dots, m_k) where $m_i \in M_i$ with addition and action of R defined componentwise is called the **direct product** of M_1, \dots, M_k , denoted $M_1 \times \dots \times M_k$.

The direct product of M_1, \dots, M_k is also referred to as the (external) **direct sum** of M_1, \dots, M_k and denoted $M_1 \oplus \dots \oplus M_k$.

Proposition 19.4.1. Let N_1, \dots, N_k be submodules of the R -module M . Then the following are equivalent:

1. The map $\pi : N_1 \times N_2 \times \dots \times N_k \rightarrow N_1 + N_2 + \dots + N_k$ defined by

$$\pi(a_1, a_2, \dots, a_k) = a_1 + a_2 + \dots + a_k$$

is an isomorphism of R -modules: $N_1 + N_2 + \dots + N_k \cong N_1 \times N_2 \times \dots \times N_k$.

2. $N_j \cap (N_1 + \dots + N_{j-1} + N_{j+1} + \dots + N_k) = \{0\}$ for all $j \in \{1, 2, \dots, k\}$.

3. Every $x \in N_1 + \dots + N_k$ can be written uniquely in the form $a_1 + a_2 + \dots + a_k$ with $a_i \in N_i$.

Proof. (1) \implies (2). Let $a_j \in N_j \cap (N_1 + \dots + N_{j-1} + N_{j+1} + \dots + N_k)$. Then

$$a_j = a_1 + \dots + a_{j-1} + a_{j+1} + \dots + a_k$$

for some $a_i \in N_i$, so $(a_1, \dots, a_{j-1}, -a_j, a_{j+1}, \dots, a_k) \in \ker \pi$, but π is an isomorphism so $\ker \pi = \{(0, \dots, 0)\}$ which implies $a_i = 0$ for each i . In particular, $a_j = 0$, so $N_j \cap (N_1 + \dots + N_{j-1} + N_{j+1} + \dots + N_k) = \{0\}$.

(2) \implies (3). Suppose $\sum_{i=1}^k a_i = \sum_{i=1}^k b_i$ for $a_i, b_i \in N_i$. Then we have that for each j ,

$$a_j - b_j = \sum_{i=1}^{j-1} (b_i - a_i) + \sum_{i=j+1}^k (b_i - a_i) \in N_j \cap (N_1 + \dots + N_{j-1} + N_{j+1} + \dots + N_k)$$

so by hypothesis, $a_j - b_j = 0$, so $a_j = b_j$. As this holds for all $j \in \{1, \dots, k\}$, the expression $\sum_{i=1}^k a_i$ is unique.

(3) \implies (1). Suppose $(a_1, \dots, a_k) \in \ker \pi$. Then $\sum_{i=1}^k a_i = 0 = \sum_{i=1}^k 0$, where $a_i \in N_i$ and $0 \in N_i$. Thus, by hypothesis on the uniqueness of the expression of elements in terms of sums of elements of the N_i , we have that $a_i = 0$ for each i . Thus, $\ker \pi = \{(0, \dots, 0)\}$, so π is injective. Thus, as π is a surjective R -module homomorphism by construction, it is also an R -module isomorphism, completing the proof. ■

If an R -module $M = N_1 + \dots + N_k$ is the sum of submodules N_1, \dots, N_k of M satisfying the equivalent conditions of the proposition above, then M is said to be the (internal) direct sum of N_1, \dots, N_k , written

$$M = N_1 \oplus \dots \oplus N_k$$

Definition 19.4.3. An R -module F is said to be free on the subset A of F if for every nonzero element x of F , there exists unique nonzero elements r_1, \dots, r_n of R and unique a_1, \dots, a_n of A such that $x = \sum_{i=1}^n r_i a_i$, for some $n \in \mathbb{Z}^+$. In this situation, we say A is a basis or set of free generators for F . If R is a commutative ring the cardinality of A is called the rank of F .

Theorem 19.4.2. For any set A there is a free R -module $F(A)$ on the set A and $F(A)$ satisfies the following universal property: if M is any R -module and $\phi : A \rightarrow M$ is any map of sets, then there is a unique R -module homomorphism $\Phi : F(A) \rightarrow M$ such that $\Phi(a) = \phi(a)$ for all $a \in A$, that is, the followign diagram commutes:

$$\begin{array}{ccc}
 A & \xrightarrow{\iota} & F(A) \\
 & \searrow \forall \phi & \downarrow \exists! \Phi \\
 & & \forall M
 \end{array}$$

when A is the finite set $\{a_1, a_2, \dots, a_n\}$, $F(A) \cong Ra_1 \oplus Ra_2 \oplus \dots \oplus Ra_n \cong R^n$.

Proof. Let $F(A) = \{0\}$ if $A = \emptyset$. If A is nonempty, let $F(A)$ be the collection of set functions $f : A \rightarrow R$ such that $f(a) = 0$ for all but finitely many $a \in A$ (i.e. finite support). Make $F(A)$ into an R -module by pointwise addition of functions and pointwise multiplication of a ring element times a function, i.e.,

$$\begin{aligned}
 (f + g)(a) &:= f(a) + g(a) \\
 (rf)(a) &:= r(f(a))
 \end{aligned}$$

for all $a \in A$, $r \in R$, and $f, g \in F(A)$. Since R is a ring addition in $F(A)$ is both associative and commutative, with additive identity the 0 map, and for every $f : A \rightarrow R$, an additive inverse $-f$. Associativity of the R action and distributivity follow from associativity and distributivity of multiplication in the ring R . Thus $F(A)$ is indeed a (left) R -module. Identify A as a subset of $F(A)$ by $a \mapsto f_a$, where f_a is the function which is 1 at a and zero elsewhere. We can, in this way, think of $F(A)$ as all finite R -linear combinations of elements of A by identifying each function with the sum $\sum_{i=1}^n r_i a_i$, where f takes on the value r_i at a_i and is zero at all other elements of A . Moreover, each element of $F(A)$ has a unique expression as such a formal sum. To establish the universal property of $F(A)$ suppose $\phi : A \rightarrow M$ is a map of the set A into the R -module M . Define $\Phi : F(A) \rightarrow M$ by

$$\Phi : \sum_{i=1}^n r_i a_i \mapsto \sum_{i=1}^n r_i \phi(a_i)$$

By the uniqueness of the expression for the elements of $F(A)$ as linear combinations of the a_i , we see that Φ is a well defined R -module homomorphism. By definition, the restriction $\Phi|_A = \phi$. Finally, since $F(A)$ is generated by A , once we know the values of an R -module homomorphism on A its values on every element of $F(A)$ are uniquely determined, so Φ is the unique extension of ϕ to all of $F(A)$. ■

Corollary 19.4.3.

1. If F_1 and F_2 are free modules on the same set A , there is a unique isomorphism between F_1 and F_2 which is the identity map on A .
2. If F is any free R -module with basis A , then $F \cong F(A)$. In particular, F enjoys the same universal property with respect to A as $F(A)$.

Proof. Let F be a free R -module with basis A . Then I shall show F enjoys the same universal property with respect to A as $F(A)$. Let M be an R module and suppose $\phi : A \rightarrow M$ is a

map of sets. Define $\Phi : F \rightarrow M$ by $\sum_{i=1}^n r_i a_i \mapsto \sum_{i=1}^n r_i \varphi(a_i)$ for all $\sum_{i=1}^n r_i a_i \in F$. Then by construction Φ is an R -linear map satisfying $\Phi \circ \iota = \varphi$. If $f : F \rightarrow M$ is another map satisfying this property, then $f(a) = \varphi(a)$ for all $a \in A$, so by linearity

$$f\left(\sum_{i=1}^n r_i a_i\right) = \sum_{i=1}^n r_i f(a_i) = \sum_{i=1}^n r_i \varphi(a_i) = \Phi\left(\sum_{i=1}^n r_i a_i\right)$$

so $f = \Phi$. Hence the map is unique and F satisfies the universal property.

Then, If we consider $\iota_F : A \rightarrow F$ the inclusion map, by the universal property there exists a unique R -linear map $\Phi_F : F(A) \rightarrow F$ making the diagram commute. Similarly, if $\iota_{F(A)} : A \rightarrow F(A)$ is the other inclusion map, again by the universal property there exists a unique R -linear map $\Phi_{F(A)} : F \rightarrow F(A)$ making the diagram commute. Now, it follows that $\Phi_{F(A)} \circ \Phi_F : F(A) \rightarrow F(A)$ makes the diagram for $F(A)$ commute, but so does $\text{Id}_{F(A)}$, so by uniqueness $\Phi_{F(A)} \circ \Phi_F = \text{Id}_{F(A)}$. Dually, $\Phi_F \circ \Phi_{F(A)} = \text{Id}_F$, so Φ_F and $\Phi_{F(A)}$ are inverse module isomorphisms. Therefore $F \cong F(A)$, and in particular, for any free modules F_1 and F_2 on A , $F_1 \cong F(A) \cong F_2$, so $F_1 \cong F_2$, and the isomorphism $\Phi : F_1 \rightarrow F_2$ is such that $\Phi \circ \iota_{F_1} = \iota_{F_2}$, so Φ is the identity restricted to A , as desired. ■

Chapter 20

§§Linear Transformations

Chapter 21

§§Matrix Theory for Free Modules

Chapter 22

§§Modules over PIDs

Chapter 23

§§Tensor Products

23.1.0 §Module Tensor Products

In this section we consider modules over unital rings (not necessarily commutative).

Motivation/Special Case

Let R be a subring of a ring S . We assume $1_R = 1_S$.

If N is a left S -module then N is also a left R -module since elements of R , being elements of S , act on N by assumption. Note that by the axioms of a left S -module $(s_1 s_2)n = s_1(s_2 n)$ for all $s_1, s_2 \in S$ and all $n \in N$, so in particular $(sr)n = s(rn)$ for all $s \in S, r \in R$ and $n \in N$.

Definition 23.1.1. Let R and S be rings and $f : R \rightarrow S$ a ring homomorphism with $f(1_R) = 1_S$. Then if N is an S -module, it can be considered as an R -module with the action $r \cdot n = f(r)n$ for all $r \in R$ and $n \in N$. In this case S is called an extension of the ring R , and the resulting R -module is said to be obtained from N by restriction of scalars from S to R .

We wish to now construct for a general R -module N an S -module that is the “best possible” target in which to try to embed N .

Construction 23.1.2. Let N be an R -module. To endow an S -module structure on N we first need a map from $S \times N$ to N . We consider the \mathbb{Z} -module structure on the set $S \times N$, considering it as the collection of all finite commuting sums of elements of the form $(s_i, n_i) \in S \times N$. This is an abelian group where there are no relations between any distinct pairs (s, n) and (s', n') , which we view as “formal products” $s \cdot n$.

Note as an S -module we must have the following relations satisfied:

$$\begin{aligned}(s_1 + s_2)n &= s_1 n + s_2 n \\ s(n_1 + n_2) &= s n_1 + s n_2\end{aligned}$$

$$(s_1 s_2)n = s_1(s_2 n)$$

for all $s_1, s_2, s \in S$ and $n_1, n_2, n \in N$. In particular, we need the relation $(sr)n = s(rn)$ to hold for all $s \in S$, $r \in R$, and $n \in N$ since $R \subseteq S$. To induce this relation on the abelian group of formal products, we take the quotient by the subgroup H generated by all elements of the form

$$\begin{aligned} (s_1 + s_2, n) - (s_1, n) - (s_2, n), \\ (s, n_1 + n_2) - (s, n_1) - (s, n_2), \\ (sr, n) - (s, rn) \end{aligned}$$

for all $s, s_1, s_2 \in S, n, n_1, n_2 \in N$ and $r \in R$, where rn is the element obtained by the R -module structure already on N .

The resulting quotient group is denoted by $S \otimes_R N$, and is called the **tensor product of S and N over R** . If $s \otimes_R n$ denotes the coset containing (s, n) in $S \otimes_R N$, then by definition of the quotient we have forced the relations

$$\begin{aligned} (s_1 + s_2) \otimes_R n &= s_1 \otimes_R n + s_2 \otimes_R n \\ s \otimes_R (n_1 + n_2) &= s \otimes_R n_1 + s \otimes_R n_2 \\ sr \otimes_R n &= s \otimes_R rn \end{aligned}$$

The elements of $S \otimes_R N$ are called **tensors** and can be written (non-uniquely in general) as finite sums of **simple tensors** of the form $s \otimes n$ for $s \in S$ and $n \in N$.

We now show that $S \otimes_R N$ is a left S -module under the action defined by:

$$s \cdot \left(\sum_{finite} s_i \otimes_R n_i \right) := \sum_{finite} (ss_i) \otimes_R n_i$$

We first check that this is well defined. First note that if $s' \in S$, then

$$\begin{aligned} (s'(s_1 + s_2), n) - (s's_1, n) - (s's_2, n) \\ (s's, n_1 + n_2) - (s's, n_1) - (s's, n_2) \\ (s'(sr), n) - (s's, rn) \end{aligned}$$

each belong to the set of generators previously defined, and hence belong to the subgroup H . This shows that multiplying generators of H on the left by s' gives another element of H . Since any element of H is a sum of these elements, it follows that for any element $\sum(s_i, n_i) \in H$, $\sum(s's_i, n_i) \in H$. Suppose now that $\sum s_i \otimes_R n_i = \sum s'_i \otimes_R n'_i$ are two representations for the same element in $S \otimes_R N$. Then $\sum(s_i, n_i) - \sum(s'_i, n'_i) \in H$, and by our previous analysis, for any $s \in S$, $\sum(ss_i, n_i) - \sum(ss'_i, n'_i) \in H$, which implies that $\sum ss_i \otimes_R n_i = \sum ss'_i \otimes_R n'_i$ in $S \otimes_R N$, so the action is well defined.

It now follows routinely that $S \otimes_R N$ is a left S -module. Indeed, one axiom of S -modules is shown below:

$$\begin{aligned} (s + s') \sum s_i \otimes_R n_i &= \sum ((s + s')s_i) \otimes_R n_i \\ &= \sum (ss_i + s's_i) \otimes_R n_i \\ &= \sum ss_i \otimes_R n_i + \sum s's_i \otimes_R n_i \\ &= s \sum s_i \otimes_R n_i + s' \sum s_i \otimes_R n_i \end{aligned}$$

Next, there is a natural map $\iota : N \rightarrow S \otimes_R N$ defined by $n \mapsto 1_S \otimes_R n$. Since $1_S \otimes_R rn = r \otimes_R n = r(1_S \otimes n)$, ι is indeed an R -module homomorphism from N to $S \otimes_R N$. Since $S \otimes_R N$ is a quotient module, this map need not be injective.

Theorem 23.1.1. *Let R be a subring of S , let N be a left R -module and let $\iota : N \rightarrow S \otimes_R N$ be the R -module homomorphism defined by $\iota(n) = 1_S \otimes_R n$. Suppose that L is any left S -module (hence also an R -module) and that $\varphi : N \rightarrow L$ is an R -module homomorphism. Then there is a unique S -module homomorphism $\Phi : S \otimes_R N \rightarrow L$ such that φ factors through Φ , i.e., $\varphi = \Phi \circ \iota$ and the diagram:*

$$\begin{array}{ccc} N & \xrightarrow{\iota} & S \otimes_R N \\ & \searrow \varphi & \downarrow \Phi \\ & & L \end{array}$$

commutes. Conversely, if $\Phi : S \otimes_R N \rightarrow L$ is an S -module homomorphism then $\varphi = \Phi \circ \iota$ is an R -module homomorphism from N to L .

Proof. Suppose $\varphi : N \rightarrow L$ is an R -module homomorphism to the S -module L . By the universal property of free modules there is a \mathbb{Z} -module homomorphism from the free \mathbb{Z} -module F on the set $S \times N$ to L that sends each generator (s, n) to $s\varphi(n)$. Since φ is an R -module homomorphism, the generators of the subgroup H in our construction all map to zero in L . Hence this \mathbb{Z} -module homomorphism factors through H , so there exists a well-defined \mathbb{Z} -module homomorphism Φ from $F/H = S \otimes_R N$ to L satisfying $\Phi(s \otimes_R n) = s\varphi(n)$. Moreover, on simple tensors we have

$$s'\Phi(s \otimes_R n) = s'(s\varphi(n)) = (s's)\varphi(n) = \Phi((s's) \otimes_R n) = \Phi(s'(s \otimes_R n))$$

Since Φ is additive it follows that Φ is an S -module homomorphism, which proves the existence statement. The module $S \otimes_R N$ is generated as an S -module by elements of the form $1_S \otimes_R n$, so any S -module homomorphism is uniquely determined by its values on these elements. Since $\Phi(1_S \otimes_R n) = \varphi(n)$, it follows that the S -module homomorphism Φ is uniquely determined by φ . The converse is immediate. ■

Corollary 23.1.2. *Let $\iota : N \rightarrow S \otimes_R N$ be the R -module homomorphism from the previous theorem. Then $N/\ker \iota$ is the unique largest quotient of N that can be embedded in any S -module. In particular, N can be embedded as an R -submodule of some left S -module if and only if ι is injective.*

Proof. By the first isomorphism theorem the quotient $N/\ker \iota$ is mapped injectively into the S -module $S \otimes_R N$. Suppose now that φ is an R -module homomorphism injecting the quotient $N/\ker \varphi$ of N into an S -module L . Then, by the previous theorem $\ker \iota$ is mapped to 0 by φ , that is $\ker \iota \subseteq \ker \varphi$. Hence $N/\ker \varphi$ is a quotient of $N/\ker \iota$ (namely, the quotient by the submodule $\ker \varphi/\ker \iota$). It follows that $N/\ker \iota$ is the unique largest quotient of N that can be embedded in any S -module. The last statement follows immediately. ■

General Construction

Now, note that in the construction of $S \otimes_R N$ as an *abelian group*, only the elements in the generating relations were involved, which in turn implies we only required S to be a *right* R -module and N to be a *left* R -module. In a similar way we can construct an abelian group $M \otimes_R N$ for any right R -module M and any left R -module N .

Secondly, observe that the S -module structure on $S \otimes_R N$ defined previously required only a *left* S -module structure on S together with a *compatibility relation*:

$$s'(sr) = (s's)r \quad \forall s, s' \in S, \forall r \in R$$

We shall first proceed with the general abelian group construction of $M \otimes_R N$ before returning to the module construction:

Construction 23.1.3. *Let N be a left R -module and M a right R -module. The quotient of the free \mathbb{Z} -module on the set $M \times N$ by the subgroup generated by all elements of the form*

$$\begin{aligned} (m_1 + m_2, n) - (m_1, n) - (m_2, n) \\ (m, n_1 + n_2) - (m, n_1) - (m, n_2) \\ (mr, n) - (m, rn) \end{aligned}$$

*for all $m, m_1, m_2 \in M, n, n_1, n_2 \in N$ and $r \in R$ is an abelian group, denoted $M \otimes_R N$, is called the **tensor product of M and N over R** . From this construction we have the following relations:*

$$\begin{aligned} (m_1 + m_2) \otimes_R n &= m_1 \otimes_R n + m_2 \otimes_R n \\ m \otimes_R (n_1 + n_2) &= m \otimes_R n_1 + m \otimes_R n_2 \\ mr \otimes_R n &= m \otimes_R rn \end{aligned}$$

Every tensor can be written (non-uniquely in general) as a finite sum of simple tensors.

Remark 23.1.1. We emphasize that each $m \otimes_R n$ represents a *coset* in some quotient group, and so we may have $m \otimes_R n = m' \otimes_R n'$ for $m \neq m'$ or $n \neq n'$. Due to this care must be taken when defining maps out of $M \otimes_R N$, since any such map must be shown to be independent of the particular choice coset of representative $m \otimes_R n$.

Another point to note is that even if M is a submodule of a larger module M' , we may have for some $m \in M$ and $n \in N$ $m \otimes_R n = 0$ in $M' \otimes_R N$ but $m \otimes_R n$ is nonzero in $M \otimes_R N$. In particular, we see that $M \otimes_R N$ need not be a subgroup of $M' \otimes_R N$ even when M is a submodule of M' .

Definition 23.1.4. *Let M be a right R -module, let N be a left R -module and let L be an abelian group (written additively). A map $\varphi : M \times N \rightarrow L$ is called **R -balanced** or **middle linear with respect to R** if*

$$\begin{aligned} \varphi(m_1 + m_2, n) &= \varphi(m_1, n) + \varphi(m_2, n) \\ \varphi(m, n_1 + n_2) &= \varphi(m, n_1) + \varphi(m, n_2) \\ \varphi(m, rn) &= \varphi(mr, n) \end{aligned}$$

for all $m, m_1, m_2 \in M, n, n_1, n_2 \in N$ and $r \in R$.

It follows that $\iota : M \times N \rightarrow M \otimes_R N$ is R -balanced.

Theorem 23.1.3 (Universal Property of Tensors and Balanced Maps). *Suppose R is a ring with 1, M is a right R -module, and N is a left R -module. Let $M \otimes_R N$ be the tensor product of M and N over R and let $\iota : M \times N \rightarrow M \otimes_R N$ be the R -balanced map defined above.*

1. *If $\Phi : M \otimes_R N \rightarrow L$ is any group homomorphism from $M \otimes_R N$ to an abelian group L then the composite map $\varphi = \Phi \circ \iota$ is an R -balanced map from $M \times N$ to L .*
2. *Conversely, suppose L is an abelian group and $\varphi : M \times N \rightarrow L$ is any R -balanced map. Then there is a unique group homomorphism $\Phi : M \otimes_R N \rightarrow L$ such that φ factors through ι , i.e., $\varphi = \Phi \circ \iota$.*

Equivalently, the correspondence $\varphi \leftrightarrow \Phi$ in the commutative diagram:

$$\begin{array}{ccc} M \times N & \xrightarrow{\iota} & M \otimes_R N \\ & \searrow \varphi & \downarrow \Phi \\ & & L \end{array}$$

establishes a bijection

$$\left\{ \begin{array}{l} R\text{-balanced maps} \\ \varphi : M \times N \rightarrow L \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} \text{group homomorphisms} \\ \Phi : M \otimes_R N \rightarrow L \end{array} \right\}$$

Proof. 1. follows immediately from the fact that ι is R -balanced. For 2., the map φ defines a unique \mathbb{Z} -module homomorphism φ from the free group on $M \times N$ to L such that $\varphi(m, n) = \varphi(m, n) \in L$. Since φ is R -balanced, φ maps each of the elements in the generating relations for $M \otimes_R N$ to 0. It follows that the kernel of φ contains the subgroup generated by these elements, hence φ induces a homomorphism Φ on the quotient group $M \otimes_R N$ to L . By definition we then have

$$\Phi(m \otimes_R n) = \varphi(m, n) = \varphi(m, n)$$

so $\varphi = \Phi \circ \iota$. The homomorphism Φ is uniquely determined by this equation since the elements $m \otimes_R n$ generate $M \otimes_R N$ as an abelian group. ■

Corollary 23.1.4. *Suppose D is an abelian group and $\iota' : M \times N \rightarrow D$ is an R -balanced map such that*

1. *the image of ι' generate3s D as an abelian group, and*
2. *every R -balanced map defined on $M \times N$ factors through ι' as in the previous theorem.*

Then there is an isomorphism $f : M \otimes_R N \cong D$ of abelian groups with $\iota' = f \circ \iota$.

Proof. Since $\iota' : M \times N \rightarrow D$ is a balanced map, the universal property of the previous theorem implies there is a unique homomorphism $f : M \otimes_R N \rightarrow D$ with $\iota' = f \circ \iota$. In particular $\iota'(m, n) = f(m \otimes_R n)$ for every $m \in M, n \in N$. By the first assumption, these elements generate D as an abelian group, so f is a surjective map. Now; the balanced map $\iota : M \times N \rightarrow M \otimes_R N$ together with the second assumption on ι' implies there is a unique homomorphism $g : D \rightarrow M \otimes_R N$ with $\iota = g \circ \iota'$. Then $m \otimes_R n = (g \circ f)(m \otimes_R n)$. Since the simple tensors $m \otimes_R n$ generate $M \otimes_R N$, it follows that $g \circ f$ is the identity map on $M \otimes_R N$ and so f is injective, hence an isomorphism. ■

We now return to giving $M \otimes_R N$ a module structure. To obtain an S -module structure on $M \otimes_R N$ we impose a structure similar to that of S on M :

Definition 23.1.5. Let R and S be any rings with 1. An abelian group M is called an (S, R) -**bi-module** if M is a left S -module, a right R -module, and $s(mr) = (sm)r$ for all $s \in S, r \in R$ and $m \in M$.

Example 23.1.1.

1. If $f : R \rightarrow S$ is any ring homomorphism with $f(1_R) = 1_S$ then S can be considered as a right R module with the action $s \cdot r = sf(r)$, and with respect to this action S becomes an (S, R) -bimodule.
2. Let I be a two-sided ideal in the ring R . Then the quotient ring R/I is an $(R/I, R)$ -bimodule. This is a special case of the previous example with the canonical projection homomorphism $R \rightarrow R/I$.
3. If R is a commutative ring, then any left (or right) R -module M can be given the structure of a right (or left) R -module by defining $mr = rm$ for all $m \in M$ and $r \in R$. This makes M into an (R, R) -bimodule.

Definition 23.1.6. Suppose M is a left (or right) R -module over the commutative ring R . Then the (R, R) -bimodule structure on M defined by letting the left and right R -actions coincide will be called the standard R -module structure on M .

Construction 23.1.7. We now continue our construction of the module structure on $M \otimes_R N$. Suppose that N is a left R -module and M is an (S, R) -bimodule. Then the (S, R) -bimodule structure on M implies that

$$s \left(\sum_{finite} m_i \otimes_R n_i \right) = \sum_{finite} (sm_i) \otimes_R n_i$$

gives a well defined action of S under which $M \otimes_R N$ is a left S -module. Note from our universal property checking that this map is well-defined is equivalent to checking that a certain map is R -balanced. Indeed, for any fixed $s \in S$ the map $(m, n) \mapsto sm \otimes_R n$ is an R -balanced map from $M \times N$ to $M \otimes_R N$. By the universal property there is a well defined group homomorphism $\lambda_s : M \otimes_R N \rightarrow M \otimes_R N$ such that $\lambda_s(m \otimes_R n) = sm \otimes_R n$. Since the right side of our action is then $\lambda_s(\sum m_i \otimes_R n_i)$, the fact that λ_s is well defined shows that this expression is indeed

independent of the representation of the tensor $\sum m_i \otimes_R n_i$ as a sum of simple tensors. Because λ_s is additive the action holds.

By a parallel argument, if M is a right R -module and N is an (R, S) -bimodule, then the tensor product $M \otimes_R N$ has the structure of a right S -module.

In the case of M and N left modules over a commutative ring R , and $S = R$, the standard R -module structure on M gives M the structure of an (R, R) -bimodule, so in this case the tensor product $M \otimes_R N$ always has the structure of a left R -module.

The corresponding map $M \times N \rightarrow M \otimes_R N$ maps $M \times N$ into an R -module and is additive in each factor. Since $r(m \otimes_R n) = rm \otimes_R n = mr \otimes_R n = m \otimes_R rn$ it also satisfies

$$r\iota(m, n) = \iota(rm, n) = \iota(m, rn)$$

Definition 23.1.8. Let R be a commutative ring with 1 and let M, N , and L be left R -modules. The map $\varphi : M \times N \rightarrow L$ is called **R -bilinear** if it is R -linear in each factor:

$$\begin{aligned}\varphi(r_1 m_1 + r_2 m_2, n) &= r_1 \varphi(m_1, n) + r_2 \varphi(m_2, n) \\ \varphi(m, r_1 n_1 + r_2 n_2) &= r_1 \varphi(m, n_1) + r_2 \varphi(m, n_2)\end{aligned}$$

for all $m, m_1, m_2 \in M, n, n_1, n_2 \in N$, and $r_1, r_2 \in R$.

Corollary 23.1.5. Suppose R is a commutative ring. Let M and N be two left R -modules and let $M \otimes_R N$ be the tensor product of M and N over R , where M is given the standard R -module structure. Then $M \otimes_R N$ is a left R -module with

$$r(m \otimes_R n) = (rm) \otimes_R n = (mr) \otimes_R n = m \otimes_R (rn)$$

and the map $\iota : M \times N \rightarrow M \otimes_R N$ with $\iota(m, n) \rightarrow m \otimes_R n$ is an R -bilinear map. If L is any left R -module then there is a bijection

$$\left\{ \begin{array}{l} R\text{-bilinear maps} \\ \varphi : M \times N \rightarrow L \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} R\text{-module homomorphisms} \\ \Phi : M \otimes_R N \rightarrow L \end{array} \right\}$$

where the correspondence between φ and Φ is given by the commutative diagram:

$$\begin{array}{ccc} M \times N & \xrightarrow{\iota} & M \otimes_R N \\ & \searrow \varphi & \downarrow \Phi \\ & & L \end{array}$$

Proof. We have seen that $M \otimes_R N$ is an R -module and that ι is R -bilinear. It remains to check that the correspondence corresponds bilinear maps with R -module homomorphisms. If $\varphi : M \times N \rightarrow L$ is bilinear then it is an R -balanced map, so the corresponding $\Phi : M \otimes_R N \rightarrow L$ is a group homomorphism. Moreover, on simple tensors $\Phi((rm) \otimes_R n) = \varphi(rm, n) = r\varphi(m, n) = r\Phi(m \otimes_R n)$ since φ is R -linear in the first variable. Since Φ is additive this extends to sums of simple tensors to show Φ is an R -module homomorphism.

Conversely, suppose $\Phi : M \otimes_R N \rightarrow L$ is an R -module homomorphism. Then it is indeed a homomorphism of abelian groups so there is a unique R -balanced map φ such that $\varphi = \Phi \circ \iota$. Since φ is R -balanced it is additive in both terms. Then observe that for any $(m, n) \in M \times N$ and $r \in R$,

$$\varphi(rm, n) = \Phi((rm) \otimes_R n) = r\Phi(m \otimes_R n) = r\varphi(m, n)$$

since Φ is an R -module homomorphism. Then φ is R -linear in the first factor, and since $\varphi(m, rn) = \varphi(rm, n)$ since φ is R -balanced, it is also R -linear in the second factor. ■

Theorem 23.1.6 (The Tensor Product of Homomorphisms). *Let M, M' be right R -modules, let N, N' be left R -modules, and suppose $\varphi : M \rightarrow M'$ and $\psi : N \rightarrow N'$ are R -module homomorphisms.*

1. *There is a unique group homomorphism, denoted $\varphi \otimes_R \psi : M \otimes_R N \rightarrow M' \otimes_R N'$ such that $(\varphi \otimes_R \psi)(m \otimes_R n) = \varphi(m) \otimes_R \psi(n)$ for all $m \in M$ and $n \in N$.*
2. *If M, M' are also (S, R) -bimodules for some ring S and φ is also an S -module homomorphism, then $\varphi \otimes_R \psi$ is a homomorphism of left S -modules.*
3. *If $\lambda : M' \rightarrow M''$ and $\mu : N' \rightarrow N''$ are R -module homomorphisms then $(\lambda \otimes_R \mu) \circ (\varphi \otimes_R \psi) = (\lambda \circ \varphi) \otimes_R (\mu \circ \psi)$.*

Proof. Observe that the map $(m, n) \mapsto \varphi(m) \otimes_R \psi(n)$ is R -balanced since φ and ψ are R -module homomorphisms, so 1. follows from the universal property.

In 2. the definition of the left action of S on M together with the assumption that φ is an S -module homomorphism imply that on simple tensors:

$$(\varphi \otimes_R \psi)(s(m \otimes_R n)) = (\varphi \otimes_R \psi)(sm \otimes_R n) = \varphi(sm) \otimes_R \psi(n) = s\varphi(m) \otimes_R \psi(n)$$

Since $\varphi \otimes_R \psi$ is additive, this extends to sums of simple tensors to show that $\varphi \otimes_R \psi$ is an S -module homomorphism.

The uniqueness condition in the universal property implies 3., completing the proof. ■

Theorem 23.1.7 (Associativity of the Tensor Product). *Suppose M is a right R -module, N is an (R, T) -bimodule, and L is a left T -module. Then there is a unique isomorphism*

$$(M \otimes_R N) \otimes_T L \cong M \otimes_R (N \otimes_T L)$$

of abelian groups such that $(m \otimes_R n) \otimes_T l \mapsto m \otimes_R (n \otimes_T l)$. If M is an (S, R) -bimodule, then this is an isomorphism of S -modules.

Proof. Note first that the (R, T) -bimodule structure on N makes $M \otimes_R N$ into a right T -module and $N \otimes_T L$ into a left R -module, so both sides of the isomorphism are well defined. For each fixed $l \in L$, the mapping $(m, n) \mapsto m \otimes_R (n \otimes_T l)$ is R -balanced, so by the universal property there is a homomorphism $M \otimes_R N \rightarrow M \otimes_R (N \otimes_T L)$ with $m \otimes_R n \mapsto m \otimes_R (n \otimes_T l)$. Then the map $(m \otimes_R n, l) \mapsto m \otimes_R (n \otimes_T l)$ from $(M \otimes_R N) \times L$ to $M \otimes_R (N \otimes_T L)$ is well defined. Moreover, it is T -balanced, so by another application of the universal property it induces a homomorphism

$(M \otimes_R N) \otimes_T L \rightarrow M \otimes_R (N \otimes_T L)$ such that $(m \otimes_R n) \otimes_T l \mapsto m \otimes_R (n \otimes_T l)$. Dually, we can construct a homomorphism in the opposite direction that is inverse to this one. This proves the group isomorphism.

Assume in addition M is an (S, R) -bimodule. Then for $s \in S$ and $t \in T$ we have

$$s((m \otimes_R n)t) = s(m \otimes_R nt) = sm \otimes_R nt = (sm \otimes_R n)t = (s(m \otimes_R n))t$$

so that $M \otimes_R N$ is an (S, T) -bimodule. Hence $(M \otimes_R N) \otimes_T L$ is a left S -module. Since $N \otimes_T L$ is a left R -module, also $M \otimes_R (N \otimes_T L)$ is a left S -module. The group isomorphism just established is notably a homomorphism of left S -modules. ■

Corollary 23.1.8. *Suppose R is commutative and M, N , and L are left R modules. Then*

$$(M \otimes_R N) \otimes_R L \cong M \otimes_R (N \otimes_R L)$$

as R -modules for the standard R -module structures on M, N , and L .

Definition 23.1.9. *Let R be a commutative ring with 1 and let M_1, M_2, \dots, M_n and L be R -modules with the standard R -module structures. A map $\varphi : M_1 \times \dots \times M_n \rightarrow L$ is called **n -multilinear over R** if it is an R -module homomorphism in each component when the other component entries are kept constant. When $n = 2$ or 3 we say φ is **bilinear** or **trilinear**, respectively.*

By the previous corollary, an n -fold tensor product may be unambiguously defined by iterating the tensor product of pairs of modules since any bracketing of $M_1 \otimes_R \dots \otimes_R M_n$ into tensor products of pairs gives an isomorphic R -module. The universal property of the tensor product of a pair of modules then implies that multilinear maps factor uniquely through the R -module $M_1 \otimes_R \dots \otimes_R M_n$.

Corollary 23.1.9. *Let R be a commutative ring and let M_1, \dots, M_n, L be R -modules. Let $M_1 \otimes_R \dots \otimes_R M_n$ denote any bracketing of the tensor product of these modules, and let*

$$\iota : M_1 \times \dots \times M_n \rightarrow M_1 \otimes_R \dots \otimes_R M_n$$

be the map defined by $\iota(m_1, \dots, m_n) = m_1 \otimes_R \dots \otimes_R m_n$. Then

1. *for every R -module homomorphism $\Phi : M_1 \otimes_R \dots \otimes_R M_n \rightarrow L$ the map $\varphi = \Phi \circ \iota$ is n -multilinear from $M_1 \times \dots \times M_n$ to L , and*
2. *if $\varphi : M_1 \times \dots \times M_n \rightarrow L$ is an n -multilinear map then there is a unique R -module homomorphism $\Phi : M_1 \otimes_R \dots \otimes_R M_n \rightarrow L$ such that $\varphi = \Phi \circ \iota$.*

Hence there is a bijection

$$\left\{ \begin{array}{c} n\text{-multilinear maps} \\ \varphi : M_1 \times \dots \times M_n \rightarrow L \end{array} \right\} \leftrightarrow \left\{ \begin{array}{c} R\text{-module homomorphisms} \\ \Phi : M_1 \otimes_R \dots \otimes_R M_n \rightarrow L \end{array} \right\}$$

where the correspondence between φ and Φ is given by the commutative diagram:

$$\begin{array}{ccc}
 M_1 \times \dots \times M_n & \xrightarrow{\iota} & M_1 \otimes_R \dots \otimes_R M_n \\
 & \searrow \varphi & \downarrow \Phi \\
 & & L
 \end{array}$$

Recall that even if $M \subseteq M'$ is a submodule, $M \otimes_R N$ is not necessarily contained in $M' \otimes_R N$. We now aim to show a sufficient condition for when this containment holds.

Theorem 23.1.10 (Tensor Products of Direct Sums). *Let M, M' be right R -modules and let N, N' be left R -modules. Then there are unique group isomorphisms:*

$$\begin{aligned}
 (M \oplus M') \otimes_R N &\cong (M \otimes_R N) \oplus (M' \otimes_R N) \\
 M \otimes_R (N \oplus N') &\cong (M \otimes_R N) \oplus (M \otimes_R N')
 \end{aligned}$$

such that $(m, m') \otimes_R n \mapsto (m \otimes_R n, m' \otimes_R n)$ and $m \otimes_R (n, n') \mapsto (m \otimes_R n, m \otimes_R n')$ respectively. If M, M' are also (S, R) -bimodules, then these are isomorphisms of left S -modules. In particular, if R is commutative, these are isomorphisms of R -modules.

Proof. The map $(M \oplus M') \times N \rightarrow (M \otimes_R N) \oplus (M' \otimes_R N)$ defined by $((m, m'), n) \mapsto (m \otimes_R n, m' \otimes_R n)$ is well defined since m and m' in $M \oplus M'$ are uniquely defined in the direct sum. The map is R -balanced, so induces a homomorphism f from $(M \oplus M') \otimes_R N$ to $(M \otimes_R N) \oplus (M' \otimes_R N)$ with

$$f((m, m') \otimes_R n) = (m \otimes_R n, m' \otimes_R n)$$

In the other direction, the R -balanced maps $M \times N \rightarrow (M \oplus M') \otimes_R N$ and $M' \times N \rightarrow (M \oplus M') \otimes_R N$ given by $(m, n) \mapsto (m, 0) \otimes_R n$ and $(m', n) \mapsto (0, m') \otimes_R n$, respectively, define homomorphisms from $M \otimes_R N$ and $M' \otimes_R N$ to $(M \oplus M') \otimes_R N$. These in turn give a homomorphism g from the direct sum $(M \otimes_R N) \oplus (M' \otimes_R N)$ to $(M \oplus M') \otimes_R N$ with

$$g((m \otimes_R n_1, m' \otimes_R n_2)) = (m, 0) \otimes_R n_1 + (0, m') \otimes_R n_2$$

f and g are inverse homomorphisms and are S -module isomorphisms when M and M' are (S, R) -bimodules. ■

This extends by induction to any finite direct sum of R -modules. The corresponding result is also true for arbitrary direct sums. For example,

$$M \otimes_R \left(\bigoplus_{i \in I} N_i \right) \cong \bigoplus_{i \in I} (M \otimes_R N_i)$$

where I is any index set.

Corollary 23.1.11 (Extension of Scalars for Free Modules). *The module obtained from the free R -module $N \cong R^n$ by extension of scalars from R to S is the free S -module S^n :*

$$S \otimes_R R^n \cong S^n$$

as left S -modules.

Corollary 23.1.12. *Let R be a commutative ring and let $M \cong R^s$ and $N \cong R^t$ be free R -modules with bases m_1, \dots, m_s and n_1, \dots, n_t , respectively. Then $M \otimes_R N$ is a free R -module of rank st , with basis $m_i \otimes_R n_j$, $1 \leq i \leq s$ and $1 \leq j \leq t$, so*

$$R^s \otimes_R R^t \cong R^{st}$$

More generally, the tensor product of two free modules of arbitrary rank over a commutative ring is free.

Proposition 23.1.13. *Suppose R is a commutative ring and M, N are left R -modules, considered with the standard R -module structures. Then there is a unique R -module isomorphism*

$$M \otimes_R N \cong N \otimes_R M$$

mapping $m \otimes_R n$ to $n \otimes_R m$.

Proof. (To be completed) ■

Proposition 23.1.14. *Let R be a commutative ring and let A and B be R -algebras. Then the multiplication $(a \otimes_R b)(a' \otimes_R b') = aa' \otimes_R bb'$ is well defined and makes $A \otimes_R B$ into an R -algebra.*

Proof. Note first that the definition of an R -algebra shows that

$$r(a \otimes_R b) = ra \otimes_R b = ar \otimes_R b = a \otimes_R rb = a \otimes_R br = (a \otimes_R b)r$$

for every $r \in R$, $a \in A$ and $b \in B$. To show that $A \otimes_R B$ is an R -algebra, our main task is to show that the specified multiplication is well defined. Consider the map $\varphi : A \times B \times A \times B \rightarrow A \otimes_R B$ defined by $f(a, b, a', b') = aa' \otimes_R bb'$ is multilinear over R . For example,

$$\begin{aligned} f(a, r_1 b_1 + r_2 b_2, a', b') &= aa' \otimes_R (r_1 b_1 + r_2 b_2) b' \\ &= aa' \otimes_R r_1 b_1 b' + aa' \otimes_R r_2 b_2 b' \\ &= r_1 f(a, b_1, a', b') + r_2 f(a, b_2, a', b') \end{aligned}$$

By a previous corollary there is a corresponding R -module homomorphism $\Phi : A \otimes_R B \otimes_R A \otimes_R B$ to $A \otimes_R B$ with $\Phi(a \otimes_R b \otimes_R a' \otimes_R b') = aa' \otimes_R bb'$. Viewing $A \otimes_R B \otimes_R A \otimes_R B$ as $(A \otimes_R B) \otimes_R (A \otimes_R B)$, we can apply the corollary in reverse to obtain a well defined R -bilinear mapping $\varphi' : (A \otimes_R B) \times (A \otimes_R B)$ to $A \otimes_R B$ with $\varphi'(a \otimes_R b, a' \otimes_R b') = aa' \otimes_R bb'$. This shows that the multiplication is indeed well defined (and also that it satisfies the distributive laws). ■

Appendices

.1.0 §Semi-Groups and Monoids

Definition .1.1. A semi-group is a set A equipped with a binary operation

$$A \times A \xrightarrow{\text{mult}} A, \quad (a_1, a_2) \mapsto a_1 \cdot a_2 \quad (.1.1)$$

which satisfies the associativity axiom:

$$\forall a_1, a_2, a_3 \in A, \quad a_1 \cdot (a_2 \cdot a_3) = (a_1 \cdot a_2) \cdot a_3 \quad (.1.2)$$

We can write this associativity axiom as the following commuting diagram:

$$\begin{array}{ccc} A \times A \times A & \xrightarrow{\text{Id}_A \times \text{mult}} & A \times A \\ \text{mult} \times \text{Id}_A \downarrow & & \downarrow \text{mult} \\ A \times A & \xrightarrow{\text{mult}} & A \end{array}$$

Definition .1.2. A semi-group is said to be a monoid if there exists an element $1 \in A$ that satisfies

$$\forall a \in A, \quad 1 \cdot a = a = a \cdot 1 \quad (.1.3)$$

An element $1 \in A$ is called the unity or identity in A .

Lemma .1.1. A monoid contains a unique identity element.

Proof. (Left to the reader) ■

Definition .1.3. An inverse of $a \in A$, a monoid, is an element $a^{-1} \in A$ such that

$$a \cdot a^{-1} = 1 = a^{-1} \cdot a \quad (.1.4)$$

Lemma .1.2. If $a \in A$ admits an inverse, then this inverse is unique.

Proof. (Left to the reader) ■

Definition .1.4. A monoid is said to be a group if every element admits an inverse.

Example .1.1.

1. $(\mathbb{Z}, +)$ is a group
2. (\mathbb{Z}, \cdot) is a monoid but not a group
3. $(\mathbb{R}, +)$ is a group
4. (\mathbb{R}, \cdot) is a monoid but not a group (0 doesn't have an inverse)
5. $(\mathbb{R} - \{0\}, \cdot)$ is a group

-
6. $\{\pm 1\} \subset \mathbb{R}$ with the operation \cdot is a group.
 7. $(\mathbb{C} - \{0\}, \cdot)$ is a group
 8. $(\{z \in \mathbb{C} - \{0\} : |z| = 1\}, \cdot)$ is a group, often denoted S^1 (the **circle group**)

Definition .1.5. A semi-group/monoid/group A is said to be **commutative** if

$$\forall a_1, a_2 \in A, \quad a_1 \cdot a_2 = a_2 \cdot a_1 \quad (.1.5)$$

We call such a structure **abelian**. We may rewrite the commutativity condition as the commutative diagram:

$$\begin{array}{ccc} A \times A & \xrightarrow{\text{mult}} & A \\ \text{swap}_A \downarrow & & \downarrow \text{Id}_A \\ A \times A & \xrightarrow{\text{mult}} & A \end{array}$$

where for any set X , $\text{swap}_X : X \times X \rightarrow X \times X$
 $(x_1, x_2) \mapsto (x_2, x_1)$