# Abstract Algebra for AM & TCS

A short introduction with applications

# Abstract Algebra for AM & TCS

A short introduction with applications

Georg Loho
University of Twente

Thomas W. Judson
Stephen F. Austin State University

September 26, 2023

**Edition**: 2023

**Website**: TODO

# Acknowledgements

I would like to acknowledge the following people in giving helpful insights and visualizations for the modified version.

- Stefano Piceghello, University of Twente

*The following acknowledgements are written by Thomas W. Judson for the original version of the book.*

I would like to acknowledge the following reviewers for their helpful comments and suggestions.

- David Anderson, University of Tennessee, Knoxville
- Robert Beezer, University of Puget Sound
- Myron Hood, California Polytechnic State University
- Herbert Kasube, Bradley University
- John Kurtzke, University of Portland
- Inessa Levi, University of Louisville
- Geoffrey Mason, University of California, Santa Cruz
- Bruce Mericle, Mankato State University
- Kimmo Rosenthal, Union College
- Mark Teply, University of Wisconsin

I would also like to thank Steve Quigley, Marnie Pommett, Cathie Griffin, Kelle Karshick, and the rest of the staff at PWS Publishing for their guidance throughout this project. It has been a pleasure to work with them.

Robert Beezer encouraged me to make *Abstract Algebra: Theory and Applications* available as an open source textbook, a decision that I have never regretted. With his assistance, the book has been rewritten in PreTeXt (pretextbook.org[1]), making it possible to quickly output print, web, PDF versions and more from the same source. The open source version of this book has received support from the National Science Foundation (Awards #DUE-1020957, #DUE–1625223, and #DUE–1821329).

[1] pretextbook.org

# History

Original Version

- Title: Abstract Algebra, Theory and Applications
- Year of version: 2022
- Original author: Thomas W. Judson
- abstract.pugetsound.edu[2]

Modified Version

- Title: Abstract Algebra for AM and TCS - A short introduction with applications
- Year: 2023
- New authors: Georg Loho
- TODO[3]

[2] abstract.pugetsound.edu
[3] TODO

# How to (read this book)

This text is a compact introduction to Abstract Algebra in 10 lectures. It is tailored to a course with students from Applied Mathematics and Technical Computer Science. Each chapter corresponds to one lecture; it starts with a list of basic learning goals. The "Additional insights" in each chapter go beyond the core scope but are interesting additions.

Towards the end of each chapter, there are sections "Core Exercises" and "Additional Exercises". The exercises in the section "Core Exercises" are important for the understanding of the course. The "Additional Exercises" are just a collection for further study.

Furthermore, towards the end of each chapter, there is the section "Material". Most importantly, this section contains links to videos which might provide additional help for understanding the content of the course. These videos cover most of the core topics but not everything; they are not maintained by us but are well-selected from other media sources.

Additionally, at the end of each chapter, there is a section with hints for selected exercises. Finally, there is a chapter "Comments to Core Exercises" in the Appendix. This is meant to be the last resort, if one has not been able to solve the exercises after weeks of trying.

# Contents

# Chapter 1

# Abstract Algebra: Getting Started

Basic learning goals

1. Basics from Linear Algebra (working with matrices, linear maps, vector spaces).
2. Computations and basic properties of modular arithmetic.
3. Structure of symmetries as functions.
4. Applications of Abstract Algebra: Cryptography, Polynomial Systems, Symmetries.

Algebra is the study of (binary) operations on sets fulfilling natural properties. We rely on the knowledge of linear algebra to motivate several constructions. Main guiding examples of sets with operations will be the real numbers with multiplication or addition, the integers with multiplication or addition and sets of (bijective) functions $\{f \mid f\colon X \to X\}$ with function composition.

## 1.1 Some Basics from Linear Algebra

We start by recalling some basic facts from linear algebra. One of the most fundamental ideas of linear algebra is that of a linear transformation. A **linear transformation** or **linear map** $T : \mathbb{R}^n \to \mathbb{R}^m$ is a map that preserves vector addition and scalar multiplication; that is, for vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ and a scalar $\alpha \in \mathbb{R}$,

$$\begin{aligned} T(\mathbf{x}+\mathbf{y}) &= T(\mathbf{x}) + T(\mathbf{y}) \\ T(\alpha\mathbf{y}) &= \alpha T(\mathbf{y}). \end{aligned}$$

An $m \times n$ matrix with entries in $\mathbb{R}$ represents a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. If we write vectors $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$ and $\mathbf{y} = (y_1, \ldots, y_n)^{\mathrm{t}}$ in $\mathbb{R}^n$ as column matrices, then an $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

maps the vectors to $\mathbb{R}^m$ linearly by matrix multiplication. Observe that if $\alpha$ is a real number,

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y} \qquad \text{and} \qquad \alpha A\mathbf{x} = A(\alpha\mathbf{x}),$$

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

We will often abbreviate the matrix $A$ by writing $(a_{ij})$.

Conversely, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map, we can associate a matrix $A$ with $T$ by considering what $T$ does to the vectors

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, \ldots, 0)^{\mathrm{t}} \\ \mathbf{e}_2 &= (0, 1, \ldots, 0)^{\mathrm{t}} \\ &\vdots \\ \mathbf{e}_n &= (0, 0, \ldots, 1)^{\mathrm{t}}. \end{aligned}$$

We can write any vector $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$ as

$$x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n.$$

Consequently, if

$$\begin{aligned} T(\mathbf{e}_1) &= (a_{11}, a_{21}, \ldots, a_{m1})^{\mathrm{t}}, \\ T(\mathbf{e}_2) &= (a_{12}, a_{22}, \ldots, a_{m2})^{\mathrm{t}}, \\ &\vdots \\ T(\mathbf{e}_n) &= (a_{1n}, a_{2n}, \ldots, a_{mn})^{\mathrm{t}}, \end{aligned}$$

then

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= \left( \sum_{k=1}^{n} a_{1k}x_k, \ldots, \sum_{k=1}^{n} a_{mk}x_k \right)^{\mathrm{t}} \\ &= A\mathbf{x}. \end{aligned}$$

**Example 1.1.1** If we let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the map given by

$$T(x_1, x_2) = (2x_1 + 5x_2, -4x_1 + 3x_2),$$

the axioms that $T$ must satisfy to be a linear transformation are easily verified. The column vectors $T\mathbf{e}_1 = (2, -4)^{\mathrm{t}}$ and $T\mathbf{e}_2 = (5, 3)^{\mathrm{t}}$ tell us that $T$ is given by the matrix

$$A = \begin{pmatrix} 2 & 5 \\ -4 & 3 \end{pmatrix}.$$

□

Inverses will play a crucial role in the following. Recall that an $n \times n$ matrix $A$ is **invertible** exactly when there exists another matrix $A^{-1}$ such

that $AA^{-1} = A^{-1}A = I$, where

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

is the $n \times n$ identity matrix. From linear algebra we know that $A$ is invertible if and only if the determinant of $A$ is nonzero. Sometimes an invertible matrix is said to be **nonsingular**.

**Example 1.1.2** If $A$ is the matrix

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix},$$

then the inverse of $A$ is

$$A^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}.$$

We are guaranteed that $A^{-1}$ exists, since $\det(A) = 2 \cdot 3 - 5 \cdot 1 = 1$ is nonzero. □

Let $A$ and $B$ be $n \times n$ matrices. From linear algebra we have the following properties of determinants.

- $\det(AB) = (\det A)(\det B)$.
- If $A$ is an invertible matrix, then $\det(A^{-1}) = 1/\det A$.
- If we define the transpose of a matrix $A = (a_{ij})$ to be $A^{\mathrm{t}} = (a_{ji})$, then $\det(A^{\mathrm{t}}) = \det A$.
- Let $T$ be the linear transformation associated with an $n \times n$ matrix $A$. Then $T$ multiplies volumes by a factor of $|\det A|$. In the case of $\mathbb{R}^2$, this means that $T$ multiplies areas by $|\det A|$.

## 1.2 Integer Equivalence Classes and Symmetries

Let us now investigate some mathematical structures that can be viewed as sets with single operations.

### 1.2.1 The Integers mod $n$

The integers mod $n$ have become indispensable in the theory and applications of algebra. In mathematics they are used in cryptography, coding theory, and the detection of errors in identification codes.

Let $r$ and $s$ be two integers and suppose that $n \in \mathbb{N}$. We say that $r$ is **congruent** to $s$ **modulo** $n$, or $r$ is congruent to $s$ mod $n$, if $r - s$ is evenly divisible by $n$; that is, $r - s = nk$ for some $k \in \mathbb{Z}$. In this case we write $r \equiv s \pmod{n}$. For example, $41 \equiv 17 \pmod{8}$ since $41 - 17 = 24$ is divisible by 8. We claim that congruence modulo $n$ forms an equivalence relation of $\mathbb{Z}$. Certainly any integer $r$ is equivalent to itself since $r - r = 0$ is divisible by $n$. We will now show that the relation is symmetric. If $r \equiv s \pmod{n}$, then $r - s = -(s - r)$ is divisible by $n$. So $s - r$ is divisible by $n$ and $s \equiv r \pmod{n}$. Now suppose that $r \equiv s \pmod{n}$ and $s \equiv t \pmod{n}$. Then there

exist integers $k$ and $l$ such that $r-s=kn$ and $s-t=ln$. To show transitivity, it is necessary to prove that $r-t$ is divisible by $n$. However,

$$r-t=r-s+s-t=kn+ln=(k+l)n,$$

and so $r-t$ is divisible by $n$.

Hence, the integers mod $n$ partition $\mathbb{Z}$ into $n$ different equivalence classes; we will denote the set of these equivalence classes by $\mathbb{Z}_n$. Consider the integers modulo 12 and the corresponding partition of the integers:

$$\begin{aligned}[0] &= \{\ldots,-12,0,12,24,\ldots\},\\ [1] &= \{\ldots,-11,1,13,25,\ldots\},\\ &\vdots\\ [11] &= \{\ldots,-1,11,23,35,\ldots\}.\end{aligned}$$

When no confusion can arise, we will use $0,1,\ldots,11$ to indicate the equivalence classes $[0],[1],\ldots,[11]$ respectively. We can do arithmetic on $\mathbb{Z}_n$. For two integers $a$ and $b$, define addition modulo $n$ to be $(a+b) \pmod n$; that is, the remainder when $a+b$ is divided by $n$. Similarly, multiplication modulo $n$ is defined as $(ab) \pmod n$, the remainder when $ab$ is divided by $n$.

**Example 1.2.1** The following examples illustrate integer arithmetic modulo $n$:

$$\begin{aligned} 7+4 &\equiv 1 \pmod 5 & \qquad 7\cdot 3 &\equiv 1 \pmod 5\\ 3+5 &\equiv 0 \pmod 8 & \qquad 3\cdot 5 &\equiv 7 \pmod 8\\ 3+4 &\equiv 7 \pmod{12} & \qquad 3\cdot 4 &\equiv 0 \pmod{12}. \end{aligned}$$

In particular, notice that it is possible that the product of two nonzero numbers modulo $n$ can be equivalent to 0 modulo $n$. □

**Example 1.2.2** Most, but not all, of the usual laws of arithmetic hold for addition and multiplication in $\mathbb{Z}_n$. For instance, it is not necessarily true that there is a multiplicative inverse. Consider the multiplication table for $\mathbb{Z}_8$ in Figure 1.2.3. Notice that 2, 4, and 6 do not have multiplicative inverses; that is, for $n=2$, 4, or 6, there is no integer $k$ such that $kn \equiv 1 \pmod 8$.

| · | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 0 | 2 | 4 | 6 | 0 | 2 | 4 | 6 |
| 3 | 0 | 3 | 6 | 1 | 4 | 7 | 2 | 5 |
| 4 | 0 | 4 | 0 | 4 | 0 | 4 | 0 | 4 |
| 5 | 0 | 5 | 2 | 7 | 4 | 1 | 6 | 3 |
| 6 | 0 | 6 | 4 | 2 | 0 | 6 | 4 | 2 |
| 7 | 0 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |

**Figure 1.2.3** Multiplication table for $\mathbb{Z}_8$

□

**Proposition 1.2.4** *Let $\mathbb{Z}_n$ be the set of equivalence classes of the integers mod $n$ and $a,b,c \in \mathbb{Z}_n$.*

1. *Addition and multiplication are commutative:*

$$a+b \equiv b+a \pmod n$$

$$ab \equiv ba \pmod{n}.$$

2. *Addition and multiplication are associative:*

$$(a+b)+c \equiv a+(b+c) \pmod{n}$$
$$(ab)c \equiv a(bc) \pmod{n}.$$

3. *There are both additive and multiplicative identities:*

$$a+0 \equiv a \pmod{n}$$
$$a \cdot 1 \equiv a \pmod{n}.$$

4. *Multiplication distributes over addition:*

$$a(b+c) \equiv ab+ac \pmod{n}.$$

5. *For every integer $a$ there is an additive inverse $-a$:*

$$a+(-a) \equiv 0 \pmod{n}.$$

6. *Let $a$ be a nonzero integer. Then $\gcd(a,n)=1$ if and only if there exists a multiplicative inverse $b$ for $a \pmod{n}$; that is, a nonzero integer $b$ such that*

$$ab \equiv 1 \pmod{n}.$$

*Proof.* We will prove (1) and (6) and leave the remaining properties as exercises.

(1) Addition and multiplication are commutative modulo $n$ since the remainder of $a+b$ divided by $n$ is the same as the remainder of $b+a$ divided by $n$.

(6) Suppose that $\gcd(a,n)=1$. Then there exist integers $r$ and $s$ such that $ar+ns=1$. Since $ns=1-ar$, it must be the case that $ar \equiv 1 \pmod{n}$. Letting $b$ be the equivalence class of $r$, $ab \equiv 1 \pmod{n}$.

Conversely, suppose that there exists an integer $b$ such that $ab \equiv 1 \pmod{n}$. Then $n$ divides $ab-1$, so there is an integer $k$ such that $ab-nk=1$. Let $d=\gcd(a,n)$. Since $d$ divides $ab-nk$, $d$ must also divide 1; hence, $d=1$. ■

### 1.2.2 Symmetries

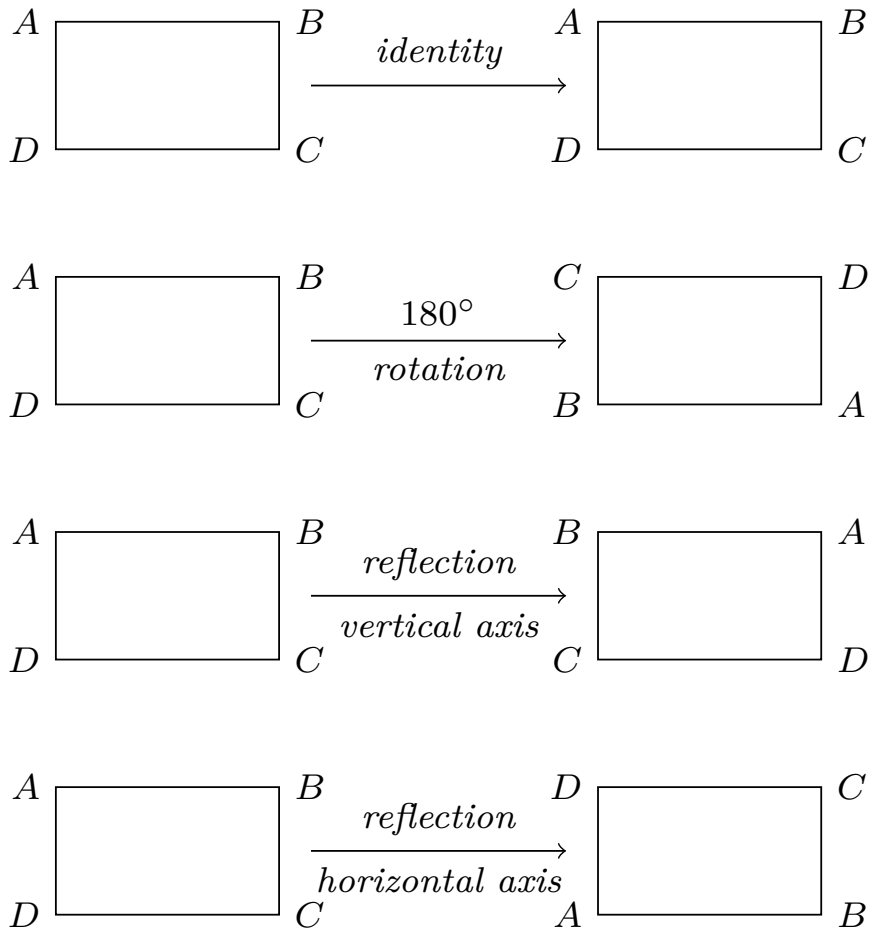
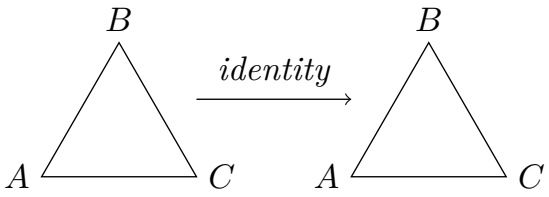
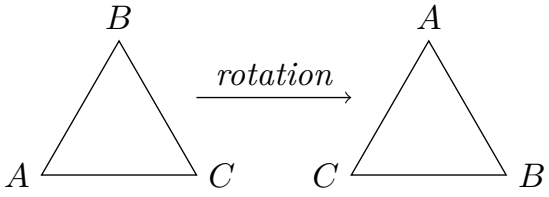



**Figure 1.2.5** Rigid motions of a rectangle

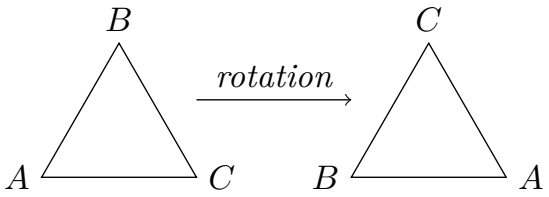
A **symmetry** of a geometric figure is a rearrangement of the figure preserving the arrangement of its sides and vertices as well as its distances and angles. A map from the plane to itself preserving the symmetry of an object is called a **rigid motion**. For example, if we look at the rectangle in Figure 1.2.5, it is easy to see that a rotation of 180° or 360° returns a rectangle in the plane with the same orientation as the original rectangle and the same relationship among the vertices. A reflection of the rectangle across either the vertical axis or the horizontal axis can also be seen to be a symmetry. However, a 90° rotation in either direction cannot be a symmetry unless the rectangle is a square.



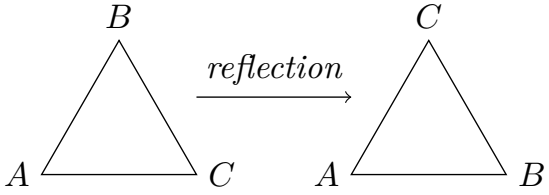

$$id = \begin{pmatrix} A & B & C \\ A & B & C \end{pmatrix}$$



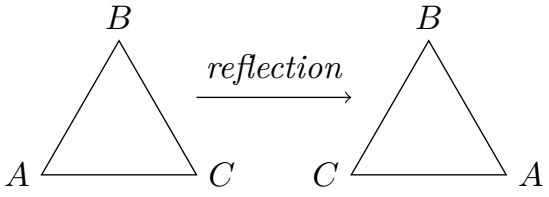

$$\rho_1 = \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}$$



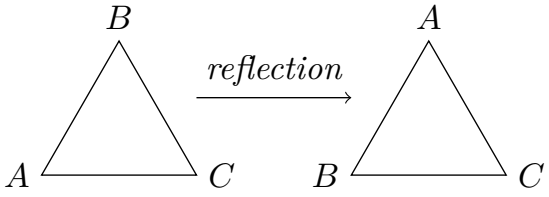

$$\rho_2 = \begin{pmatrix} A & B & C \\ C & A & B \end{pmatrix}$$




$$\mu_1 = \begin{pmatrix} A & B & C \\ A & C & B \end{pmatrix}$$




$$\mu_2 = \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}$$




$$\mu_3 = \begin{pmatrix} A & B & C \\ B & A & C \end{pmatrix}$$

**Figure 1.2.6** Symmetries of a triangle

Let us find the symmetries of the equilateral triangle $\triangle ABC$. To find a symmetry of $\triangle ABC$, we must first examine the permutations of the vertices $A$, $B$, and $C$ and then ask if a permutation extends to a symmetry of the triangle. Recall that a **permutation** of a set $S$ is a one-to-one and onto map $\pi : S \to S$. The three vertices have $3! = 6$ permutations, so the triangle has at most six symmetries. To see that there are six permutations, observe there are three different possibilities for the first vertex, and two for the second, and the remaining vertex is determined by the placement of the first two. So we have $3 \cdot 2 \cdot 1 = 3! = 6$ different arrangements. To denote the permutation of

the vertices of an equilateral triangle that sends $A$ to $B$, $B$ to $C$, and $C$ to $A$, we write the array

$$\begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}.$$

Notice that this particular permutation corresponds to the rigid motion of rotating the triangle by 120° in a clockwise direction. In fact, every permutation gives rise to a symmetry of the triangle. All of these symmetries are shown in Figure 1.2.6.

A natural question to ask is what happens if one motion of the triangle $\triangle ABC$ is followed by another. Which symmetry is $\mu_1\rho_1$; that is, what happens when we do the permutation $\rho_1$ and then the permutation $\mu_1$? *Remember that we are composing functions here. Although we usually multiply left to right, we compose functions right to left.* We have

$$\begin{aligned}(\mu_1\rho_1)(A) &= \mu_1(\rho_1(A)) = \mu_1(B) = C \\ (\mu_1\rho_1)(B) &= \mu_1(\rho_1(B)) = \mu_1(C) = B \\ (\mu_1\rho_1)(C) &= \mu_1(\rho_1(C)) = \mu_1(A) = A.\end{aligned}$$

This is the same symmetry as $\mu_2$. Suppose we do these motions in the opposite order, $\rho_1$ then $\mu_1$. It is easy to determine that this is the same as the symmetry $\mu_3$; hence, $\rho_1\mu_1 \neq \mu_1\rho_1$. A multiplication table for the symmetries of an equilateral triangle $\triangle ABC$ is given in Figure 1.2.7.

Notice that in the multiplication table for the symmetries of an equilateral triangle, for every motion of the triangle $\alpha$ there is another motion $\beta$ such that $\alpha\beta = \mathrm{id}$; that is, for every motion there is another motion that takes the triangle back to its original orientation.

| $\circ$ | id | $\rho_1$ | $\rho_2$ | $\mu_1$ | $\mu_2$ | $\mu_3$ |
|---|---|---|---|---|---|---|
| id | id | $\rho_1$ | $\rho_2$ | $\mu_1$ | $\mu_2$ | $\mu_3$ |
| $\rho_1$ | $\rho_1$ | $\rho_2$ | id | $\mu_3$ | $\mu_1$ | $\mu_2$ |
| $\rho_2$ | $\rho_2$ | id | $\rho_1$ | $\mu_2$ | $\mu_3$ | $\mu_1$ |
| $\mu_1$ | $\mu_1$ | $\mu_2$ | $\mu_3$ | id | $\rho_1$ | $\rho_2$ |
| $\mu_2$ | $\mu_2$ | $\mu_3$ | $\mu_1$ | $\rho_2$ | id | $\rho_1$ |
| $\mu_3$ | $\mu_3$ | $\mu_1$ | $\mu_2$ | $\rho_1$ | $\rho_2$ | id |

**Figure 1.2.7** Symmetries of an equilateral triangle

## 1.3 Selected Applications of Algebra

### 1.3.1 Cryptography

In all kinds of communication, it can be desirable to hide the exchanged messages from potential listeners. This is the fundamental problem which cryptography tries to solve by modifying exchanged messages in such a way that they can only be understood by the intended recipients. To break it down, one considers a single message M as a piece of data. Depending on the encoding, one can even think of it just as a (long) number, e.g., as a long binary number composed of 0s and 1s representing bits encoding other symbols. Now, the crucial idea is to apply a function to the message M which is hard to invert - unless one has a crucial additional piece of information, the key.

This leads to the idea of a one-way function. That is a function which is easy to evaluate on every input but hard to invert on a random element of the codomain. The flavours of 'easy' and 'hard' depend heavily on the context.

Cryptography always has to keep up with the computational power and the algorithms for attacking encryption. For example, it is currently a big question how cryptography will change with the rise of quantum computers which might be capable of solving much 'harder' computations.

The first impression could be that cryptography is most important in a hostile context. Though, it is actually important in the daily (virtual) communication to ensure privacy, just to name one important occurence. Note that most email communication is not encrypted and therefore easily readable for anyone who can access the channel through which the email was sent.

Algebra serves as a toolbox for functions which are hard to invert. During this course, we will come across two algebraic problems which give rise to one-way functions. The first problem is the Discrete Logarithm Problem which is at the heart of RSA (named after Ron Rivest, Adi Shamir and Leonard Adleman). We will see algebraic structures, where the inverse of the exponentiation map is hard to compute. The second problem is the Word Problem. This plays a prominent role in several questions of computational complexity. It is related to the group isomorphism problem - that is similar to the graph isomorphism problem but provably harder. The basic question is to find a sequence of allowed moves to transform a given word into the empty word. This idea is already featured in Exercise 1.4.7.

### 1.3.2 Bijective Maps and Symmetries

Assume we have some objects which come in some order, like appointments in a calender or food orders. If this order is changed but the objects stay the same then we have a re-ordering. This process is actually nothing else than applying a permutation, a bijective map of the objects; (see Functions describe the world (Youtube)[1] at second 42 for an even more general statement about the world).

Now, assume that we have some objects in some space, like a room with furniture or planets in space. If these objects are moved around, e.g., when you rearrange the room or by the movement of the planets on their orbits, then the objects itself are not changed but only their position. Like the reorderings above, these geometric transformations are essentially bijective maps of the objects in space.

All the examples above constituent bijective functions on some fixed set of objects or of some space. The fundamental properties of these transformations are captured by the notion of a group. In terms of these transformations, the group axioms say that:

1. There is the option to leave the configuration fixed. (Identity)
2. Each transformation is reversible. (Inverse)
3. The order in which we apply the transformations can matter but not how we assemble in which order they are applied. (Associativity)

These very ideas are needed to implement how motion is rendered in computer games, how engineering applications like robot arms are modeled or how similarity tests generalizing the graph isomorphism problem are captured. Group theory provides the language and the formal tools to find efficient solutions for these questions in practice.

[1] www.youtube.com/watch?v=zHU1xH6Ogs4

### 1.3.3 Solution of polynomial systems

One of the big topics in linear algebra was solving linear equality systems. Recall that this meant to find simultaneous solutions of several polynomials in which each monomial was just a single variable. While this already has many applications, linear polynomials do not suffice to describe more complicated models.

One of the most basic examples of non-linear polynomials arises when one wants to describe the trajectory of a thrown ball. This is described by a parabola - this means a polynomial of degree 2! Finding the point where the ball again hits the ground can be described by a quadratic equation of the form $ax^2 + bx + c = 0$. The zeros, namely those $x$ where the equation holds, are given by an explicit formula for quadratic polynomials.

Though, imagine you have a more complicated problem leading to a more complicated curve that can only be described by a curve of degree 5 or higher. Then the famous Abel-Ruffini theorem states that, in general, there is not an explicit formula for the zeros. On one hand, Abstract Algebra provides the framework to prove this theorem. On the other hand, Abstract Algebra helps to understand the sets of zeros nevertheless.

Finally, assume that we have not only one polynomial but a system of polynomial equations. For example, this could be used to describe the intersection of two orbits in space - to see if a planet and a comet hit along their trajectories; see Orbits and Kepler's Law (NASA)[2] for more background on these kinds of orbits (this word will appear in a different context later!). Assuming that one of the orbits is an ellipse and the other one a parabola, we get a system of the form

$$\frac{x^2}{p^2} + \frac{y^2}{q^2} = r$$
$$ax^2 + bx + c = y$$

The description of the solutions of polynomial systems leads to algebraic geometry; this goes beyond the scope of this course. Though, we will discuss the basic structure for this in the context of rings later. Note that polynomial equality systems are often solved with numerical methods in practice.

### 1.3.4 Further applications

We finish the overview of applications by briefly mentioning two further topics which belong to the greater context of machine learning.

The first one is Topological Data Analysis. Here the basic idea is to consider data points from 'far away' to exhibit their overall structure.

The second application is Homotopy Type Theory. This is a powerful framework with a new approach to automated proof systems based on the idea that one can continuously transform a statement to a true statement; see the Homotopy Type Theory Homepage[3] for further resources.

## 1.4 Core Exercises

1. Check your notes and literature from Linear Algebra; do you know how to work with matrices, linear maps, vector spaces?
2. Describe the symmetries of a pentagon or of a cube.

[2] solarsystem.nasa.gov/resources/310/orbits-and-keplers-laws/
[3] homotopytypetheory.org/

**3.** For each of the following, find an element $z \in \mathbb{Z}$ fulfilling the respective equation.

(a) $5 \cdot z \equiv 1 \pmod 7$

(b) $23 \cdot z \equiv 1 \pmod{97}$

**4.** Show that addition and multiplication mod $n$ are well defined operations. That is, show that the operations do not depend on the choice of the representative from the equivalence classes mod $n$.

**5.** Give an example of two square matrices $A, B$ such that

$$A \cdot B \neq B \cdot A \ .$$

**6.** Let $X$ be an arbitrary set and let $F$ be the set of functions on $X$, that is $\{f \mid f\colon X \to X\}$. Recall that function composition is associative, in particular the expression $f^n := f \circ f \circ \cdots \circ f$ is well-defined.

(a) Give an example of a binary operation which is not associative.

(b) Think of possibilities to define $f^n$ assuming that function composition would not be associative. [Note that function composition is a binary operation on the set of functions.]

**7.** **Riddle for Word Problem.**Let $S$ be a set of symbols. We consider words formed from symbols in $S$, that is, sequences of the form

$$s_1^{\sigma_1} \dots s_n^{\sigma_n}$$

where $s_1, \dots, s_n \in S$ and $\sigma_1, \dots, \sigma_n \in \{-1, 1\}$. Note that there is also the empty word $\epsilon$ consisting of no symbols. Let $R$ be a subset of the words. In this setting, we play the following game.

Given a word $w$ from $S$, there are three rules to modify $w$:

(i) For an arbitrary $x \in S$, a consecutive occurrence $xx^{-1}$ or $x^{-1}x$ can be replaced by the empty word.

(ii) A word in $R$ can be replaced by the empty word.

(iii) The empty word can be replaced by a word in $R$ or by an expression $xx^{-1}$ or $x^{-1}x$ for an arbitrary $x \in S$.

For example, let $S = \{a, b\}$ and $R = \{aa, bbbb, abab\}$. Then the word $babb^{-1}abbb$ can be modified in the following way

$$babb^{-1}abbb \quad \to \quad baabbb \quad \to \quad bbbb \quad \to \quad \epsilon$$

or

$$babb^{-1}abbb \to babb^{-1}ab^{-1}bbbb \to babb^{-1}ab^{-1} \to a^{-1}ababb^{-1}ab^{-1} \to a^{-1}b^{-1}ab^{-1}$$

and there are many other possibilities. [Note that $a$ can be thought of as the reflection of a square and $b$ as a rotation. ]

(a) For the sets $S$ and $R$ from before, decide if $abaaabb^{-1}abbb$ can be modified to the empty word.

(b) Let $S = \{a, b, c\}$ and $R = \{bc^{-1}, ac, baa, bcba\}$, decide if $cbaabcabc^{-1}$ can be modified to the empty word.

(c) Now let $R = \emptyset$ and $S$ be an arbitrary finite set. Show that for each word $x$ over $S$ there is another word $y$ such that their concatenation $xy$ can be modified to the empty word.

## 1.5 Additional Exercises

**1.** Find all $x \in \mathbb{Z}$ satisfying each of the following equations.

(a) $3x \equiv 2 \pmod 7$

(b) $5x + 1 \equiv 13 \pmod{23}$

(c) $5x + 1 \equiv 13 \pmod{26}$

(d) $9x \equiv 3 \pmod 5$

(e) $5x \equiv 1 \pmod 6$

(f) $3x \equiv 1 \pmod 6$

**2.** Show that

$$0 + a \equiv a + 0 \equiv a \pmod n$$

for all $a \in \mathbb{Z}_n$.

**3.** Prove that there is a multiplicative identity for the integers modulo $n$:

$$a \cdot 1 \equiv a \pmod n.$$

**4.** For each $a \in \mathbb{Z}_n$ find an element $b \in \mathbb{Z}_n$ such that

$$a + b \equiv b + a \equiv 0 \pmod n.$$

**5.** Show that addition and multiplication mod $n$ are associative operations.

**6.** Show that multiplication distributes over addition modulo $n$:

$$a(b + c) \equiv ab + ac \pmod n.$$

## 1.6 Material

1. Essence of Linear Algebra[1]
2. Matrix Multiplication[2]
3. Symmetry group regular triangle[3]
4. More details on modular arithmetic[4] in the sections 'Modular Arithmetic as Remainders', 'Congruence', 'Addition', 'Multiplication' of this resource

## 1.7 Hints to Selected Exercises

### 1.5 · Additional Exercises

**1.5.1.** (a) $3 + 7\mathbb{Z} = \{\ldots, -4, 3, 10, \ldots\}$; (c) $18 + 26\mathbb{Z}$; (e) $5 + 6\mathbb{Z}$.

[1] www.3blue1brown.com/lessons/eola-preview
[2] www.3blue1brown.com/lessons/matrix-multiplication
[3] www.socratica.com/lesson/symmetry-groups-of-triangles
[4] brilliant.org/wiki/modular-arithmetic/

# Chapter 2

# Groups

Basic learning goals

1. Definition of a group and a subgroup as well as the related terminology.
2. Basic properties of groups and ability to derive further properties from them.
3. Recognizing a (sub-)group structure.
4. Cayley table, generated (sub-)groups, direct products of groups, examples of groups (in particular matrices and modular arithmetic).

We begin our study of algebraic structures by investigating sets associated with a single operation that satisfies certain reasonable axioms; that is, we want to define an operation on a set in a way that will generalize such familiar structures as the integers $\mathbb{Z}$ together with the single operation of addition, or invertible $2 \times 2$ matrices together with the single operation of matrix multiplication, or permutations (of a finite set) with the single operation of function composition. The integers and the $2 \times 2$ matrices and permutations, together with their respective single operation, are examples of algebraic structures known as groups.

The theory of groups occupies a central position in mathematics. Modern group theory arose from an attempt to find the roots of a polynomial in terms of its coefficients. Groups now play a central role in such areas as coding theory, counting, and the study of symmetries; many areas of biology, chemistry, and physics have benefited from group theory.

## 2.1 Definitions and Examples

The integers mod $n$ and the symmetries of a triangle or a rectangle are examples of groups. A **binary operation** or **law of composition** on a set $G$ is a function $G \times G \to G$ that assigns to each pair $(a, b) \in G \times G$ a unique element $a \circ b$, or $ab$ in $G$, called the composition of $a$ and $b$. A **group** $(G, \circ)$ is a set $G$ together with a law of composition $(a, b) \mapsto a \circ b$ that satisfies the following axioms.

- The law of composition is **associative**. That is,

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

  for $a, b, c \in G$.

- There exists an element $e \in G$, called the **identity (or neutral) element**, such that for any element $a \in G$

$$e \circ a = a \circ e = a.$$

- For each element $a \in G$, there exists an **inverse element** in G, denoted by $a^{-1}$, such that

$$a \circ a^{-1} = a^{-1} \circ a = e.$$

A group $G$ with the property that $a \circ b = b \circ a$ for all $a, b \in G$ is called **abelian** or **commutative**. Groups not satisfying this property are said to be **nonabelian** or **noncommutative**.

**Example 2.1.1** The integers $\mathbb{Z} = \{\ldots, -1, 0, 1, 2, \ldots\}$ form a group under the operation of addition. The binary operation on two integers $m, n \in \mathbb{Z}$ is just their sum. Since the integers under addition already have a well-established notation, we will use the operator $+$ instead of $\circ$; that is, we shall write $m + n$ instead of $m \circ n$. The identity is 0, and the inverse of $n \in \mathbb{Z}$ is written as $-n$ instead of $n^{-1}$. Notice that the set of integers under addition have the additional property that $m + n = n + m$ and therefore form an abelian group. □

Most of the time we will write $ab$ instead of $a \circ b$; however, if the group already has a natural operation such as addition in the integers, we will use that operation. That is, if we are adding two integers, we still write $m+n$, $-n$ for the inverse, and 0 for the identity as usual. We also write $m - n$ instead of $m + (-n)$.

It is often convenient to describe a group in terms of an addition or multiplication table. Such a table is called a **Cayley table**.

**Example 2.1.2** The integers mod $n$ form a group under addition modulo $n$. Consider $\mathbb{Z}_5$, consisting of the equivalence classes of the integers 0, 1, 2, 3, and 4. We define the group operation on $\mathbb{Z}_5$ by modular addition. We write the binary operation on the group additively; that is, we write $m + n$. The element 0 is the identity of the group and each element in $\mathbb{Z}_5$ has an inverse. For instance, $2 + 3 = 3 + 2 = 0$. Figure 2.1.3 is a Cayley table for $\mathbb{Z}_5$. By Proposition 1.2.4, $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ is a group under the binary operation of addition mod $n$.

| + | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 | 4 |
| 1 | 1 | 2 | 3 | 4 | 0 |
| 2 | 2 | 3 | 4 | 0 | 1 |
| 3 | 3 | 4 | 0 | 1 | 2 |
| 4 | 4 | 0 | 1 | 2 | 3 |

**Figure 2.1.3** Cayley table for $(\mathbb{Z}_5, +)$

□

**Example 2.1.4** Not every set with a binary operation is a group. For example, if we let modular multiplication be the binary operation on $\mathbb{Z}_n$, then $\mathbb{Z}_n$ fails to be a group. The element 1 acts as a group identity since $1 \cdot k = k \cdot 1 = k$ for any $k \in \mathbb{Z}_n$; however, a multiplicative inverse for 0 does not exist since $0 \cdot k = k \cdot 0 = 0$ for every $k$ in $\mathbb{Z}_n$. Even if we consider the set $\mathbb{Z}_n \setminus \{0\}$, we still may not have a group. For instance, let $2 \in \mathbb{Z}_6$. Then 2 has no multiplicative inverse since

$$0 \cdot 2 = 0 \qquad 1 \cdot 2 = 2$$

$$2 \cdot 2 = 4 \qquad 3 \cdot 2 = 0$$
$$4 \cdot 2 = 2 \qquad 5 \cdot 2 = 4.$$

By Proposition 1.2.4, every nonzero $k$ does have an inverse in $\mathbb{Z}_n$ if $k$ is relatively prime to $n$. Denote the set of all such nonzero elements in $\mathbb{Z}_n$ by $U(n)$. Then $U(n)$ is a group called the **group of units** of $\mathbb{Z}_n$. Figure 2.1.5 is a Cayley table for the group $U(8)$.

| $\cdot$ | 1 | 3 | 5 | 7 |
|---|---|---|---|---|
| 1 | 1 | 3 | 5 | 7 |
| 3 | 3 | 1 | 7 | 5 |
| 5 | 5 | 7 | 1 | 3 |
| 7 | 7 | 5 | 3 | 1 |

**Figure 2.1.5** Multiplication table for $U(8)$

□

**Example 2.1.6** The symmetries of an equilateral triangle described in Section 1.2 form a nonabelian group. As we observed, it is not necessarily true that $\alpha\beta = \beta\alpha$ for two symmetries $\alpha$ and $\beta$. Using Figure 1.2.7, which is a Cayley table for this group, we can easily check that the symmetries of an equilateral triangle are indeed a group. We will denote this group by either $S_3$ or $D_3$, for reasons that will be explained later. □

**Example 2.1.7** We use $\mathbb{M}_2(\mathbb{R})$ to denote the set of all $2 \times 2$ matrices. Let $GL_2(\mathbb{R})$ be the subset of $\mathbb{M}_2(\mathbb{R})$ consisting of invertible matrices; that is, a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is in $GL_2(\mathbb{R})$ if there exists a matrix $A^{-1}$ such that $AA^{-1} = A^{-1}A = I$, where $I$ is the $2 \times 2$ identity matrix. For $A$ to have an inverse is equivalent to requiring that the determinant of $A$ be nonzero; that is, $\det A = ad - bc \neq 0$. The set of invertible matrices forms a group called the **general linear group**. The identity of the group is the identity matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The inverse of $A \in GL_2(\mathbb{R})$ is

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

The product of two invertible matrices is again invertible. Matrix multiplication is associative, satisfying the other group axiom. For matrices it is not true in general that $AB = BA$; hence, $GL_2(\mathbb{R})$ is another example of a nonabelian group. □

**Example 2.1.8** Let

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad I = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
$$J = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \qquad K = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

where $i^2 = -1$. Then the relations $I^2 = J^2 = K^2 = -1$, $IJ = K$, $JK =$

$I$, $KI = J$, $JI = -K$, $KJ = -I$, and $IK = -J$ hold. The set $Q_8 = \{\pm 1, \pm I, \pm J, \pm K\}$ is a group called the **quaternion group**. Notice that $Q_8$ is noncommutative. □

**Example 2.1.9** Let $\mathbb{C}^*$ be the set of nonzero complex numbers. Under the operation of multiplication $\mathbb{C}^*$ forms a group. The identity is 1. If $z = a + bi$ is a nonzero complex number, then

$$z^{-1} = \frac{a - bi}{a^2 + b^2}$$

is the inverse of $z$. It is easy to see that the remaining group axioms hold. □

A group is **finite**, or has **finite order**, if it contains a finite number of elements; otherwise, the group is said to be **infinite** or to have **infinite order**. The **order** of a finite group is the number of elements that it contains. If $G$ is a group containing $n$ elements, we write $|G| = n$. The group $\mathbb{Z}_5$ is a finite group of order 5; the integers $\mathbb{Z}$ form an infinite group under addition, and we sometimes write $|\mathbb{Z}| = \infty$.

### 2.1.1 Basic Properties of Groups

**Proposition 2.1.10**

1. *The identity element in a group $G$ is unique; that is, there exists only one element $e \in G$ such that $eg = ge = g$ for all $g \in G$.*
2. *If $g$ is any element in a group $G$, then the inverse of $g$, denoted by $g^{-1}$, is unique.*

**Proposition 2.1.11**

1. *Let $G$ be a group. If $a, b \in G$, then $(ab)^{-1} = b^{-1}a^{-1}$.*
2. *Let $G$ be a group. For any $a \in G$, $(a^{-1})^{-1} = a$.*

*Proof.*

1. Let $a, b \in G$. Then $abb^{-1}a^{-1} = aea^{-1} = aa^{-1} = e$. Similarly, $b^{-1}a^{-1}ab = e$. But by the previous proposition, inverses are unique; hence, $(ab)^{-1} = b^{-1}a^{-1}$.
2. Observe that $a^{-1}(a^{-1})^{-1} = e$. Consequently, multiplying both sides of this equation by $a$, we have

$$(a^{-1})^{-1} = e(a^{-1})^{-1} = aa^{-1}(a^{-1})^{-1} = ae = a.$$

■

We can use exponential notation for groups just as we do in ordinary algebra. If $G$ is a group and $g \in G$, then we define $g^0 = e$. For $n \in \mathbb{N}$, we define

$$g^n = \underbrace{g \cdot g \cdots g}_{n \text{ times}}$$

and

$$g^{-n} = \underbrace{g^{-1} \cdot g^{-1} \cdots g^{-1}}_{n \text{ times}}.$$

While composition is actually just a binary operation, the evaluation of these expressions is well-defined because of associativity: it does not matter how we put brackets in these compositions of several elements.

**Theorem 2.1.12** *In a group, the usual laws of exponents hold; that is, for all* $g, h \in G$,

1. $g^m g^n = g^{m+n}$ *for all* $m, n \in \mathbb{Z}$;
2. $(g^m)^n = g^{mn}$ *for all* $m, n \in \mathbb{Z}$;
3. $(gh)^n = (h^{-1}g^{-1})^{-n}$ *for all* $n \in \mathbb{Z}$. *Furthermore, if* $G$ *is abelian, then* $(gh)^n = g^n h^n$.

Notice that $(gh)^n \neq g^n h^n$ in general, since the group may not be abelian. If the group is $\mathbb{Z}$ or $\mathbb{Z}_n$, we write the group operation additively and the exponential operation multiplicatively; that is, we write $ng$ instead of $g^n$. The laws of exponents now become

1. $mg + ng = (m + n)g$ for all $m, n \in \mathbb{Z}$;
2. $m(ng) = (mn)g$ for all $m, n \in \mathbb{Z}$;
3. $m(g + h) = mg + mh$ for all $n \in \mathbb{Z}$.

It is important to realize that the last statement can be made only because $\mathbb{Z}$ and $\mathbb{Z}_n$ are commutative groups.

### 2.1.2 Historical Note

Although the first clear axiomatic definition of a group was not given until the late 1800s, group-theoretic methods had been employed before this time in the development of many areas of mathematics, including geometry and the theory of algebraic equations.

Joseph-Louis Lagrange used group-theoretic methods in a 1770–1771 memoir to study methods of solving polynomial equations. Later, Évariste Galois (1811–1832) succeeded in developing the mathematics necessary to determine exactly which polynomial equations could be solved in terms of the coefficients of the polynomial. Galois' primary tool was group theory.

The study of geometry was revolutionized in 1872 when Felix Klein proposed that geometric spaces should be studied by examining those properties that are invariant under a transformation of the space. Sophus Lie, a contemporary of Klein, used group theory to study solutions of partial differential equations. One of the first modern treatments of group theory appeared in William Burnside's *The Theory of Groups of Finite Order*, first published in 1897.

## 2.2 Subgroups

### 2.2.1 Definitions and Examples

Sometimes we wish to investigate smaller groups sitting inside a larger group. The set of even integers $2\mathbb{Z} = \{\ldots, -2, 0, 2, 4, \ldots\}$ is a group under the operation of addition. This smaller group sits naturally inside of the group of integers under addition. We define a **subgroup** $H$ of a group $G$ to be a subset $H$ of $G$ such that when the group operation of $G$ is restricted to $H$, $H$ is a group in its own right. Observe that every group $G$ with at least two elements will always have at least two subgroups, the subgroup consisting of the identity element alone and the entire group itself. The subgroup $H = \{e\}$ of a group $G$ is called the **trivial subgroup**. A subgroup that is a proper subset of $G$ is called a **proper subgroup**. In many of the examples that we have

investigated up to this point, there exist other subgroups besides the trivial and improper subgroups.

**Example 2.2.1** Consider the set of nonzero real numbers, $\mathbb{R}^*$, with the group operation of multiplication. The identity of this group is 1 and the inverse of any element $a \in \mathbb{R}^*$ is just $1/a$. We will show that

$$\mathbb{Q}^* = \{p/q : p \text{ and } q \text{ are nonzero integers}\}$$

is a subgroup of $\mathbb{R}^*$. The identity of $\mathbb{R}^*$ is 1; however, $1 = 1/1$ is the quotient of two nonzero integers. Hence, the identity of $\mathbb{R}^*$ is in $\mathbb{Q}^*$. Given two elements in $\mathbb{Q}^*$, say $p/q$ and $r/s$, their product $pr/qs$ is also in $\mathbb{Q}^*$. The inverse of any element $p/q \in \mathbb{Q}^*$ is again in $\mathbb{Q}^*$ since $(p/q)^{-1} = q/p$. Since multiplication in $\mathbb{R}^*$ is associative, multiplication in $\mathbb{Q}^*$ is associative. □

**Example 2.2.2** Recall that $\mathbb{C}^*$ is the multiplicative group of nonzero complex numbers. Let $H = \{1, -1, i, -i\}$. Then $H$ is a subgroup of $\mathbb{C}^*$. It is quite easy to verify that $H$ is a group under multiplication and that $H \subset \mathbb{C}^*$. Furthermore, the **circle group**,

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$$

is also a subgroup of $\mathbb{C}^*$. □

**Example 2.2.3** Let $SL_2(\mathbb{R})$ be the subset of $GL_2(\mathbb{R})$ consisting of matrices of determinant one; that is, a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is in $SL_2(\mathbb{R})$ exactly when $ad - bc = 1$. To show that $SL_2(\mathbb{R})$ is a subgroup of the general linear group, we must show that it is a group under matrix multiplication. The $2 \times 2$ identity matrix is in $SL_2(\mathbb{R})$, as is the inverse of the matrix $A$:

$$A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

It remains to show that multiplication is closed; that is, that the product of two matrices of determinant one also has determinant one. This is shown in Subsection 2.2.3. The group $SL_2(\mathbb{R})$ is called the **special linear group**. □

**Example 2.2.4** It is important to realize that a subset $H$ of a group $G$ can be a group without being a subgroup of $G$. For $H$ to be a subgroup of $G$, it must inherit the binary operation of $G$. The set of all $2 \times 2$ matrices, $\mathbb{M}_2(\mathbb{R})$, forms a group under the operation of addition. The $2 \times 2$ general linear group is a subset of $\mathbb{M}_2(\mathbb{R})$ and is a group under matrix multiplication, but it is not a subgroup of $\mathbb{M}_2(\mathbb{R})$. If we add two invertible matrices, we do not necessarily obtain another invertible matrix. Observe that

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

but the zero matrix is not in $GL_2(\mathbb{R})$. □

**Example 2.2.5** One way of telling whether or not two groups are the same is by examining their subgroups. Other than the trivial subgroup and the group itself, the group $\mathbb{Z}_4$ has a single subgroup consisting of the elements 0 and 2. From the group $\mathbb{Z}_2$, we can form another group of four elements as follows. As a set this group is $\mathbb{Z}_2 \times \mathbb{Z}_2$. We perform the group operation

coordinatewise; that is, $(a,b)+(c,d) = (a+c, b+d)$. Figure 2.2.6 is an addition table for $\mathbb{Z}_2 \times \mathbb{Z}_2$. Since there are three nontrivial proper subgroups of $\mathbb{Z}_2 \times \mathbb{Z}_2$, $H_1 = \{(0,0),(0,1)\}$, $H_2 = \{(0,0),(1,0)\}$, and $H_3 = \{(0,0),(1,1)\}$, $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ must be different groups.

| $+$ | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(0,0)$ | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
| $(0,1)$ | $(0,1)$ | $(0,0)$ | $(1,1)$ | $(1,0)$ |
| $(1,0)$ | $(1,0)$ | $(1,1)$ | $(0,0)$ | $(0,1)$ |
| $(1,1)$ | $(1,1)$ | $(1,0)$ | $(0,1)$ | $(0,0)$ |

**Figure 2.2.6** Addition table for $\mathbb{Z}_2 \times \mathbb{Z}_2$

□

Suppose that $G$ is a group and let $\{g_i : i \in I\}$ be a set of elements in $G$, where $i$ ranges over some (not necessarily finite) index set $I$ (that is a set whose elements are used to label elements of another set). The smallest subgroup of $G$ containing all of the $g_i$'s is the subgroup of $G$ **generated** by the $g_i$'s. If this subgroup of $G$ is in fact all of $G$, then $G$ is generated by the set $\{g_i : i \in I\}$. In this case the $g_i$'s are said to be the **generators** of $G$. If there is a finite set $\{g_i : i \in I\}$ that generates $G$, then $G$ is **finitely generated**.

### 2.2.2 Some Subgroup Theorems

Let us examine some criteria for determining exactly when a subset of a group is a subgroup.

**Proposition 2.2.7** *A subset $H$ of $G$ is a subgroup if and only if it satisfies the following conditions.*

1. *The identity $e$ of $G$ is in $H$.*
2. *If $h_1, h_2 \in H$, then $h_1 h_2 \in H$.*
3. *If $h \in H$, then $h^{-1} \in H$.*

*Proof.* First suppose that $H$ is a subgroup of $G$. We must show that the three conditions hold. Since $H$ is a group, it must have an identity $e_H$. We must show that $e_H = e$, where $e$ is the identity of $G$. We know that $e_H e_H = e_H$ and that $e e_H = e_H e = e_H$; hence, $e e_H = e_H e_H$. By right-hand cancellation, $e = e_H$. The second condition holds since a subgroup $H$ is a group. To prove the third condition, let $h \in H$. Since $H$ is a group, there is an element $h' \in H$ such that $hh' = h'h = e$. By the uniqueness of the inverse in $G$, $h' = h^{-1}$.

Conversely, if the three conditions hold, we must show that $H$ is a group under the same operation as $G$; however, these conditions plus the associativity of the binary operation are exactly the axioms stated in the definition of a group. ■

**Proposition 2.2.8** *Let $H$ be a subset of a group $G$. Then $H$ is a subgroup of $G$ if and only if $H \neq \emptyset$, and whenever $g, h \in H$ then $gh^{-1}$ is in $H$.*

*Proof.* First assume that $H$ is a subgroup of $G$. We wish to show that $gh^{-1} \in H$ whenever $g$ and $h$ are in $H$. Since $h$ is in $H$, its inverse $h^{-1}$ must also be in $H$. Because of the closure of the group operation, $gh^{-1} \in H$.

Conversely, suppose that $H \subset G$ such that $H \neq \emptyset$ and $gh^{-1} \in H$ whenever $g, h \in H$. If $g \in H$, then $gg^{-1} = e$ is in $H$. If $g \in H$, then $eg^{-1} = g^{-1}$ is also in $H$. Now let $h_1, h_2 \in H$. We must show that their product is also in $H$. However, $h_1(h_2^{-1})^{-1} = h_1 h_2 \in H$. Hence, $H$ is a subgroup of $G$. ■

### 2.2.3 Some Matrix Groups

The set of all $n \times n$ invertible matrices forms a group called the **general linear group**. We will denote this group by $GL_n(\mathbb{R})$. The general linear group has several important subgroups. The multiplicative properties of the determinant imply that the set of matrices with determinant one is a subgroup of the general linear group. Stated another way, suppose that $\det(A) = 1$ and $\det(B) = 1$. Then $\det(AB) = \det(A)\det(B) = 1$ and $\det(A^{-1}) = 1/\det A = 1$. This subgroup is called the **special linear group** and is denoted by $SL_n(\mathbb{R})$.

**Example 2.2.9** Given a $2 \times 2$ matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the determinant of $A$ is $ad - bc$. The group $GL_2(\mathbb{R})$ consists of those matrices in which $ad - bc \neq 0$. The inverse of $A$ is

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

If $A$ is in $SL_2(\mathbb{R})$, then

$$A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Geometrically, $SL_2(\mathbb{R})$ is the group that preserves the areas of parallelograms. Let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

be in $SL_2(\mathbb{R})$. In Figure 2.2.10, the unit square corresponding to the vectors $\mathbf{x} = (1,0)^t$ and $\mathbf{y} = (0,1)^t$ is taken by $A$ to the parallelogram with sides $(1,0)^t$ and $(1,1)^t$; that is, $A\mathbf{x} = (1,0)^t$ and $A\mathbf{y} = (1,1)^t$. Notice that these two parallelograms have the same area. □

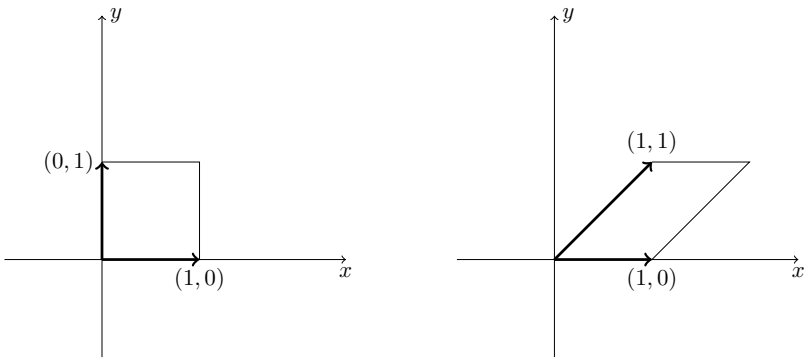



**Figure 2.2.10** $SL_2(\mathbb{R})$ acting on the unit square

Let $\texttt{Triang}_n(\mathbb{R})$ be the set of invertible square upper triangular $(n \times n)$-matrices with real entries. By writing out the definition of matrix multiplication, we see that the product of two matrices in $\texttt{Triang}_n(\mathbb{R})$ is again in this set. Furthermore, computing the inverse of a matrix by Gaussian elimination shows that the inverse of an upper triangular matrix is again upper triangular. Finally, the identity matrix is also in $\texttt{Triang}_n(\mathbb{R})$. Hence, Proposition 2.2.7 yields that $\texttt{Triang}_n(\mathbb{R})$ is a subgroup of $GL_n(\mathbb{R})$.

Now, let $\texttt{Elem}(n)$ be the set $(n \times n)$-matrices corresponding to elementary

row operations (as obtained from left multiplication). From Gaussian elimination, we know that each matrix in $GL_n(\mathbb{R})$ can be transformed to an upper triangular matrix by multiplication from the left of elementary matrices. Therefore, $GL_n(\mathbb{R})$ is generated by $\texttt{Triang}_n(\mathbb{R}) \cup \texttt{Elem}(n)$. Actually, by also considering multiplication from the right, which corresponds to elementary column operations, we see that already $\texttt{Elem}(n)$ forms generators for $GL_n(\mathbb{R})$.

## 2.3 Additional insights

### 2.3.1 Equations in groups

**Proposition 2.3.1**

1. *Equations in groups.*

   *Let $G$ be a group and $a$ and $b$ be any two elements in $G$. Then the equations $ax = b$ and $xa = b$ have unique solutions in $G$.*

2. *Right and left cancellation laws.*

   *If $G$ is a group and $a, b, c \in G$, then $ba = ca$ implies $b = c$ and $ab = ac$ implies $b = c$.*

*Proof.* Suppose that $ax = b$. We must show that such an $x$ exists. We can multiply both sides of $ax = b$ by $a^{-1}$ to find $x = ex = a^{-1}ax = a^{-1}b$.

To show uniqueness, suppose that $x_1$ and $x_2$ are both solutions of $ax = b$; then $ax_1 = b = ax_2$. So $x_1 = a^{-1}ax_1 = a^{-1}ax_2 = x_2$. The proof for the existence and uniqueness of the solution of $xa = b$ is similar. ■

### 2.3.2 Multiplicative Group of Complex Numbers

The **complex numbers** are defined as

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\},$$

where $i^2 = -1$. If $z = a + bi$, then $a$ is the **real part** of $z$ and $b$ is the **imaginary part** of $z$.

To add two complex numbers $z = a + bi$ and $w = c + di$, we just add the corresponding real and imaginary parts:

$$z + w = (a + bi) + (c + di) = (a + c) + (b + d)i.$$

Remembering that $i^2 = -1$, we multiply complex numbers just like polynomials. The product of $z$ and $w$ is

$$(a + bi)(c + di) = ac + bdi^2 + adi + bci = (ac - bd) + (ad + bc)i.$$

Every nonzero complex number $z = a+bi$ has a multiplicative inverse; that is, there exists a $z^{-1} \in \mathbb{C}^*$ such that $zz^{-1} = z^{-1}z = 1$. If $z = a + bi$, then

$$z^{-1} = \frac{a - bi}{a^2 + b^2}.$$

The **complex conjugate** of a complex number $z = a + bi$ is defined to be $\overline{z} = a - bi$. The **absolute value** or **modulus** of $z = a + bi$ is $|z| = \sqrt{a^2 + b^2}$.

**Example 2.3.2** Let $z = 2 + 3i$ and $w = 1 - 2i$. Then

$$z + w = (2 + 3i) + (1 - 2i) = 3 + i$$

and

$$zw = (2 + 3i)(1 - 2i) = 8 - i.$$

Also,

$$\begin{aligned} z^{-1} &= \frac{2}{13} - \frac{3}{13}i \\ |z| &= \sqrt{13} \\ \overline{z} &= 2 - 3i. \end{aligned}$$

□

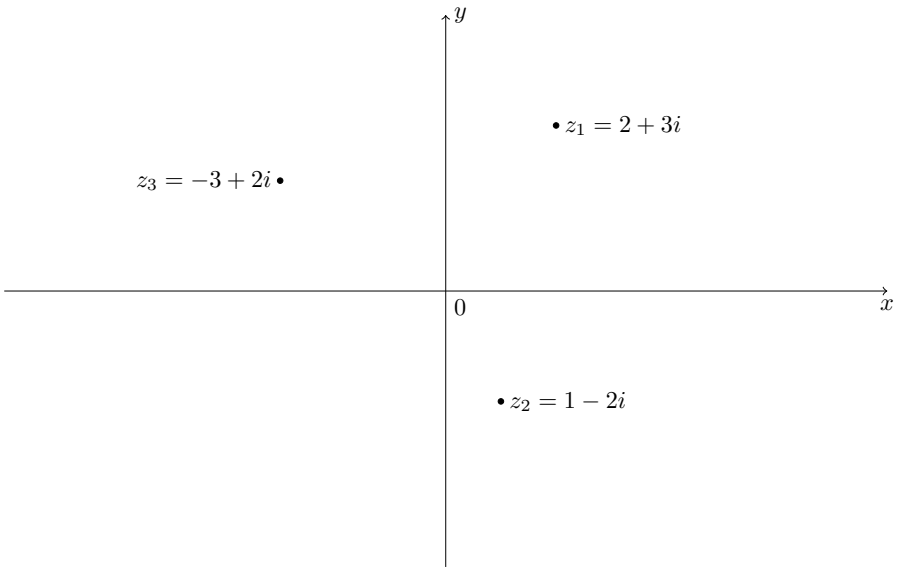



**Figure 2.3.3** Rectangular coordinates of a complex number

There are several ways of graphically representing complex numbers. We can represent a complex number $z = a + bi$ as an ordered pair on the $xy$ plane where $a$ is the $x$ (or real) coordinate and $b$ is the $y$ (or imaginary) coordinate. This is called the **rectangular** or **Cartesian** representation. The rectangular representations of $z_1 = 2 + 3i$, $z_2 = 1 - 2i$, and $z_3 = -3 + 2i$ are depicted in Figure 2.3.3.

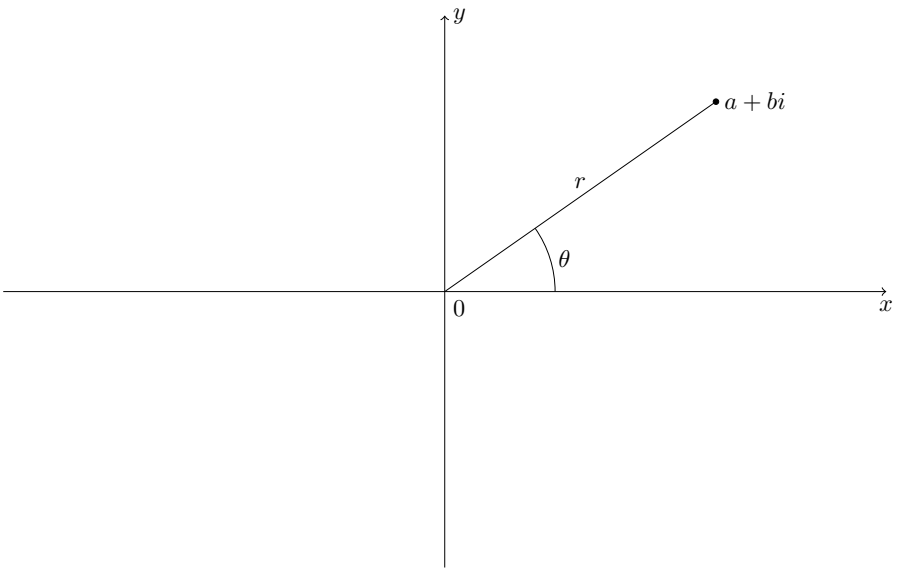



**Figure 2.3.4** Polar coordinates of a complex number

Nonzero complex numbers can also be represented using **polar coordinates**. To specify any nonzero point on the plane, it suffices to give an angle $\theta$ from the positive $x$ axis in the counterclockwise direction and a distance $r$ from the origin, as in Figure 2.3.4. We can see that

$$z = a + bi = r(\cos\theta + i\sin\theta).$$

Hence,

$$r = |z| = \sqrt{a^2 + b^2}$$

and

$$\begin{aligned} a &= r\cos\theta \\ b &= r\sin\theta. \end{aligned}$$

We sometimes abbreviate $r(\cos\theta + i\sin\theta)$ as $r\,\mathrm{cis}\,\theta$. To assure that the representation of $z$ is well-defined, we also require that $0^\circ \leq \theta < 360^\circ$. If the measurement is in radians, then $0 \leq \theta < 2\pi$.

**Example 2.3.5** Suppose that $z = 2\,\mathrm{cis}\,60^\circ$. Then

$$a = 2\cos 60^\circ = 1$$

and

$$b = 2\sin 60^\circ = \sqrt{3}.$$

Hence, the rectangular representation is $z = 1 + \sqrt{3}\,i$.

Conversely, if we are given a rectangular representation of a complex number, it is often useful to know the number's polar representation. If $z = 3\sqrt{2} - 3\sqrt{2}\,i$, then

$$r = \sqrt{a^2 + b^2} = \sqrt{36} = 6$$

and

$$\theta = \arctan\left(\frac{b}{a}\right) = \arctan(-1) = 315^\circ,$$

so $3\sqrt{2} - 3\sqrt{2}\,i = 6\,\mathrm{cis}\,315^\circ$. □

The polar representation of a complex number makes it easy to find products and powers of complex numbers. The proof of the following proposition is straightforward and is left as an exercise.

**Proposition 2.3.6** *Let $z = r\operatorname{cis}\theta$ and $w = s\operatorname{cis}\phi$ be two nonzero complex numbers. Then*

$$zw = rs\operatorname{cis}(\theta + \phi).$$

**Example 2.3.7** If $z = 3\operatorname{cis}(\pi/3)$ and $w = 2\operatorname{cis}(\pi/6)$, then $zw = 6\operatorname{cis}(\pi/2) = 6i$. □

**Theorem 2.3.8 DeMoivre.** *Let $z = r\operatorname{cis}\theta$ be a nonzero complex number. Then*

$$[r\operatorname{cis}\theta]^n = r^n\operatorname{cis}(n\theta)$$

*for $n = 1, 2, \ldots$.*

*Proof.* We will use induction on $n$. For $n = 1$ the theorem is trivial. Assume that the theorem is true for all $k$ such that $1 \leq k \leq n$. Then

$$\begin{aligned} z^{n+1} &= z^n z \\ &= r^n(\cos n\theta + i\sin n\theta)r(\cos\theta + i\sin\theta) \\ &= r^{n+1}[(\cos n\theta\cos\theta - \sin n\theta\sin\theta) + i(\sin n\theta\cos\theta + \cos n\theta\sin\theta)] \\ &= r^{n+1}[\cos(n\theta + \theta) + i\sin(n\theta + \theta)] \\ &= r^{n+1}[\cos(n+1)\theta + i\sin(n+1)\theta]. \end{aligned}$$

■

**Example 2.3.9** Suppose that $z = 1 + i$ and we wish to compute $z^{10}$. Rather than computing $(1+i)^{10}$ directly, it is much easier to switch to polar coordinates and calculate $z^{10}$ using DeMoivre's Theorem:

$$\begin{aligned} z^{10} &= (1+i)^{10} \\ &= \left(\sqrt{2}\operatorname{cis}\left(\frac{\pi}{4}\right)\right)^{10} \\ &= (\sqrt{2}\,)^{10}\operatorname{cis}\left(\frac{5\pi}{2}\right) \\ &= 32\operatorname{cis}\left(\frac{\pi}{2}\right) \\ &= 32i. \end{aligned}$$

□

The multiplicative group of the complex numbers, $\mathbb{C}^*$, possesses some interesting subgroups. Whereas $\mathbb{Q}^*$ and $\mathbb{R}^*$ have no interesting subgroups of finite order, $\mathbb{C}^*$ has many. An important example is the **circle group**,

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}.$$

The following proposition is a direct result of Proposition 2.3.6.

**Proposition 2.3.10** *The circle group is a subgroup of $\mathbb{C}^*$.*

## 2.4 Core Exercises

**1.** Recall from Example 2.1.7 the general linear group $GL_2(\mathbb{R})$ and from Example 2.2.3 the special linear group $SL_2(\mathbb{R})$.

Write out the proof that $SL_2(\mathbb{R})$ is a subgroup of $GL_2(\mathbb{R})$. Furthermore, prove that the matrices with integer entries and determinant 1 denoted by $SL_2(\mathbb{Z})$ form a subgroup of $SL_2(\mathbb{R})$.

**2.** If $(G, \cdot)$ and $(H, \circ)$ are groups, then we can make the Cartesian product of $G$ and $H$ into a new group. As a set, our group is just the ordered pairs $(g, h) \in G \times H$ where $g \in G$ and $h \in H$. We can define a binary operation on $G \times H$ by

$$(g_1, h_1)(g_2, h_2) = (g_1 \cdot g_2, h_1 \circ h_2);$$

that is, we just multiply elements in the first coordinate as we do in $G$ and elements in the second coordinate as we do in $H$. We have specified the particular operations $\cdot$ and $\circ$ in each group here for the sake of clarity; we usually just write $(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)$.

Prove the following statement: Let $G$ and $H$ be groups. The set $G \times H$ is a group under the operation $(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)$ where $g_1, g_2 \in G$ and $h_1, h_2 \in H$.

The group $G \times H$ is called the **external direct product** of $G$ and $H$. Notice that there is nothing special about the fact that we have used only two groups to build a new group. The direct product

$$\prod_{i=1}^{n} G_i = G_1 \times G_2 \times \cdots \times G_n$$

of the groups $G_1, G_2, \ldots, G_n$ is defined in exactly the same manner. If $G = G_1 = G_2 = \cdots = G_n$, we often write $G^n$ instead of $G_1 \times G_2 \times \cdots \times G_n$.

**3.** Prove Proposition 2.1.10.

**4.** Give an example of a nonabelian group $G$ and elements $a, b$ in this group demonstrating the properties shown in Proposition 2.1.11, i.e., that $(ab)^{-1} = b^{-1}a^{-1}$.

**5.** Prove Theorem 2.1.12.

**6.** Let $G$ be a group and $a$ be any element in $G$. Prove the following: The set

$$\langle a \rangle = \{a^k : k \in \mathbb{Z}\}$$

is a subgroup of $G$. Furthermore, $\langle a \rangle$ is the smallest subgroup of $G$ that contains $a$.

**7.** Assume that you are given the Cayley table of a group $G$ and a subset $T$ of the elements. Compare how to apply Proposition 2.2.7 or Proposition 2.2.8 to the Cayley table to determine the subgroup of $G$ generated by $T$. Apply your thoughts to the following example that is, determine in the following examples the subgroup of $G$ generated by $T$.

(a) Let $G$ be given by

| $\circ$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $a$ | $b$ | $a$ | $d$ | $c$ |
| $b$ | $a$ | $b$ | $c$ | $d$ |
| $c$ | $d$ | $c$ | $b$ | $a$ |
| $d$ | $c$ | $d$ | $a$ | $b$ |

and $T = \{d\}$.

(b) Let $G$ be given by

| $\circ$ | id | $\rho_1$ | $\rho_2$ | $\mu_1$ | $\mu_2$ | $\mu_3$ |
|---|---|---|---|---|---|---|
| id | id | $\rho_1$ | $\rho_2$ | $\mu_1$ | $\mu_2$ | $\mu_3$ |
| $\rho_1$ | $\rho_1$ | $\rho_2$ | id | $\mu_3$ | $\mu_1$ | $\mu_2$ |
| $\rho_2$ | $\rho_2$ | id | $\rho_1$ | $\mu_2$ | $\mu_3$ | $\mu_1$ |
| $\mu_1$ | $\mu_1$ | $\mu_2$ | $\mu_3$ | id | $\rho_1$ | $\rho_2$ |
| $\mu_2$ | $\mu_2$ | $\mu_3$ | $\mu_1$ | $\rho_2$ | id | $\rho_1$ |
| $\mu_3$ | $\mu_3$ | $\mu_1$ | $\mu_2$ | $\rho_1$ | $\rho_2$ | id |

and $T = \{\rho_1\}$.

(c) Let $G$ be as before and $T = \{\mu_1, \rho_2\}$.

**8. Programming problem.** Write a program to check if a binary operation on a set gives rise to a group. Assume that the input is given as a set $S$ of elements together with a dictionary which assigns a value to each pair of elements in $S$.

## 2.5 Additional Exercises

**1.** Which of the following multiplication tables defined on the set $G = \{a, b, c, d\}$ form a group? Support your answer in each case.

(a)

| $\circ$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $a$ | $a$ | $c$ | $d$ | $a$ |
| $b$ | $b$ | $b$ | $c$ | $d$ |
| $c$ | $c$ | $d$ | $a$ | $b$ |
| $d$ | $d$ | $a$ | $b$ | $c$ |

(b)

| $\circ$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $a$ | $a$ | $b$ | $c$ | $d$ |
| $b$ | $b$ | $a$ | $d$ | $c$ |
| $c$ | $c$ | $d$ | $a$ | $b$ |
| $d$ | $d$ | $c$ | $b$ | $a$ |

(c)

| $\circ$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $a$ | $a$ | $b$ | $c$ | $d$ |
| $b$ | $b$ | $c$ | $d$ | $a$ |
| $c$ | $c$ | $d$ | $a$ | $b$ |
| $d$ | $d$ | $a$ | $b$ | $c$ |

(d)

| $\circ$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $a$ | $a$ | $b$ | $c$ | $d$ |
| $b$ | $b$ | $a$ | $c$ | $d$ |
| $c$ | $c$ | $b$ | $a$ | $d$ |
| $d$ | $d$ | $d$ | $b$ | $c$ |

**2.** Write out Cayley tables for groups formed by the symmetries of a rectangle and for $(\mathbb{Z}_4, +)$. How many elements are in each group? Are the groups the same? Why or why not?

**3.** Prove that the set of matrices of the form

$$\begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}$$

is a group under matrix multiplication. This group, known as the **Heisenberg group**, is important in quantum physics. Matrix multiplication in the Heisenberg group is defined by

$$\begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x' & y' \\ 0 & 1 & z' \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x+x' & y+y'+xz' \\ 0 & 1 & z+z' \\ 0 & 0 & 1 \end{pmatrix}.$$

**4.** Prove or disprove that every group containing six elements is abelian.

## 2.6 Material

1. Group Definition (Quick)[1]
2. Group Definition (Expanded)[2]
3. Cayley Tables[3]
4. Motivation for definition of groups[4]
5. Subgroups[5]
6. Group Check (Integer edition)[6]
7. Direct Products of Groups[7]

## 2.7 Hints to Selected Exercises

### 2.4 · Core Exercises

**2.4.5.**

(a) How is $g^n$ defined? How does associativity come into play?

(b) Try to prove the statements first for non-negative exponents.

(c) Mathematical induction might help.

### 2.5 · Additional Exercises

**2.5.1.** (a) Not a group; (c) a group.

**2.5.4.** There is a nonabelian group containing six elements.

[1] www.socratica.com/lesson/group-definition-quick
[2] www.socratica.com/lesson/group-definition
[3] www.socratica.com/lesson/group-multiplication-tables-cayley-tables
[4] www.socratica.com/lesson/groups-motivation-for-definition
[5] www.socratica.com/lesson/subgroups
[6] www.socratica.com/lesson/group-or-not-group-integer-edition
[7] www.socratica.com/lesson/direct-products-of-groups

# Chapter 3

# Cyclic Groups and Permutation Groups

Basic learning goals

1. Definition of cyclic (sub-)group and important examples.
2. Basic properties of cyclic groups and their elements, ability to compute with them.
3. Discrete Logarithm Problem and ability to perform related computations.
4. Definition of permutation group and ability to make basic computations (products, cycle notation, order)
5. Computation of representation by transpositions and if a permutation is odd or even.

Two fundamental families of groups are the cyclic groups and permutations groups. They are crucial to the understanding of general (finite) groups. On the one hand, every element of a group gives rise to a cyclic group and this sheds light on the properties of the whole group. On the other hand, each (finite) group can be considered as a subgroup of a symmetric group.

## 3.1 Cyclic (Sub-)Groups

The groups $\mathbb{Z}$ and $\mathbb{Z}_n$, which are among the most familiar and easily understood groups, are both examples of what are called cyclic groups. In this chapter we will study the properties of cyclic groups and cyclic subgroups, which play a fundamental part in the classification of all abelian groups.

Often a subgroup will depend entirely on a single element of the group; that is, knowing that particular element will allow us to compute any other element in the subgroup.

**Example 3.1.1** Suppose that we consider $3 \in \mathbb{Z}$ and look at all multiples (both positive and negative) of 3. As a set, this is

$$3\mathbb{Z} = \{\ldots, -3, 0, 3, 6, \ldots\}.$$

It is easy to see that $3\mathbb{Z}$ is a subgroup of the integers. This subgroup is completely determined by the element 3 since we can obtain all of the other elements of the group by taking multiples of 3. Every element in the subgroup

is "generated" by 3. □

**Example 3.1.2** If $H = \{2^n : n \in \mathbb{Z}\}$, then $H$ is a subgroup of the multiplicative group of nonzero rational numbers, $\mathbb{Q}^*$. If $a = 2^m$ and $b = 2^n$ are in $H$, then $ab^{-1} = 2^m 2^{-n} = 2^{m-n}$ is also in $H$. By Proposition 2.2.8, $H$ is a subgroup of $\mathbb{Q}^*$ determined by the element 2. □

**Theorem 3.1.3** *Let $G$ be a group and $a$ be any element in $G$.*
*The set*

$$\langle a \rangle = \{a^k : k \in \mathbb{Z}\}$$

*is a subgroup of $G$. Furthermore, $\langle a \rangle$ is the smallest subgroup of $G$ that contains $a$.*

*Proof.* This was already proven in Exercise 2.4.6. ■

**Remark 3.1.4** If we are using the "+" notation, as in the case of the integers under addition, we write $\langle a \rangle = \{na : n \in \mathbb{Z}\}$.

For $a \in G$, we call $\langle a \rangle$ the **cyclic subgroup** generated by $a$. If $G$ contains some element $a$ such that $G = \langle a \rangle$, then $G$ is a **cyclic group**. In this case $a$ is a **generator** of $G$. If $a$ is an element of a group $G$, we define the **order** of $a$ to be the smallest positive integer $n$ such that $a^n = e$, and we write $|a| = n$. If there is no such integer $n$, we say that the order of $a$ is infinite and write $|a| = \infty$ to denote the order of $a$.

**Example 3.1.5** Notice that a cyclic group can have more than a single generator. Both 1 and 5 generate $\mathbb{Z}_6$; hence, $\mathbb{Z}_6$ is a cyclic group. Not every element in a cyclic group is necessarily a generator of the group. The order of $2 \in \mathbb{Z}_6$ is 3. The cyclic subgroup generated by 2 is $\langle 2 \rangle = \{0, 2, 4\}$. □

The groups $\mathbb{Z}$ and $\mathbb{Z}_n$ are cyclic groups. The elements 1 and $-1$ are generators for $\mathbb{Z}$. We can certainly generate $\mathbb{Z}_n$ with 1 although there may be other generators of $\mathbb{Z}_n$, as in the case of $\mathbb{Z}_6$.

**Example 3.1.6** The group of units, $U(9)$, in $\mathbb{Z}_9$ is a cyclic group. As a set, $U(9)$ is $\{1, 2, 4, 5, 7, 8\}$. The element 2 is a generator for $U(9)$ since

$$\begin{aligned} 2^1 &= 2 & 2^2 &= 4 \\ 2^3 &= 8 & 2^4 &= 7 \\ 2^5 &= 5 & 2^6 &= 1. \end{aligned}$$

□

**Example 3.1.7** Not every group is a cyclic group. Consider the symmetry group of an equilateral triangle $S_3$. The multiplication table for this group is Figure 1.2.7. The subgroups of $S_3$ are shown in Figure 3.1.8. Notice that every subgroup is cyclic; however, no single element generates the entire group.

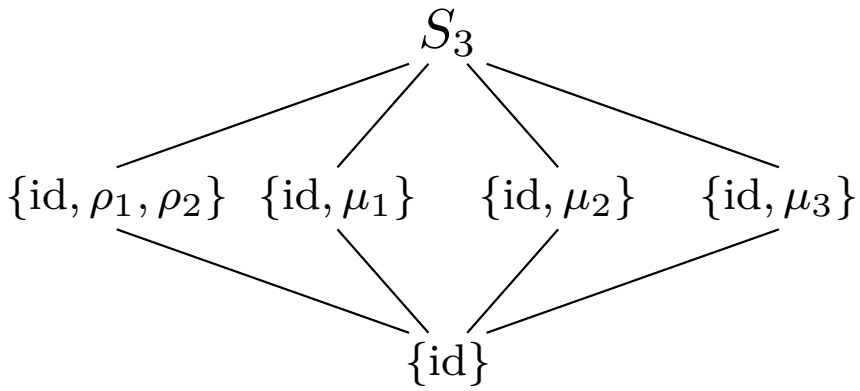



**Figure 3.1.8** Subgroups of $S_3$

□

**Theorem 3.1.9** *Every cyclic group is abelian.*

### 3.1.1 Subgroups of Cyclic Groups

We can ask some interesting questions about cyclic subgroups of a group and subgroups of a cyclic group. If $G$ is a group, which subgroups of $G$ are cyclic? If $G$ is a cyclic group, what type of subgroups does $G$ possess?

**Theorem 3.1.10** *Every subgroup of a cyclic group is cyclic.*

*Proof.* The main tools used in this proof are the division algorithm and the Principle of Well-Ordering. Let $G$ be a cyclic group generated by $a$ and suppose that $H$ is a subgroup of $G$. If $H = \{e\}$, then trivially $H$ is cyclic. Suppose that $H$ contains some other element $g$ distinct from the identity. Then $g$ can be written as $a^n$ for some integer $n$. Since $H$ is a subgroup, $g^{-1} = a^{-n}$ must also be in $H$. Since either $n$ or $-n$ is positive, we can assume that $H$ contains positive powers of $a$ and $n > 0$. Let $m$ be the smallest natural number such that $a^m \in H$. Such an $m$ exists by the Principle of Well-Ordering.

We claim that $h = a^m$ is a generator for $H$. We must show that every $h' \in H$ can be written as a power of $h$. Since $h' \in H$ and $H$ is a subgroup of $G$, $h' = a^k$ for some integer $k$. Using the division algorithm, we can find numbers $q$ and $r$ such that $k = mq + r$ where $0 \leq r < m$; hence,

$$a^k = a^{mq+r} = (a^m)^q a^r = h^q a^r.$$

So $a^r = a^k h^{-q}$. Since $a^k$ and $h^{-q}$ are in $H$, $a^r$ must also be in $H$. However, $m$ was the smallest positive number such that $a^m$ was in $H$; consequently, $r = 0$ and so $k = mq$. Therefore,

$$h' = a^k = a^{mq} = h^q$$

and $H$ is generated by $h$. ■

**Corollary 3.1.11** *The subgroups of $\mathbb{Z}$ are exactly $n\mathbb{Z}$ for $n = 0, 1, 2, \ldots$.*

**Proposition 3.1.12** *Let $G$ be a cyclic group of order $n$ and suppose that $a$ is a generator for $G$. Then $a^k = e$ if and only if $n$ divides $k$.*

**Theorem 3.1.13** *Let $G$ be a cyclic group of order $n$ and suppose that $a \in G$ is a generator of the group. If $b = a^k$, then the order of $b$ is $n/d$, where $d = \gcd(k, n)$.*

*Proof.* We wish to find the smallest integer $m$ such that $e = b^m = a^{km}$. By Proposition 3.1.12, this is the smallest integer $m$ such that $n$ divides $km$ or, equivalently, $n/d$ divides $m(k/d)$. Since $d$ is the greatest common divisor of $n$ and $k$, $n/d$ and $k/d$ are relatively prime. Hence, for $n/d$ to divide $m(k/d)$ it must divide $m$. The smallest such $m$ is $n/d$. ■

**Corollary 3.1.14** *The generators of* $(\mathbb{Z}_n, +)$ *are the integers* $r$ *such that* $1 \leq r < n$ *and* $\gcd(r, n) = 1$.

**Example 3.1.15** Let us examine the group $\mathbb{Z}_{16}$. The numbers 1, 3, 5, 7, 9, 11, 13, and 15 are the elements of $\mathbb{Z}_{16}$ that are relatively prime to 16. Each of these elements generates $\mathbb{Z}_{16}$. For example,

$$\begin{array}{lll} 1 \cdot 9 = 9 & 2 \cdot 9 = 2 & 3 \cdot 9 = 11 \\ 4 \cdot 9 = 4 & 5 \cdot 9 = 13 & 6 \cdot 9 = 6 \\ 7 \cdot 9 = 15 & 8 \cdot 9 = 8 & 9 \cdot 9 = 1 \\ 10 \cdot 9 = 10 & 11 \cdot 9 = 3 & 12 \cdot 9 = 12 \\ 13 \cdot 9 = 5 & 14 \cdot 9 = 14 & 15 \cdot 9 = 7. \end{array}$$

□

## 3.2 Discrete Logarithm Problem

One widely used cryptosystem is RSA. It is treated in more detail in the parallel course 'Discrete Mathematics'. We discuss the fundamental algorithmic problem behind it: the Discrete Logarithm Problem.

We start by looking at ordinary integers. Let $z$ and $n$ be (positive) integers. Suppose we know $z$ and $z^n$. Then it is 'easy' to figure out what $n$ is. Indeed, we could start by guessing a value $n_0$ for $n$ and then adapting our guess depending on the comparison if $z^{n_0}$ is bigger or smaller than the desired $z^n$.

Now, we consider an analogous problem for $\mathbb{Z}_k$. Assume we are given two integers $a, b \in \mathbb{Z}$ and the guarantee that $a \equiv b^n \pmod{k}$ for some non-negative integer $n$. Again, the problem is to find the exponent $n$, the discrete logarithm of $a$ with basis $b$. Then the idea from before does not work anymore as there is no order on $\mathbb{Z}_k$.

To go one step further, let us restate the latter problem in a (slightly) more general form using the terminology introduced in this chapter. Let $G$ be an arbitrary group. As we are stating an algorithmic problem, it is actually crucial in which form this group is given, but we ignore this for now. Furthermore, let $g, h$ be elements of the group with the guarantee that $h = g^n$ for some integer $n$. Note that the problem modulo $n$ described before is indeed for the group $U(n)$ of units of $\mathbb{Z}_n$ and not for $(\mathbb{Z}_n, +)$.

**Example 3.2.1** Let $a = 13$, $b = 7$ and $k = 15$. Then

$$b^2 = 7^2 \equiv 4 \pmod{15}$$

and further

$$b^3 \equiv 4 \cdot 7 = 28 \equiv 13 \pmod{15} \ .$$

Hence, the discrete logarithm of 13 with basis 7 modulo 15 is 3. □

One source of groups, which has several advantages over RSA based on modular arithmetic, is the family of elliptic curves. Those arise as solution sets of equations of the form $y^2 = x^3 + ax + b$. They have a remarkable group structure but this goes beyond the scope of this course.

## 3.3 Permutation Groups

Permutation groups are central to the study of geometric symmetries and to Galois theory, the study of finding solutions of polynomial equations. They also provide abundant examples of nonabelian groups.

Let us recall for a moment the symmetries of the equilateral triangle $\triangle ABC$ from Chapter 2. The symmetries actually consist of permutations of the three vertices, where a **permutation** of the set $S = \{A, B, C\}$ is a one-to-one and onto map $\pi : S \to S$. The three vertices have the following six permutations.

$$\begin{pmatrix} A & B & C \\ A & B & C \end{pmatrix} \quad \begin{pmatrix} A & B & C \\ C & A & B \end{pmatrix} \quad \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}$$
$$\begin{pmatrix} A & B & C \\ A & C & B \end{pmatrix} \quad \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix} \quad \begin{pmatrix} A & B & C \\ B & A & C \end{pmatrix}$$

We have used the array

$$\begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}$$

to denote the permutation that sends $A$ to $B$, $B$ to $C$, and $C$ to $A$. That is,

$$A \mapsto B$$
$$B \mapsto C$$
$$C \mapsto A.$$

The symmetries of a triangle form a group. In this chapter we will study groups of this type.

In general, the permutations of a set $X$ form a group $S_X$. If $X$ is a finite set, we can assume $X = \{1, 2, \ldots, n\}$. In this case we write $S_n$ instead of $S_X$. The following theorem says that $S_n$ is a group. We call this group the **symmetric group** on $n$ letters.

**Theorem 3.3.1** *The symmetric group on $n$ letters, $S_n$, is a group with $n!$ elements, where the binary operation is the composition of maps.*

*Proof.* The identity of $S_n$ is just the identity map that sends 1 to 1, 2 to 2, ..., $n$ to $n$. If $f : S_n \to S_n$ is a permutation, then $f^{-1}$ exists, since $f$ is one-to-one and onto; hence, every permutation has an inverse. Composition of maps is associative, which makes the group operation associative. We leave the proof that $|S_n| = n!$ as an exercise. ■

A subgroup of $S_n$ is called a **permutation group**.

**Example 3.3.2** Consider the subgroup $G$ of $S_5$ consisting of the identity permutation id and the permutations

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 5 & 4 \end{pmatrix}$$
$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 4 & 5 \end{pmatrix}$$
$$\mu = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 5 & 4 \end{pmatrix}.$$

The following table tells us how to multiply elements in the permutation group $G$.

| $\circ$ | id | $\sigma$ | $\tau$ | $\mu$ |
|---|---|---|---|---|
| id | id | $\sigma$ | $\tau$ | $\mu$ |
| $\sigma$ | $\sigma$ | id | $\mu$ | $\tau$ |
| $\tau$ | $\tau$ | $\mu$ | id | $\sigma$ |
| $\mu$ | $\mu$ | $\tau$ | $\sigma$ | id |

□

**Remark 3.3.3** Though it is natural to multiply elements in a group from left to right, functions are composed from right to left. Let $\sigma$ and $\tau$ be permutations on a set $X$. To compose $\sigma$ and $\tau$ as functions, we calculate $(\sigma \circ \tau)(x) = \sigma(\tau(x))$. That is, we do $\tau$ first, then $\sigma$. There are several ways to approach this inconsistency. *We will adopt the convention of multiplying permutations right to left. To compute $\sigma\tau$, do $\tau$ first and then $\sigma$.* That is, by $\sigma\tau(x)$ we mean $\sigma(\tau(x))$. (Another way of solving this problem would be to write functions on the right; that is, instead of writing $\sigma(x)$, we could write $(x)\sigma$. We could also multiply permutations left to right to agree with the usual way of multiplying elements in a group. Certainly all of these methods have been used.

**Example 3.3.4** Permutation multiplication is not usually commutative. Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$$
$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

Then

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix},$$

but

$$\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}.$$

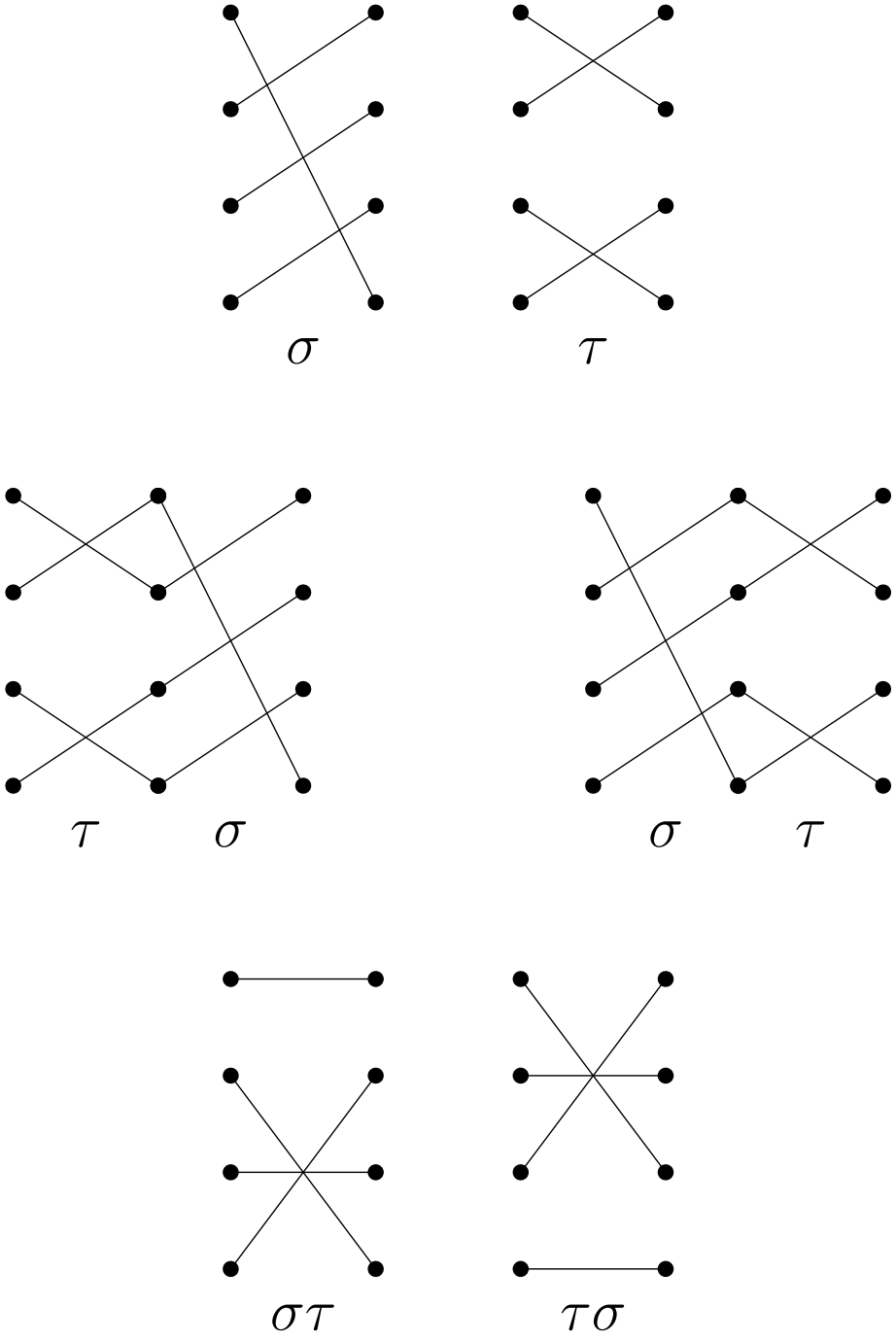



**Figure 3.3.5** The permutations in Example 3.3.4 as bipartite graphs

□

### 3.3.1 Cycle Notation

The notation that we have used to represent permutations up to this point is cumbersome, to say the least. To work effectively with permutation groups, we need a more streamlined method of writing down and manipulating permutations.

A permutation $\sigma \in S_X$ is a **cycle of length** $k$ if there exist elements

$a_1, a_2, \ldots, a_k \in X$ such that

$$\begin{aligned} \sigma(a_1) &= a_2 \\ \sigma(a_2) &= a_3 \\ &\vdots \\ \sigma(a_k) &= a_1 \end{aligned}$$

and $\sigma(x) = x$ for all other elements $x \in X$. We will write $(a_1, a_2, \ldots, a_k)$ to denote the cycle $\sigma$. Cycles are the building blocks of all permutations.

**Example 3.3.6** The permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 3 & 5 & 1 & 4 & 2 & 7 \end{pmatrix} = (1\,6\,2\,3\,5\,4)$$

is a cycle of length 6, whereas

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 3 & 5 & 6 \end{pmatrix} = (2\,4\,3)$$

is a cycle of length 3.

Not every permutation is a cycle. Consider the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 1 & 3 & 6 & 5 \end{pmatrix} = (1\,2\,4\,3)(5\,6).$$

This permutation actually contains a cycle of length 2 and a cycle of length 4.

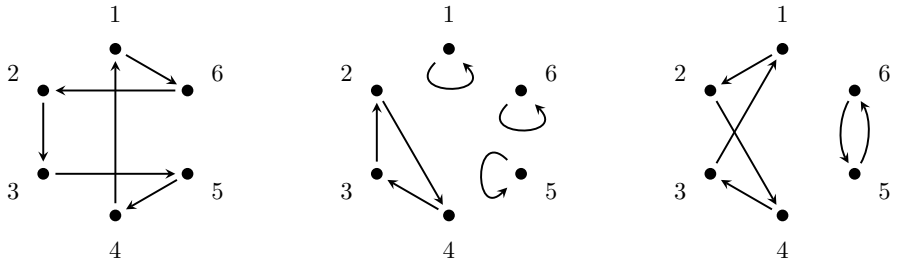



**Figure 3.3.7** The permutations in Example 3.3.6 as directed graphs

□

**Example 3.3.8** It is very easy to compute products of cycles. Suppose that

$$\sigma = (1\,3\,5\,2) \quad \text{and} \quad \tau = (2\,5\,6).$$

If we think of $\sigma$ as

$$1 \mapsto 3, \qquad 3 \mapsto 5, \qquad 5 \mapsto 2, \qquad 2 \mapsto 1,$$

and $\tau$ as

$$2 \mapsto 5, \qquad 5 \mapsto 6, \qquad 6 \mapsto 2,$$

then for $\sigma\tau$ remembering that we apply $\tau$ first and then $\sigma$, it must be the case that

$$1 \mapsto 3, \qquad 3 \mapsto 5, \qquad 5 \mapsto 6, \qquad 6 \mapsto 2 \mapsto 1,$$

or $\sigma\tau = (1\,3\,5\,6)$. If $\mu = (1\,6\,3\,4)$, then $\sigma\mu = (1\,6\,5\,2)(3\,4)$. □

Two cycles in $S_X$, $\sigma = (a_1, a_2, \ldots, a_k)$ and $\tau = (b_1, b_2, \ldots, b_l)$, are **disjoint**

if $a_i \neq b_j$ for all $i$ and $j$.

**Example 3.3.9** The cycles $(1\,3\,5)$ and $(2\,7)$ are disjoint; however, the cycles $(1\,3\,5)$ and $(3\,4\,7)$ are not. Calculating their products, we find that

$$(1\,3\,5)(2\,7) = (1\,3\,5)(2\,7)$$
$$(1\,3\,5)(3\,4\,7) = (1\,3\,4\,7\,5).$$

The product of two cycles that are not disjoint may reduce to something less complicated; the product of disjoint cycles cannot be simplified. □

**Proposition 3.3.10** *Let $\sigma$ and $\tau$ be two disjoint cycles in $S_X$. Then $\sigma\tau = \tau\sigma$.*

**Theorem 3.3.11** *Every permutation in $S_n$ can be written as the product of disjoint cycles.*

*Proof.* We can assume that $X = \{1, 2, \ldots, n\}$. If $\sigma \in S_n$ and we define $X_1$ to be $\{\sigma(1), \sigma^2(1), \ldots\}$, then the set $X_1$ is finite since $X$ is finite. Now let $i$ be the first integer in $X$ that is not in $X_1$ and define $X_2$ by $\{\sigma(i), \sigma^2(i), \ldots\}$. Again, $X_2$ is a finite set. Continuing in this manner, we can define finite disjoint sets $X_3, X_4, \ldots$. Since $X$ is a finite set, we are guaranteed that this process will end and there will be only a finite number of these sets, say $r$. If $\sigma_i$ is the cycle defined by

$$\sigma_i(x) = \begin{cases} \sigma(x) & x \in X_i \\ x & x \notin X_i \end{cases},$$

then $\sigma = \sigma_1\sigma_2\cdots\sigma_r$. Since the sets $X_1, X_2, \ldots, X_r$ are disjoint, the cycles $\sigma_1, \sigma_2, \ldots, \sigma_r$ must also be disjoint. ■

**Example 3.3.12** Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 3 & 1 & 5 & 2 \end{pmatrix}$$
$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 5 & 6 & 4 \end{pmatrix}.$$

Using cycle notation, we can write

$$\sigma = (1\,6\,2\,4)$$
$$\tau = (1\,3)(4\,5\,6)$$
$$\sigma\tau = (1\,3\,6)(2\,4\,5)$$
$$\tau\sigma = (1\,4\,3)(2\,5\,6).$$

□

**Remark 3.3.13** From this point forward we will find it convenient to use cycle notation to represent permutations. When using cycle notation, we often denote the identity permutation by (1).

### 3.3.2 Transpositions

The simplest permutation is a cycle of length 2. Such cycles are called **transpositions**. Since

$$(a_1, a_2, \ldots, a_n) = (a_1, a_n)(a_1, a_{n-1})\cdots(a_1, a_3)(a_1, a_2),$$

any cycle can be written as the product of transpositions, leading to the following proposition.

**Proposition 3.3.14** *Any permutation of a finite set containing at least two elements can be written as the product of transpositions.*

**Example 3.3.15** Consider the permutation

$$(1\,6)(2\,5\,3) = (1\,6)(2\,3)(2\,5) = (1\,6)(4\,5)(2\,3)(4\,5)(2\,5).$$

As we can see, there is no unique way to represent permutation as the product of transpositions. For instance, we can write the identity permutation as $(1\,2)(1\,2)$, as $(1\,3)(2\,4)(1\,3)(2\,4)$, and in many other ways. However, as it turns out, no permutation can be written as the product of both an even number of transpositions and an odd number of transpositions. For instance, we could represent the permutation $(1\,6)$ by

$$(2\,3)(1\,6)(2\,3)$$

or by

$$(3\,5)(1\,6)(1\,3)(1\,6)(1\,3)(3\,5)(5\,6),$$

but $(1\,6)$ will always be the product of an odd number of transpositions. □

**Lemma 3.3.16** *If the identity is written as the product of $r$ transpositions,*

$$\mathrm{id} = \tau_1\tau_2\cdots\tau_r,$$

*then $r$ is an even number.*

*Proof.* We will employ induction on $r$. A transposition cannot be the identity; hence, $r > 1$. If $r = 2$, then we are done. Suppose that $r > 2$. In this case the product of the last two transpositions, $\tau_{r-1}\tau_r$, must be one of the following cases:

$$\begin{aligned}(a,b)(a,b) &= \mathrm{id}\\ (b,c)(a,b) &= (a,c)(b,c)\\ (c,d)(a,b) &= (a,b)(c,d)\\ (a,c)(a,b) &= (a,b)(b,c),\end{aligned}$$

where $a$, $b$, $c$, and $d$ are distinct.

The first equation simply says that a transposition is its own inverse. If this case occurs, delete $\tau_{r-1}\tau_r$ from the product to obtain

$$\mathrm{id} = \tau_1\tau_2\cdots\tau_{r-3}\tau_{r-2}.$$

By induction $r-2$ is even; hence, $r$ must be even.

In each of the other three cases, we can replace $\tau_{r-1}\tau_r$ with the right-hand side of the corresponding equation to obtain a new product of $r$ transpositions for the identity. In this new product the last occurrence of $a$ will be in the next-to-the-last transposition. We can continue this process with $\tau_{r-2}\tau_{r-1}$ to obtain either a product of $r-2$ transpositions or a new product of $r$ transpositions where the last occurrence of $a$ is in $\tau_{r-2}$. If the identity is the product of $r-2$ transpositions, then again we are done, by our induction hypothesis; otherwise, we will repeat the procedure with $\tau_{r-3}\tau_{r-2}$.

At some point either we will have two adjacent, identical transpositions canceling each other out or $a$ will be shuffled so that it will appear only in the first transposition. However, the latter case cannot occur, because the identity would not fix $a$ in this instance. Therefore, the identity permutation must be the product of $r-2$ transpositions and, again by our induction hypothesis, we are done. ■

**Theorem 3.3.17** *If a permutation $\sigma$ can be expressed as the product of an even number of transpositions, then any other product of transpositions equaling $\sigma$ must also contain an even number of transpositions. Similarly, if $\sigma$ can be expressed as the product of an odd number of transpositions, then any other product of transpositions equaling $\sigma$ must also contain an odd number of transpositions.*

*Proof.* Suppose that

$$\sigma = \sigma_1\sigma_2\cdots\sigma_m = \tau_1\tau_2\cdots\tau_n,$$

where $m$ is even. We must show that $n$ is also an even number. The inverse of $\sigma$ is $\sigma_m\cdots\sigma_1$. Since

$$\mathrm{id} = \sigma\sigma_m\cdots\sigma_1 = \tau_1\cdots\tau_n\sigma_m\cdots\sigma_1,$$

$n$ must be even by Lemma 3.3.16. The proof for the case in which $\sigma$ can be expressed as an odd number of transpositions is left as an exercise. ■

In light of Theorem 3.3.17, we define a permutation to be **even** if it can be expressed as an even number of transpositions and **odd** if it can be expressed as an odd number of transpositions.

### 3.3.3 Historical Note

Lagrange first thought of permutations as functions from a set to itself, but it was Cauchy who developed the basic theorems and notation for permutations. He was the first to use cycle notation. Augustin-Louis Cauchy (1789–1857) was born in Paris at the height of the French Revolution. His family soon left Paris for the village of Arcueil to escape the Reign of Terror. One of the family's neighbors there was Pierre-Simon Laplace (1749–1827), who encouraged him to seek a career in mathematics. Cauchy began his career as a mathematician by solving a problem in geometry given to him by Lagrange. Cauchy wrote over 800 papers on such diverse topics as differential equations, finite groups, applied mathematics, and complex analysis. He was one of the mathematicians responsible for making calculus rigorous. Perhaps more theorems and concepts in mathematics have the name Cauchy attached to them than that of any other mathematician.

## 3.4 Additional insights

### 3.4.1 The Circle Group and the Roots of Unity

Recall the **circle group**,

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}.$$

The following proposition is a direct result of Proposition 2.3.6.

**Proposition 3.4.1** *The circle group is a subgroup of $\mathbb{C}^*$.*

Although the circle group has infinite order, it has many interesting finite subgroups. Suppose that $H = \{1, -1, i, -i\}$. Then $H$ is a subgroup of the circle group. Also, 1, $-1$, $i$, and $-i$ are exactly those complex numbers that satisfy the equation $z^4 = 1$. The complex numbers satisfying the equation $z^n = 1$ are called the **$n$th roots of unity**.

**Theorem 3.4.2** *If $z^n = 1$, then the $n$th roots of unity are*

$$z = \text{cis}\left(\frac{2k\pi}{n}\right),$$

*where* $k = 0, 1, \ldots, n-1$. *Furthermore, the* $n$*th roots of unity form a cyclic subgroup of* $\mathbb{T}$ *of order* $n$

*Proof.* By DeMoivre's Theorem,

$$z^n = \text{cis}\left(n\frac{2k\pi}{n}\right) = \text{cis}(2k\pi) = 1.$$

The $z$'s are distinct since the numbers $2k\pi/n$ are all distinct and are greater than or equal to 0 but less than $2\pi$. The fact that these are all of the roots of the equation $z^n = 1$ follows because a polynomial of degree $n$ can have at most $n$ roots as we will see later. We will leave the proof that the $n$th roots of unity form a cyclic subgroup of $\mathbb{T}$ as an exercise. ■

A generator for the group of the $n$th roots of unity is called a **primitive $n$th root of unity**.

**Example 3.4.3** The 8th roots of unity can be represented as eight equally spaced points on the unit circle (Figure 3.4.4). The primitive 8th roots of unity are

$$\omega = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i$$
$$\omega^3 = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i$$
$$\omega^5 = -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i$$
$$\omega^7 = \frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i.$$

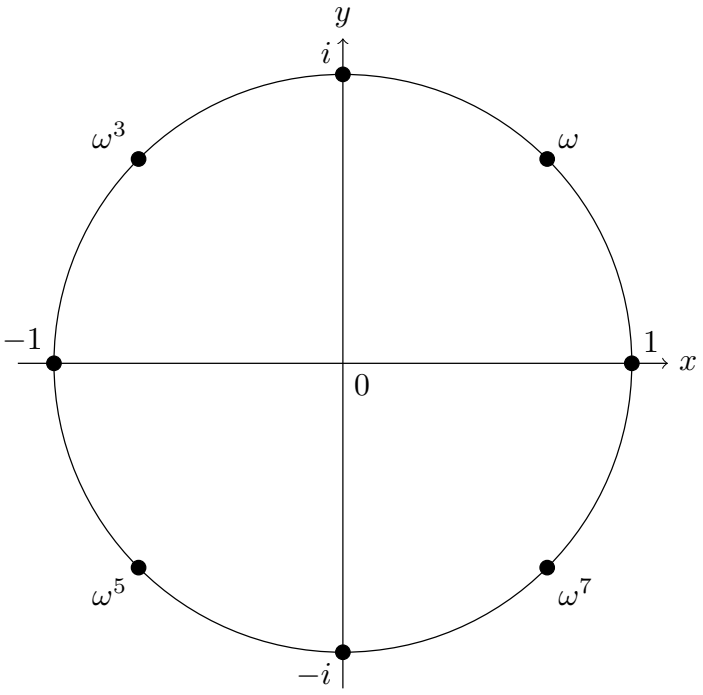



**Figure 3.4.4** 8th roots of unity

□

### 3.4.2 The Alternating Groups

One of the most important subgroups of $S_n$ is the set of all even permutations, $A_n$. The group $A_n$ is called the **alternating group on $n$ letters**.

**Theorem 3.4.5** *The set $A_n$ is a subgroup of $S_n$.*

*Proof.* Since the product of two even permutations must also be an even permutation, $A_n$ is closed. The identity is an even permutation and therefore is in $A_n$. If $\sigma$ is an even permutation, then

$$\sigma = \sigma_1\sigma_2\cdots\sigma_r,$$

where $\sigma_i$ is a transposition and $r$ is even. Since the inverse of any transposition is itself,

$$\sigma^{-1} = \sigma_r\sigma_{r-1}\cdots\sigma_1$$

is also in $A_n$. ■

**Proposition 3.4.6** *The number of even permutations in $S_n$, $n \geq 2$, is equal to the number of odd permutations; hence, the order of $A_n$ is $n!/2$.*

*Proof.* Let $A_n$ be the set of even permutations in $S_n$ and $B_n$ be the set of odd permutations. If we can show that there is a bijection between these sets, they must contain the same number of elements. Fix a transposition $\sigma$ in $S_n$. Since $n \geq 2$, such a $\sigma$ exists. Define

$$\lambda_\sigma : A_n \to B_n$$

by

$$\lambda_\sigma(\tau) = \sigma\tau.$$

Suppose that $\lambda_\sigma(\tau) = \lambda_\sigma(\mu)$. Then $\sigma\tau = \sigma\mu$ and so

$$\tau = \sigma^{-1}\sigma\tau = \sigma^{-1}\sigma\mu = \mu.$$

Therefore, $\lambda_\sigma$ is one-to-one. We will leave the proof that $\lambda_\sigma$ is surjective to the reader. ■

**Example 3.4.7** The group $A_4$ is the subgroup of $S_4$ consisting of even permutations. There are twelve elements in $A_4$:

| | | | |
|---|---|---|---|
| (1) | (1 2)(3 4) | (1 3)(2 4) | (1 4)(2 3) |
| (1 2 3) | (1 3 2) | (1 2 4) | (1 4 2) |
| (1 3 4) | (1 4 3) | (2 3 4) | (2 4 3). |

One of the end-of-chapter exercises will be to write down all the subgroups of $A_4$. You will find that there is no subgroup of order 6. Does this surprise you?

□

### 3.4.3 Properties of Dihedral groups

**Theorem 3.4.8** *The group $D_n$, $n \geq 3$, consists of all products of the two elements $r$ and $s$, where $r$ has order $n$ and $s$ has order 2, and these two elements satisfy the relation $srs = r^{-1}$.*

*Proof.* The possible motions of a regular $n$-gon are either reflections or rotations (Figure 3.5.2). There are exactly $n$ possible rotations:

$$\text{id}, \frac{360^\circ}{n}, 2 \cdot \frac{360^\circ}{n}, \ldots, (n-1) \cdot \frac{360^\circ}{n}.$$

We will denote the rotation $360^\circ/n$ by $r$. The rotation $r$ generates all of the other rotations. That is,

$$r^k = k \cdot \frac{360^\circ}{n}.$$

Label the $n$ reflections $s_1, s_2, \ldots, s_n$, where $s_k$ is the reflection that leaves vertex $k$ fixed. There are two cases of reflections, depending on whether $n$ is even or odd. If there are an even number of vertices, then two vertices are left fixed by a reflection, and $s_1 = s_{n/2+1}, s_2 = s_{n/2+2}, \ldots, s_{n/2} = s_n$. If there are an odd number of vertices, then only a single vertex is left fixed by a reflection and $s_1, s_2, \ldots, s_n$ are distinct (Figure 3.5.3). In either case, the order of each $s_k$ is two. Let $s = s_1$. Then $s^2 = 1$ and $r^n = 1$. Since any rigid motion $t$ of the $n$-gon replaces the first vertex by the vertex $k$, the second vertex must be replaced by either $k+1$ or by $k-1$. If the second vertex is replaced by $k+1$, then $t = r^k$. If the second vertex is replaced by $k-1$, then $t = r^k s$.[1] Hence, $r$ and $s$ generate $D_n$. That is, $D_n$ consists of all finite products of $r$ and $s$,

$$D_n = \{1, r, r^2, \ldots, r^{n-1}, s, rs, r^2 s, \ldots, r^{n-1} s\}.$$

We will leave the proof that $srs = r^{-1}$ as an exercise. ■

## 3.5 Core Exercises

**1.** Prove Theorem 3.1.9.

**2.** Prove Proposition 3.1.12.

**3.** Prove Proposition 3.3.10.

**4.** Recall the definition of (external) direct product from Exercise 2.4.2. Prove the following about the orders.

(a) Let $(g, h) \in G \times H$. If $g$ and $h$ have finite orders $r$ and $s$ respectively, then the order of $(g, h)$ in $G \times H$ is the least common multiple of $r$ and $s$.

(b) Let $(g_1, \ldots, g_n) \in \prod G_i$. If $g_i$ has finite order $r_i$ in $G_i$, then the order of $(g_1, \ldots, g_n)$ in $\prod G_i$ is the least common multiple of $r_1, \ldots, r_n$.

(c) Argue that the order of a permutation is the least common multiple of the lengths of its cycles.

**5.** Find $(a_1, a_2, \ldots, a_n)^{-1}$.

**6.** Discuss how to solve the logarithm problem efficiently for integers. What complexity does your algorithm have? (We are not expecting a clear answer from you here, it is more about getting a feeling for the problem).

Now, we turn to the discrete logarithm problem for finite groups. In each of the following examples, determine the exponent needed to obtain the result by exponentiation in the group from the given basis.

(a) We consider the group $U(17)$ of units of $\mathbb{Z}_{17}$. What is the discrete

[1] Since we are in an abstract group, we will adopt the convention that group elements are multiplied left to right.

logarithm of

$$3 \qquad \text{with basis} \qquad 11 \ ?$$

(b) We consider the group $S_8$ of permutations on 8 elements. What is the discrete logarithm of

$$(15423)(768) \qquad \text{with basis} \qquad (24513)(768) \ ?$$

(c) We consider the group $GL_2(\mathbb{R})$. What is the discrete logarithm of

$$\begin{pmatrix} 36 & 28 \\ 28 & 36 \end{pmatrix} \qquad \text{with basis} \qquad \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \ ?$$

**7. Dihedral groups.** Another special type of permutation group is the dihedral group. Recall the symmetry group of an equilateral triangle in Chapter 2. Such groups consist of the rigid motions of a regular $n$-sided polygon or $n$-gon. For $n = 3, 4, \ldots$, we define the **nth dihedral group** to be the group of rigid motions of a regular $n$-gon. We will denote this group by $D_n$. We can number the vertices of a regular $n$-gon by $1, 2, \ldots, n$ (Figure 3.5.1). Notice that there are exactly $n$ choices to replace the first vertex. If we replace the first vertex by $k$, then the second vertex must be replaced either by vertex $k + 1$ or by vertex $k - 1$; hence, there are $2n$ possible rigid motions of the $n$-gon. The two basic symmetries are be the rotation $r$ by $360°/n$ and a reflection $s$. From them, we get all other symmetries. We summarize this in the following observations. For a proof, see Subsection 3.4.3.

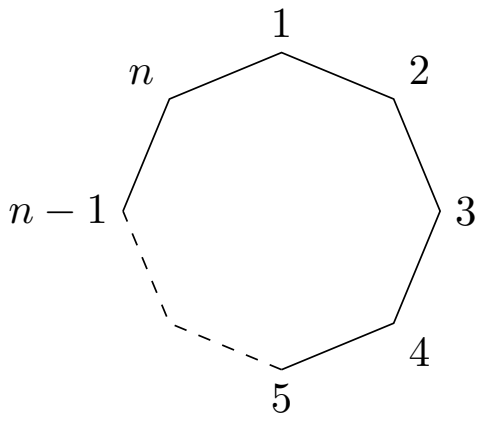



**Figure 3.5.1** A regular $n$-gon

(I) The dihedral group, $D_n$, is a subgroup of $S_n$ of order $2n$.

(II) The group $D_n$, $n \geq 3$, consists of all products of the two elements $r$ and $s$, where $r$ has order $n$ and $s$ has order 2, and these two elements satisfy the relation $srs = r^{-1}$.

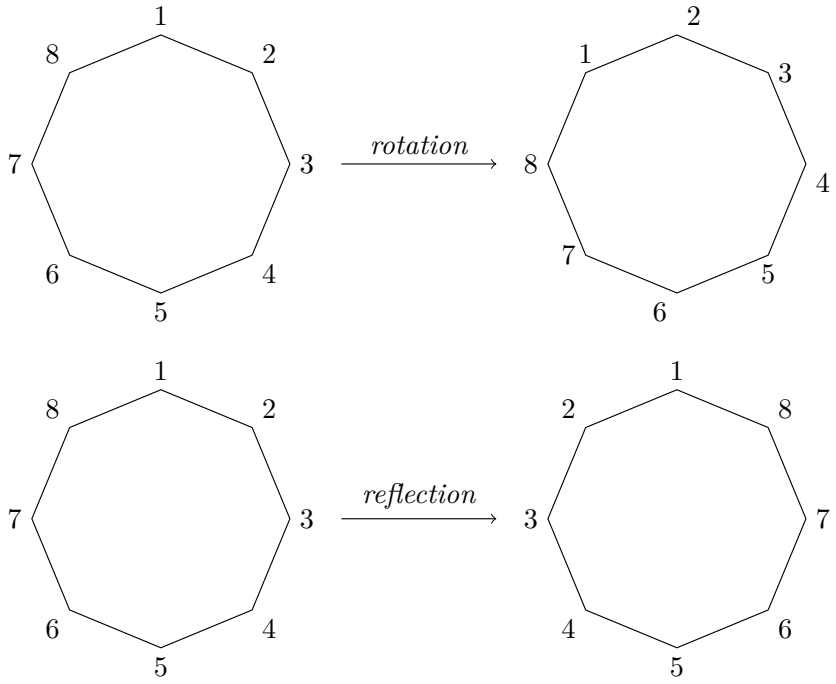



**Figure 3.5.2** Rotations and reflections of a regular $n$-gon

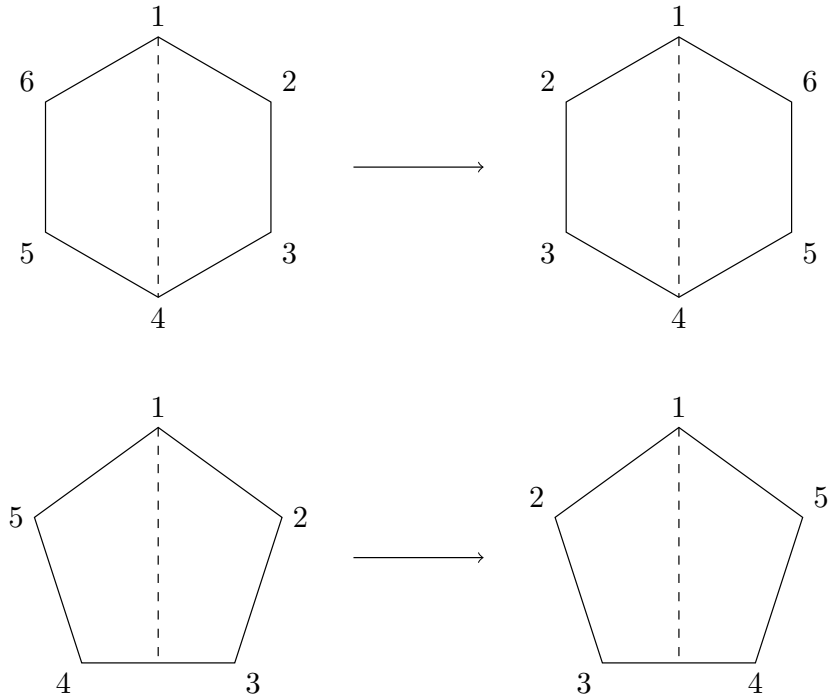



**Figure 3.5.3** Types of reflections of a regular $n$-gon

(a) What is the order of $rsr^2srs^3$ in $D_{12}$?

(b) What is the cycle decomposition of $r^3$ and of $s$ in $D_{12}$?

(c) List all cyclic subgroups of $D_5$ and $D_6$. Argue that you have found all cyclic subgroups.

**8. Programming problem.** Write a program to compute the multiplicative inverse of an element $z$ in $\mathbb{Z}_k$. Make sure to check that $z$ actually has a multiplicative inverse.

Note that there are (at least) two different approaches for solving this algorithmic problem. Compare their efficiency.

## 3.6 Additional Exercises

**1.** Prove or disprove each of the following statements.

(a) All of the generators of $\mathbb{Z}_{60}$ are prime.

(b) $U(8)$ is cyclic.

(c) $\mathbb{Q}$ is cyclic.

(d) If every proper subgroup of a group $G$ is cyclic, then $G$ is a cyclic group.

(e) A group with a finite number of subgroups is finite.

**2.** Find the order of each of the following elements.

(a) $5 \in \mathbb{Z}_{12}$

(b) $\sqrt{3} \in \mathbb{R}$

(c) $\sqrt{3} \in \mathbb{R}^*$

(d) $-i \in \mathbb{C}^*$

(e) $72 \in \mathbb{Z}_{240}$

(f) $312 \in \mathbb{Z}_{471}$

**3.** List all of the elements in each of the following subgroups.

(a) The subgroup of $\mathbb{Z}$ generated by 7

(b) The subgroup of $\mathbb{Z}_{24}$ generated by 15

(c) All subgroups of $\mathbb{Z}_{12}$

(d) All subgroups of $\mathbb{Z}_{60}$

(e) All subgroups of $\mathbb{Z}_{13}$

(f) All subgroups of $\mathbb{Z}_{48}$

(g) The subgroup generated by 3 in $U(20)$

(h) The subgroup generated by 5 in $U(18)$

(i) The subgroup of $\mathbb{R}^*$ generated by 7

(j) The subgroup of $\mathbb{C}^*$ generated by $i$ where $i^2 = -1$

(k) The subgroup of $\mathbb{C}^*$ generated by $2i$

(l) The subgroup of $\mathbb{C}^*$ generated by $(1+i)/\sqrt{2}$

(m) The subgroup of $\mathbb{C}^*$ generated by $(1+\sqrt{3}\,i)/2$

**4.** Find the subgroups of $GL_2(\mathbb{R})$ generated by each of the following matrices.

(a) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$

(b) $\begin{pmatrix} 0 & 1/3 \\ 3 & 0 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{pmatrix}$

**5.** Write the following permutations in cycle notation.

(a) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 4 & 2 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 3 & 2 & 5 \end{pmatrix}$

**6.** Compute each of the following.

(a) $(1\,3\,4\,5)(2\,3\,4)$

(b) $(1\,2)(1\,2\,5\,3)$

(c) $(1\,4\,3)(2\,3)(2\,4)$

(d) $(1\,4\,2\,3)(3\,4)(5\,6)(1\,3\,2\,4)$

(e) $(1\,2\,5\,4)(1\,3)(2\,5)$

(f) $(1\,2\,5\,4)(1\,3)(2\,5)^2$

(g) $(1\,2\,5\,4)^{-1}(1\,2\,3)(4\,5)(1\,2\,5\,4)$

(h) $(1\,2\,5\,4)^2(1\,2\,3)(4\,5)$

(i) $(1\,2\,3)(4\,5)(1\,2\,5\,4)^{-2}$

(j) $(1\,2\,5\,4)^{100}$

(k) $|(1\,2\,5\,4)|$

(l) $|(1\,2\,5\,4)^2|$

(m) $(1\,2)^{-1}$

(n) $(1\,2\,5\,3\,7)^{-1}$

(o) $[(1\,2)(3\,4)(1\,2)(4\,7)]^{-1}$

(p) $[(1\,2\,3\,5)(4\,6\,7)]^{-1}$

**7.** Express the following permutations as products of transpositions and identify them as even or odd.

(a) $(1\,4\,3\,5\,6)$

(b) $(1\,5\,6)(2\,3\,4)$

(c) $(1\,4\,2\,6)(1\,4\,2)$

(d) $(1\,7\,2\,5\,4)(1\,4\,2\,3)(1\,5\,4\,6\,3\,2)$

(e) $(1\,4\,2\,6\,3\,7)$

**8.** Show that a 3-cycle is an even permutation.

**9.** Find an element of largest order in $S_n$ for $n = 3, \ldots, 10$.

**10.** Prove that if $G$ is a cyclic group of order $m$ and $d \mid m$, then $G$ must have a subgroup of order $d$.

## 3.7 Material

1. Cyclic groups[1]
2. Order of an element[2]
3. Symmetric group[3]
4. Cycle notation for permutation[4]

[1] www.socratica.com/lesson/cyclic-groups
[2] www.socratica.com/lesson/order-of-an-element
[3] www.socratica.com/lesson/symmetric-groups
[4] www.socratica.com/lesson/cycle-notation-for-permutations

5. Dihedral Groups[5]
6. Matrix Groups[6]

Not core topics of the course but for further interest:

1. Book on 'Cryptography'[7]
2. : a little more on elliptic curves[8]

## 3.8 Hints to Selected Exercises

### 3.5 · Core Exercises

**3.5.5.** $(a_1, a_2, \ldots, a_n)^{-1} = (a_1, a_n, a_{n-1}, \ldots, a_2)$

### 3.6 · Additional Exercises

**3.6.1.** (a) False; (c) false; (e) true.

**3.6.2.** (a) 12; (c) infinite; (e) 10.

**3.6.3.** (a) $7\mathbb{Z} = \{\ldots, -7, 0, 7, 14, \ldots\}$; (b) $\{0, 3, 6, 9, 12, 15, 18, 21\}$; (c) $\{0\}$, $\{0, 6\}$, $\{0, 4, 8\}$, $\{0, 3, 6, 9\}$, $\{0, 2, 4, 6, 8, 10\}$; (g) $\{1, 3, 7, 9\}$; (j) $\{1, -1, i, -i\}$.

**3.6.4.** (a)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

(c)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

**3.6.5.** (a) (1 2 4 5 3); (c) (1 3)(2 5).

**3.6.6.** (a) (1 3 5)(2 4); (c) (1 4)(2 3); (e) (1 3 2 4); (g) (1 3 4)(2 5); (n) (1 7 3 5 2).

**3.6.7.** (a) (1 6)(1 5)(1 3)(1 4); (c) (1 6)(1 4)(1 2).

[5] www.socratica.com/lesson/dihedral-groups
[6] www.socratica.com/lesson/matrix-groups
[7] link.springer.com/content/pdf/10.1007/978-3-319-94818-8.pdf?pdf=button
[8] brilliant.org/wiki/elliptic-curves/

# Chapter 4

# Homomorphisms

Basic learning goals

1. Recognizing and checking group homomorphisms.
2. Basic properties of group homomorphisms.
3. Recognizing, checking and basic properties of group isomorphisms.
4. Characterization of cyclic and abelian groups (Fundamental Theorem of Finite Abelian Groups).
5. Normal subgroups, their interplay with homomorphisms.
6. Free groups and generated groups.

One of the most fundamental ideas of algebra is the concept of a homomorphism. In the study of groups, we already identified structures of certain operations on sets. Now, the next crucial insights come from the study of structure-preserving maps. This generalizes the idea of linear maps between vector spaces as considered in linear algebra. Recall that linear maps take the sum of two vectors to the sum of their images. This idea is relaxed to maps taking composition in one group to the composition in another group.

## 4.1 Group Homomorphisms

A **homomorphism** between groups $(G, \cdot)$ and $(H, \circ)$ is a map $\phi : G \to H$ such that

$$\phi(g_1 \cdot g_2) = \phi(g_1) \circ \phi(g_2)$$

for $g_1, g_2 \in G$. The range of $\phi$ in $H$ is called the **homomorphic image** of $\phi$.

**Example 4.1.1** First note that a vector space is an abelian group. For the special vector space $\mathbb{R}^k$, one can directly see this from Exercise 2.4.2 and the basic insight that $(\mathbb{R}, +)$ is an abelian group.

Now, recall the map from Example 1.1.1 given by $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ with

$$T(x_1, x_2) = (2x_1 + 5x_2, -4x_1 + 3x_2) \ .$$

It maps the abelian group $(\mathbb{R}^2, +)$ to itself while preserving the additive structure, that is

$$T(x_1 + y_1, x_2 + y_2) = T(x_1, x_2) + T(y_1, y_2) \ .$$

More generally, recall from the introduction of Chapter 1 that a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ fulfills $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$. Hence, a linear map is a

homomorphism between the additive structure of two vector spaces. Note that we did not require compatibility with scalar multiplication as we only care about the additive structure of a vector space here. □

More generally, we use homomorphisms to study relationships between groups.

**Example 4.1.2** The symmetric group $S_n$ and the group $\mathbb{Z}_2$ are related by the fact that $S_n$ can be divided into even and odd permutations that exhibit a group structure like that $\mathbb{Z}_2$, as shown in the following multiplication table.

| | even | odd |
|---|---|---|
| even | even | odd |
| odd | odd | even |

□

**Example 4.1.3** Let $G$ be a group and $g \in G$. Define a map $\phi : \mathbb{Z} \to G$ by $\phi(n) = g^n$. Then $\phi$ is a group homomorphism, since

$$\phi(m+n) = g^{m+n} = g^m g^n = \phi(m)\phi(n).$$

This homomorphism maps $\mathbb{Z}$ onto the cyclic subgroup of $G$ generated by $g$. □

**Example 4.1.4** Let $G = GL_2(\mathbb{R})$. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is in $G$, then the determinant is nonzero; that is, $\det(A) = ad - bc \neq 0$. Also, for any two elements $A$ and $B$ in $G$, $\det(AB) = \det(A)\det(B)$. Using the determinant, we can define a homomorphism $\phi : GL_2(\mathbb{R}) \to \mathbb{R}^*$ by $A \mapsto \det(A)$. □

**Example 4.1.5** Recall that the circle group $\mathbb{T}$ consists of all complex numbers $z$ such that $|z| = 1$. We can define a homomorphism $\phi$ from the additive group of real numbers $\mathbb{R}$ to $\mathbb{T}$ by $\phi : \theta \mapsto \cos\theta + i\sin\theta$. Indeed,

$$\begin{aligned} \phi(\alpha+\beta) &= \cos(\alpha+\beta) + i\sin(\alpha+\beta) \\ &= (\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\sin\alpha\cos\beta + \cos\alpha\sin\beta) \\ &= (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) \\ &= \phi(\alpha)\phi(\beta). \end{aligned}$$

Geometrically, we are simply wrapping the real line around the circle in a group-theoretic fashion. □

The following proposition lists some basic properties of group homomorphisms.

**Proposition 4.1.6** *Let $\phi : G_1 \to G_2$ be a homomorphism of groups.*

1. *If $e$ is the identity of $G_1$, then $\phi(e)$ is the identity of $G_2$.*
2. *For any element $g \in G_1$, $\phi(g^{-1}) = [\phi(g)]^{-1}$.*
3. *For any element $g \in G_1$ and $n \in \mathbb{Z}$, $\phi(g^n) = [\phi(g)]^n$.*
4. *If $H_1$ is a subgroup of $G_1$, then $\phi(H_1)$ is a subgroup of $G_2$.*

Let $\phi : G \to H$ be a group homomorphism and suppose that $e$ is the identity of $H$. By Proposition 4.1.6, $\phi^{-1}(\{e\})$ is a subgroup of $G$. This subgroup is called the **kernel** of $\phi$ and will be denoted by $\ker\phi$.

**Example 4.1.7** Let us examine the homomorphism $\phi : GL_2(\mathbb{R}) \to \mathbb{R}^*$ defined by $A \mapsto \det(A)$. Since 1 is the identity of $\mathbb{R}^*$, the kernel of this homomorphism is all $2 \times 2$ matrices having determinant one. That is, $\ker \phi = SL_2(\mathbb{R})$. □

**Example 4.1.8** The kernel of the group homomorphism $\phi : \mathbb{R} \to \mathbb{C}^*$ defined by $\phi(\theta) = \cos\theta + i \sin\theta$ is $\{2\pi n : n \in \mathbb{Z}\}$. Notice that $\ker \phi \cong \mathbb{Z}$. □

**Example 4.1.9** Suppose that we wish to determine all possible homomorphisms $\phi$ from $\mathbb{Z}_7$ to $\mathbb{Z}_{12}$. Since the kernel of $\phi$ must be a subgroup of $\mathbb{Z}_7$, there are only two possible kernels, $\{0\}$ and all of $\mathbb{Z}_7$. The image of a subgroup of $\mathbb{Z}_7$ must be a subgroup of $\mathbb{Z}_{12}$. Hence, there is no injective homomorphism; otherwise, $\mathbb{Z}_{12}$ would have a subgroup of order 7, which is impossible. Consequently, the only possible homomorphism from $\mathbb{Z}_7$ to $\mathbb{Z}_{12}$ is the one mapping all elements to zero. □

**Example 4.1.10** Let $G$ be a group. Suppose that $g \in G$ and $\phi$ is the homomorphism from $\mathbb{Z}$ to $G$ given by $\phi(n) = g^n$. If the order of $g$ is infinite, then the kernel of this homomorphism is $\{0\}$ since $\phi$ maps $\mathbb{Z}$ onto the cyclic subgroup of $G$ generated by $g$. However, if the order of $g$ is finite, say $n$, then the kernel of $\phi$ is $n\mathbb{Z}$. □

## 4.2 Isomorphisms

Many groups may appear to be different at first glance, but can be shown to be the same by a simple renaming of the group elements. For example, $\mathbb{Z}_4$ and the subgroup of the circle group $\mathbb{T}$ generated by $i$ can be shown to be the same by demonstrating a one-to-one correspondence between the elements of the two groups and between the group operations. In such a case we say that the groups are isomorphic. This is exactly the case when there is a bijective homomorphism between the groups, i.e., whose kernel is trivial.

Two groups $(G, \cdot)$ and $(H, \circ)$ are **isomorphic** if there exists a group homomorphism $\phi : G \to H$ that is a one-to-one and onto map. If $G$ is isomorphic to $H$, we write $G \cong H$. The map $\phi$ is called an **isomorphism**.

**Example 4.2.1** To show that $\mathbb{Z}_4 \cong \langle i \rangle$, define a map $\phi : \mathbb{Z}_4 \to \langle i \rangle$ by $\phi(n) = i^n$. We must show that $\phi$ is bijective and preserves the group operation. The map $\phi$ is one-to-one and onto because

$$\begin{aligned} \phi(0) &= 1 \\ \phi(1) &= i \\ \phi(2) &= -1 \\ \phi(3) &= -i. \end{aligned}$$

Since

$$\phi(m+n) = i^{m+n} = i^m i^n = \phi(m)\phi(n),$$

the group operation is preserved. □

**Example 4.2.2** We can define an isomorphism $\phi$ from the additive group of real numbers $(\mathbb{R}, +)$ to the multiplicative group of positive real numbers $(\mathbb{R}^+, \cdot)$ with the exponential map; that is,

$$\phi(x+y) = e^{x+y} = e^x e^y = \phi(x)\phi(y).$$

Of course, we must still show that $\phi$ is one-to-one and onto, but this can be determined using calculus. □

**Example 4.2.3** The integers are isomorphic to the subgroup of $\mathbb{Q}^*$ consisting of elements of the form $2^n$. Define a map $\phi : \mathbb{Z} \to \mathbb{Q}^*$ by $\phi(n) = 2^n$. Then

$$\phi(m+n) = 2^{m+n} = 2^m 2^n = \phi(m)\phi(n).$$

By definition the map $\phi$ is onto the subset $\{2^n : n \in \mathbb{Z}\}$ of $\mathbb{Q}^*$. To show that the map is injective, assume that $m \neq n$. If we can show that $\phi(m) \neq \phi(n)$, then we are done. Suppose that $m > n$ and assume that $\phi(m) = \phi(n)$. Then $2^m = 2^n$ or $2^{m-n} = 1$, which is impossible since $m - n > 0$. □

**Example 4.2.4** The groups $\mathbb{Z}_8$ and $\mathbb{Z}_{12}$ cannot be isomorphic since they have different orders; however, it is true that $U(8) \cong U(12)$. We know that

$$U(8) = \{1, 3, 5, 7\}$$
$$U(12) = \{1, 5, 7, 11\}.$$

An isomorphism $\phi : U(8) \to U(12)$ is then given by

$$1 \mapsto 1$$
$$3 \mapsto 5$$
$$5 \mapsto 7$$
$$7 \mapsto 11.$$

The map $\phi$ is not the only possible isomorphism between these two groups. We could define another isomorphism $\psi$ by $\psi(1) = 1$, $\psi(3) = 11$, $\psi(5) = 5$, $\psi(7) = 7$. In fact, both of these groups are isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$ (see Example 2.2.5 in Chapter 2). □

**Example 4.2.5** Even though $S_3$ and $\mathbb{Z}_6$ possess the same number of elements, we would suspect that they are not isomorphic, because $\mathbb{Z}_6$ is abelian and $S_3$ is nonabelian. To demonstrate that this is indeed the case, suppose that $\phi : \mathbb{Z}_6 \to S_3$ is an isomorphism. Let $a, b \in S_3$ be two elements such that $ab \neq ba$. Since $\phi$ is an isomorphism, there exist elements $m$ and $n$ in $\mathbb{Z}_6$ such that

$$\phi(m) = a \quad \text{and} \quad \phi(n) = b.$$

However,

$$ab = \phi(m)\phi(n) = \phi(m+n) = \phi(n+m) = \phi(n)\phi(m) = ba,$$

which contradicts the fact that $a$ and $b$ do not commute. □

**Theorem 4.2.6** *Let $\phi : G \to H$ be an isomorphism of two groups. Then the following statements are true.*

1. *$\phi^{-1} : H \to G$ is an isomorphism.*
2. *$|G| = |H|$.*
3. *If $G$ is abelian, then $H$ is abelian.*
4. *If $G$ is cyclic, then $H$ is cyclic.*
5. *If $G$ has a subgroup of order $n$, then $H$ has a subgroup of order $n$.*

*Proof.* Assertions (1) and (2) follow from the fact that $\phi$ is a bijection.

(3) Suppose that $h_1$ and $h_2$ are elements of $H$. Since $\phi$ is onto, there exist

elements $g_1, g_2 \in G$ such that $\phi(g_1) = h_1$ and $\phi(g_2) = h_2$. Therefore,

$$h_1h_2 = \phi(g_1)\phi(g_2) = \phi(g_1g_2) = \phi(g_2g_1) = \phi(g_2)\phi(g_1) = h_2h_1.$$

(4) Let $g$ be a generator of $G$. Then

$$\phi(G) = \phi\left(\{g^k \mid k \in \mathbb{Z}\}\right) = \{\phi(g^k) \mid k \in \mathbb{Z}\} \ .$$

Since $\phi(G) = H$, we get from Proposition 4.1.6 that

$$H = \{\phi(g^k) \mid k \in \mathbb{Z}\} = \{\phi(g)^k \mid k \in \mathbb{Z}\} \ .$$

Therefore, also $H$ is cyclic and the generators of $G$ and $H$ are in bijection.

(5) Let $G_1$ be a subgroup of $G$ of order $n$. Then $\phi(G_1)$ is a subgroup of $H$ by Proposition 4.1.6. Since $\phi$ is a bijection, this subgroup has cardinality $|\phi(G_1)| = |G_1| = n$. ■

We are now in a position to characterize all cyclic groups.

**Theorem 4.2.7** *All cyclic groups of infinite order are isomorphic to $\mathbb{Z}$.*

*Proof.* Let $G$ be a cyclic group with infinite order and suppose that $a$ is a generator of $G$. Define a map $\phi : \mathbb{Z} \to G$ by $\phi : n \mapsto a^n$. Then

$$\phi(m + n) = a^{m+n} = a^m a^n = \phi(m)\phi(n).$$

To show that $\phi$ is injective, suppose that $m$ and $n$ are two elements in $\mathbb{Z}$, where $m \neq n$. We can assume that $m > n$. We must show that $a^m \neq a^n$. Let us suppose the contrary; that is, $a^m = a^n$. In this case $a^{m-n} = e$, where $m-n > 0$, which contradicts the fact that $a$ has infinite order. Our map is onto since any element in $G$ can be written as $a^n$ for some integer $n$ and $\phi(n) = a^n$. ■

**Theorem 4.2.8** *If $G$ is a cyclic group of order $n$, then $G$ is isomorphic to $\mathbb{Z}_n$.*

*Proof.* Let $G$ be a cyclic group of order $n$ generated by $a$ and define a map $\phi : \mathbb{Z}_n \to G$ by $\phi : k \mapsto a^k$, where $0 \leq k < n$. One can check right from the definitions that $\phi$ is an isomorphism. ■

The main goal in group theory is to classify all groups; however, it makes sense to consider two groups to be the same if they are isomorphic. We state this result in the following theorem.

**Theorem 4.2.9** *The isomorphism of groups determines an equivalence relation on the class of all groups.*

Hence, we can modify our goal of classifying all groups to classifying all groups **up to isomorphism**; that is, we will consider two groups to be the same if they are isomorphic.

We finish the discussion of isomorphisms with an important class of groups which has a nice representative up to isomorphism. The Fundamental Theorem of Finite Abelian Groups tells us that every finite abelian group is isomorphic to a direct product of cyclic groups whose order is a prime power.

**Theorem 4.2.10 Fundamental Theorem of Finite Abelian Groups.** *Every finite abelian group $G$ is isomorphic to a direct product of cyclic groups of the form*

$$\mathbb{Z}_{p_1^{\alpha_1}} \times \mathbb{Z}_{p_2^{\alpha_2}} \times \cdots \times \mathbb{Z}_{p_n^{\alpha_n}}$$

*here the $p_i$'s are primes (not necessarily distinct).*

**Example 4.2.11** Suppose that we wish to classify all abelian groups of order $540 = 2^2 \cdot 3^3 \cdot 5$. The Fundamental Theorem of Finite Abelian Groups tells us

that we have the following six possibilities.

- $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$;
- $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_9 \times \mathbb{Z}_5$;
- $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{27} \times \mathbb{Z}_5$;
- $\mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$;
- $\mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_9 \times \mathbb{Z}_5$;
- $\mathbb{Z}_4 \times \mathbb{Z}_{27} \times \mathbb{Z}_5$.

□

## 4.3 Normal subgroups and kernels

A subgroup $H$ of a group $G$ is **normal** in G if

$$gHg^{-1} := \{ghg^{-1} \mid h \in H\} = H \text{ for all } g \in G.$$

**Example 4.3.1** Let $G$ be an abelian group. Every subgroup $H$ of $G$ is a normal subgroup. Since $ghg^{-1} = h$ for all $g \in G$ and $h \in H$ by commutativity, it will always be the case that $gHg^{-1} = H$. □

**Proposition 4.3.2** *Let $\phi : G_1 \to G_2$ be a homomorphism of groups and $H_2$ is a subgroup of $G_2$.*

1. *$\phi^{-1}(H_2) = \{g \in G_1 : \phi(g) \in H_2\}$ is a subgroup of $G_1$.*
2. *If $H_2$ is normal in $G_2$, then $\phi^{-1}(H_2)$ is normal in $G_1$.*

*Proof.* Let $H_2$ be a subgroup of $G_2$ and define $H_1$ to be $\phi^{-1}(H_2)$; that is, $H_1$ is the set of all $g \in G_1$ such that $\phi(g) \in H_2$. The identity is in $H_1$ since $\phi(e) = e'$. If $a$ and $b$ are in $H_1$, then $\phi(ab^{-1}) = \phi(a)[\phi(b)]^{-1}$ is in $H_2$ since $H_2$ is a subgroup of $G_2$. Therefore, $ab^{-1} \in H_1$ and $H_1$ is a subgroup of $G_1$.

If $H_2$ is normal in $G_2$, we must show that $g^{-1}hg \in H_1$ for $h \in H_1$ and $g \in G_1$. But

$$\phi(g^{-1}hg) = [\phi(g)]^{-1}\phi(h)\phi(g) \in H_2,$$

since $H_2$ is a normal subgroup of $G_2$. Therefore, $g^{-1}hg \in H_1$. ■

Note that the trivial subgroup consisting only of the identity element is normal. Together with Proposition 4.3.2 this yields the following theorem, which says that with every homomorphism of groups we can naturally associate a normal subgroup.

**Theorem 4.3.3** *Let $\phi : G \to H$ be a group homomorphism. Then the kernel of $\phi$ is a normal subgroup of $G$.*

Let $G$ be a group and $T$ be an arbitrary subset. Then the **normal closure** of $T$ is the smallest normal subgroup of $G$ that contains $T$. Note that this is well-defined as there is always a normal subgroup containing $T$, at least the group $G$ itself.

We list a few insights about the normal closure. The proof is omitted as this statement is not so central for the course but it gives background on the construction which we will exhibit in Section 6.3.

**Proposition 4.3.4**

1. *The normal closure is the intersection of all normal subgroups of $G$ containing $T$.*

2. *The normal closure is generated by* $\{g^{-1}tg \mid g \in G, t \in T\}$ *that means that it equals*

$$\left\{g_1^{-1}t_1^{\sigma_1}g_1g_2^{-1}\dots t_n^{\sigma_n}g_n \mid n \geq 0, g_i \in G, t_i \in T, \sigma_i \in \{1,-1\}\forall i \in \{1,\dots,n\}\right\}.$$

## 4.4 Free groups

The **free group** on a given set $S$ is formed by all words over the alphabet $S$, that is all expressions of the form

$$s_1^{\sigma_1}\dots s_n^{\sigma_n}$$

where $s_1,\dots,s_n \in S$ and $\sigma_1,\dots,\sigma_n \in \{-1,1\}$. The binary operation of the group is concatenation of words. More precisely, the elements of the group are the equivalence classes of words under the identification of subwords $xx^{-1}$ and $x^{-1}x$ for $x \in S$ with the empty word $\epsilon$. We will denote the free group by $\mathcal{F}(S)$. The neutral element of this group is the empty word $\epsilon$. For example, for $S = \{a,b\}$, the words $aa^{-1}$ or $b^{-1}b$ are in the same equivalence class as $\epsilon$ and the words $ab^{-1}aa^{-1}b$ and $aa^{-1}a^{-1}aa$ are in the same equivalence class as $a$ (among many others). The proof of the following statement implies that the free group is actually a group.

Recall from Subsection 2.2.1 what it means for a group to be generated. With this, the free group on $S$ is generated by $S$. The number of generators of a free group is the **rank** of the free group.

**Proposition 4.4.1** *Let $H$ be the subgroup of a group $G$ that is generated by $\{g_i \in G : i \in I\}$. Then $h \in H$ exactly when it is a product of the form*

$$h = g_{i_1}^{\alpha_1}\cdots g_{i_n}^{\alpha_n},$$

*where the $g_{i_k}$s are not necessarily distinct.*

*Proof.* Let $K$ be the set of all products of the form $g_{i_1}^{\alpha_1}\cdots g_{i_n}^{\alpha_n}$, where the $g_{i_k}$s are not necessarily distinct. Certainly $K$ is a subset of $H$. We need only show that $K$ is a subgroup of $G$. If this is the case, then $K = H$, since $H$ is the smallest subgroup containing all the $g_i$s.

Clearly, the set $K$ is closed under the group operation. Since $g_i^0 = 1$, the identity is in $K$. It remains to show that the inverse of an element $g = g_{i_1}^{k_1}\cdots g_{i_n}^{k_n}$ in $K$ must also be in $K$. However,

$$g^{-1} = (g_{i_1}^{k_1}\cdots g_{i_n}^{k_n})^{-1} = (g_{i_n}^{-k_n}\cdots g_{i_1}^{-k_1}).$$

■

**Example 4.4.2** The free group on one element $a$ denoted by $\mathcal{F}(\{a\})$ is isomorphic with $\mathbb{Z}$. □

Let $G$ be a group and $\{g_s \colon s \in S\} \subseteq G$ be a subset generating it, e.g., the set of all elements of $G$. Then there is a **canonical homomorphism** such that

$$s \mapsto g_s \tag{4.4.1}$$

for all $s \in S$. Taking a suitable subset of the elements in the kernel yields a so-called **presentation** of $G$. We will discuss this further in Section 6.3.

## 4.5 Core Exercises

**1.** Prove Proposition 4.1.6.

**2.** Consider the map $\psi\colon S_3 \to GL_3(\mathbb{R})$ given by $\pi \mapsto M$ where

$$M_{ij} = \begin{cases} 1 & \text{for } i = \pi(j), \\ 0 & \text{otherwise.} \end{cases}$$

For example, one obtains

$$\psi\left((2\,1\,3)\right) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} .$$

(a) Determine explicitly the image of $\psi$.

(b) Show that $\psi$ is a group homomorphism.

(c) What is the kernel of $\psi$? Is it an isomorphism?

(d) Give a generalization of the statement for $S_n$ for arbirary natural numbers $n$.

(e) What does the determinant of the image of a permutation say about the permutation?

**3.** Which of the following maps are homomorphisms? If the map is a homomorphism, what is the kernel?

(a) $\phi : \mathbb{R}^* \to GL_2(\mathbb{R})$ defined by

$$\phi(a) = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$$

(b) $\phi : \mathbb{R} \to GL_2(\mathbb{R})$ defined by

$$\phi(a) = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$$

(c) $\phi : GL_2(\mathbb{R}) \to \mathbb{R}$ defined by

$$\phi\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = a + d$$

(d) $\phi : GL_2(\mathbb{R}) \to \mathbb{R}^*$ defined by

$$\phi\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = ad - bc$$

(e) $\phi : \mathbb{M}_2(\mathbb{R}) \to \mathbb{R}$ defined by

$$\phi\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = b,$$

where $\mathbb{M}_2(\mathbb{R})$ is the additive group of $2 \times 2$ matrices with entries in $\mathbb{R}$.

**4.**

(a) List all subgroups of $S_3$. Which ones are normal? What is the normal closure for each of those that are not normal?

(b) List all subgroups of $D_4$. Which ones are normal? What is the normal closure for each of those that are not normal?

(c) List all subgroups of $Q_8$. Which ones are normal? What is the normal closure for each of those that are not normal?

**5.**

(a) Why are $\mathbb{Z}_8$ and $\mathbb{Z}_2 \times \mathbb{Z}_4$ not isomorphic?

(b) List all abelian groups of order 24.

**6.** Let $\mathrm{Sym}(X)$ be the group of all bijective maps on $X$. We consider a map $\phi\colon G \to \mathrm{Sym}(X)$. Show that $\phi$ is a group homomorphism if and only if it fulfills

(a) $\phi(e)(x) = x$ for all $x \in X$;

(b) $\phi(g_1 g_2)(x) = \phi(g_1)\,(\phi(g_2)(x))$ for all $x \in X$ and all $g_1, g_2 \in G$.

**7.** Prove the following: let $G$ be a group and $\{g_s\colon s \in S\} \subseteq G$ be a subset generating it, e.g., the set of all elements of $G$. Then the map from (4.4.1) defined by

$$s \mapsto g_s$$

for all $s \in S$ gives rise to a well-defined group homomorphism from $\mathcal{F}(S)$ to $G$; that is, each element in an equivalence class forming $\mathcal{F}(S)$ has the same image.

**8.** **Riddle with nails and a picture.**(Maybe you want to start with (c))

(a) To get more familiar with free groups, consider the free group on the alphabet $\{a, b\}$. Write explicitly the words with up to 4 letters in the equivalence class of the empty word.

Observe that there are two natural group homomorphisms

$$\psi_b\colon \mathcal{F}(\{a,b\}) \to \mathcal{F}(\{a\}) \text{ and } \psi_a\colon \mathcal{F}(\{a,b\}) \to \mathcal{F}(\{b\})\ ,$$

such that $\psi_b$ maps $b$ to the empty word $\epsilon$ and such that $\psi_a$ maps $a$ to the empty word $\epsilon$. Convince yourself that they actually give rise to group homomorphisms.

(b) Extending the consideration from before, we turn to the free group on the alphabet $\{a, b, c\}$. Again, write explicitly the words with up to 4 letters in the equivalence class of the empty word.

Observe that there are three natural group homomorphisms

$$\begin{aligned}\psi_c\colon \mathcal{F}(\{a,b,c\}) &\to \mathcal{F}(\{a,b\})\,,\\ \psi_b\colon \mathcal{F}(\{a,b,c\}) &\to \mathcal{F}(\{a,c\})\,,\\ \psi_a\colon \mathcal{F}(\{a,b,c\}) &\to \mathcal{F}(\{b,c\})\,,\end{aligned}$$

such that $\psi_c$ maps $c$ to the empty word $\epsilon$, such that $\psi_b$ maps $b$ to the empty word $\epsilon$ and such that $\psi_a$ maps $a$ to the empty word $\epsilon$. Convince yourself that they actually give rise to group homomorphisms.

(c) Now, we are ready to tackle a 'real-world' problem. We want to hang up a picture on the wall but in a peculiar way. The picture is

supposed to hang on a string which is hold by 3 nails in such a way that: whatever nail we pull out of the wall, the picture will fall to the ground. Figure 4.5.1 displays such a configuration in the case we only use two nails. Assume that three nails are already in the wall and you have a sufficiently long string; find a way to wrap the string around the nails and attach it to the picture in such a way that the picture is held in place, but falls off when you pull out any of the nails.

(d) How is the question about the picture related to the group homomorphisms in the beginning of the exercise? How does this question generalize to more nails or letters?

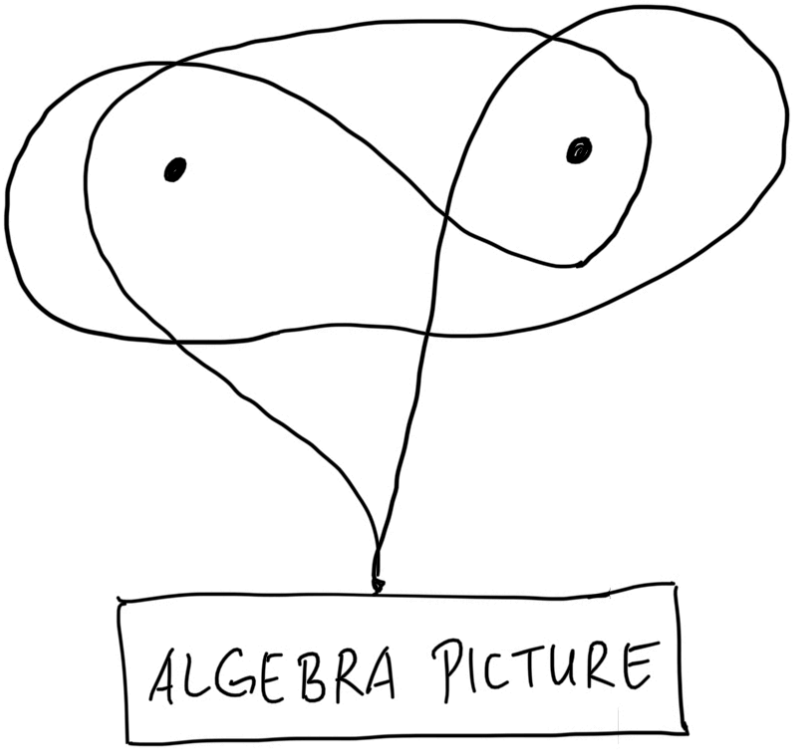



**Figure 4.5.1** A picture hung up on a wall with two nails

**9. Programming Problem.** Assume you are given two groups (e.g. via their Cayley table as dictionaries) and a map between them (e.g. again as a dictionary). Write a program to check that the map is a group homomorphism.

## 4.6 Additional Exercises

**1.** Let $\phi : G \to H$ be a group homomorphism. Show that $\phi$ is one-to-one if and only if $\phi^{-1}(e) = \{e\}$.

**2.** Show that $\left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{R} \right\}$ forms a subgroup of $SL_2(\mathbb{R})$. Furthermore,

[1] en.wikipedia.org/wiki/Associahedron

show that there is a homomorphism from this subgroup to the real numbers with addition. What is its kernel? Is it an isomorphism?

**3.** Let $A$ be an $m \times n$ matrix. Show that matrix multiplication, $x \mapsto Ax$, defines a homomorphism $\phi : \mathbb{R}^n \to \mathbb{R}^m$.

**4.** Let $\phi : \mathbb{Z} \to \mathbb{Z}$ be given by $\phi(n) = 7n$. Prove that $\phi$ is a group homomorphism. Find the kernel and the image of $\phi$.

**5.** Describe all of the homomorphisms from $\mathbb{Z}_{24}$ to $\mathbb{Z}_{18}$.

**6.** Describe all of the homomorphisms from $\mathbb{Z}$ to $\mathbb{Z}_{12}$.

**7.** Prove that $\mathbb{Z} \cong n\mathbb{Z}$ for $n \neq 0$.

**8.** Prove that $\mathbb{C}^*$ is isomorphic to the subgroup of $GL_2(\mathbb{R})$ consisting of matrices of the form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

**9.** Prove or disprove: $U(8) \cong \mathbb{Z}_4$.

**10.** Prove that $U(8)$ is isomorphic to the group of matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

**11.** Show that $U(5)$ is isomorphic to $U(10)$, but $U(12)$ is not.

**12.** Show that the $n$th roots of unity are isomorphic to $\mathbb{Z}_n$.

**13.** Show that any cyclic group of order $n$ is isomorphic to $\mathbb{Z}_n$.

**14.** Prove that $\mathbb{Q}$ is not isomorphic to $\mathbb{Z}$.

**15.** Let $G = \mathbb{R} \setminus \{-1\}$ and define a binary operation on $G$ by

$$a * b = a + b + ab.$$

Prove that $G$ is a group under this operation. Show that $(G, *)$ is isomorphic to the multiplicative group of nonzero real numbers.

## 4.7 Material

1. Group Homomorphisms (Quick)[1]
2. Group Homomorphisms (Expanded)[2]
3. Isomorphisms for Groups[3]
4. Kernel of Group Homomorphisms[4]
5. As motivation for group homomorphism: Linear transformations[5]

[1] www.socratica.com/lesson/group-homomorphisms-quick
[2] www.socratica.com/lesson/group-homomorphisms
[3] www.socratica.com/lesson/isomorphisms-for-groups
[4] www.socratica.com/lesson/kernel-of-group-homomorphisms
[5] www.3blue1brown.com/lessons/linear-transformations

# 4.8 Hints to Selected Exercises

## 4.5 · Core Exercises

**4.5.8. Riddle with nails and a picture.**

(a) There are 41 of these words.

## 4.6 · Additional Exercises

**4.6.4.** Since $\phi(m+n) = 7(m+n) = 7m+7n = \phi(m)+\phi(n)$, $\phi$ is a homomorphism.

**4.6.5.** For any homomorphism $\phi : \mathbb{Z}_{24} \to \mathbb{Z}_{18}$, the kernel of $\phi$ must be a subgroup of $\mathbb{Z}_{24}$ and the image of $\phi$ must be a subgroup of $\mathbb{Z}_{18}$. Now use the fact that a generator must map to a generator.

**4.6.7.** Every infinite cyclic group is isomorphic to $\mathbb{Z}$ by Theorem 4.2.7.

**4.6.8.** Define $\phi : \mathbb{C}^* \to GL_2(\mathbb{R})$ by

$$\phi(a+bi) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

**4.6.9.** False.

**4.6.12.** Define a map from $\mathbb{Z}_n$ into the $n$th roots of unity by $k \mapsto \text{cis}(2k\pi/n)$.

**4.6.14.** Assume that $\mathbb{Q}$ is cyclic and try to find a generator.

# Chapter 5

# Cosets and Group actions

Basic learning goals

1. Definition and properties of cosets.
2. Lagrange's Theorem and some of its implications.
3. Definition of group action and examples.
4. Terminology for group actions.
5. Orbit-Stabilizer Theorem.
6. Cayley's Theorem.

Lagrange's Theorem, one of the most important results in finite group theory, states that the order of a subgroup must divide the order of the group. This theorem provides a powerful tool for analyzing finite groups; it gives us an idea of exactly what type of subgroups we might expect a finite group to possess. Central to understanding Lagranges's Theorem is the notion of a coset.

Furthermore, we consider the elements of a group as bijective functions in the context of group actions. While this is a versatile tool in itself, it also shows that each group can actually be considered a permutation group.

## 5.1 Cosets

Let $G$ be a group and $H$ a subgroup of $G$. Define a **left coset** of $H$ with **representative** $g \in G$ to be the set

$$gH = \{gh : h \in H\}.$$

**Right cosets** can be defined similarly by

$$Hg = \{hg : h \in H\}.$$

If left and right cosets coincide or if it is clear from the context to which type of coset that we are referring, we will use the word *coset* without specifying left or right.

**Example 5.1.1** Let $H$ be the subgroup of $\mathbb{Z}_6$ consisting of the elements 0 and

3. The cosets are

$$0 + H = 3 + H = \{0, 3\}$$
$$1 + H = 4 + H = \{1, 4\}$$
$$2 + H = 5 + H = \{2, 5\}.$$

We will always write the cosets of subgroups of $\mathbb{Z}$ and $\mathbb{Z}_n$ with the additive notation we have used for cosets here. In a commutative group, left and right cosets are always identical. □

**Example 5.1.2** Let $H$ be the subgroup of $S_3$ defined by the permutations $\{(1), (1\,2\,3), (1\,3\,2)\}$. The left cosets of $H$ are

$$(1)H = (1\,2\,3)H = (1\,3\,2)H = \{(1), (1\,2\,3), (1\,3\,2)\}$$
$$(1\,2)H = (1\,3)H = (2\,3)H = \{(1\,2), (1\,3), (2\,3)\}.$$

The right cosets of $H$ are exactly the same as the left cosets:

$$H(1) = H(1\,2\,3) = H(1\,3\,2) = \{(1), (1\,2\,3), (1\,3\,2)\}$$
$$H(1\,2) = H(1\,3) = H(2\,3) = \{(1\,2), (1\,3), (2\,3)\}.$$

It is not always the case that a left coset is the same as a right coset. Let $K$ be the subgroup of $S_3$ defined by the permutations $\{(1), (1\,2)\}$. Then the left cosets of $K$ are

$$(1)K = (1\,2)K = \{(1), (1\,2)\}$$
$$(1\,3)K = (1\,2\,3)K = \{(1\,3), (1\,2\,3)\}$$
$$(2\,3)K = (1\,3\,2)K = \{(2\,3), (1\,3\,2)\};$$

however, the right cosets of $K$ are

$$K(1) = K(1\,2) = \{(1), (1\,2)\}$$
$$K(1\,3) = K(1\,3\,2) = \{(1\,3), (1\,3\,2)\}$$
$$K(2\,3) = K(1\,2\,3) = \{(2\,3), (1\,2\,3)\}.$$

□

The following lemma is quite useful when dealing with cosets.

**Lemma 5.1.3** *Let $H$ be a subgroup of a group $G$ and suppose that $g_1, g_2 \in G$. The following conditions are equivalent.*

1. $g_1 H = g_2 H$;
2. $H g_1^{-1} = H g_2^{-1}$;
3. $g_1 H \subset g_2 H$;
4. $g_2 \in g_1 H$;
5. $g_1^{-1} g_2 \in H$.

In all of our examples the cosets of a subgroup $H$ partition the larger group $G$. The following theorem proclaims that this will always be the case.

**Theorem 5.1.4** *Let $H$ be a subgroup of a group $G$. Then the left cosets of $H$ in $G$ partition $G$. That is, the group $G$ is the disjoint union of the left cosets of $H$ in $G$.*

*Proof.* Let $g_1H$ and $g_2H$ be two cosets of $H$ in $G$. We must show that either $g_1H \cap g_2H = \emptyset$ or $g_1H = g_2H$. Suppose that $g_1H \cap g_2H \neq \emptyset$ and $a \in g_1H \cap g_2H$. Then by the definition of a left coset, $a = g_1h_1 = g_2h_2$ for some elements $h_1$ and $h_2$ in $H$. Hence, $g_1 = g_2h_2h_1^{-1}$ or $g_1 \in g_2H$. By Lemma 5.1.3, $g_1H = g_2H$. ■

**Remark 5.1.5** There is nothing special in this theorem about left cosets. Right cosets also partition $G$; the proof of this fact is exactly the same as the proof for left cosets except that all group multiplications are done on the opposite side of $H$.

**Theorem 5.1.6** *Let $H$ be a subgroup of a group $G$. The number of left cosets of $H$ in $G$ is the same as the number of right cosets of $H$ in $G$.*

*Proof.* Let $\mathcal{L}_H$ and $\mathcal{R}_H$ denote the set of left and right cosets of $H$ in $G$, respectively. If we can define a bijective map $\phi : \mathcal{L}_H \to \mathcal{R}_H$, then the theorem will be proved. If $gH \in \mathcal{L}_H$, let $\phi(gH) = Hg^{-1}$. By Lemma 5.1.3, the map $\phi$ is well-defined; that is, if $g_1H = g_2H$, then $Hg_1^{-1} = Hg_2^{-1}$. To show that $\phi$ is one-to-one, suppose that

$$Hg_1^{-1} = \phi(g_1H) = \phi(g_2H) = Hg_2^{-1}.$$

Again by Lemma 5.1.3, $g_1H = g_2H$. The map $\phi$ is onto since $\phi(g^{-1}H) = Hg$. ■

Let $G$ be a group and $H$ be a subgroup of $G$. Define the **index** of $H$ in $G$ to be the number of left cosets of $H$ in $G$. We will denote the index by $[G : H]$.

**Example 5.1.7** Let $G = \mathbb{Z}_6$ and $H = \{0, 3\}$. Then $[G : H] = 3$. □

**Example 5.1.8** Suppose that $G = S_3$, $H = \{(1), (1\,2\,3), (1\,3\,2)\}$, and $K = \{(1), (1\,2)\}$. Then $[G : H] = 2$ and $[G : K] = 3$. □

## 5.2 Lagrange's Theorem

**Proposition 5.2.1** *For $g \in G$, define a function*

$$\lambda_g : G \to G \text{ with } \lambda_g(a) = ga \text{ for all } a \in G.$$

*Then $\lambda_g$ is a bijection, hence a permutation of $G$.*

*Proof.* To show that $\lambda_g$ is one-to-one, suppose that $\lambda_g(a) = \lambda_g(b)$. Then

$$ga = \lambda_g(a) = \lambda_g(b) = gb.$$

Multiplication from the left by $g^{-1}$ yields $a = b$. To show that $\lambda_g$ is onto, we must prove that for each $a \in G$, there is a $b$ such that $\lambda_g(b) = a$. This is given by $b = g^{-1}a$. ■

**Corollary 5.2.2** *Let $H$ be a subgroup of $G$ and $g \in G$. Then the number of elements in $H$ is the same as the number of elements in $gH$.*

*Proof.* The restriction of the bijection $\lambda_g$ is again one-to-one, so it is a bijection to its image $gH$. ■

Now we are readily equipped to show an important theorem for finite groups with several implications.

**Theorem 5.2.3 Lagrange.** *Let $G$ be a finite group and let $H$ be a subgroup of $G$. Then $|G|/|H| = [G : H]$ is the number of distinct left cosets of $H$ in $G$.*

*In particular, the number of elements in $H$ must divide the number of elements in $G$.*

*Proof.* We know from Theorem 5.1.4 that the group $G$ is partitioned into $[G : H]$ distinct left cosets. Each left coset has $|H|$ elements by Corollary 5.2.2. Therefore, $|G| = [G : H]|H|$. ■

**Corollary 5.2.4** *Suppose that $G$ is a finite group and $g \in G$. Then the order of $g$ must divide the number of elements in $G$.*

**Corollary 5.2.5** *Let $|G| = p$ with $p$ a prime number. Then $G$ is cyclic and any $g \in G$ such that $g \neq e$ is a generator. In particular, $G$ is isomorphic to $\mathbb{Z}_p$.*

*Proof.* Let $g$ be in $G$ such that $g \neq e$. Then by Corollary 5.2.4, the order of $g$ must divide the order of the group. Since $|\langle g \rangle| > 1$, it must be $p$. Hence, $g$ generates $G$. ■

**Corollary 5.2.6** *Let $H$ and $K$ be subgroups of a finite group $G$ such that $G \supset H \supset K$. Then*

$$[G : K] = [G : H][H : K].$$

*Proof.* Observe that

$$[G : K] = \frac{|G|}{|K|} = \frac{|G|}{|H|} \cdot \frac{|H|}{|K|} = [G : H][H : K].$$

■

### 5.2.1 Fermat's and Euler's Theorems

The **Euler $\phi$-function** is the map $\phi : \mathbb{N} \to \mathbb{N}$ defined by $\phi(n) = 1$ for $n = 1$, and, for $n > 1$, $\phi(n)$ is the number of positive integers $m$ with $1 \leq m < n$ and $\gcd(m, n) = 1$.

From Proposition 1.2.4, we know that the order of $U(n)$, the group of units in $\mathbb{Z}_n$, is $\phi(n)$. For example, $|U(12)| = \phi(12) = 4$ since the numbers that are relatively prime to 12 are 1, 5, 7, and 11. For any prime $p$, $\phi(p) = p - 1$. We state these results in the following theorem.

**Theorem 5.2.7** *Let $U(n)$ be the group of units in $\mathbb{Z}_n$. Then $|U(n)| = \phi(n)$.*

The following theorem is an important result in number theory, due to Leonhard Euler.

**Theorem 5.2.8 Euler's Theorem.** *Let $a$ and $n$ be integers such that $n > 0$ and $\gcd(a, n) = 1$. Then $a^{\phi(n)} \equiv 1 \pmod{n}$.*

*Proof.* By Theorem 5.2.7 the order of $U(n)$ is $\phi(n)$. Consequently, $a^{\phi(n)} = 1$ for all $a \in U(n)$; or $a^{\phi(n)} - 1$ is divisible by $n$. Therefore, $a^{\phi(n)} \equiv 1 \pmod{n}$. ■

If we consider the special case of Euler's Theorem in which $n = p$ is prime and recall that $\phi(p) = p - 1$, we obtain the following result, due to Pierre de Fermat.

**Theorem 5.2.9 Fermat's Little Theorem.** *Let $p$ be any prime number and suppose that $p \nmid a$ ($p$ does not divide $a$). Then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

*Furthermore, for any integer $b$, $b^p \equiv b \pmod{p}$.*

### 5.2.2 Historical Note

Joseph-Louis Lagrange (1736–1813), born in Turin, Italy, was of French and Italian descent. His talent for mathematics became apparent at an early age. Leonhard Euler recognized Lagrange's abilities when Lagrange, who was only 19, communicated to Euler some work that he had done in the calculus of variations. That year he was also named a professor at the Royal Artillery School in Turin. At the age of 23 he joined the Berlin Academy. Frederick the Great had written to Lagrange proclaiming that the "greatest king in Europe" should have the "greatest mathematician in Europe" at his court. For 20 years Lagrange held the position vacated by his mentor, Euler. His works include contributions to number theory, group theory, physics and mechanics, the calculus of variations, the theory of equations, and differential equations. Along with Laplace and Lavoisier, Lagrange was one of the people responsible for designing the metric system. During his life Lagrange profoundly influenced the development of mathematics, leaving much to the next generation of mathematicians in the form of examples and new problems to be solved.

## 5.3 Group Actions

Group actions generalize group multiplication. If $G$ is a group and $X$ is an arbitrary set, a group action of an element $g \in G$ and $x \in X$ is a product, $gx$, living in $X$. This is how groups often appear in reality as transformations or reorderings or relabelings. Furthermore, many problems in algebra are best be attacked via group actions; the proofs of the Sylow theorems and of Burnside's Counting Theorem (which are both beyond the scope of this course) are most easily understood when they are formulated in terms of group actions.

Let $X$ be a set and $G$ be a group. A **(left) action** of $G$ on $X$ is a map $G \times X \to X$ given by $(g, x) \mapsto gx$, where

1. $ex = x$ for all $x \in X$;
2. $(g_1 g_2)x = g_1(g_2 x)$ for all $x \in X$ and all $g_1, g_2 \in G$.

Under these considerations $X$ is called a **$G$-set**. Notice that we are not requiring $X$ to be related to $G$ in any way. It is true that every group $G$ acts on every set $X$ by the trivial action $(g, x) \mapsto x$; however, group actions are more interesting if the set $X$ is somehow related to the group $G$.

Recall Exercise 4.5.6. In the new terminology, this actually means that an action of a group $G$ on a set $X$ is a group homomorphism from $G$ to $S_X$.

**Example 5.3.1** Let $G = GL_2(\mathbb{R})$ and $X = \mathbb{R}^2$. Then $G$ acts on $X$ by left multiplication. If $v \in \mathbb{R}^2$ and $I$ is the identity matrix, then $Iv = v$. If $A$ and $B$ are $2 \times 2$ invertible matrices, then $(AB)v = A(Bv)$ since matrix multiplication is associative. □

**Example 5.3.2** Let $G = D_4$ be the symmetry group of a square. If $X = \{1, 2, 3, 4\}$ is the set of vertices of the square, then we can consider $D_4$ to consist of the following permutations:

$$\{(1), (1\,3), (2\,4), (1\,4\,3\,2), (1\,2\,3\,4), (1\,2)(3\,4), (1\,4)(2\,3), (1\,3)(2\,4)\}.$$

The elements of $D_4$ act on $X$ as functions. The permutation $(1\,3)(2\,4)$ acts on vertex 1 by sending it to vertex 3, on vertex 2 by sending it to vertex 4, and so on. It is easy to see that the axioms of a group action are satisfied. □

In general, if $X$ is any set and $G$ is a subgroup of $S_X$, the group of all permutations acting on $X$, then $X$ is a $G$-set under the group action

$$(\sigma, x) \mapsto \sigma(x)$$

for $\sigma \in G$ and $x \in X$.

**Example 5.3.3** If we let $X = G$, then every group $G$ acts on itself by the left multiplication; that is, $(g, x) \mapsto \lambda_g(x) = gx$ (see also Proposition 5.2.1)

$$e \cdot x = \lambda_e x = ex = x$$
$$(gh) \cdot x = \lambda_{gh} x = \lambda_g \lambda_h x = \lambda_g(hx) = g \cdot (h \cdot x).$$

If $H$ is a subgroup of $G$, then $G$ is an $H$-set under left multiplication by elements of $H$. □

**Example 5.3.4** Let $G$ be a group and suppose that $X = G$. If $H$ is a subgroup of $G$, then $G$ is an $H$-set under **conjugation**; that is, we can define an action of $H$ on $G$,

$$H \times G \to G,$$

via

$$(h, g) \mapsto hgh^{-1}$$

for $h \in H$ and $g \in G$. Clearly, the first axiom for a group action holds. Observing that

$$\begin{aligned}(h_1h_2, g) &= h_1h_2g(h_1h_2)^{-1} \\ &= h_1(h_2gh_2^{-1})h_1^{-1} \\ &= (h_1, (h_2, g)),\end{aligned}$$

we see that the second condition is also satisfied. □

**Example 5.3.5** Let $H$ be a subgroup of $G$ and $\mathcal{L}_H$ the set of left cosets of $H$. The set $\mathcal{L}_H$ is a $G$-set under the action

$$(g, xH) \mapsto gxH.$$

Again, it is easy to see that the first axiom is true. Since $(gg')xH = g(g'xH)$, the second axiom is also true. □

If $G$ acts on a set $X$ and $x, y \in X$, then $x$ is said to be **$G$-equivalent** to $y$ if there exists a $g \in G$ such that $gx = y$. We write $x \sim_G y$ or $x \sim y$ if two elements are $G$-equivalent.

**Proposition 5.3.6** *Let $X$ be a $G$-set. Then $G$-equivalence is an equivalence relation on $X$.*

Hence, if $X$ is a $G$-set, then $G$-equivalence gives rise to a partition of $X$. Each of the blocks is called an **orbit** of $X$ under $G$. We will denote the orbit that contains an element $x$ of $X$ by $\mathcal{O}_x$.

**Example 5.3.7** Let $G$ be the permutation group defined by

$$G = \{(1), (1\,2\,3), (1\,3\,2), (4\,5), (1\,2\,3)(4\,5), (1\,3\,2)(4\,5)\}$$

and $X = \{1, 2, 3, 4, 5\}$. Then $X$ is a $G$-set. The orbits are $\mathcal{O}_1 = \mathcal{O}_2 = \mathcal{O}_3 = \{1, 2, 3\}$ and $\mathcal{O}_4 = \mathcal{O}_5 = \{4, 5\}$. □

Now suppose that $G$ is a group acting on a set $X$ and let $g$ be an element of $G$. The **fixed point set** of $g$ in $X$, denoted by $X_g$, is the set of all $x \in X$ such that $gx = x$. We can also study the group elements $g$ that fix a given

$x \in X$. This set is more than a subset of $G$, it is a subgroup. This subgroup is called the **stabilizer subgroup** or **isotropy subgroup** of $x$. We will denote the stabilizer subgroup of $x$ by $G_x$.

**Remark 5.3.8** It is important to remember that $X_g \subset X$ and $G_x \subset G$.

**Example 5.3.9** Let $X = \{1, 2, 3, 4, 5, 6\}$ and suppose that $G$ is the permutation group given by the permutations

$$\{(1), (1\,2)(3\,4\,5\,6), (3\,5)(4\,6), (1\,2)(3\,6\,5\,4)\}.$$

Then the fixed point sets of $X$ under the action of $G$ are

$$\begin{aligned} X_{(1)} &= X, \\ X_{(3\,5)(4\,6)} &= \{1, 2\}, \\ X_{(1\,2)(3\,4\,5\,6)} &= X_{(1\,2)(3\,6\,5\,4)} = \emptyset, \end{aligned}$$

and the stabilizer subgroups are

$$\begin{aligned} G_1 = G_2 &= \{(1), (3\,5)(4\,6)\}, \\ G_3 = G_4 = G_5 &= G_6 = \{(1)\}. \end{aligned}$$

It is easily seen that $G_x$ is a subgroup of $G$ for each $x \in X$. □

**Proposition 5.3.10** *Let $G$ be a group acting on a set $X$ and $x \in X$. The stabilizer group of $x$, $G_x$, is a subgroup of $G$.*

*Proof.* Clearly, $e \in G_x$ since the identity fixes every element in the set $X$. Let $g, h \in G_x$. Then $gx = x$ and $hx = x$. So $(gh)x = g(hx) = gx = x$; hence, the product of two elements in $G_x$ is also in $G_x$. Finally, if $g \in G_x$, then $x = ex = (g^{-1}g)x = (g^{-1})gx = g^{-1}x$. So $g^{-1}$ is in $G_x$. ■

We will denote the number of elements in the fixed point set of an element $g \in G$ by $|X_g|$ and denote the number of elements in the orbit of $x \in X$ by $|\mathcal{O}_x|$. The next theorem demonstrates the relationship between orbits of an element $x \in X$ and the left cosets of $G_x$ in $G$.

**Theorem 5.3.11 (Orbit-Stabilizer Theorem).** *Let $G$ be a finite group and $X$ a finite $G$-set. If $x \in X$, then $|\mathcal{O}_x| = [G : G_x]$.*

*Proof.* We know that $|G|/|G_x|$ is the number of left cosets of $G_x$ in $G$ by Lagrange's Theorem (Theorem 5.2.3). We will define a bijective map $\phi$ between the orbit $\mathcal{O}_x$ of $X$ and the set of left cosets $\mathcal{L}_{G_x}$ of $G_x$ in $G$. Let $y \in \mathcal{O}_x$. Then there exists a $g$ in $G$ such that $gx = y$. Define $\phi$ by $\phi(y) = gG_x$. To show that $\phi$ is one-to-one, assume that $\phi(y_1) = \phi(y_2)$. Then

$$\phi(y_1) = g_1G_x = g_2G_x = \phi(y_2),$$

where $g_1x = y_1$ and $g_2x = y_2$. Since $g_1G_x = g_2G_x$, there exists a $g \in G_x$ such that $g_2 = g_1g$,

$$y_2 = g_2x = g_1gx = g_1x = y_1;$$

consequently, the map $\phi$ is one-to-one. Finally, we must show that the map $\phi$ is onto. Let $gG_x$ be a left coset. If $gx = y$, then $\phi(y) = gG_x$. ■

### 5.3.1 Cayley's Theorem

Cayley proved that if $G$ is a group, it is isomorphic to a group of permutations on some set; hence, every group is a permutation group. Cayley's Theorem is

what we call a representation theorem. The aim of representation theory is to find an isomorphism of some group $G$ that we wish to study into a group that we know a great deal about, such as a group of permutations or matrices.

**Example 5.3.12** Consider the group $\mathbb{Z}_3$. The Cayley table for $\mathbb{Z}_3$ is as follows.

| $+$ | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0 | 1 | 2 |
| 1 | 1 | 2 | 0 |
| 2 | 2 | 0 | 1 |

The addition table of $\mathbb{Z}_3$ suggests that it is the same as the permutation group $G = \{(0), (012), (021)\}$. The isomorphism here is

$$0 \mapsto \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} = (0)$$
$$1 \mapsto \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix} = (0\,1\,2)$$
$$2 \mapsto \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix} = (0\,2\,1).$$

□

**Theorem 5.3.13 Cayley.** *Every group is isomorphic to a group of permutations.*

*Proof.* Let $G$ be a group. We must find a group of permutations $\overline{G}$ that is isomorphic to $G$.

We saw already in Proposition 5.2.1 how to associate a permutation with each element of $G$. Let

$$\overline{G} = \{\lambda_g : g \in G\}.$$

We must show that $\overline{G}$ is a group under composition of functions and find an isomorphism between $G$ and $\overline{G}$. We have closure under composition of functions since

$$(\lambda_g \circ \lambda_h)(a) = \lambda_g(ha) = gha = \lambda_{gh}(a).$$

Also,

$$\lambda_e(a) = ea = a$$

and

$$(\lambda_{g^{-1}} \circ \lambda_g)(a) = \lambda_{g^{-1}}(ga) = g^{-1}ga = a = \lambda_e(a).$$

We can define an isomorphism from $G$ to $\overline{G}$ by $\phi : g \mapsto \lambda_g$. The group operation is preserved since

$$\phi(gh) = \lambda_{gh} = \lambda_g\lambda_h = \phi(g)\phi(h).$$

It is also one-to-one, because if $\phi(g)(a) = \phi(h)(a)$, then

$$ga = \lambda_g a = \lambda_h a = ha.$$

Hence, $g = h$. That $\phi$ is onto follows from the fact that $\phi(g) = \lambda_g$ for any $\lambda_g \in \overline{G}$. ■

The isomorphism $g \mapsto \lambda_g$ is known as the **left regular representation** of $G$.

### 5.3.2 Historical Note

Arthur Cayley was born in England in 1821, though he spent much of the first part of his life in Russia, where his father was a merchant. Cayley was educated at Cambridge, where he took the first Smith's Prize in mathematics. A lawyer for much of his adult life, he wrote several papers in his early twenties before entering the legal profession at the age of 25. While practicing law he continued his mathematical research, writing more than 300 papers during this period of his life. These included some of his best work. In 1863 he left law to become a professor at Cambridge. Cayley wrote more than 900 papers in fields such as group theory, geometry, and linear algebra. His legal knowledge was very valuable to Cambridge; he participated in the writing of many of the university's statutes. Cayley was also one of the people responsible for the admission of women to Cambridge.

## 5.4 Additional insights

### 5.4.1 The converse of Lagrange's Theorem is false

The group $A_4$ has order 12; however, it can be shown that it does not possess a subgroup of order 6. According to Lagrange's Theorem, subgroups of a group of order 12 can have orders of either 1, 2, 3, 4, or 6. However, we are not guaranteed that subgroups of every possible order exist. To prove that $A_4$ has no subgroup of order 6, we will assume that it does have such a subgroup $H$ and show that a contradiction must occur. Since $A_4$ contains eight 3-cycles, we know that $H$ must contain a 3-cycle. We will show that if $H$ contains one 3-cycle, then it must contain more than 6 elements.

**Proposition 5.4.1** *The group $A_4$ has no subgroup of order 6.*

*Proof.* Since $[A_4 : H] = 2$, there are only two cosets of $H$ in $A_4$. Inasmuch as one of the cosets is $H$ itself, right and left cosets must coincide; therefore, $gH = Hg$ or $gHg^{-1} = H$ for every $g \in A_4$. Since there are eight 3-cycles in $A_4$, at least one 3-cycle must be in $H$. Without loss of generality, assume that $(1\,2\,3)$ is in $H$. Then $(1\,2\,3)^{-1} = (1\,3\,2)$ must also be in $H$. Since $ghg^{-1} \in H$ for all $g \in A_4$ and all $h \in H$ and

$$(1\,2\,4)(1\,2\,3)(1\,2\,4)^{-1} = (1\,2\,4)(1\,2\,3)(1\,4\,2) = (2\,4\,3)$$
$$(2\,4\,3)(1\,2\,3)(2\,4\,3)^{-1} = (2\,4\,3)(1\,2\,3)(2\,3\,4) = (1\,4\,2)$$

we can conclude that $H$ must have at least seven elements

$$(1), (1\,2\,3), (1\,3\,2), (2\,4\,3), (2\,4\,3)^{-1} = (2\,3\,4), (1\,4\,2), (1\,4\,2)^{-1} = (1\,2\,4).$$

Therefore, $A_4$ has no subgroup of order 6. ■

In fact, we can say more about when two cycles have the same length.

**Theorem 5.4.2** *Two cycles $\tau$ and $\mu$ in $S_n$ have the same length if and only if there exists a $\sigma \in S_n$ such that $\mu = \sigma\tau\sigma^{-1}$.*

*Proof.* Suppose that

$$\tau = (a_1, a_2, \ldots, a_k)$$
$$\mu = (b_1, b_2, \ldots, b_k).$$

Define $\sigma$ to be the permutation

$$\begin{aligned}\sigma(a_1) &= b_1\\ \sigma(a_2) &= b_2\\ &\vdots\\ \sigma(a_k) &= b_k.\end{aligned}$$

Then $\mu = \sigma\tau\sigma^{-1}$.

Conversely, suppose that $\tau = (a_1, a_2, \ldots, a_k)$ is a $k$-cycle and $\sigma \in S_n$. If $\sigma(a_i) = b$ and $\sigma(a_{(i \bmod k)+1}) = b'$, then $\mu(b) = b'$. Hence,

$$\mu = (\sigma(a_1), \sigma(a_2), \ldots, \sigma(a_k)).$$

Since $\sigma$ is one-to-one and onto, $\mu$ is a cycle of the same length as $\tau$. ■

### 5.4.2 The Class Equation

Let $X$ be a finite $G$-set and $X_G$ be the set of fixed points in $X$; that is,

$$X_G = \{x \in X : gx = x \text{ for all } g \in G\}.$$

Since the orbits of the action partition $X$,

$$|X| = |X_G| + \sum_{i=k}^{n} |\mathcal{O}_{x_i}|,$$

where $x_k, \ldots, x_n$ are representatives from the distinct nontrivial orbits of $X$.

Now consider the special case in which $G$ acts on itself by conjugation, $(g, x) \mapsto gxg^{-1}$. The **center** of $G$,

$$Z(G) = \{x : xg = gx \text{ for all } g \in G\},$$

is the set of points that are fixed by conjugation. The nontrivial orbits of the action are called the **conjugacy classes** of $G$. If $x_1, \ldots, x_k$ are representatives from each of the nontrivial conjugacy classes of $G$ and $|\mathcal{O}_{x_1}| = n_1, \ldots, |\mathcal{O}_{x_k}| = n_k$, then

$$|G| = |Z(G)| + n_1 + \cdots + n_k.$$

The stabilizer subgroups of each of the $x_i$'s, $C(x_i) = \{g \in G : gx_i = x_i g\}$, are called the **centralizer subgroups** of the $x_i$'s. From Theorem 5.3.11, we obtain the **class equation**:

$$|G| = |Z(G)| + [G : C(x_1)] + \cdots + [G : C(x_k)].$$

One of the consequences of the class equation is that the order of each conjugacy class must divide the order of $G$.

**Example 5.4.3** It is easy to check that the conjugacy classes in $S_3$ are the following:

$$\{(1)\}, \quad \{(1\,2\,3), (1\,3\,2)\}, \quad \{(1\,2), (1\,3), (2\,3)\}.$$

The class equation is $6 = 1 + 2 + 3$. □

**Example 5.4.4** The center of $D_4$ is $\{(1), (1\,3)(2\,4)\}$, and the conjugacy classes are

$$\{(1\,3), (2\,4)\}, \quad \{(1\,4\,3\,2), (1\,2\,3\,4)\}, \quad \{(1\,2)(3\,4), (1\,4)(2\,3)\}.$$

Thus, the class equation for $D_4$ is $8 = 2 + 2 + 2 + 2$. □

**Example 5.4.5** For $S_n$ it takes a bit of work to find the conjugacy classes. We begin with cycles. Suppose that $\sigma = (a_1, \ldots, a_k)$ is a cycle and let $\tau \in S_n$. By Theorem 5.4.2,

$$\tau\sigma\tau^{-1} = (\tau(a_1), \ldots, \tau(a_k)).$$

Consequently, any two cycles of the same length are conjugate. Now let $\sigma = \sigma_1\sigma_2\cdots\sigma_r$ be a cycle decomposition, where the length of each cycle $\sigma_i$ is $r_i$. Then $\sigma$ is conjugate to every other $\tau \in S_n$ whose cycle decomposition has the same lengths.

The number of conjugate classes in $S_n$ is the number of ways in which $n$ can be partitioned into sums of positive integers. In the case of $S_3$ for example, we can partition the integer 3 into the following three sums:

$$\begin{aligned} 3 &= 1 + 1 + 1 \\ 3 &= 1 + 2 \\ 3 &= 3; \end{aligned}$$

therefore, there are three conjugacy classes. There are variations to problem of finding the number of such partitions for any positive integer $n$ that are what computer scientists call **NP-complete**. This effectively means that the problem cannot be solved for a large $n$ because the computations would be too time-consuming for even the largest computer. □

**Theorem 5.4.6** *Let $G$ be a group of order $p^n$ where $p$ is prime. Then $G$ has a nontrivial center.*

*Proof.* We apply the class equation

$$|G| = |Z(G)| + n_1 + \cdots + n_k.$$

Since each $n_i > 1$ and $n_i \mid |G|$, it follows that $p$ must divide each $n_i$. Also, $p \mid |G|$; hence, $p$ must divide $|Z(G)|$. Since the identity is always in the center of $G$, $|Z(G)| \geq 1$. Therefore, $|Z(G)| \geq p$, and there exists some $g \in Z(G)$ such that $g \neq 1$. ■

**Corollary 5.4.7** *Let $G$ be a group of order $p^2$ where $p$ is prime. Then $G$ is abelian.*

*Proof.* By Theorem 5.4.6, $|Z(G)| = p$ or $p^2$. Suppose that $|Z(G)| = p$. Then $Z(G)$ and $G/Z(G)$ both have order $p$ and must both be cyclic groups. Choosing a generator $aZ(G)$ for $G/Z(G)$, we can write any element $gZ(G)$ in the quotient group as $a^mZ(G)$ for some integer $m$; hence, $g = a^mx$ for some $x$ in the center of $G$. Similarly, if $hZ(G) \in G/Z(G)$, there exists a $y$ in $Z(G)$ such that $h = a^ny$ for some integer $n$. Since $x$ and $y$ are in the center of $G$, they commute with all other elements of $G$; therefore,

$$gh = a^mxa^ny = a^{m+n}xy = a^nya^mx = hg,$$

and $G$ must be abelian. Hence, $|Z(G)| = p^2$. ■

### 5.4.3 The Sylow Theorems

We will use what we have learned about group actions to prove the Sylow Theorems. Recall for a moment what it means for $G$ to act on itself by conjugation and how conjugacy classes are distributed in the group according to the class equation, discussed in Subsection 5.4.2. A group $G$ acts on itself by conjugation via the map $(g, x) \mapsto gxg^{-1}$. Let $x_1, \ldots, x_k$ be representatives from each

of the distinct conjugacy classes of $G$ that consist of more than one element. Then the class equation can be written as

$$|G| = |Z(G)| + [G : C(x_1)] + \cdots + [G : C(x_k)],$$

where $Z(G) = \{g \in G : gx = xg \text{ for all } x \in G\}$ is the center of $G$ and $C(x_i) = \{g \in G : gx_i = x_i g\}$ is the centralizer subgroup of $x_i$.

We begin our investigation of the Sylow Theorems by examining subgroups of order $p$, where $p$ is prime. A group $G$ is a **$p$-group** if every element in $G$ has as its order a power of $p$, where $p$ is a prime number. A subgroup of a group $G$ is a **$p$-subgroup** if it is a $p$-group.

**Theorem 5.4.8 Cauchy.** *Let $G$ be a finite group and $p$ a prime such that $p$ divides the order of $G$. Then $G$ contains a subgroup of order $p$.*

*Proof.* We will use induction on the order of $G$. If $|G| = p$, then clearly $G$ itself is the required subgroup. We now assume that every group of order $k$, where $p \leq k < n$ and $p$ divides $k$, has an element of order $p$. Assume that $|G| = n$ and $p \mid n$ and consider the class equation of $G$:

$$|G| = |Z(G)| + [G : C(x_1)] + \cdots + [G : C(x_k)].$$

We have two cases.

*Case 1.* Suppose the order of one of the centralizer subgroups, $C(x_i)$, is divisible by $p$ for some $i$, $i = 1, \ldots, k$. In this case, by our induction hypothesis, we are done. Since $C(x_i)$ is a proper subgroup of $G$ and $p$ divides $|C(x_i)|$, $C(x_i)$ must contain an element of order $p$. Hence, $G$ must contain an element of order $p$.

*Case 2.* Suppose the order of no centralizer subgroup is divisible by $p$. Then $p$ divides $[G : C(x_i)]$, the order of each conjugacy class in the class equation; hence, $p$ must divide the center of $G$, $Z(G)$. Since $Z(G)$ is abelian, it must have a subgroup of order $p$ by the Fundamental Theorem of Finite Abelian Groups. Therefore, the center of $G$ contains an element of order $p$. ■

**Corollary 5.4.9** *Let $G$ be a finite group. Then $G$ is a $p$-group if and only if $|G| = p^n$.*

**Example 5.4.10** Let us consider the group $A_5$. We know that $|A_5| = 60 = 2^2 \cdot 3 \cdot 5$. By Cauchy's Theorem, we are guaranteed that $A_5$ has subgroups of orders 2, 3 and 5. The Sylow Theorems will give us even more information about the possible subgroups of $A_5$. □

We are now ready to state and prove the first of the Sylow Theorems. The proof is very similar to the proof of Cauchy's Theorem.

**Theorem 5.4.11 First Sylow Theorem.** *Let $G$ be a finite group and $p$ a prime such that $p^r$ divides $|G|$. Then $G$ contains a subgroup of order $p^r$.*

*Proof.* We induct on the order of $G$ once again. If $|G| = p$, then we are done. Now suppose that the order of $G$ is $n$ with $n > p$ and that the theorem is true for all groups of order less than $n$, where $p$ divides $n$. We shall apply the class equation once again:

$$|G| = |Z(G)| + [G : C(x_1)] + \cdots + [G : C(x_k)].$$

First suppose that $p$ does not divide $[G : C(x_i)]$ for some $i$. Then $p^r \mid |C(x_i)|$, since $p^r$ divides $|G| = |C(x_i)| \cdot [G : C(x_i)]$. Now we can apply the induction hypothesis to $C(x_i)$.

Hence, we may assume that $p$ divides $[G : C(x_i)]$ for all $i$. Since $p$ divides

$|G|$, the class equation says that $p$ must divide $|Z(G)|$; hence, by Cauchy's Theorem, $Z(G)$ has an element of order $p$, say $g$. Let $N$ be the group generated by $g$. Clearly, $N$ is a normal subgroup of $Z(G)$ since $Z(G)$ is abelian; therefore, $N$ is normal in $G$ since every element in $Z(G)$ commutes with every element in $G$. Now consider the factor group $G/N$ of order $|G|/p$. By the induction hypothesis, $G/N$ contains a subgroup $H$ of order $p^{r-1}$. The inverse image of $H$ under the canonical homomorphism $\phi : G \to G/N$ is a subgroup of order $p^r$ in $G$. ■

A **Sylow $p$-subgroup** $P$ of a group $G$ is a maximal $p$-subgroup of $G$. To prove the other two Sylow Theorems, we need to consider conjugate subgroups as opposed to conjugate elements in a group. For a group $G$, let $\mathcal{S}$ be the collection of all subgroups of $G$. For any subgroup $H$, $\mathcal{S}$ is a $H$-set, where $H$ acts on $\mathcal{S}$ by conjugation. That is, we have an action

$$H \times \mathcal{S} \to \mathcal{S}$$

defined by

$$h \cdot K \mapsto hKh^{-1}$$

for $K$ in $\mathcal{S}$.

The set

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

is a subgroup of $G$ called the the **normalizer** of $H$ in $G$. Notice that $H$ is a normal subgroup of $N(H)$. In fact, $N(H)$ is the largest subgroup of $G$ in which $H$ is normal.

**Lemma 5.4.12** *Let $P$ be a Sylow $p$-subgroup of a finite group $G$ and let $x$ have as its order a power of $p$. If $x^{-1}Px = P$, then $x \in P$.*

*Proof.* Certainly $x \in N(P)$, and the cyclic subgroup, $\langle xP \rangle \subset N(P)/P$, has as its order a power of $p$. By the Correspondence Theorem there exists a subgroup $H$ of $N(P)$ containing $P$ such that $H/P = \langle xP \rangle$. Since $|H| = |P| \cdot |\langle xP \rangle|$, the order of $H$ must be a power of $p$. However, $P$ is a Sylow $p$-subgroup contained in $H$. Since the order of $P$ is the largest power of $p$ dividing $|G|$, $H = P$. Therefore, $H/P$ is the trivial subgroup and $xP = P$, or $x \in P$. ■

**Lemma 5.4.13** *Let $H$ and $K$ be subgroups of $G$. The number of distinct $H$-conjugates of $K$ is $[H : N(K) \cap H]$.*

*Proof.* We define a bijection between the conjugacy classes of $K$ and the right cosets of $N(K) \cap H$ by $h^{-1}Kh \mapsto (N(K) \cap H)h$. To show that this map is a bijection, let $h_1, h_2 \in H$ and suppose that $(N(K) \cap H)h_1 = (N(K) \cap H)h_2$. Then $h_2h_1^{-1} \in N(K)$. Therefore, $K = h_2h_1^{-1}Kh_1h_2^{-1}$ or $h_1^{-1}Kh_1 = h_2^{-1}Kh_2$, and the map is an injection. It is easy to see that this map is surjective; hence, we have a one-to-one and onto map between the $H$-conjugates of $K$ and the right cosets of $N(K) \cap H$ in $H$. ■

**Theorem 5.4.14 Second Sylow Theorem.** *Let $G$ be a finite group and $p$ a prime dividing $|G|$. Then all Sylow $p$-subgroups of $G$ are conjugate. That is, if $P_1$ and $P_2$ are two Sylow $p$-subgroups, there exists a $g \in G$ such that $gP_1g^{-1} = P_2$.*

*Proof.* Let $P$ be a Sylow $p$-subgroup of $G$ and suppose that $|G| = p^r m$ with $|P| = p^r$. Let

$$\mathcal{S} = \{P = P_1, P_2, \ldots, P_k\}$$

consist of the distinct conjugates of $P$ in $G$. By Lemma 5.4.13, $k = [G : N(P)]$. Notice that

$$|G| = p^r m = |N(P)| \cdot [G : N(P)] = |N(P)| \cdot k.$$

Since $p^r$ divides $|N(P)|$, $p$ cannot divide $k$.

Given any other Sylow $p$-subgroup $Q$, we must show that $Q \in \mathcal{S}$. Consider the $Q$-conjugacy classes of each $P_i$. Clearly, these conjugacy classes partition $\mathcal{S}$. The size of the partition containing $P_i$ is $[Q : N(P_i) \cap Q]$ by Lemma 5.4.13, and Lagrange's Theorem tells us that $|Q| = [Q : N(P_i) \cap Q]|N(P_i) \cap Q|$. Thus, $[Q : N(P_i) \cap Q]$ must be a divisor of $|Q| = p^r$. Hence, the number of conjugates in every equivalence class of the partition is a power of $p$. However, since $p$ does not divide $k$, one of these equivalence classes must contain only a single Sylow $p$-subgroup, say $P_j$. In this case, $x^{-1}P_j x = P_j$ for all $x \in Q$. By Lemma 5.4.12, $P_j = Q$. ■

**Theorem 5.4.15 Third Sylow Theorem.** *Let $G$ be a finite group and let $p$ be a prime dividing the order of $G$. Then the number of Sylow $p$-subgroups is congruent to* 1 (mod $p$) *and divides* $|G|$.

*Proof.* Let $P$ be a Sylow $p$-subgroup acting on the set of Sylow $p$-subgroups,

$$\mathcal{S} = \{P = P_1, P_2, \ldots, P_k\},$$

by conjugation. From the proof of the Second Sylow Theorem, the only $P$-conjugate of $P$ is itself and the order of the other $P$-conjugacy classes is a power of $p$. Each $P$-conjugacy class contributes a positive power of $p$ toward $|\mathcal{S}|$ except the equivalence class $\{P\}$. Since $|\mathcal{S}|$ is the sum of positive powers of $p$ and 1, $|\mathcal{S}| \equiv 1 \pmod{p}$.

Now suppose that $G$ acts on $\mathcal{S}$ by conjugation. Since all Sylow $p$-subgroups are conjugate, there can be only one orbit under this action. For $P \in \mathcal{S}$,

$$|\mathcal{S}| = |\text{orbit of } P| = [G : N(P)]$$

by Lemma 5.4.13. But $[G : N(P)]$ is a divisor of $|G|$; consequently, the number of Sylow $p$-subgroups of a finite group must divide the order of the group. ■

### 5.4.4 Historical Note

Peter Ludvig Mejdell Sylow was born in 1832 in Christiania, Norway (now Oslo). After attending Christiania University, Sylow taught high school. In 1862 he obtained a temporary appointment at Christiania University. Even though his appointment was relatively brief, he influenced students such as Sophus Lie (1842–1899). Sylow had a chance at a permanent chair in 1869, but failed to obtain the appointment. In 1872, he published a 10-page paper presenting the theorems that now bear his name. Later Lie and Sylow collaborated on a new edition of Abel's works. In 1898, a chair at Christiania University was finally created for Sylow through the efforts of his student and colleague Lie. Sylow died in 1918.

### 5.4.5 Burnside's Counting Theorem

Suppose that we wish to color the vertices of a square with two different colors, say black and white. We might suspect that there would be $2^4 = 16$ different colorings. However, some of these colorings are equivalent. If we color the first vertex black and the remaining vertices white, it is the same as coloring the second vertex black and the remaining ones white since we could obtain the second coloring simply by rotating the square 90° (Figure 5.4.16).

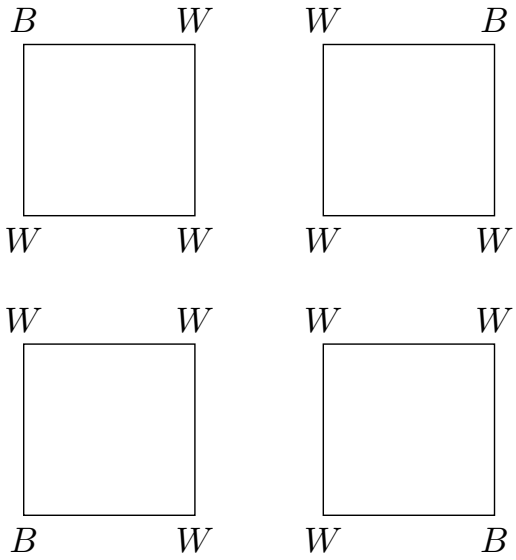



**Figure 5.4.16** Equivalent colorings of square

Burnside's Counting Theorem offers a method of computing the number of distinguishable ways in which something can be done. In addition to its geometric applications, the theorem has interesting applications to areas in switching theory and chemistry. The proof of Burnside's Counting Theorem depends on the following lemma.

**Lemma 5.4.17** *Let $X$ be a $G$-set and suppose that $x \sim y$. Then $G_x$ is isomorphic to $G_y$. In particular, $|G_x| = |G_y|$.*

*Proof.* Let $G$ act on $X$ by $(g, x) \mapsto g \cdot x$. Since $x \sim y$, there exists a $g \in G$ such that $g \cdot x = y$. Let $a \in G_x$. Since

$$gag^{-1} \cdot y = ga \cdot g^{-1}y = ga \cdot x = g \cdot x = y,$$

we can define a map $\phi : G_x \to G_y$ by $\phi(a) = gag^{-1}$. The map $\phi$ is a homomorphism since

$$\phi(ab) = gabg^{-1} = gag^{-1}gbg^{-1} = \phi(a)\phi(b).$$

Suppose that $\phi(a) = \phi(b)$. Then $gag^{-1} = gbg^{-1}$ or $a = b$; hence, the map is injective. To show that $\phi$ is onto, let $b$ be in $G_y$; then $g^{-1}bg$ is in $G_x$ since

$$g^{-1}bg \cdot x = g^{-1}b \cdot gx = g^{-1}b \cdot y = g^{-1} \cdot y = x;$$

and $\phi(g^{-1}bg) = b$. ■

**Theorem 5.4.18 Burnside.** *Let $G$ be a finite group acting on a set $X$ and let $k$ denote the number of orbits of $X$. Then*

$$k = \frac{1}{|G|} \sum_{g \in G} |X_g|.$$

*Proof.* We look at all the fixed points $x$ of all the elements in $g \in G$; that is, we look at all $g$'s and all $x$'s such that $gx = x$. If viewed in terms of fixed point sets, the number of all $g$'s fixing $x$'s is

$$\sum_{g \in G} |X_g|.$$

However, if viewed in terms of the stabilizer subgroups, this number is

$$\sum_{x \in X} |G_x|;$$

hence, $\sum_{g \in G} |X_g| = \sum_{x \in X} |G_x|$. By Lemma 5.4.17,

$$\sum_{y \in \mathcal{O}_x} |G_y| = |\mathcal{O}_x| \cdot |G_x|.$$

By Theorem 5.3.11 and Lagrange's Theorem, this expression is equal to $|G|$. Summing over all of the $k$ distinct orbits, we conclude that

$$\sum_{g \in G} |X_g| = \sum_{x \in X} |G_x| = k \cdot |G|.$$

■

**Example 5.4.19** Let $X = \{1, 2, 3, 4, 5\}$ and suppose that $G$ is the permutation group $G = \{(1), (1\,3), (1\,3)(2\,5), (2\,5)\}$. The orbits of $X$ are $\{1, 3\}$, $\{2, 5\}$, and $\{4\}$. The fixed point sets are

$$\begin{aligned} X_{(1)} &= X \\ X_{(1\,3)} &= \{2, 4, 5\} \\ X_{(1\,3)(2\,5)} &= \{4\} \\ X_{(2\,5)} &= \{1, 3, 4\}. \end{aligned}$$

Burnside's Theorem says that

$$k = \frac{1}{|G|} \sum_{g \in G} |X_g| = \frac{1}{4}(5 + 3 + 1 + 3) = 3.$$

□

**A Geometric Example.** Before we apply Burnside's Theorem to switching-theory problems, let us examine the number of ways in which the vertices of a square can be colored black or white. Notice that we can sometimes obtain equivalent colorings by simply applying a rigid motion to the square. For instance, as we have pointed out, if we color one of the vertices black and the remaining three white, it does not matter which vertex was colored black since a rotation will give an equivalent coloring.

The symmetry group of a square, $D_4$, is given by the following permutations:

| | | | |
|---|---|---|---|
| $(1)$ | $(1\,3)$ | $(2\,4)$ | $(1\,4\,3\,2)$ |
| $(1\,2\,3\,4)$ | $(1\,2)(3\,4)$ | $(1\,4)(2\,3)$ | $(1\,3)(2\,4)$ |

The group $G$ acts on the set of vertices $\{1, 2, 3, 4\}$ in the usual manner. We can describe the different colorings by mappings from $X$ into $Y = \{B, W\}$ where $B$ and $W$ represent the colors black and white, respectively. Each map $f : X \to Y$ describes a way to color the corners of the square. Every $\sigma \in D_4$ induces a permutation $\widetilde{\sigma}$ of the possible colorings given by $\widetilde{\sigma}(f) = f \circ \sigma$ for $f : X \to Y$. For example, suppose that $f$ is defined by

$$f(1) = B$$

$$f(2) = W$$
$$f(3) = W$$
$$f(4) = W$$

and $\sigma = (12)(34)$. Then $\widetilde{\sigma}(f) = f \circ \sigma$ sends vertex 2 to $B$ and the remaining vertices to $W$. The set of all such $\widetilde{\sigma}$ is a permutation group $\widetilde{G}$ on the set of possible colorings. Let $\widetilde{X}$ denote the set of all possible colorings; that is, $\widetilde{X}$ is the set of all possible maps from $X$ to $Y$. Now we must compute the number of $\widetilde{G}$-equivalence classes.

1. $\widetilde{X}_{(1)} = \widetilde{X}$ since the identity fixes every possible coloring. $|\widetilde{X}| = 2^4 = 16$.
2. $\widetilde{X}_{(1\,2\,3\,4)}$ consists of all $f \in \widetilde{X}$ such that $f$ is unchanged by the permutation $(1\,2\,3\,4)$. In this case $f(1) = f(2) = f(3) = f(4)$, so that all values of $f$ must be the same; that is, either $f(x) = B$ or $f(x) = W$ for every vertex $x$ of the square. So $|\widetilde{X}_{(1\,2\,3\,4)}| = 2$.
3. $|\widetilde{X}_{(1\,4\,3\,2)}| = 2$.
4. For $\widetilde{X}_{(1\,3)(2\,4)}$, $f(1) = f(3)$ and $f(2) = f(4)$. Thus, $|\widetilde{X}_{(1\,3)(2\,4)}| = 2^2 = 4$.
5. $|\widetilde{X}_{(1\,2)(3\,4)}| = 4$.
6. $|\widetilde{X}_{(1\,4)(2\,3)}| = 4$.
7. For $\widetilde{X}_{(1\,3)}$, $f(1) = f(3)$ and the other corners can be of any color; hence, $|\widetilde{X}_{(1\,3)}| = 2^3 = 8$.
8. $|\widetilde{X}_{(2\,4)}| = 8$.

By Burnside's Theorem, we can conclude that there are exactly

$$\frac{1}{8}(2^4 + 2^1 + 2^2 + 2^1 + 2^2 + 2^2 + 2^3 + 2^3) = 6$$

ways to color the vertices of the square.

**Proposition 5.4.20** *Let $G$ be a permutation group of $X$ and $\widetilde{X}$ the set of functions from $X$ to $Y$. Then $G$ induces a group $\widetilde{G}$ that permutes the elements of $\widetilde{X}$, where $\widetilde{\sigma} \in \widetilde{G}$ is defined by $\widetilde{\sigma}(f) = f \circ \sigma$ for $\sigma \in G$ and $f \in \widetilde{X}$. Furthermore, if $n$ is the number of cycles in the cycle decomposition of $\sigma$, then $|\widetilde{X}_\sigma| = |Y|^n$.*

*Proof.* Let $\sigma \in G$ and $f \in \widetilde{X}$. Since $\sigma$ permutes the elements of $X$, $f \circ \sigma$ must also be in $\widetilde{X}$. Suppose that $g$ is another function from $X$ to $Y$ such that $\widetilde{\sigma}(f) = \widetilde{\sigma}(g)$. Then for each $x \in X$,

$$f(\sigma(x)) = \widetilde{\sigma}(f)(x) = \widetilde{\sigma}(g)(x) = g(\sigma(x)).$$

Since $\sigma$ is a permutation of $X$, every element $x'$ in $X$ is the image of some $x$ in $X$ under $\sigma$; hence, $f$ and $g$ agree on all elements of $X$. Therefore, $f = g$ and $\widetilde{\sigma}$ is injective. The map $\sigma \mapsto \widetilde{\sigma}$ is onto, since the two sets are the same size.

Suppose that $\sigma$ is a permutation of $X$ with cycle decomposition $\sigma = \sigma_1\sigma_2 \cdots \sigma_n$. Any $f$ in $\widetilde{X}_\sigma$ must have the same value on each cycle of $\sigma$. Since there are $n$ cycles and $|Y|$ possible values for each cycle, $|\widetilde{X}_\sigma| = |Y|^n$. ■

**Example 5.4.21** Let $X = \{1, 2, \ldots, 7\}$ and suppose that $Y = \{A, B, C\}$. If $g$ is the permutation of $X$ given by $(1\,3)(2\,4\,5) = (1\,3)(2\,4\,5)(6)(7)$, then $n = 4$.

Any $f \in \widetilde{X}_g$ must have the same value on each cycle in $g$. There are $|Y| = 3$ such choices for any value, so $|\widetilde{X}_g| = 3^4 = 81$. □

**Example 5.4.22** Suppose that we wish to color the vertices of a square using four different colors. By Proposition 5.4.20, we can immediately decide that there are

$$\frac{1}{8}(4^4 + 4^1 + 4^2 + 4^1 + 4^2 + 4^2 + 4^3 + 4^3) = 55$$

possible ways. □

**Switching Functions.** In switching theory we are concerned with the design of electronic circuits with binary inputs and outputs. The simplest of these circuits is a switching function that has $n$ inputs and a single output (Figure 5.4.23). Large electronic circuits can often be constructed by combining smaller modules of this kind. The inherent problem here is that even for a simple circuit a large number of different switching functions can be constructed. With only four inputs and a single output, we can construct 65,536 different switching functions. However, we can often replace one switching function with another merely by permuting the input leads to the circuit (Figure 5.4.24).

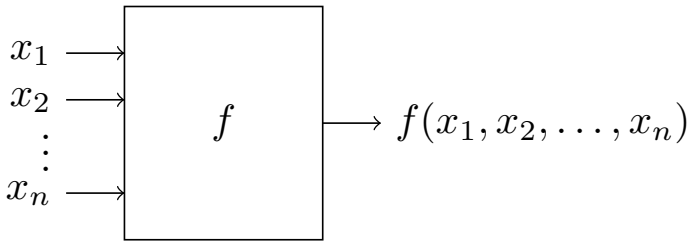



**Figure 5.4.23** A switching function of $n$ variables

We define a **switching** or **Boolean function** of $n$ variables to be a function from $\mathbb{Z}_2^n$ to $\mathbb{Z}_2$. Since any switching function can have two possible values for each binary $n$-tuple and there are $2^n$ binary $n$-tuples, $2^{2^n}$ switching functions are possible for $n$ variables. In general, allowing permutations of the inputs greatly reduces the number of different kinds of modules that are needed to build a large circuit.

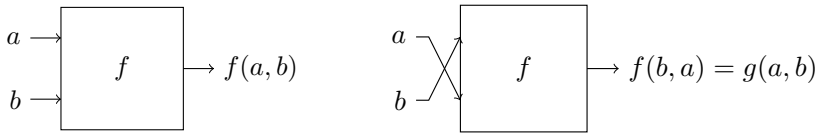



**Figure 5.4.24** Switching functions in two variables

The possible switching functions with two input variables $a$ and $b$ are listed in Table 5.4.25. Two switching functions $f$ and $g$ are equivalent if $g$ can be obtained from $f$ by a permutation of the input variables. For example, $g(a, b, c) = f(b, c, a)$. In this case $g \sim f$ via the permutation $(a, c, b)$. In the case of switching functions of two variables, the permutation $(a, b)$ reduces 16 possible switching functions to 12 equivalent functions since

$$f_2 \sim f_4$$
$$f_3 \sim f_5$$
$$f_{10} \sim f_{12}$$
$$f_{11} \sim f_{13}.$$

**Table 5.4.25 Switching functions in two variables**

| Inputs | | Outputs | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $f_0$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | $f_7$ |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| Inputs | | Outputs | | | | | | | |
| | | $f_8$ | $f_9$ | $f_{10}$ | $f_{11}$ | $f_{12}$ | $f_{13}$ | $f_{14}$ | $f_{15}$ |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |

For three input variables there are $2^{2^3} = 256$ possible switching functions; in the case of four variables there are $2^{2^4} = 65{,}536$. The number of equivalence classes is too large to reasonably calculate directly. It is necessary to employ Burnside's Theorem.

Consider a switching function with three possible inputs, $a$, $b$, and $c$. As we have mentioned, two switching functions $f$ and $g$ are equivalent if a permutation of the input variables of $f$ gives $g$. It is important to notice that a permutation of the switching functions is not simply a permutation of the input values $\{a, b, c\}$. A switching function is a set of output values for the inputs $a$, $b$, and $c$, so when we consider equivalent switching functions, we are permuting $2^3$ possible outputs, not just three input values. For example, each binary triple $(a, b, c)$ has a specific output associated with it. The permutation $(acb)$ changes outputs as follows:

$$\begin{aligned}
(0,0,0) &\mapsto (0,0,0)\\
(0,0,1) &\mapsto (0,1,0)\\
(0,1,0) &\mapsto (1,0,0)\\
&\vdots\\
(1,1,0) &\mapsto (1,0,1)\\
(1,1,1) &\mapsto (1,1,1).
\end{aligned}$$

Let $X$ be the set of output values for a switching function in $n$ variables. Then $|X| = 2^n$. We can enumerate these values as follows:

$$\begin{aligned}
(0,\ldots,0,1) &\mapsto 0\\
(0,\ldots,1,0) &\mapsto 1\\
(0,\ldots,1,1) &\mapsto 2\\
&\vdots\\
(1,\ldots,1,1) &\mapsto 2^n - 1.
\end{aligned}$$

Now let us consider a circuit with four input variables and a single output. Suppose that we can permute the leads of any circuit according to the following permutation group:

$$\begin{gathered}
(a), \quad (a,c), \quad (b,d), \quad (a,d,c,b),\\
(a,b,c,d), \quad (a,b)(c,d), \quad (a,d)(b,c), \quad (a,c)(b,d).
\end{gathered}$$

The permutations of the four possible input variables induce the permutations of the output values in Table 5.4.26.

Hence, there are

$$\frac{1}{8}(2^{16} + 2 \cdot 2^{12} + 2 \cdot 2^6 + 3 \cdot 2^{10}) = 9616$$

possible switching functions of four variables under this group of permutations. This number will be even smaller if we consider the full symmetric group on four letters.

**Table 5.4.26 Permutations of switching functions in four variables**

| Group Permutation | Switching Function Permutation | Number of Cycles |
|---|---|---|
| $(a)$ | $(0)$ | 16 |
| $(a, c)$ | $(2, 8)(3, 9)(6, 12)(7, 13)$ | 12 |
| $(b, d)$ | $(1, 4)(3, 6)(9, 12)(11, 14)$ | 12 |
| $(a, d, c, b)$ | $(1, 2, 4, 8)(3, 6.12, 9)(5, 10)(7, 14, 13, 11)$ | 6 |
| $(a, b, c, d)$ | $(1, 8, 4, 2)(3, 9, 12, 6)(5, 10)(7, 11, 13, 14)$ | 6 |
| $(a, b)(c, d)$ | $(1, 2)(4, 8)(5, 10)(6, 9)(7, 11)(13, 14)$ | 10 |
| $(a, d)(b, c)$ | $(1, 8)(2, 4)(3, 12)(5, 10)(7, 14)(11, 13)$ | 10 |
| $(a, c)(b, d)$ | $(1, 4)(2, 8)(3, 12)(6, 9)(7, 13)(11, 14)$ | 10 |

## 5.5 Core Exercises

**1.** List the left and right cosets of the subgroups in each of the following, and determine the index of the respective subgroup.

(a) $\langle 8 \rangle$ in $\mathbb{Z}_{24}$

(b) $\langle 3 \rangle$ in $U(8)$

(c) $D_4$ in $S_4$

(d) $\mathbb{T}$ in $\mathbb{C}^*$

(e) $H = \{(1), (1\,2\,3), (1\,3\,2)\}$ in $S_3$

(f) $H = \{(1), (1\,2\,3), (1\,3\,2)\}$ in $S_4$

**2.** Give two examples of a group and a subgroup, and demonstrate what Lagrange's Theorem states there.

**3.** Consider the subgroup $H = 3\mathbb{Z}$ of $G = \mathbb{Z}$. It acts on $\mathbb{Z}$ by addition from the left; note that this is a special case of Example 5.3.3.

(a) What are the cosets of $H$?

(b) What are the orbits of the action of $H$ on $G$?

(c) Consider the results of (a) and (b) and observe what happened here. Can you generalize the observation to the more general case of the subgroup $n\mathbb{Z}$ of $\mathbb{Z}$ for an arbitrary $n \in \mathbb{Z}$?

(d) Can you generalize the observation of (c) to an arbitrary group $G$ with subgroup $H$?

**4.** Let $G = S_4$ and $H = S_3$ be the subgroup fixing the element 4. Then $G$ acts on the cosets of $H$ as a special case of Example 5.3.5.

(a) What are the orbits of this group action? For each coset, what is

the stabilizer subgroup?

(b) As discussed in the beginning of Section 5.3 this group action gives rise to a homomorphism from $G$ to the group of bijections of the set $\{gH \mid g \in H\}$. What is the kernel of this group homomorphism?

(c) Now consider the general case of an arbitrary group $G$ with a subgroup $H$. What is the kernel of the respective group homomorphism given by the action of $G$ on the cosets of $H$?

**5.** Compute all $X_g$ fixed point sets and all stabilizer sugroups $G_x$ for each of the following permutation groups; check the Orbit-Stabilizer Theorem for them.

(a) $X = \{1, 2, 3\}$, $G = S_3 = \{(1), (1\,2), (1\,3), (2\,3), (1\,2\,3), (1\,3\,2)\}$

(b) $X = \{1, 2, 3, 4, 5, 6\}$, $G = \{(1), (1\,2), (3\,4\,5), (3\,5\,4), (1\,2)(3\,4\,5), (1\,2)(3\,5\,4)\}$

**6.** Let $G$ be the additive group of real numbers. Let the action of $\theta \in G$ on the real plane $\mathbb{R}^2$ be given by rotating the plane counterclockwise about the origin through $\theta$ radians. Let $P$ be a point on the plane other than the origin.

(a) Show that $\mathbb{R}^2$ is a $G$-set.

(b) Describe geometrically the orbit containing $P$.

(c) Find the group $G_P$.

**7.** Prove Lemma 5.1.3.

**8.** Prove Proposition 5.3.6.

**9.** **Programming problem.** Assume you are given a group action in the following form: a group $G$ via its Cayley table (e.g. as a dictionary encoding the map $G \times G \to G$); a set $X$; the group action as a map $G \times X \to X$ (may also be given by a dictionary).

Write a program which returns the orbits of $X$ under the group action as a list of sets.

**10.** **Connection to Linear Algebra.** We consider some constructions from linear algebra in the larger context of group actions.

(a) Show that the direct product $GL_3(\mathbb{R}) \times GL_3(\mathbb{R})$ acts on the set of square matrices $\mathbb{R}^{3\times 3}$ by $(P, Q)\, M = PMQ^{-1}$. Name at least one representative for each orbit.

(b) Show that $SL_3(\mathbb{R})$, the group of invertible matrices with determinant 1, acts on the set of square matrices $\mathbb{R}^{3\times 3}$ by left multiplication $Q\, M = Q \cdot M$. Name at least one representative for each orbit.

(c) Generalize (a) and (b) to matrices of bigger sizes.

## 5.6 Additional Exercises

**1.** Suppose that $G$ is a finite group with an element $g$ of order 5 and an element $h$ of order 7. Why must $|G| \geq 35$?

**2.** Suppose that $G$ is a finite group with 60 elements. What are the orders of possible subgroups of $G$?

**3.** Prove or disprove: Every subgroup of the integers has finite index.

**4.** Prove or disprove: Every subgroup of the integers has finite order.

**5.** Verify Euler's Theorem for $n = 15$ and $a = 4$.

**6.** Use Fermat's Little Theorem to show that if $p = 4n + 3$ is prime, there is no solution to the equation $x^2 \equiv -1 \pmod{p}$.

**7.** Show that the integers have infinite index in the additive group of rational numbers.

**8.** Show that the additive group of real numbers has infinite index in the additive group of the complex numbers.

**9.** If $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$, show that right cosets are identical to left cosets. That is, show that $gH = Hg$ for all $g \in G$.

**10.** What fails in the proof of Theorem 5.1.6 if $\phi : \mathcal{L}_H \to \mathcal{R}_H$ is defined by $\phi(gH) = Hg$?

**11.** Describe the left cosets of $SL_2(\mathbb{R})$ in $GL_2(\mathbb{R})$. What is the index of $SL_2(\mathbb{R})$ in $GL_2(\mathbb{R})$?

**12.** Examples 5.3.1–5.3.5 in the first section each describe an action of a group $G$ on a set $X$, which will give rise to the equivalence relation defined by $G$-equivalence. For each example, compute the equivalence classes of the equivalence relation, the $G$-**equivalence classes**.

**13.** Compute the $G$-equivalence classes of $X$ for each of the $G$-sets in Exercise 5.5.5. For each $x \in X$ verify that $|G| = |\mathcal{O}_x| \cdot |G_x|$.

**14.** Suppose that $g^n = e$. Show that the order of $g$ divides $n$.

**15.** The **cycle structure** of a permutation $\sigma$ is defined as the unordered list of the sizes of the cycles in the cycle decomposition $\sigma$. For example, the permutation $\sigma = (1\,2)(3\,4\,5)(7\,8)(9)$ has cycle structure $(2, 3, 2, 1)$ which can also be written as $(1, 2, 2, 3)$.

Show that any two permutations $\alpha, \beta \in S_n$ have the same cycle structure if and only if there exists a permutation $\gamma$ such that $\beta = \gamma\alpha\gamma^{-1}$. If $\beta = \gamma\alpha\gamma^{-1}$ for some $\gamma \in S_n$, then $\alpha$ and $\beta$ are **conjugate**.

**16.** If $|G| = 2n$, prove that the number of elements of order 2 is odd. Use this result to show that $G$ must contain a subgroup of order 2.

**17.** Suppose that $[G : H] = 2$. If $a$ and $b$ are not in $H$, show that $ab \in H$.

**18.** If $[G : H] = 2$, prove that $gH = Hg$.

**19.** Let $H$ and $K$ be subgroups of a group $G$. Prove that $gH \cap gK$ is a coset of $H \cap K$ in $G$.

**20.** Let $H$ and $K$ be subgroups of a group $G$. Define a relation $\sim$ on $G$ by $a \sim b$ if there exists an $h \in H$ and a $k \in K$ such that $hak = b$. Show that this relation is an equivalence relation. The corresponding equivalence classes are called **double cosets**. Compute the double cosets of $H = \{(1), (1\,2\,3), (1\,3\,2)\}$ in $A_4$.

**21.** Let $G$ be a cyclic group of order $n$. Show that there are exactly $\phi(n)$ generators for $G$.

**22.** Let $G = A_4$ and suppose that $G$ acts on itself by conjugation; that is, $(g, h) \mapsto ghg^{-1}$.

(a) Determine the conjugacy classes (orbits) of each element of $G$.

(b) Determine all of the isotropy subgroups for each element of $G$.

**23.** A group acts **faithfully** on a $G$-set $X$ if the identity is the only element of $G$ that leaves every element of $X$ fixed. Show that $G$ acts faithfully on $X$ if and only if no two distinct elements of $G$ have the same action on

each element of $X$.

## 5.7 Material

1. Cosets and Lagrange's Theorem[1]
2. More general intuition for groups and what 'act' means: Groups and Monsters[2]

## 5.8 Hints to Selected Exercises

### 5.5 · Core Exercises

**5.5.1.** (a) $\langle 8\rangle$, $1+\langle 8\rangle$, $2+\langle 8\rangle$, $3+\langle 8\rangle$, $4+\langle 8\rangle$, $5+\langle 8\rangle$, $6+\langle 8\rangle$, and $7+\langle 8\rangle$

**5.5.10. Connection to Linear Algebra.**

(a) Basis exchange.

(b) Gaussian elimination. The matrices for elementary row operations and row exchange are in the special linear group.

### 5.6 · Additional Exercises

**5.6.1.** The order of $g$ and the order $h$ must both divide the order of $G$.

**5.6.2.** The possible orders must divide 60.

**5.6.3.** This is true for every proper nontrivial subgroup.

**5.6.4.** False.

**5.6.5.** $4^{\phi(15)} \equiv 4^8 \equiv 1 \pmod{15}$.

**5.6.9.** Let $g_1 \in gH$. Show that $g_1 \in Hg$ and thus $gH \subset Hg$.

**5.6.12.** Example 5.3.1: $0$, $\mathbb{R}^2 \setminus \{0\}$. Example 5.3.2: $X = \{1,2,3,4\}$.

**5.6.13.** (a) $\mathcal{O}_1 = \mathcal{O}_2 = \mathcal{O}_3 = \{1,2,3\}$.

**5.6.19.** Show that $g(H \cap K) = gH \cap gK$.

[1] www.socratica.com/lesson/cosets-and-lagranges-theorem
[2] www.3blue1brown.com/lessons/groups-and-monsters

# Chapter 6

# Quotients and Motivation for Rings

Basic learning goals

1. Equivalent characterizations of normal subgroups.
2. Definitions and computation of factor groups.
3. Isomorphism theorem and examples.
4. Presentation of a group and related terminology.
5. Word problem and basic computations.
6. Motivating examples (integers) and ideas (solving polynomial systems) for rings.

Normal subgroups play an important role in the study of group homomorphisms. Based on their properties, one can decompose groups into 'factors'; it turns out that the cosets of a normal subgroup inherit a group structure. These groups are called factor or quotient groups. Considering factor groups of free groups leads to the concept of group presentations, yet another way to encode a group. We use this to state the word problem. We embed this problem in the bigger context of complexity theory as well as cryptography, and we give some instructive examples. The chapter finishes with some preparations and intuition for rings.

## 6.1 Normal Subgroups and Factor Groups

### 6.1.1 Normal Subgroups

Recall from Section 4.3 that a subgroup $H$ of a group $G$ is **normal** in G if $gHg^{-1} = H$ for all $g \in G$.

The following theorem gives equivalent characterizations of normal subgroups, in particular relating to cosets.

**Theorem 6.1.1** *Let $G$ be a group and $N$ be a subgroup of $G$. Then the following statements are equivalent.*

1. *For all $g \in G$, $gN = Ng$.*
2. *For all $g \in G$, $gNg^{-1} \subset N$.*

*3. For all $g \in G$, $gNg^{-1} = N$.*

Hence, a normal subgroup group equivalently fulfills $gH = Hg$ for all $g \in G$. That is, a normal subgroup of a group $G$ is one in which the right and left cosets are precisely the same.

We have already seen in Example 4.3.1 that all subgroups of an abelian group are normal. Recall that it becomes more subtle when we look at non-abelian groups.

**Example 6.1.2** Let $H$ be the subgroup of $S_3$ consisting of elements (1) and (12). Since

$$(123)H = \{(1\,2\,3), (1\,3)\} \quad \text{and} \quad H(1\,2\,3) = \{(1\,2\,3), (2\,3)\},$$

$H$ cannot be a normal subgroup of $S_3$. However, the subgroup $N$, consisting of the permutations (1), (1 2 3), and (1 3 2), is normal since the cosets of $N$ are

$$N = \{(1), (1\,2\,3), (1\,3\,2)\}$$
$$(1\,2)N = N(1\,2) = \{(1\,2), (1\,3), (2\,3)\}.$$

□

## 6.1.2 Factor Groups

If $N$ is a normal subgroup of a group $G$, then the cosets of $N$ in $G$ form a group $G/N$ under the operation $(aN)(bN) = abN$. This group is called the **factor** or **quotient group** of $G$ and $N$. Our first task is to prove that $G/N$ is indeed a group.

**Theorem 6.1.3** *Let $N$ be a normal subgroup of a group $G$. The cosets of $N$ in $G$ form a group $G/N$ of order $[G : N]$.*

*Proof.* The group operation on $G/N$ is $(aN)(bN) = abN$. This operation must be shown to be well-defined; that is, group multiplication must be independent of the choice of coset representative. Let $aN = bN$ and $cN = dN$. We must show that

$$(aN)(cN) = acN = bdN = (bN)(dN).$$

Then $a = bn_1$ and $c = dn_2$ for some $n_1$ and $n_2$ in $N$. Hence,

$$\begin{aligned} acN &= bn_1dn_2N \\ &= bn_1dN \\ &= bn_1Nd \\ &= bNd \\ &= bdN. \end{aligned}$$

The remainder of the theorem is easy: $eN = N$ is the identity and $g^{-1}N$ is the inverse of $gN$. The order of $G/N$ is, of course, the number of cosets of $N$ in $G$. ■

It is very important to remember that the elements in a factor group are *sets of elements* in the original group.

**Example 6.1.4** Consider the normal subgroup of $S_3$, $N = \{(1), (1\,2\,3), (1\,3\,2)\}$. The cosets of $N$ in $S_3$ are $N$ and $(12)N$. The factor group $S_3/N$ has the following multiplication table.

| | $N$ | $(1\,2)N$ |
|---|---|---|
| $N$ | $N$ | $(1\,2)N$ |
| $(1\,2)N$ | $(1\,2)N$ | $N$ |

This group is isomorphic to $\mathbb{Z}_2$. At first, multiplying cosets seems both complicated and strange; however, notice that $S_3/N$ is a smaller group. The factor group displays a certain amount of information about $S_3$. Actually, $N$ is the group of even permutations ($A_3$), and the coset $(1\,2)N = \{(1\,2), (1\,3), (2\,3)\}$ is the set of odd permutations. The information captured in $G/N$ is parity; that is, multiplying two even or two odd permutations results in an even permutation, whereas multiplying an odd permutation by an even permutation yields an odd permutation. □

**Example 6.1.5** Consider the normal subgroup $3\mathbb{Z}$ of $\mathbb{Z}$. The cosets of $3\mathbb{Z}$ in $\mathbb{Z}$ are

$$\begin{aligned} 0 + 3\mathbb{Z} &= \{\ldots, -3, 0, 3, 6, \ldots\} \\ 1 + 3\mathbb{Z} &= \{\ldots, -2, 1, 4, 7, \ldots\} \\ 2 + 3\mathbb{Z} &= \{\ldots, -1, 2, 5, 8, \ldots\}. \end{aligned}$$

The group $\mathbb{Z}/3\mathbb{Z}$ is given by the Cayley table below.

| $+$ | $0+3\mathbb{Z}$ | $1+3\mathbb{Z}$ | $2+3\mathbb{Z}$ |
|---|---|---|---|
| $0+3\mathbb{Z}$ | $0+3\mathbb{Z}$ | $1+3\mathbb{Z}$ | $2+3\mathbb{Z}$ |
| $1+3\mathbb{Z}$ | $1+3\mathbb{Z}$ | $2+3\mathbb{Z}$ | $0+3\mathbb{Z}$ |
| $2+3\mathbb{Z}$ | $2+3\mathbb{Z}$ | $0+3\mathbb{Z}$ | $1+3\mathbb{Z}$ |

In general, the subgroup $n\mathbb{Z}$ of $\mathbb{Z}$ is normal since $\mathbb{Z}$ is abelian. The cosets of $\mathbb{Z}/n\mathbb{Z}$ are

$$\begin{gathered} n\mathbb{Z} \\ 1 + n\mathbb{Z} \\ 2 + n\mathbb{Z} \\ \vdots \\ (n-1) + n\mathbb{Z}. \end{gathered}$$

The sum of the cosets $k + n\mathbb{Z}$ and $l + n\mathbb{Z}$ is $k + l + n\mathbb{Z}$. Notice that we have written our cosets additively, because the group operation is integer addition. This just yields the group $(\mathbb{Z}_n, +)$. □

**Example 6.1.6** Consider the dihedral group $D_n$, generated by the two elements $r$ and $s$, satisfying the relations

$$\begin{aligned} r^n &= \mathrm{id} \\ s^2 &= \mathrm{id} \\ srs &= r^{-1}. \end{aligned}$$

The element $r$ actually generates the cyclic subgroup of rotations, $R_n$, of $D_n$. Since $srs^{-1} = srs = r^{-1} \in R_n$, the group of rotations is a normal subgroup of $D_n$; therefore, $D_n/R_n$ is a group. Since there are exactly two elements in this group, it must be isomorphic to $\mathbb{Z}_2$. □

## 6.2 The Isomorphism Theorem(s)

Although it is not evident at first, factor groups correspond exactly to homomorphic images, and we can use factor groups to study homomorphisms. We already know that with every group homomorphism $\phi : G \to H$ we can associate a normal subgroup of $G$, the kernel $\ker \phi$. The converse is also true; that is, every normal subgroup of a group $G$ gives rise to a homomorphism of groups.

Let $H$ be a normal subgroup of $G$. Define the **natural** or **canonical homomorphism**

$$\phi : G \to G/H$$

by

$$\phi(g) = gH.$$

This is indeed a homomorphism, since

$$\phi(g_1 g_2) = g_1 g_2 H = g_1 g_2 HH = g_1 H g_2 H = \phi(g_1)\phi(g_2),$$

where we use that $H = HH$ (closedness) and $g_2 H = H g_2$ (normality). The kernel of this homomorphism is $H$. The following theorem describes the fundamental relationship between group homomorphisms, normal subgroups, and factor groups. There are two more isomorphism theorems refining this relation but they are outside the scope of this course and can be found in Subsection 6.5.1.

**Theorem 6.2.1 First Isomorphism Theorem.** *If $\psi : G \to H$ is a group homomorphism with $K = \ker \psi$, then $K$ is normal in $G$. Let $\phi : G \to G/K$ be the canonical homomorphism. Then there exists a unique isomorphism $\eta : G/K \to \psi(G)$ such that $\psi = \eta\phi$.*

*Proof.* We already know that $K$ is normal in $G$. Define $\eta : G/K \to \psi(G)$ by $\eta(gK) = \psi(g)$. We first show that $\eta$ is a well-defined map. If $g_1 K = g_2 K$, then for some $k \in K$, $g_1 k = g_2$; consequently,

$$\eta(g_1 K) = \psi(g_1) = \psi(g_1)\psi(k) = \psi(g_1 k) = \psi(g_2) = \eta(g_2 K).$$

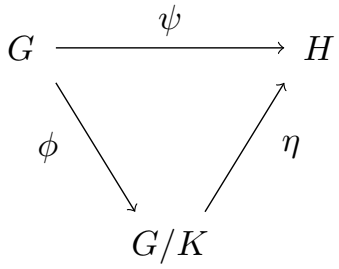
Thus, $\eta$ does not depend on the choice of coset representatives and the map $\eta : G/K \to \psi(G)$ is uniquely defined since $\psi = \eta\phi$. We must also show that $\eta$ is a homomorphism. Indeed,

$$\begin{aligned} \eta(g_1 K g_2 K) &= \eta(g_1 g_2 K) \\ &= \psi(g_1 g_2) \\ &= \psi(g_1)\psi(g_2) \\ &= \eta(g_1 K)\eta(g_2 K). \end{aligned}$$

Clearly, $\eta$ is onto $\psi(G)$. To show that $\eta$ is one-to-one, suppose that $\eta(g_1 K) = \eta(g_2 K)$. Then $\psi(g_1) = \psi(g_2)$. This implies that $\psi(g_1^{-1} g_2) = e$, or $g_1^{-1} g_2$ is in the kernel of $\psi$; hence, $g_1^{-1} g_2 K = K$; that is, $g_1 K = g_2 K$. ■

**Corollary 6.2.2** *Normal subgroups are exactly the kernels of group homomorphisms.*

Mathematicians often use diagrams called **commutative diagrams** to describe such theorems. The following diagram "commutes" since $\psi = \eta\phi$.




**Example 6.2.3** Let $G$ be a cyclic group with generator $g$. Define a map $\phi : \mathbb{Z} \to G$ by $n \mapsto g^n$. This map is a surjective homomorphism since

$$\phi(m+n) = g^{m+n} = g^m g^n = \phi(m)\phi(n).$$

Clearly $\phi$ is onto. If $|g| = m$, then $g^m = e$. Hence, $\ker \phi = m\mathbb{Z}$ and $\mathbb{Z}/\ker \phi = \mathbb{Z}/m\mathbb{Z} \cong G$. On the other hand, if the order of $g$ is infinite, then $\ker \phi = 0$ and $\phi$ is an isomorphism of $G$ and $\mathbb{Z}$. Hence, two cyclic groups are isomorphic exactly when they have the same order. Up to isomorphism, the only cyclic groups are $\mathbb{Z}$ and $\mathbb{Z}_n$. □

## 6.3 Group presentations and the word problem

A presentation of a group $G$ consists of a set of generators $S$ and a set of relations $R$, a subset of the words over the alphabet $S$. This means that every element in $G$ can be written as a product of (negative) powers of the generators. Furthermore, $R$ specifies which products of the generators equal to the neutral element.

Recall free groups from Section 4.4. A group $G$ is described by the presentation $\langle S \mid R \rangle$ if $G$ is isomorphic to the quotient $F(S)/N$ where $F(S)$ is the free group on $S$ and $N$ is the smallest normal subgroup of $F(S)$ containing $R$ (see Proposition 4.3.4 for the normal closure). Recall the definition of a free group and the canonical homomorphism $\phi$ to $G$ from Section 4.4. Then the images of the elements in $S$ actually form a generating set for $G$ and $N$ is the kernel of $\phi$. By the First Isomorphism Theorem (Theorem 6.2.1), this idea yields a presentation for every group.

There are several related notions arising from this concept. A presentation is

1. finitely generated if $S$ is finite,
2. finitely related if $R$ is finite,
3. finite if both are finite.

A group is finitely generated / related / presented if it has a presentation with the respective property.

**Example 6.3.1** The presentation of the free group $\mathcal{F}(S)$ on a set $S$ is $\langle S|\emptyset\rangle$, it has no relations. □

**Example 6.3.2** By Theorem 4.2.8, we know that there is exactly one cyclic group of order $n$ up to isomorphism, i.e., we can think of it as $(\mathbb{Z}_n, +)$. It has the presentation $\langle a|a^n\rangle$. Note that it has also the presentation $\langle a, b|a^n, ab^{-1}\rangle$ but $b$ is redundant in all words here. □

**Example 6.3.3** We have seen the dihedral groups as symmetry groups of regular $n$-gons and they were further discussed in Exercise 3.5.7. There, it was

exhibited that all elements of the group $D_n$ arise from two special symmetries $r$ and $s$ (for the proof see Subsection 3.4.3). It implies that $D_n$ has the presentation $\langle r, s \mid r^n, s^2, srsr\rangle$. □

We saw in (4.4.1) that there is a particularly simple way of describing group homomorphisms whose domain is a free group. It was enough to describe the image of the generators of the free group and then this extends to a group homomorphism. For a group given by a presentation there is an extension of this idea: we can also just prescribe the images of the generators as long as all the words in $R$ are still mapped to the identity.

### 6.3.1 The word problem

Recall the rewriting question from Exercise 1.4.7. This is actually a special instance of the so-called **Word Problem**. That is, one is given a (finitely presented) group via a group presentation $\langle S \mid R\rangle$ where $S$ is a finite set and $R$ is a finite set of words over $S$. Now, for a fixed word $w$ over $S$, one is asked to determine if $w$ represents the identity.

Actually, this problem is not so much about the group structure itself but more about the structure of the equivalence classes of words representing the same element. This amounts to use the rewriting rules which were already stated before:

Given a word $w$ from $S$, there are three rules to modify $w$:

(i) For an arbitrary $x \in S$, a consecutive occurrence $xx^{-1}$ or $x^{-1}x$ can be replaced by the empty word.

(ii) A word in $R$ can be replaced by the empty word.

(iii) The empty word can be replaced by a word in $R$ or by an expression $xx^{-1}$ or $x^{-1}x$ for an arbitrary $x \in S$.

While the examples in Exercise 1.4.7 could be solved by hand, the problem is arbitrarily hard in general.

**Theorem 6.3.4 Novikov–Boone Theorem.** *There is a finite presentation $\langle S \mid R\rangle$ such that no algorithm can decide whether a word $w$ represents the identity element.*

We briefly sketch a general approach for solving the word problem.

The substitution rules define an abstract rewriting system[1]. For every word $w$ one needs to define a normal form $\bar{w}$. Then, to solve the word problem, we just derive the normal form and compare it to the empty word. Defining normal forms is in general not trivial and specific to the problem (e.g. compare with the graph isomorphism problem); note that a normal form may not even be computable in general. Intuitively, $\bar{w}$ should be one of the words of shortest length which we can obtain by rewriting.

Let $w \to w'$ mean that we can rewrite $w$ to obtain $w'$, and let $w \rightsquigarrow w'$ mean that we can rewrite $w$ to obtain $w'$ using only the first two rules or that we just leave $w$ fixed. We want the following properties.

1. Normal forms are unique.
2. $\bar{\epsilon} = \epsilon$.
3. If $w \to v$ then $\bar{w} = \bar{v}$.
4. $w \rightsquigarrow \bar{w}$.

[1] en.wikipedia.org/wiki/Abstract_rewriting_system

Then this yields that $x \to y$ implies $x \rightsquigarrow \bar{x} = \bar{y}$ and $y \rightsquigarrow \bar{y}$. Now, to check if a word $z$ can be rewritten to obtain the empty word $\epsilon$ it suffices to compute $\bar{z}$.

A similar idea appears in the famous Knuth-Bendix algorithm[2] which has many further use cases.

The word problem is closely related to two other problems which are in general also undecidable.

1. Conjugacy Problem: given a finitely presented group $G$ and two words $v$ and $w$, determine if they represent two conjugate elements of $G$.

2. Group Isomorphism Problem: given two finite presentations $P_1$ and $P_2$, determine if they encode isomorphic groups.

For further insights see An Introduction to Combinatorial Group Theory and the Word Problem[3].

As already discussed before, a fundamental ingredient for cryptosystems are one-way functions. Computational problems arising for (finitely presented) groups like the word problem give rise to good candidates for this. For further details, see, e.g.,

1. Using decision problems in public key cryptography[4]

2. Aspects of Nonabelian Group Based Cryptography: A Survey and Open Problems[5]

3. A Public-Key Cryptosystem Based on the Word Problem[6]

## 6.4 Integers and polynomials

### 6.4.1 Prime Numbers

Let $p$ be an integer such that $p > 1$. We say that $p$ is a **prime number**, or simply $p$ is **prime**, if the only positive numbers that divide $p$ are 1 and $p$ itself. An integer $n > 1$ that is not prime is said to be **composite**.

**Lemma 6.4.1 Euclid.** *Let $a$ and $b$ be integers and $p$ be a prime number. If $p \mid ab$, then either $p \mid a$ or $p \mid b$.*

*Proof.* Suppose that $p$ does not divide $a$. We must show that $p \mid b$. Since $\gcd(a, p) = 1$, there exist integers $r$ and $s$ such that $ar + ps = 1$. So

$$b = b(ar + ps) = (ab)r + p(bs).$$

Since $p$ divides both $ab$ and itself, $p$ must divide $b = (ab)r + p(bs)$. ■

**Theorem 6.4.2 Euclid.** *There exist an infinite number of primes.*

*Proof.* We will prove this theorem by contradiction. Suppose that there are only a finite number of primes, say $p_1, p_2, \ldots, p_n$. Let $P = p_1 p_2 \cdots p_n + 1$. Then $P$ must be divisible by some $p_i$ for $1 \leq i \leq n$. In this case, $p_i$ must divide $P - p_1 p_2 \cdots p_n = 1$, which is a contradiction. Hence, either $P$ is prime or there exists an additional prime number $p \neq p_i$ that divides $P$. ■

[2] en.wikipedia.org/wiki/Knuth%E2%80%93Bendix_completion_algorithm
[3] www.tandfonline.com/doi/abs/10.1080/0025570X.1997.11996491
[4] arxiv.org/abs/math/0703656
[5] arxiv.org/abs/1103.4093
[6] link.springer.com/chapter/10.1007/3-540-39568-7_3

**Theorem 6.4.3 Fundamental Theorem of Arithmetic.** *Let $n$ be an integer such that $n > 1$. Then*

$$n = p_1 p_2 \cdots p_k,$$

*where $p_1, \ldots, p_k$ are primes (not necessarily distinct). Furthermore, this factorization is unique; that is, if*

$$n = q_1 q_2 \cdots q_l,$$

*then $k = l$ and the $q_i$'s are just the $p_i$'s rearranged.*

*Proof. Uniqueness.* To show uniqueness we will use induction on $n$. The theorem is certainly true for $n = 2$ since in this case $n$ is prime. Now assume that the result holds for all integers $m$ such that $1 \leq m < n$, and

$$n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l,$$

where $p_1 \leq p_2 \leq \cdots \leq p_k$ and $q_1 \leq q_2 \leq \cdots \leq q_l$. By Lemma 6.4.1, $p_1 \mid q_i$ for some $i = 1, \ldots, l$ and $q_1 \mid p_j$ for some $j = 1, \ldots, k$. Since all of the $p_i$'s and $q_i$'s are prime, $p_1 = q_i$ and $q_1 = p_j$. Hence, $p_1 = q_1$ since $p_1 \leq p_j = q_1 \leq q_i = p_1$. By the induction hypothesis,

$$n' = p_2 \cdots p_k = q_2 \cdots q_l$$

has a unique factorization. Hence, $k = l$ and $q_i = p_i$ for $i = 1, \ldots, k$.

*Existence.* To show existence, suppose that there is some integer that cannot be written as the product of primes. Let $S$ be the set of all such numbers. By the Principle of Well-Ordering, $S$ has a smallest number, say $a$. If the only positive factors of $a$ are $a$ and 1, then $a$ is prime, which is a contradiction. Hence, $a = a_1 a_2$ where $1 < a_1 < a$ and $1 < a_2 < a$. Neither $a_1 \in S$ nor $a_2 \in S$, since $a$ is the smallest element in $S$. So

$$a_1 = p_1 \cdots p_r$$
$$a_2 = q_1 \cdots q_s.$$

Therefore,

$$a = a_1 a_2 = p_1 \cdots p_r q_1 \cdots q_s.$$

So $a \notin S$, which is a contradiction. ■

There is a nice application of the former result to encode tuples of non-negative integers. Fix a positive integer $k$ and assume we want to encode all tuples $\mathbb{Z}_{>0}^k$. Denoting the first $k$ prime numbers by $p_1, p_2, \ldots, p_k$, this is in bijection with the set $\{p_1^{t_1} \cdots p_k^{t_k} \mid (t_1, \ldots, t_k) \in \mathbb{Z}_{>0}^k\}$ which is itself just a subset of $\mathbb{Z}_{>0}$.

### 6.4.2 Diophantine Equations and Divisibility

A **Diophantine Equation** is an equation (usually given by a polynomial with integer coefficients) for which only integers are allowed as solutions.

A famous example are the equations

$$x^n + y^n = z^n$$

for a fixed positive integer. For $n = 2$ the solutions form the triples of side integer lengths of a right triangle, the so-called Pythagorean triples. For $n > 2$,

this leads to the famous Fermat's Last Theorem finally proven in 1994 (see The Proof of Fermat's Last Theorem by R.Taylor and A.Wiles[1]).

Another example is the **Pell Equation**. Let $d$ be a square-free positive integer. Then one asks for the integer solution of the equation

$$x^2 - dy^2 = 1 \ .$$

It has the remarkable property that if $(p, q), (r, s)$ are solution pairs, so that

$$(p + \sqrt{d}q)(p - \sqrt{d}q) = (r + \sqrt{d}s)(r - \sqrt{d}s) = 1$$

Then also

$$(p + \sqrt{d}q)(r + \sqrt{d}s)(p - \sqrt{d}q)(r - \sqrt{d}s) =$$
$$(pr + dqs + (ps + qr)\sqrt{d})(pr + dqs - (ps + qr)\sqrt{d}) = 1\,,$$

so that $(pr + dqs, ps + qr)$ is a solution as well. This gives rise to an explicit way to write all solutions and this all relies on the observation that these pairs are the **units** for the algebraic structure $\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\}$. For further details on Pell Equations, we refer to Pell Equation[2] and the references herein.

### 6.4.3 Solving polynomial systems

We already saw groups as generalizations of vector spaces in Chapter 2 and group homomorphisms as generalizations of linear maps in Chapter 4. Another core topic of Linear Algebra is the theory for solving systems of linear equations. As sketched in Subsection 1.3.3, this is a special case of the important general question for solving systems of polynomial equations. We start by recalling some of these ideas and point towards structures for studying their solutions.

**Systems of linear equations.** A matrix $A \in \mathbb{R}^{m\times n}$ gives rise to the homogeneous system.

$$Ax = 0 \ . \tag{6.4.1}$$

Let $x_0 \in \mathbb{R}^n$ be a solution of this system. Then, for an arbitrary vector $y \in \mathbb{R}^m$, the vector $x_0$ is also a solution to the single linear equation

$$y^\top Ax = 0$$

arising as linear combination of the equations in (6.4.1). In particular, a vector $x^* \in \mathbb{R}^n$ is a solution of (6.4.1) if and only if all polynomials in

$$\left\{ \sum_{i=1}^{m} y_i \sum_{j=1}^{n} a_{ij}x_j \mid y_1, \ldots, y_m \in \mathbb{R} \right\} \ , \tag{6.4.2}$$

linear combinations of the linear forms of the equality system, vanish when evaluated at $x^*$.

[1] www.ams.org/notices/199507/faltings.pdf
[2] mathworld.wolfram.com/PellEquation.html

**Systems of polynomial equations.** Recall the exemplary system from Subsection 1.3.3 for the intersection of an ellipse and a parabola

$$\frac{x^2}{p^2} + \frac{y^2}{q^2} = r \tag{6.4.3}$$

$$ax^2 + bx + c = y \tag{6.4.4}$$

. Note that a solution $(x^*, y^*) \in \mathbb{R}^2$ to this system is actually a solution to each equation of the form

$$g(x,y) \cdot \left( \frac{x^2}{p^2} + \frac{y^2}{q^2} - r \right) + h(x,y) \cdot \left( ax^2 + bx + c - y \right) = 0$$

for arbitary real polynomials $g(x,y)$ and $h(x,y)$ in the variables $x$ and $y$.

**Vanishing sets of polynomials.** Let $p_1, p_2, \ldots, p_m$ be polynomials in the variables $x_1, \ldots, x_n$ with real coefficients. Then

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= 0 \\ p_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ p_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

is a system of polynomial equations. Similarly to above, a point $(x_1^*, \ldots, x_n^*)$ is a solution if it is a solution to each equation

$$h_1(x_1, \ldots, x_n) \cdot p_1(x_1, \ldots, x_n) + \cdots + h_m(x_1, \ldots, x_n) \cdot p_m(x_1, \ldots, x_n) = 0 \tag{6.4.5}$$

with polynomials $h_1(x_1, \ldots, x_n), h_2(x_1, \ldots, x_n), \ldots, h_m(x_1, \ldots, x_n)$. The set of such polynomials in (6.4.5) forms an ideal in the ring of polynomials - this will be an important notion in the next chapters.

**Solving polynomial systems.** To first exhibit the structure of the system, one can use Groebner basis methods. This usually involves Buchberger's Algorithm[3] which is similar to the algorithms in Subsection 6.3.1. Furthermore, numerical methods are helpful tools, like Newton's method[4] or the Homotopy continuation method[5].

## 6.5 Additional insights

### 6.5.1 Isomorphisms of factor groups

**Theorem 6.5.1 Second Isomorphism Theorem.** *Let $H$ be a subgroup of a group $G$ (not necessarily normal in $G$) and $N$ a normal subgroup of $G$. Then $HN$ is a subgroup of $G$, $H \cap N$ is a normal subgroup of $H$, and*

$$H/H \cap N \cong HN/N.$$

*Proof.* We will first show that $HN = \{hn : h \in H, n \in N\}$ is a subgroup of $G$.

[3] en.wikipedia.org/wiki/Buchberger%27s_algorithm
[4] en.wikipedia.org/wiki/Newton%27s_method
[5] en.wikipedia.org/wiki/System_of_polynomial_equations#Homotopy_continuation_method

Suppose that $h_1n_1, h_2n_2 \in HN$. Since $N$ is normal, $(h_2)^{-1}n_1h_2 \in N$. So

$$(h_1n_1)(h_2n_2) = h_1h_2((h_2)^{-1}n_1h_2)n_2$$

is in $HN$. The inverse of $hn \in HN$ is in $HN$ since

$$(hn)^{-1} = n^{-1}h^{-1} = h^{-1}(hn^{-1}h^{-1}).$$

Next, we prove that $H \cap N$ is normal in $H$. Let $h \in H$ and $n \in H \cap N$. Then $h^{-1}nh \in H$ since each element is in $H$. Also, $h^{-1}nh \in N$ since $N$ is normal in $G$; therefore, $h^{-1}nh \in H \cap N$.

Now define a map $\phi$ from $H$ to $HN/N$ by $h \mapsto hN$. The map $\phi$ is onto, since any coset $hnN = hN$ is the image of $h$ in $H$. We also know that $\phi$ is a homomorphism because

$$\phi(hh') = hh'N = hNh'N = \phi(h)\phi(h').$$

By the First Isomorphism Theorem, the image of $\phi$ is isomorphic to $H/\ker\phi$; that is,

$$HN/N = \phi(H) \cong H/\ker\phi.$$

Since

$$\ker\phi = \{h \in H : h \in N\} = H \cap N,$$

$HN/N = \phi(H) \cong H/H \cap N$. ■

**Theorem 6.5.2 Correspondence Theorem.** *Let $N$ be a normal subgroup of a group $G$. Then $H \mapsto H/N$ is a one-to-one correspondence between the set of subgroups $H$ of $G$ containing $N$ and the set of subgroups of $G/N$. Furthermore, the normal subgroups of $G$ containing $N$ correspond to normal subgroups of $G/N$.*

*Proof.* Let $H$ be a subgroup of $G$ containing $N$. Since $N$ is normal in $H$, $H/N$ is a factor group. Let $aN$ and $bN$ be elements of $H/N$. Then $(aN)(b^{-1}N) = ab^{-1}N \in H/N$; hence, $H/N$ is a subgroup of $G/N$.

Let $S$ be a subgroup of $G/N$. This subgroup is a set of cosets of $N$. If $H = \{g \in G : gN \in S\}$, then for $h_1, h_2 \in H$, we have that $(h_1N)(h_2N) = h_1h_2N \in S$ and $h_1^{-1}N \in S$. Therefore, $H$ must be a subgroup of $G$. Clearly, $H$ contains $N$. Therefore, $S = H/N$. Consequently, the map $H \mapsto H/N$ is onto.

Suppose that $H_1$ and $H_2$ are subgroups of $G$ containing $N$ such that $H_1/N = H_2/N$. If $h_1 \in H_1$, then $h_1N \in H_1/N$. Hence, $h_1N = h_2N \subset H_2$ for some $h_2$ in $H_2$. However, since $N$ is contained in $H_2$, we know that $h_1 \in H_2$ or $H_1 \subset H_2$. Similarly, $H_2 \subset H_1$. Since $H_1 = H_2$, the map $H \mapsto H/N$ is one-to-one.

Suppose that $H$ is normal in $G$ and $N$ is a subgroup of $H$. Then it is easy to verify that the map $G/N \to G/H$ defined by $gN \mapsto gH$ is a homomorphism. The kernel of this homomorphism is $H/N$, which proves that $H/N$ is normal in $G/N$.

Conversely, suppose that $H/N$ is normal in $G/N$. The homomorphism given by

$$G \to G/N \to \frac{G/N}{H/N}$$

has kernel $H$. Hence, $H$ must be normal in $G$. ■

Notice that in the course of the proof of Theorem 6.5.2, we have also proved the following theorem.

**Theorem 6.5.3 Third Isomorphism Theorem.** *Let $G$ be a group and $N$ and $H$ be normal subgroups of $G$ with $N \subset H$. Then*

$$G/H \cong \frac{G/N}{H/N}.$$

**Example 6.5.4** By the Third Isomorphism Theorem,

$$\mathbb{Z}/m\mathbb{Z} \cong (\mathbb{Z}/mn\mathbb{Z})/(m\mathbb{Z}/mn\mathbb{Z}).$$

Since $|\mathbb{Z}/mn\mathbb{Z}| = mn$ and $|\mathbb{Z}/m\mathbb{Z}| = m$, we have $|m\mathbb{Z}/mn\mathbb{Z}| = n$. □

### 6.5.2 The Fundamental Theorem of Finite Abelian Groups

Recall the Fundamental Theorem of Finite Abelian Groups (Theorem 4.2.10). It states that every finite abelian group is isomorphic to a direct product of cyclic groups whose order is a prime power. Letting $p$ be prime, we define a group $G$ to be a **$p$-group** if every element in $G$ has as its order a power of $p$. For example, both $\mathbb{Z}_2 \times \mathbb{Z}_2$ and $\mathbb{Z}_4$ are 2-groups, whereas $\mathbb{Z}_{27}$ is a 3-group.

The proof of the Fundamental Theorem of Finite Abelian Groups depends on several lemmas.

**Lemma 6.5.5** *Let $G$ be a finite abelian group of order $n$. If $p$ is a prime that divides $n$, then $G$ contains an element of order $p$.*

*Proof.* We will prove this lemma by induction. If $n = 1$, then there is nothing to show. Now suppose that the lemma is true for all groups of order $k$, where $k < n$. Furthermore, let $p$ be a prime that divides $n$.

If $G$ has no proper nontrivial subgroups, then $G = \langle a \rangle$, where $a$ is any element other than the identity. By Exercise 3.6.10, the order of $G$ must be prime. Since $p$ divides $n$, we know that $p = n$, and $G$ contains $p - 1$ elements of order $p$.

Now suppose that $G$ contains a nontrivial proper subgroup $H$. Then $1 < |H| < n$. If $p \mid |H|$, then $H$ contains an element of order $p$ by induction and the lemma is true. Suppose that $p$ does not divide the order of $H$. Since $G$ is abelian, it must be the case that $H$ is a normal subgroup of $G$, and $|G| = |H| \cdot |G/H|$. Consequently, $p$ must divide $|G/H|$. Since $|G/H| < |G| = n$, we know that $G/H$ contains an element $aH$ of order $p$ by the induction hypothesis. Thus,

$$H = (aH)^p = a^p H,$$

and $a^p \in H$ but $a \notin H$. If $|H| = r$, then $p$ and $r$ are relatively prime, and there exist integers $s$ and $t$ such that $sp + tr = 1$. Furthermore, the order of $a^p$ must divide $r$, and $(a^p)^r = (a^r)^p = 1$.

We claim that $a^r$ has order $p$. We must show that $a^r \neq 1$. Suppose $a^r = 1$. Then

$$\begin{aligned} a &= a^{sp+tr} \\ &= a^{sp}a^{tr} \\ &= (a^p)^s(a^r)^t \\ &= (a^p)^s 1 \\ &= (a^p)^s. \end{aligned}$$

Since $a^p \in H$, it must be the case that $a = (a^p)^s \in H$, which is a contradiction. Therefore, $a^r \neq 1$ is an element of order $p$ in $G$. ■

Lemma 6.5.5 is a special case of Cauchy's Theorem (Theorem 5.4.8), which states that if $G$ is a finite group and $p$ a prime such that $p$ divides the order of $G$, then $G$ contains a subgroup of order $p$. We will prove Cauchy's Theorem in Subsection 5.4.3.

**Lemma 6.5.6** *A finite abelian group is a $p$-group if and only if its order is a power of $p$.*

*Proof.* If $|G| = p^n$ then by Lagrange's theorem, then the order of any $g \in G$ must divide $p^n$, and therefore must be a power of $p$. Conversely, if $|G|$ is not a power of $p$, then it has some other prime divisor $q$, so by Lemma 6.5.5, $G$ has an element of order $q$ and thus is not a $p$-group. ■

**Lemma 6.5.7** *Let $G$ be a finite abelian group of order $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, where where $p_1, \ldots, p_k$ are distinct primes and $\alpha_1, \alpha_2, \ldots, \alpha_k$ are positive integers. Then $G$ is the internal direct product of subgroups $G_1, G_2, \ldots, G_k$, where $G_i$ is the subgroup of $G$ consisting of all elements of order $p_i^r$ for some integer $r$.*

*Proof.* Since $G$ is an abelian group, we are guaranteed that $G_i$ is a subgroup of $G$ for $i = 1, \ldots, k$. Since the identity has order $p_i^0 = 1$, we know that $1 \in G_i$. If $g \in G_i$ has order $p_i^r$, then $g^{-1}$ must also have order $p_i^r$. Finally, if $h \in G_i$ has order $p_i^s$, then

$$(gh)^{p_i^t} = g^{p_i^t} h^{p_i^t} = 1 \cdot 1 = 1,$$

where $t$ is the maximum of $r$ and $s$.

We must show that

$$G = G_1 G_2 \cdots G_k$$

and $G_i \cap G_j = \{1\}$ for $i \neq j$. Suppose that $g_1 \in G_1$ is in the subgroup generated by $G_2, G_3, \ldots, G_k$. Then $g_1 = g_2 g_3 \cdots g_k$ for $g_i \in G_i$. Since $g_i$ has order $p^{\alpha_i}$, we know that $g_i^{p^{\alpha_i}} = 1$ for $i = 2, 3, \ldots, k$, and $g_1^{p_2^{\alpha_2} \cdots p_k^{\alpha_k}} = 1$. Since the order of $g_1$ is a power of $p_1$ and $\gcd(p_1, p_2^{\alpha_2} \cdots p_k^{\alpha_k}) = 1$, it must be the case that $g_1 = 1$ and the intersection of $G_1$ with any of the subgroups $G_2, G_3, \ldots, G_k$ is the identity. A similar argument shows that $G_i \cap G_j = \{1\}$ for $i \neq j$.

Next, we must show that it possible to write every $g \in G$ as a product $g_1 \cdots g_k$, where $g_i \in G_i$. Since the order of $g$ divides the order of $G$, we know that

$$|g| = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$$

for some integers $\beta_1, \ldots, \beta_k$. Letting $a_i = |g|/p_i^{\beta_i}$, the $a_i$'s are relatively prime; hence, there exist integers $b_1, \ldots, b_k$ such that $a_1 b_1 + \cdots + a_k b_k = 1$. Consequently,

$$g = g^{a_1 b_1 + \cdots + a_k b_k} = g^{a_1 b_1} \cdots g^{a_k b_k}.$$

Since

$$g^{(a_i b_i) p_i^{\beta_i}} = g^{b_i |g|} = e,$$

it follows that $g^{a_i b_i}$ must be in $G_i$. Let $g_i = g^{a_i b_i}$. Then $g = g_1 \cdots g_k \in G_1 G_2 \cdots G_k$. Therefore, $G = G_1 G_2 \cdots G_k$ is an internal direct product of subgroups. ■

If remains for us to determine the possible structure of each $p_i$-group $G_i$ in Lemma 6.5.7.

**Lemma 6.5.8** *Let $G$ be a finite abelian $p$-group and suppose that $g \in G$ has maximal order. Then $G$ is isomorphic to $\langle g \rangle \times H$ for some subgroup $H$ of $G$.*

*Proof.* By Lemma 6.5.6, we may assume that the order of $G$ is $p^n$. We shall induct on $n$. If $n = 1$, then $G$ is cyclic of order $p$ and must be generated by $g$. Suppose now that the statement of the lemma holds for all integers $k$ with

$1 \leq k < n$ and let $g$ be of maximal order in $G$, say $|g| = p^m$. Then $a^{p^m} = e$ for all $a \in G$. Now choose $h$ in $G$ such that $h \notin \langle g \rangle$, where $h$ has the smallest possible order. Certainly such an $h$ exists; otherwise, $G = \langle g \rangle$ and we are done. Let $H = \langle h \rangle$.

We claim that $\langle g \rangle \cap H = \{e\}$. It suffices to show that $|H| = p$. Since $|h^p| = |h|/p$, the order of $h^p$ is smaller than the order of $h$ and must be in $\langle g \rangle$ by the minimality of $h$; that is, $h^p = g^r$ for some number $r$. Hence,

$$(g^r)^{p^{m-1}} = (h^p)^{p^{m-1}} = h^{p^m} = e,$$

and the order of $g^r$ must be less than or equal to $p^{m-1}$. Therefore, $g^r$ cannot generate $\langle g \rangle$. Notice that $p$ must occur as a factor of $r$, say $r = ps$, and $h^p = g^r = g^{ps}$. Define $a$ to be $g^{-s}h$. Then $a$ cannot be in $\langle g \rangle$; otherwise, $h$ would also have to be in $\langle g \rangle$. Also,

$$a^p = g^{-sp}h^p = g^{-r}h^p = h^{-p}h^p = e.$$

We have now formed an element $a$ with order $p$ such that $a \notin \langle g \rangle$. Since $h$ was chosen to have the smallest order of all of the elements that are not in $\langle g \rangle$, $|H| = p$.

Now we will show that the order of $gH$ in the factor group $G/H$ must be the same as the order of $g$ in $G$. If $|gH| < |g| = p^m$, then

$$H = (gH)^{p^{m-1}} = g^{p^{m-1}}H;$$

hence, $g^{p^{m-1}}$ must be in $\langle g \rangle \cap H = \{e\}$, which contradicts the fact that the order of $g$ is $p^m$. Therefore, $gH$ must have maximal order in $G/H$. By the Correspondence Theorem and our induction hypothesis,

$$G/H \cong \langle gH \rangle \times K/H$$

for some subgroup $K$ of $G$ containing $H$. We claim that $\langle g \rangle \cap K = \{e\}$. If $b \in \langle g \rangle \cap K$, then $bH \in \langle gH \rangle \cap K/H = \{H\}$ and $b \in \langle g \rangle \cap H = \{e\}$. It follows that $G = \langle g \rangle K$ implies that $G \cong \langle g \rangle \times K$. ■

The proof of the Fundamental Theorem of Finite Abelian Groups follows very quickly from Lemma 6.5.8. Suppose that $G$ is a finite abelian group and let $g$ be an element of maximal order in $G$. If $\langle g \rangle = G$, then we are done; otherwise, $G \cong \mathbb{Z}_{|g|} \times H$ for some subgroup $H$ contained in $G$ by the lemma. Since $|H| < |G|$, we can apply mathematical induction.

We finish by stating the more general theorem for all finitely generated abelian groups.

**Theorem 6.5.9 The Fundamental Theorem of Finitely Generated Abelian Groups.** *Every finitely generated abelian group $G$ is isomorphic to a direct product of cyclic groups of the form*

$$\mathbb{Z}_{p_1^{\alpha_1}} \times \mathbb{Z}_{p_2^{\alpha_2}} \times \cdots \times \mathbb{Z}_{p_n^{\alpha_n}} \times \mathbb{Z} \times \cdots \times \mathbb{Z},$$

*where the $p_i$'s are primes (not necessarily distinct).*

## 6.6 Core Exercises

**1.** Prove Theorem 6.1.1.

**2.** Consider Exercise 4.5.4 again. Compute the Cayley tables of the factor groups of the normal subgroups. Identify which Cayley tables 'look like' groups, you have seen before (i.e. check the First Isomorphism Theorem on them).

**3.** Let $T$ be the group of nonsingular upper triangular $2 \times 2$ matrices with entries in $\mathbb{R}$; that is, matrices of the form

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix},$$

where $a$, $b$, $c \in \mathbb{R}$ and $ac \neq 0$. Let $U$ consist of matrices of the form

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},$$

where $x \in \mathbb{R}$.

(a) Show that $U$ is a subgroup of $T$.

(b) Prove that $U$ is abelian.

(c) Prove that $U$ is normal in $T$.

(d) Show that $T/U$ is abelian.

(e) Is $T$ normal in $GL_2(\mathbb{R})$?

**4.** Let $G$ be a finite group and $N$ a normal subgroup of $G$. If $H$ is a subgroup of $G/N$, prove that $\phi^{-1}(H)$ is a subgroup in $G$ of order $|H| \cdot |N|$, where $\phi : G \to G/N$ is the canonical homomorphism.

**5.** Given a homomorphism $\phi : G \to H$ define a relation $\sim$ on $G$ by $a \sim b$ if $\phi(a) = \phi(b)$ for $a, b \in G$. Show this relation is an equivalence relation and describe the equivalence classes.

**6.**

(a) Look at Exercise 1.4.7 again. If you haven't solved it yet, solve it now. Note that this was already an instance of the word problem.

(b) Consider $\langle a, b \mid a^2, b^3, (ab)^3 \rangle$. Is *bbbabaaabaaab* equivalent to the empty word?

(c) Consider the group homomorphism from $\langle a, b \mid a^2, b^3, (ab)^5 \rangle$ to $S_5$ which maps $a \mapsto (1\,2)(3\,4)$, $b \mapsto (1\,3\,5)$. Check that $(ab)^5$ equals the identity so that it actually gives rise to a group homomorphism. What is its kernel?

(d) Consider $\langle a, b, c \mid a^2, b^2, c^2, (ac)^2, (ab)^3, (bc)^3 \rangle$. Do you find a homomorphism from this group to $S_4$? Is *acbabcbabbbaccaaabcccaba* equivalent to the empty word?

**7.** Determine all pairs of integral solutions of the equation

$$x^2 - 5y^2 = 1.$$

**8.** Revisit those exercises from the previous chapters that you struggled with. Try them again with the new knowledge you have gained.

**9.** **Programming Problem.** Write a program to solve Exercise 6.6.2.

## 6.7 Additional Exercises

**1.** If a group $G$ has exactly one subgroup $H$ of order $k$, prove that $H$ is normal in $G$.

**2.** Let $G_1$ and $G_2$ be groups, and let $H_1$ and $H_2$ be normal subgroups of $G_1$ and $G_2$ respectively. Let $\phi : G_1 \to G_2$ be a homomorphism. Show that $\phi$ induces a homomorphism $\overline{\phi} : (G_1/H_1) \to (G_2/H_2)$ if $\phi(H_1) \subset H_2$.

**3.** If $H$ and $K$ are normal subgroups of $G$ and $H \cap K = \{e\}$, prove that $G$ is isomorphic to a subgroup of $G/H \times G/K$.

**4.** Let $\phi : G_1 \to G_2$ be a surjective group homomorphism. Let $H_1$ be a normal subgroup of $G_1$ and suppose that $\phi(H_1) = H_2$. Prove or disprove that $G_1/H_1 \cong G_2/H_2$.

**5.** Show that the intersection of two normal subgroups is a normal subgroup.

**6.** If $G$ is abelian, prove that $G/H$ must also be abelian.

**7.** Prove or disprove: If $H$ is a normal subgroup of $G$ such that $H$ and $G/H$ are abelian, then $G$ is abelian.

**8.** If $G$ is cyclic, prove that $G/H$ must also be cyclic.

**9.** Prove or disprove: If $H$ and $G/H$ are cyclic, then $G$ is cyclic.

**10.** Prove or disprove: $\mathbb{Q}/\mathbb{Z} \cong \mathbb{Q}$.

**11.** Let $\mathrm{Aut}(G)$ be the set of all automorphisms of $G$; that is, isomorphisms from $G$ to itself. Prove this set forms a group and is a subgroup of the group of permutations of $G$; that is, $\mathrm{Aut}(G) \leq S_G$.

**12.** An **inner automorphism** of $G$,

$$i_g : G \to G,$$

is defined by the map

$$i_g(x) = gxg^{-1},$$

for $g \in G$. Show that $i_g \in \mathrm{Aut}(G)$.

**13.** The set of all inner automorphisms is denoted by $\mathrm{Inn}(G)$. Show that $\mathrm{Inn}(G)$ is a subgroup of $\mathrm{Aut}(G)$.

**14.** Find an automorphism of a group $G$ that is not an inner automorphism.

## 6.8 Material

1. Normal Subgroups and Factor Groups[1]
2. Beyond the scope of the course: Simple Groups[2]

About the word problem The word problem and the isomorphism problem for groups (John Stillwell)[3]

## 6.9 Hints to Selected Exercises

### 6.7 · Additional Exercises

[1] www.socratica.com/lesson/normal-subgroups-and-factor-groups
[2] www.socratica.com/lesson/simple-groups
[3] www.ams.org/journals/bull/1982-06-01/S0273-0979-1982-14963-1/S0273-0979-1982-14963-1.pdf

**6.7.4.** Find a counterexample.

**6.7.8.** If $a \in G$ is a generator for $G$, then $aH$ is a generator for $G/H$.

# Chapter 7

# Rings

Basic learning goals

1. Terminology and classes of rings.
2. Examples and basic properties of rings.
3. Ring homomorphisms and ideals.
4. Basic properties of ring homomorphisms.
5. Factorrings, in particular for prime and maximal ideals.

Up to this point we have studied sets with a single binary operation satisfying certain axioms, but we are often more interested in working with sets that have two binary operations. We have seen this with the operations of addition and multiplication for the integers and for polynomials in Section 6.4. The two operations are related to one another by the distributive property. If we consider a set with two such related binary operations satisfying certain axioms, we have an algebraic structure called a ring. Prominent examples are real numbers, complex numbers, matrices, and functions.

## 7.1 Rings

A nonempty set $R$ is a **ring** if it has two closed binary operations, addition and multiplication, satisfying the following conditions.

1. $a + b = b + a$ for $a, b \in R$.
2. $(a + b) + c = a + (b + c)$ for $a, b, c \in R$.
3. There is an element 0 in $R$ such that $a + 0 = a$ for all $a \in R$.
4. For every element $a \in R$, there exists an element $-a$ in $R$ such that $a + (-a) = 0$.
5. $(ab)c = a(bc)$ for $a, b, c \in R$.
6. For $a, b, c \in R$,

$$a(b + c) = ab + ac$$
$$(a + b)c = ac + bc.$$

This last condition, the distributive axiom, relates the binary operations of addition and multiplication. Notice that the first four axioms simply require that a ring be an abelian group under addition, so we could also have defined a ring to be an abelian group $(R, +)$ together with a second binary operation satisfying the fifth and sixth conditions given above.

If there is an element $1 \in R$ such that $1 \neq 0$ and $1a = a1 = a$ for each element $a \in R$, we say that $R$ is a ring with **unity** or **identity**. A ring $R$ for which $ab = ba$ for all $a, b$ in $R$ is called a **commutative ring**. A commutative ring $R$ with identity is called an **integral domain** if, for every $a, b \in R$ such that $ab = 0$, $a = 0$ or $b = 0$. A **division ring** is a ring $R$, with an identity, in which every nonzero element in $R$ is a **unit**; that is, for each $a \in R$ with $a \neq 0$, there exists a unique element $a^{-1}$ such that $a^{-1}a = aa^{-1} = 1$. A commutative division ring is called a **field**. The relationship among rings, integral domains, division rings, and fields is shown in Figure 7.1.1.

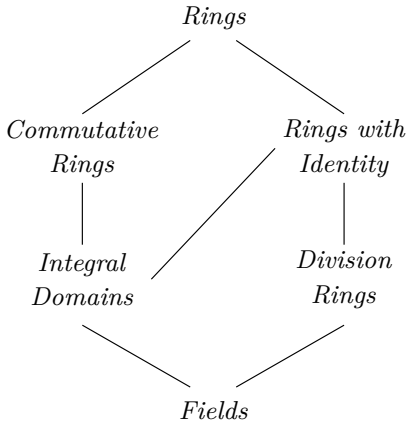



**Figure 7.1.1** Types of rings

**Example 7.1.2** As we have mentioned previously, the integers form a ring. In fact, $\mathbb{Z}$ is an integral domain. Certainly if $ab = 0$ for two integers $a$ and $b$, $a = 0$ or $b = 0$. However, $\mathbb{Z}$ is not a field. There is no integer that is the multiplicative inverse of 2, since $1/2$ is not an integer. The only integers with multiplicative inverses are 1 and $-1$. □

**Example 7.1.3** Under the ordinary operations of addition and multiplication, all of the familiar number systems are rings: the rationals, $\mathbb{Q}$; the real numbers, $\mathbb{R}$; and the complex numbers, $\mathbb{C}$. Each of these rings is a field. □

**Example 7.1.4** We can define the product of two elements $a$ and $b$ in $\mathbb{Z}_n$ by $ab \pmod{n}$. For instance, in $\mathbb{Z}_{12}$, $5 \cdot 7 \equiv 11 \pmod{12}$. This product makes the abelian group $\mathbb{Z}_n$ into a ring. Certainly $\mathbb{Z}_n$ is a commutative ring; however, it may fail to be an integral domain. If we consider $3 \cdot 4 \equiv 0 \pmod{12}$ in $\mathbb{Z}_{12}$, it is easy to see that a product of two nonzero elements in the ring can be equal to zero. □

A nonzero element $a$ in a commutative ring $R$ is called a **zero divisor** if there is a nonzero element $b$ in $R$ such that $ab = 0$. In the previous example, 3 and 4 are zero divisors in $\mathbb{Z}_{12}$.

**Example 7.1.5** In calculus the continuous real-valued functions on an interval $[a, b]$ form a commutative ring. We add or multiply two functions by adding or multiplying the values of the functions. If $f(x) = x^2$ and $g(x) = \cos x$, then $(f + g)(x) = f(x) + g(x) = x^2 + \cos x$ and $(fg)(x) = f(x)g(x) = x^2 \cos x$. □

**Example 7.1.6** The $2\times 2$ matrices with entries in $\mathbb{R}$ form a ring under the usual operations of matrix addition and multiplication. This ring is noncommutative, since it is usually the case that $AB \neq BA$. Also, notice that we can have $AB = 0$ when $A$ nor $B$ is zero. These ideas also extend to square matrices of bigger size. □

**Example 7.1.7** For an example of a noncommutative division ring, let

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{i} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{j} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

where $i^2 = -1$. These elements satisfy the following relations:

$$\begin{aligned} \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 &= -1 \\ \mathbf{ij} &= \mathbf{k} \\ \mathbf{jk} &= \mathbf{i} \\ \mathbf{ki} &= \mathbf{j} \\ \mathbf{ji} &= -\mathbf{k} \\ \mathbf{kj} &= -\mathbf{i} \\ \mathbf{ik} &= -\mathbf{j}. \end{aligned}$$

Let $\mathbb{H}$ consist of elements of the form $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, where $a, b, c, d$ are real numbers. Equivalently, $\mathbb{H}$ can be considered to be the set of all $2 \times 2$ matrices of the form

$$\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix},$$

where $\alpha = a + di$ and $\beta = b + ci$ are complex numbers. We can define addition and multiplication on $\mathbb{H}$ either by the usual matrix operations or in terms of the generators 1, **i**, **j**, and **k**:

$$\begin{aligned} &(a_1 + b_1\mathbf{i} + c_1\mathbf{j} + d_1\mathbf{k}) + (a_2 + b_2\mathbf{i} + c_2\mathbf{j} + d_2\mathbf{k}) \\ &= (a_1 + a_2) + (b_1 + b_2)\mathbf{i} + (c_1 + c_2)\mathbf{j} + (d_1 + d_2)\mathbf{k} \end{aligned}$$

and

$$(a_1 + b_1\mathbf{i} + c_1\mathbf{j} + d_1\mathbf{k})(a_2 + b_2\mathbf{i} + c_2\mathbf{j} + d_2\mathbf{k}) = \alpha + \beta\mathbf{i} + \gamma\mathbf{j} + \delta\mathbf{k},$$

where

$$\begin{aligned} \alpha &= a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2 \\ \beta &= a_1b_2 + a_2b_1 + c_1d_2 - d_1c_2 \\ \gamma &= a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2 \\ \delta &= a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2. \end{aligned}$$

Though multiplication looks complicated, it is actually a straightforward computation if we remember that we just add and multiply elements in $\mathbb{H}$ like polynomials and keep in mind the relationships between the generators **i**, **j**, and **k**. The ring $\mathbb{H}$ is called the ring of **quaternions**.

To show that the quaternions are a division ring, we must be able to find an inverse for each nonzero element. Notice that

$$(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})(a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}) = a^2 + b^2 + c^2 + d^2.$$

This element can be zero only if $a$, $b$, $c$, and $d$ are all zero. So if $a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k} \neq 0$,

$$(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \left( \frac{a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}}{a^2 + b^2 + c^2 + d^2} \right) = 1.$$

□

**Proposition 7.1.8** *Let $R$ be a ring with $a, b \in R$. Then*

1. $a0 = 0a = 0$;
2. $a(-b) = (-a)b = -ab$;
3. $(-a)(-b) = ab$.

*Proof.* To prove (1), observe that

$$a0 = a(0 + 0) = a0 + a0;$$

hence, $a0 = 0$. Similarly, $0a = 0$. For (2), we have $ab + a(-b) = a(b - b) = a0 = 0$; consequently, $-ab = a(-b)$. Similarly, $-ab = (-a)b$. Part (3) follows directly from (2) since $(-a)(-b) = -(a(-b)) = -(-ab) = ab$. ■

Just as we have subgroups of groups, we have an analogous class of substructures for rings. A **subring** $S$ of a ring $R$ is a subset $S$ of $R$ such that $S$ is also a ring under the inherited operations from $R$.

**Example 7.1.9** The ring $n\mathbb{Z}$ is a subring of $\mathbb{Z}$. Notice that even though the original ring may have an identity, we do not require that its subring have an identity. We have the following chain of subrings:

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

□

The following proposition gives us some easy criteria for determining whether or not a subset of a ring is indeed a subring. The proof is omitted as it is very similar to the proof of Proposition 2.2.8.

**Proposition 7.1.10** *Let $R$ be a ring and $S$ a subset of $R$. Then $S$ is a subring of $R$ if and only if the following conditions are satisfied.*

1. $S \neq \emptyset$.
2. $rs \in S$ *for all* $r, s \in S$.
3. $r - s \in S$ *for all* $r, s \in S$.

**Example 7.1.11** Let $R = \mathbb{M}_2(\mathbb{R})$ be the ring of $2 \times 2$ matrices with entries in $\mathbb{R}$. If $T$ is the set of upper triangular matrices in $R$; i.e.,

$$T = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : a, b, c \in \mathbb{R} \right\},$$

then $T$ is a subring of $R$. If

$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix}$$

are in $T$, then clearly $A - B$ is also in $T$. Also,

$$AB = \begin{pmatrix} aa' & ab' + bc' \\ 0 & cc' \end{pmatrix}$$

is in $T$. □

## 7.2 Integral Domains and Fields

Let us briefly recall some definitions. If $R$ is a commutative ring and $r$ is a nonzero element in $R$, then $r$ is said to be a **zero divisor** if there is some nonzero element $s \in R$ such that $rs = 0$. A commutative ring with identity is said to be an **integral domain** if it has no zero divisors. If an element $a$ in a ring $R$ with identity has a multiplicative inverse, we say that $a$ is a **unit**. If every nonzero element in a ring $R$ is a unit, then $R$ is called a **division ring**. A commutative division ring is called a **field**.

**Example 7.2.1** If $i^2 = -1$, then the set $\mathbb{Z}[i] = \{m+ni : m, n \in \mathbb{Z}\}$ forms a ring known as the **Gaussian integers**. It is easily seen that the Gaussian integers are a subring of the complex numbers since they are closed under addition and multiplication. Let $\alpha = a + bi$ be a unit in $\mathbb{Z}[i]$. Then $\overline{\alpha} = a - bi$ is also a unit since if $\alpha\beta = 1$, then $\overline{\alpha}\overline{\beta} = 1$. If $\beta = c + di$, then

$$1 = \alpha\beta\overline{\alpha}\overline{\beta} = (a^2 + b^2)(c^2 + d^2).$$

Therefore, $a^2 + b^2$ must either be 1 or $-1$; or, equivalently, $a + bi = \pm 1$ or $a + bi = \pm i$. Therefore, units of this ring are $\pm 1$ and $\pm i$; hence, the Gaussian integers are not a field. We will leave it as an exercise to prove that the Gaussian integers are an integral domain. □

**Example 7.2.2** The set of matrices

$$F = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

with entries in $\mathbb{Z}_2$ forms a field. □

**Example 7.2.3** The set $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is a field. The inverse of an element $a + b\sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$ is

$$\frac{a}{a^2 - 2b^2} + \frac{-b}{a^2 - 2b^2}\sqrt{2}.$$

□

We have the following alternative characterization of integral domains.

**Proposition 7.2.4 Cancellation Law.** *Let $D$ be a commutative ring with identity. Then $D$ is an integral domain if and only if for all nonzero elements $a \in D$ with $ab = ac$, we have $b = c$.*

*Proof.* Let $D$ be an integral domain. Then $D$ has no zero divisors. Let $ab = ac$ with $a \neq 0$. Then $a(b - c) = 0$. Hence, $b - c = 0$ and $b = c$.

Conversely, let us suppose that cancellation is possible in $D$. That is, suppose that $ab = ac$ implies $b = c$. Let $ab = 0$. If $a \neq 0$, then $ab = a0$ or $b = 0$. Therefore, $a$ cannot be a zero divisor. ■

For any nonnegative integer $n$ and any element $r$ in a ring $R$ we write $r + \cdots + r$ ($n$ times) as $nr$. We define the **characteristic** of a ring $R$ to be the least positive integer $n$ such that $nr = 0$ for all $r \in R$. If no such integer exists, then the characteristic of $R$ is defined to be 0. We will denote the characteristic of $R$ by $\operatorname{char} R$.

**Example 7.2.5** For every prime $p$, the congruence classes $\mathbb{Z}_p$ form a field of characteristic $p$. By Proposition 1.2.4, every nonzero element in $\mathbb{Z}_p$ has an inverse; hence, $\mathbb{Z}_p$ is a field. If $a$ is any nonzero element in the field, then $pa = 0$, since the order of any nonzero element in the abelian group $\mathbb{Z}_p$ is $p$.

□

**Lemma 7.2.6** *Let $R$ be a ring with identity. If 1 has order $n$, then the characteristic of $R$ is $n$.*

*Proof.* If 1 has order $n$, then $n$ is the least positive integer such that $n1 = 0$. Thus, for all $r \in R$,

$$nr = n(1r) = (n1)r = 0r = 0.$$

On the other hand, if no positive $n$ exists such that $n1 = 0$, then the characteristic of $R$ is zero. ■

**Theorem 7.2.7** *The characteristic of an integral domain is either prime or zero.*

*Proof.* Let $D$ be an integral domain and suppose that the characteristic of $D$ is $n$ with $n \neq 0$. If $n$ is not prime, then $n = ab$, where $1 < a < n$ and $1 < b < n$. By Lemma 7.2.6, we need only consider the case $n1 = 0$. Since $0 = n1 = (ab)1 = (a1)(b1)$ and there are no zero divisors in $D$, either $a1 = 0$ or $b1 = 0$. Hence, the characteristic of $D$ must be less than $n$, which is a contradiction. Therefore, $n$ must be prime. ■

## 7.3 Ring Homomorphisms and Ideals

In the study of groups, a homomorphism is a map that preserves the operation of the group. Similarly, a homomorphism between rings preserves the operations of addition and multiplication in the ring. More specifically, if $R$ and $S$ are rings, then a **ring homomorphism** is a map $\phi : R \to S$ satisfying

$$\begin{aligned} \phi(a+b) &= \phi(a) + \phi(b) \\ \phi(ab) &= \phi(a)\phi(b) \end{aligned}$$

for all $a, b \in R$. If $\phi : R \to S$ is a one-to-one and onto homomorphism, then $\phi$ is called an **isomorphism** of rings.

The set of elements that a ring homomorphism maps to 0 plays a fundamental role in the theory of rings. For any ring homomorphism $\phi : R \to S$, we define the **kernel** of a ring homomorphism to be the set

$$\ker \phi = \{r \in R : \phi(r) = 0\}.$$

**Example 7.3.1** For any integer $n$ we can define a ring homomorphism $\phi : \mathbb{Z} \to \mathbb{Z}_n$ by $a \mapsto a \pmod n$. This is indeed a ring homomorphism, since

$$\begin{aligned} \phi(a+b) &= (a+b) \pmod n \\ &= a \pmod n + b \pmod n \\ &= \phi(a) + \phi(b) \end{aligned}$$

and

$$\begin{aligned} \phi(ab) &= ab \pmod n \\ &= a \pmod n \cdot b \pmod n \\ &= \phi(a)\phi(b). \end{aligned}$$

The kernel of the homomorphism $\phi$ is $n\mathbb{Z}$. □

**Example 7.3.2** Let $C[a, b]$ be the ring of continuous real-valued functions on an interval $[a, b]$ as in Example 7.1.5. For a fixed $\alpha \in [a, b]$, we can define a ring homomorphism $\phi_\alpha : C[a, b] \to \mathbb{R}$ by $\phi_\alpha(f) = f(\alpha)$. This is a ring

homomorphism since

$$\phi_\alpha(f+g) = (f+g)(\alpha) = f(\alpha) + g(\alpha) = \phi_\alpha(f) + \phi_\alpha(g)$$
$$\phi_\alpha(fg) = (fg)(\alpha) = f(\alpha)g(\alpha) = \phi_\alpha(f)\phi_\alpha(g).$$

Ring homomorphisms of the type $\phi_\alpha$ are called **evaluation homomorphisms**. □

In the next proposition we will examine some fundamental properties of ring homomorphisms. The proof of the proposition is left as an exercise.

**Proposition 7.3.3** *Let $\phi : R \to S$ be a ring homomorphism.*

1. *If $R$ is a commutative ring, then $\phi(R)$ is a commutative ring.*
2. *$\phi(0) = 0$.*
3. *Let $1_R$ and $1_S$ be the identities for $R$ and $S$, respectively. If $\phi$ is onto, then $\phi(1_R) = 1_S$.*
4. *If $R$ is a field and $\phi(R) \neq \{0\}$, then $\phi(R)$ is a field.*

In group theory we found that normal subgroups play a special role. These subgroups have nice characteristics that make them more interesting to study than arbitrary subgroups. In ring theory the objects corresponding to normal subgroups are a special class of subrings called ideals. An **ideal** in a ring $R$ is a subring $I$ of $R$ such that if $a$ is in $I$ and $r$ is in $R$, then both $ar$ and $ra$ are in $I$; that is, $rI \subset I$ and $Ir \subset I$ for all $r \in R$.

**Example 7.3.4** Every ring $R$ has at least two ideals, $\{0\}$ and $R$. These ideals are called the **trivial ideals**. □

Let $R$ be a ring with identity and suppose that $I$ is an ideal in $R$ such that 1 is in $I$. Since for any $r \in R$, $r1 = r \in I$ by the definition of an ideal, $I = R$.

**Example 7.3.5** If $a$ is any element in a commutative ring $R$ with identity, then the set

$$\langle a \rangle = \{ar : r \in R\}$$

is an ideal in $R$. Certainly, $\langle a \rangle$ is nonempty since both $0 = a0$ and $a = a1$ are in $\langle a \rangle$. The sum of two elements in $\langle a \rangle$ is again in $\langle a \rangle$ since $ar + ar' = a(r + r')$. The inverse of $ar$ is $-ar = a(-r) \in \langle a \rangle$. Finally, if we multiply an element $ar \in \langle a \rangle$ by an arbitrary element $s \in R$, we have $s(ar) = a(sr)$. Therefore, $\langle a \rangle$ satisfies the definition of an ideal. □

If $R$ is a commutative ring with identity, then an ideal of the form $\langle a \rangle = \{ar : r \in R\}$ is called a **principal ideal**.

**Theorem 7.3.6** *Every ideal in the ring of integers $\mathbb{Z}$ is a principal ideal.*

**Example 7.3.7** The set $n\mathbb{Z}$ is ideal in the ring of integers. If $na$ is in $n\mathbb{Z}$ and $b$ is in $\mathbb{Z}$, then $nab$ is in $n\mathbb{Z}$ as required. In fact, by Theorem 7.3.6, these are the only ideals of $\mathbb{Z}$. □

**Proposition 7.3.8** *The kernel of any ring homomorphism $\phi : R \to S$ is an ideal in $R$.*

*Proof.* We know from group theory that $\ker \phi$ is an additive subgroup of $R$. Suppose that $r \in R$ and $a \in \ker \phi$. Then we must show that $ar$ and $ra$ are in $\ker \phi$. However,

$$\phi(ar) = \phi(a)\phi(r) = 0\phi(r) = 0$$

and

$$\phi(ra) = \phi(r)\phi(a) = \phi(r)0 = 0.$$

■

**Theorem 7.3.9** *Let $I$ be an ideal of $R$. The factor group $R/I$ is a ring with multiplication defined by*

$$(r + I)(s + I) = rs + I.$$

*Proof.* We already know that $R/I$ is an abelian group under addition. Let $r+I$ and $s + I$ be in $R/I$. We must show that the product $(r + I)(s + I) = rs + I$ is independent of the choice of coset; that is, if $r' \in r + I$ and $s' \in s + I$, then $r's'$ must be in $rs + I$. Since $r' \in r + I$, there exists an element $a$ in $I$ such that $r' = r + a$. Similarly, there exists a $b \in I$ such that $s' = s + b$. Notice that

$$r's' = (r + a)(s + b) = rs + as + rb + ab$$

and $as + rb + ab \in I$ since $I$ is an ideal; consequently, $r's' \in rs + I$. We will leave as an exercise the verification of the associative law for multiplication and the distributive laws. ■

The ring $R/I$ in Theorem 7.3.9 is called the **factor** or **quotient ring**. Just as with group homomorphisms and normal subgroups, there is a relationship between ring homomorphisms and ideals.

**Theorem 7.3.10** *Let $I$ be an ideal of $R$. The map $\phi : R \to R/I$ defined by $\phi(r) = r + I$ is a ring homomorphism of $R$ onto $R/I$ with kernel $I$.*

*Proof.* Certainly $\phi : R \to R/I$ is a surjective abelian group homomorphism. It remains to show that $\phi$ works correctly under ring multiplication. Let $r$ and $s$ be in $R$. Then

$$\phi(r)\phi(s) = (r + I)(s + I) = rs + I = \phi(rs),$$

which completes the proof of the theorem. ■

The map $\phi : R \to R/I$ is often called the **natural** or **canonical homomorphism**. In ring theory we have isomorphism theorems relating ideals and ring homomorphisms similar to the isomorphism theorems for groups that relate normal subgroups and homomorphisms in Chapter 4. We will prove only the First Isomorphism Theorem for rings in this chapter and leave the proofs of the other two theorems as exercises. All of the proofs are similar to the proofs of the isomorphism theorems for groups.

**Theorem 7.3.11 First Isomorphism Theorem.** *Let $\psi : R \to S$ be a ring homomorphism. Then $\ker \psi$ is an ideal of $R$. If $\phi : R \to R/\ker \psi$ is the canonical homomorphism, then there exists a unique isomorphism $\eta : R/\ker \psi \to \psi(R)$ such that $\psi = \eta\phi$.*

*Proof.* Let $K = \ker \psi$. By the First Isomorphism Theorem for groups, there exists a well-defined group homomorphism $\eta : R/K \to \psi(R)$ defined by $\eta(r + K) = \psi(r)$ for the additive abelian groups $R$ and $R/K$. To show that this is a ring homomorphism, we need only show that $\eta((r + K)(s + K)) = \eta(r + K)\eta(s + K)$; but

$$\begin{aligned}\eta((r + K)(s + K)) &= \eta(rs + K) \\ &= \psi(rs) \\ &= \psi(r)\psi(s) \\ &= \eta(r + K)\eta(s + K).\end{aligned}$$

■

## 7.4 Maximal and Prime Ideals

In this particular section we are especially interested in certain ideals of commutative rings. These ideals give us special types of factor rings. More specifically, we would like to characterize those ideals $I$ of a commutative ring $R$ such that $R/I$ is an integral domain or a field.

A proper ideal $M$ of a ring $R$ is a **maximal ideal** of $R$ if the ideal $M$ is not a proper subset of any ideal of $R$ except $R$ itself. That is, $M$ is a maximal ideal if for any ideal $I$ properly containing $M$, $I = R$. The following theorem completely characterizes maximal ideals for commutative rings with identity in terms of their corresponding factor rings.

**Theorem 7.4.1** *Let $R$ be a commutative ring with identity and $M$ an ideal in $R$. Then $M$ is a maximal ideal of $R$ if and only if $R/M$ is a field.*

*Proof.* Let $M$ be a maximal ideal in $R$. If $R$ is a commutative ring, then $R/M$ must also be a commutative ring. Clearly, $1 + M$ acts as an identity for $R/M$. We must also show that every nonzero element in $R/M$ has an inverse. If $a + M$ is a nonzero element in $R/M$, then $a \notin M$. Define $I$ to be the set $\{ra + m : r \in R \text{ and } m \in M\}$. We will show that $I$ is an ideal in $R$. The set $I$ is nonempty since $0a + 0 = 0$ is in $I$. If $r_1a + m_1$ and $r_2a + m_2$ are two elements in $I$, then

$$(r_1a + m_1) - (r_2a + m_2) = (r_1 - r_2)a + (m_1 - m_2)$$

is in $I$. Also, for any $r \in R$ it is true that $rI \subset I$; hence, $I$ is closed under multiplication and satisfies the necessary conditions to be an ideal. Therefore, by Proposition 7.1.10 and the definition of an ideal, $I$ is an ideal properly containing $M$. Since $M$ is a maximal ideal, $I = R$; consequently, by the definition of $I$ there must be an $m$ in $M$ and an element $b$ in $R$ such that $1 = ab + m$. Therefore,

$$1 + M = ab + M = ba + M = (a + M)(b + M).$$

Conversely, suppose that $M$ is an ideal and $R/M$ is a field. Since $R/M$ is a field, it must contain at least two elements: $0 + M = M$ and $1 + M$. Hence, $M$ is a proper ideal of $R$. Let $I$ be any ideal properly containing $M$. We need to show that $I = R$. Choose $a$ in $I$ but not in $M$. Since $a + M$ is a nonzero element in a field, there exists an element $b + M$ in $R/M$ such that $(a + M)(b + M) = ab + M = 1 + M$. Consequently, there exists an element $m \in M$ such that $ab + m = 1$ and 1 is in $I$. Therefore, $r1 = r \in I$ for all $r \in R$. Consequently, $I = R$. ■

**Example 7.4.2** Let $p\mathbb{Z}$ be an ideal in $\mathbb{Z}$, where $p$ is prime. Then $p\mathbb{Z}$ is a maximal ideal since $\mathbb{Z}/p\mathbb{Z} \cong \mathbb{Z}_p$ is a field. □

A proper ideal $P$ in a commutative ring $R$ is called a **prime ideal** if whenever $ab \in P$, then $a \in P$ or $b \in P$.[1]

**Example 7.4.3** It is easy to check that the set $P = \{0, 2, 4, 6, 8, 10\}$ is an ideal in $\mathbb{Z}_{12}$. This ideal is prime. In fact, it is a maximal ideal. □

Compare the definition of a prime ideal with Lemma 6.4.1. Let $p \in \mathbb{Z}$ be a prime number and let $I = p\mathbb{Z}$ be the ideal generated by $p$. The condition $p \mid ab$ is equivalent with $ab \in p\mathbb{Z}$. Now, Euclid's Lemma just shows that $I$ is indeed a prime ideal.

[1] It is possible to define prime ideals in a noncommutative ring. See [1] or [3].

**Proposition 7.4.4** *Let $R$ be a commutative ring with identity 1, where $1 \neq 0$. Then $P$ is a prime ideal in $R$ if and only if $R/P$ is an integral domain.*

*Proof.* First let us assume that $P$ is an ideal in $R$ and $R/P$ is an integral domain. Suppose that $ab \in P$. If $a + P$ and $b + P$ are two elements of $R/P$ such that $(a + P)(b + P) = 0 + P = P$, then $a + P = P$ or $b + P = P$. This means that $a$ is in $P$ or $b$ is in $P$, which shows that $P$ must be prime.

Conversely, suppose that $P$ is prime and

$$(a + P)(b + P) = ab + P = 0 + P = P.$$

Then $ab \in P$. If $a \notin P$, then $b$ must be in $P$ by the definition of a prime ideal; hence, $b + P = 0 + P$ and $R/P$ is an integral domain. ■

**Example 7.4.5** Every ideal in $\mathbb{Z}$ is of the form $n\mathbb{Z}$. The factor ring $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$ is an integral domain only when $n$ is prime. It is actually a field. Hence, the nonzero prime ideals in $\mathbb{Z}$ are the ideals $p\mathbb{Z}$, where $p$ is prime. This example really justifies the use of the word "prime" in our definition of prime ideals. □

Since every field is an integral domain, we have the following corollary.

**Corollary 7.4.6** *Every maximal ideal in a commutative ring with identity is also a prime ideal.*

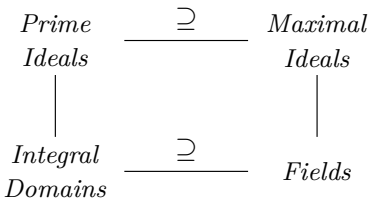



**Figure 7.4.7** Special ideals in commutative rings with identity and their factor rings

### 7.4.1 Historical Note

Amalie Emmy Noether, one of the outstanding mathematicians of the twentieth century, was born in Erlangen, Germany in 1882. She was the daughter of Max Noether (1844–1921), a distinguished mathematician at the University of Erlangen. Together with Paul Gordon (1837–1912), Emmy Noether's father strongly influenced her early education. She entered the University of Erlangen at the age of 18. Although women had been admitted to universities in England, France, and Italy for decades, there was great resistance to their presence at universities in Germany. Noether was one of only two women among the university's 986 students. After completing her doctorate under Gordon in 1907, she continued to do research at Erlangen, occasionally lecturing when her father was ill.

Noether went to Göttingen to study in 1916. David Hilbert and Felix Klein tried unsuccessfully to secure her an appointment at Göttingen. Some of the faculty objected to women lecturers, saying, "What will our soldiers think when they return to the university and are expected to learn at the feet of a woman?" Hilbert, annoyed at the question, responded, "Meine Herren, I do not see that the sex of a candidate is an argument against her admission as a Privatdozent. After all, the Senate is not a bathhouse." At the end of World War I, attitudes changed and conditions greatly improved for women. After Noether passed her habilitation examination in 1919, she was given a title and was paid a small sum for her lectures.

In 1922, Noether became a Privatdozent at Göttingen. Over the next 11 years she used axiomatic methods to develop an abstract theory of rings and ideals. Though she was not good at lecturing, Noether was an inspiring teacher. One of her many students was B. L. van der Waerden, author of the first text treating abstract algebra from a modern point of view. Some of the other mathematicians Noether influenced or closely worked with were Alexandroff, Artin, Brauer, Courant, Hasse, Hopf, Pontryagin, von Neumann, and Weyl. One of the high points of her career was an invitation to address the International Congress of Mathematicians in Zurich in 1932. In spite of all the recognition she received from her colleagues, Noether's abilities were never recognized as they should have been during her lifetime. She was never promoted to full professor by the Prussian academic bureaucracy.

In 1933, Noether, who was Jewish, was banned from participation in all academic activities in Germany. She emigrated to the United States, took a position at Bryn Mawr College, and became a member of the Institute for Advanced Study at Princeton. Noether died suddenly on April 14, 1935. After her death she was eulogized by such notable scientists as Albert Einstein.

## 7.5 Additional insights

### 7.5.1 Theorem by Wedderburn

**Theorem 7.5.1** *Every finite integral domain is a field.*

*Proof.* Let $D$ be a finite integral domain and $D^*$ be the set of nonzero elements of $D$. We must show that every element in $D^*$ has an inverse. For each $a \in D^*$ we can define a map $\lambda_a : D^* \to D^*$ by $\lambda_a(d) = ad$. This map makes sense, because if $a \neq 0$ and $d \neq 0$, then $ad \neq 0$. The map $\lambda_a$ is one-to-one, since for $d_1, d_2 \in D^*$,

$$ad_1 = \lambda_a(d_1) = \lambda_a(d_2) = ad_2$$

implies $d_1 = d_2$ by left cancellation. Since $D^*$ is a finite set, the map $\lambda_a$ must also be onto; hence, for some $d \in D^*$, $\lambda_a(d) = ad = 1$. Therefore, $a$ has a left inverse. Since $D$ is commutative, $d$ must also be a right inverse for $a$. Consequently, $D$ is a field. ■

### 7.5.2 Two-sided vs. one-sided ideals

**Remark 7.5.2** In our definition of an ideal we have required that $rI \subset I$ and $Ir \subset I$ for all $r \in R$. Such ideals are sometimes referred to as **two-sided ideals**. We can also consider **one-sided ideals**; that is, we may require only that either $rI \subset I$ or $Ir \subset I$ for $r \in R$ hold but not both. Such ideals are called **left ideals** and **right ideals**, respectively. Of course, in a commutative ring any ideal must be two-sided. In this text we will concentrate on two-sided ideals.

### 7.5.3 More Isomorphism Theorems for Rings

**Theorem 7.5.3 Second Isomorphism Theorem.** *Let $I$ be a subring of a ring $R$ and $J$ an ideal of $R$. Then $I \cap J$ is an ideal of $I$ and*

$$I/I \cap J \cong (I + J)/J.$$

**Theorem 7.5.4 Third Isomorphism Theorem.** *Let $R$ be a ring and $I$*

*and $J$ be ideals of $R$ where $J \subset I$. Then*

$$R/I \cong \frac{R/J}{I/J}.$$

**Theorem 7.5.5 Correspondence Theorem.** *Let $I$ be an ideal of a ring $R$. Then $S \mapsto S/I$ is a one-to-one correspondence between the set of subrings $S$ containing $I$ and the set of subrings of $R/I$. Furthermore, the ideals of $R$ containing $I$ correspond to ideals of $R/I$.*

## 7.6 Core Exercises

**1.** Which of the following sets are rings with respect to the usual operations of addition and multiplication? If the set is a ring, is it also a field?

(a) $7\mathbb{Z}$

(b) $\mathbb{Z}_{18}$

(c) $\mathbb{Q}(\sqrt{2}\,) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$

(d) $\mathbb{Q}(\sqrt{2}, \sqrt{3}\,) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : a, b, c, d \in \mathbb{Q}\}$

(e) $\mathbb{Z}[\sqrt{3}\,] = \{a + b\sqrt{3} : a, b \in \mathbb{Z}\}$

(f) $R = \{a + b\sqrt[3]{3} : a, b \in \mathbb{Q}\}$

(g) $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z} \text{ and } i^2 = -1\}$

(h) $\mathbb{Q}(\sqrt[3]{3}\,) = \{a + b\sqrt[3]{3} + c\sqrt[3]{9} : a, b, c \in \mathbb{Q}\}$

**2.** Find all of the ideals in each of the following rings. Which of these ideals are maximal and which are prime?

(a) $\mathbb{Z}_{18}$

(b) $\mathbb{Z}_{25}$

(c) $\mathbb{M}_2(\mathbb{R})$, the $2 \times 2$ matrices with entries in $\mathbb{R}$

(d) $\mathbb{M}_2(\mathbb{Z})$, the $2 \times 2$ matrices with entries in $\mathbb{Z}$

(e) $\mathbb{Q}$

**3.** Let $\{0, 1, a, b\}$ be the ground set of a field with four elements. Below are the addition and multiplication table of the field. Fill the tables according to the axioms; there is a unique way in this case.

| $+$ | 0 | 1 | $a$ | $b$ |
|---|---|---|---|---|
| 0 | | | | |
| 1 | | | | |
| $a$ | | | | |
| $b$ | | | | |

**Figure 7.6.1** Addition table

| $\cdot$ | 1 | $a$ | $b$ |
|---|---|---|---|
| 1 | | | |
| $a$ | | | |
| $b$ | | | |

**Figure 7.6.2** Multiplication table

**4.** Define a map $\phi : \mathbb{C} \to \mathbb{M}_2(\mathbb{R})$ by

$$\phi(a + bi) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Show that $\phi$ is an isomorphism of $\mathbb{C}$ with its image in $\mathbb{M}_2(\mathbb{R})$.

**5.** Prove Proposition 7.3.3.

**6.** Prove Theorem 7.3.6

**7.** For each of the following rings $R$ with ideal $I$, give an addition table and a multiplication table for $R/I$.

(a) $R = \mathbb{Z}$ and $I = 6\mathbb{Z}$

(b) $R = \mathbb{Z}_{12}$ and $I = \{0, 3, 6, 9\}$

**8.** Find all homomorphisms $\phi : \mathbb{Z}/6\mathbb{Z} \to \mathbb{Z}/15\mathbb{Z}$.

**9.** Let $X$ be a finite set.

(a) Show that the power set $\mathcal{P}(X)$ forms a ring with symmetric difference as addition and intersection as multiplication.

(b) Does this ring have an identity? What is the characteristic of this ring?

(c) Show that $\mathcal{P}(X)$ with the given ring structure is isomorphic to the ring $\mathbb{Z}_2^{|X|}$ with component-wise addition and multiplication.

(d) Let $|X| = 3$. Determine the ideals of this ring.

**10.** Let $(H, +)$ be an abelian group and let $\text{End}(H)$ be the set of group homomorphisms from $H$ to itself.

(a) Prove that $\text{End}(H)$ forms a ring where the addition and multiplication of two elements $\phi, \psi \in \text{End}(H)$ is defined by $(\phi + \psi)(x) = \phi(x) + \psi(x)$ for all $x \in H$ and $(\phi \cdot \psi)(x) = \phi \circ \psi(x)$ for all $x \in H$, where $\circ$ denotes function composition.

(b) Give an example of an abelian group and the ring formed by the group homomorphisms on itself.

(c) ***Advanced*** Check that $H$ forms a **module** over $\text{End}(H)$; a module is a generalization of a vector space, where the field is replaced by a ring.

## 7.7 Additional Exercises

**1.** Let $R$ be the ring of $2 \times 2$ matrices of the form

$$\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix},$$

where $a, b \in \mathbb{R}$. Show that although $R$ is a ring that has no identity, we

can find a subring $S$ of $R$ with an identity.

**2.** List or characterize all of the units in each of the following rings.

(a) $\mathbb{Z}_{10}$

(b) $\mathbb{Z}_{12}$

(c) $\mathbb{Z}_7$

(d) $\mathbb{M}_2(\mathbb{Z})$, the $2 \times 2$ matrices with entries in $\mathbb{Z}$

(e) $\mathbb{M}_2(\mathbb{Z}_2)$, the $2 \times 2$ matrices with entries in $\mathbb{Z}_2$

**3.** Prove that $\mathbb{R}$ is not isomorphic to $\mathbb{C}$.

**4.** Prove or disprove: The ring $\mathbb{Q}(\sqrt{2}\,) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is isomorphic to the ring $\mathbb{Q}(\sqrt{3}\,) = \{a + b\sqrt{3} : a, b \in \mathbb{Q}\}$.

**5.** What is the characteristic of the field formed by the set of matrices

$$F = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

with entries in $\mathbb{Z}_2$?

**6.** Prove that the Gaussian integers, $\mathbb{Z}[i]$, are an integral domain.

**7.** Prove that $\mathbb{Z}[\sqrt{3}\,i] = \{a + b\sqrt{3}\,i : a, b \in \mathbb{Z}\}$ is an integral domain.

**8.** If $R$ is a field, show that the only two ideals of $R$ are $\{0\}$ and $R$ itself.

**9.** Let $a$ be any element in a ring $R$ with identity. Show that $(-1)a = -a$.

**10.** Prove that the associative law for multiplication and the distributive laws hold in $R/I$.

**11.** Prove the Second Isomorphism Theorem for rings: Let $I$ be a subring of a ring $R$ and $J$ an ideal in $R$. Then $I \cap J$ is an ideal in $I$ and

$$I/I \cap J \cong I + J/J.$$

**12.** Prove the Third Isomorphism Theorem for rings: Let $R$ be a ring and $I$ and $J$ be ideals of $R$, where $J \subset I$. Then

$$R/I \cong \frac{R/J}{I/J}.$$

**13.** Prove the Correspondence Theorem: Let $I$ be an ideal of a ring $R$. Then $S \to S/I$ is a one-to-one correspondence between the set of subrings $S$ containing $I$ and the set of subrings of $R/I$. Furthermore, the ideals of $R$ correspond to ideals of $R/I$.

**14.** Let $R$ be a ring and $S$ a subset of $R$. Show that $S$ is a subring of $R$ if and only if each of the following conditions is satisfied.

(a) $S \neq \emptyset$.

(b) $rs \in S$ for all $r, s \in S$.

(c) $r - s \in S$ for all $r, s \in S$.

**15.** Let $R$ be a ring with a collection of subrings $\{R_\alpha\}$. Prove that $\bigcap R_\alpha$ is a subring of $R$. Give an example to show that the union of two subrings is not necessarily a subring.

**16.** Let $\{I_\alpha\}_{\alpha\in A}$ be a collection of ideals in a ring $R$. Prove that $\bigcap_{\alpha\in A} I_\alpha$ is also an ideal in $R$. Give an example to show that if $I_1$ and $I_2$ are ideals in $R$, then $I_1 \cup I_2$ may not be an ideal.

**17.** Let $R$ be an integral domain. Show that if the only ideals in $R$ are $\{0\}$ and $R$ itself, $R$ must be a field.

**18.** Let $R$ be a commutative ring. An element $a$ in $R$ is **nilpotent** if $a^n = 0$ for some positive integer $n$. Show that the set of all nilpotent elements forms an ideal in $R$.

**19.** A ring $R$ is a **Boolean ring** if for every $a \in R$, $a^2 = a$. Show that every Boolean ring is a commutative ring.

**20.** Let $p$ be prime. Prove that

$$\mathbb{Z}_{(p)} = \{a/b : a, b \in \mathbb{Z} \text{ and } \gcd(b, p) = 1\}$$

is a ring. The ring $\mathbb{Z}_{(p)}$ is called the **ring of integers localized at** $p$.

**21.** Let $R$ and $S$ be arbitrary rings. Show that their Cartesian product is a ring if we define addition and multiplication in $R \times S$ by

(a) $(r, s) + (r', s') = (r + r', s + s')$

(b) $(r, s)(r', s') = (rr', ss')$

## 7.8 Material

1. Ring Definition (Quick)[1]
2. Ring Definition (Expanded)[2]
3. Ring Examples[3]
4. Units in a ring[4]
5. Ideals in rings[5]
6. Integral Domains[6]

## 7.9 Hints to Selected Exercises

### 7.6 · Core Exercises

**7.6.3.** $1 + 1 = 0$

### 7.7 · Additional Exercises

**7.7.2.** (a) $\{1, 3, 7, 9\}$; (c) $\{1, 2, 3, 4, 5, 6\}$; (e)

$$\left\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}, \right\}.$$

**7.7.3.** Assume there is an isomorphism $\phi : \mathbb{C} \to \mathbb{R}$ with $\phi(i) = a$.

[1] www.socratica.com/lesson/ring-definition-quick
[2] www.socratica.com/lesson/ring-definition
[3] www.socratica.com/lesson/ring-examples
[4] www.socratica.com/lesson/units-in-a-ring
[5] www.socratica.com/lesson/ideals-in-rings
[6] www.socratica.com/lesson/integral-domains

**7.7.4.** False. Assume there is an isomorphism $\phi : \mathbb{Q}(\sqrt{2}\,) \to \mathbb{Q}(\sqrt{3}\,)$ such that $\phi(\sqrt{2}\,) = a$.

**7.7.8.** If $I \neq \{0\}$, show that $1 \in I$.

**7.7.17.** Let $a \in R$ with $a \neq 0$. Then the principal ideal generated by $a$ is $R$. Thus, there exists a $b \in R$ such that $ab = 1$.

**7.7.19.** Compute $(a+b)^2$ and $(-ab)^2$.

**7.7.20.** Let $a/b, c/d \in \mathbb{Z}_{(p)}$. Then $a/b + c/d = (ad+bc)/bd$ and $(a/b)\cdot(c/d) = (ac)/(bd)$ are both in $\mathbb{Z}_{(p)}$, since $\gcd(bd, p) = 1$.

# Chapter 8

# Polynomials and Remainders

Basic learning goals

1. Ring structures of polynomials with different coefficient rings.
2. Polynomial Division and Euclidean Algorithm for polynomials.
3. Properties of ideals in polynomial rings.
4. Basic properties of zeros of polynomials.
5. Computations with Chinese Remainder Theorem.

Most people are fairly familiar with polynomials by the time they begin to study abstract algebra. When we examine polynomial expressions such as

$$\begin{aligned} p(x) &= x^3 - 3x + 2 \\ q(x) &= 3x^2 - 6x + 5, \end{aligned}$$

we have a pretty good idea of what $p(x) + q(x)$ and $p(x)q(x)$ mean. We just add and multiply polynomials as functions; that is,

$$\begin{aligned} (p+q)(x) &= p(x) + q(x) \\ &= (x^3 - 3x + 2) + (3x^2 - 6x + 5) \\ &= x^3 + 3x^2 - 9x + 7 \end{aligned}$$

and

$$\begin{aligned} (pq)(x) &= p(x)q(x) \\ &= (x^3 - 3x + 2)(3x^2 - 6x + 5) \\ &= 3x^5 - 6x^4 - 4x^3 + 24x^2 - 27x + 10. \end{aligned}$$

It is probably no surprise that polynomials form a ring. In this chapter we shall emphasize the algebraic structure of polynomials by studying polynomial rings. We can prove many results for polynomial rings that are similar to the theorems we proved for the integers. Analogs of prime numbers, the division algorithm, and the Euclidean algorithm exist for polynomials.

## 8.1 The Division Algorithm for Integers

This section is included to refresh the knowledge about the division algorithm for integers.

**Theorem 8.1.1 Division Algorithm.** *Let $a$ and $b$ be integers, with $b > 0$. Then there exist unique integers $q$ and $r$ such that*

$$a = bq + r$$

*where $0 \leq r < b$.*

*Proof.* This is a perfect example of the existence-and-uniqueness type of proof. We must first prove that the numbers $q$ and $r$ actually exist. Then we must show that if $q'$ and $r'$ are two other such numbers, then $q = q'$ and $r = r'$.

*Existence of $q$ and $r$.* Let

$$S = \{a - bk : k \in \mathbb{Z} \text{ and } a - bk \geq 0\}.$$

If $0 \in S$, then $b$ divides $a$, and we can let $q = a/b$ and $r = 0$. If $0 \notin S$, we can use the Well-Ordering Principle. We must first show that $S$ is nonempty. If $a > 0$, then $a - b \cdot 0 \in S$. If $a < 0$, then $a - b(2a) = a(1 - 2b) \in S$. In either case $S \neq \emptyset$. By the Well-Ordering Principle, $S$ must have a smallest member, say $r = a - bq$. Therefore, $a = bq + r$, $r \geq 0$. We now show that $r < b$. Suppose that $r > b$. Then

$$a - b(q + 1) = a - bq - b = r - b > 0.$$

In this case we would have $a - b(q+1)$ in the set $S$. But then $a - b(q+1) < a - bq$, which would contradict the fact that $r = a - bq$ is the smallest member of $S$. So $r \leq b$. Since $0 \notin S$, $r \neq b$ and so $r < b$.

*Uniqueness of $q$ and $r$.* Suppose there exist integers $r$, $r'$, $q$, and $q'$ such that

$$a = bq + r, 0 \leq r < b \quad \text{and} \quad a = bq' + r', 0 \leq r' < b.$$

Then $bq + r = bq' + r'$. Assume that $r' \geq r$. From the last equation we have $b(q - q') = r' - r$; therefore, $b$ must divide $r' - r$ and $0 \leq r' - r \leq r' < b$. This is possible only if $r' - r = 0$. Hence, $r = r'$ and $q = q'$. ■

Let $a$ and $b$ be integers. If $b = ak$ for some integer $k$, we write $a \mid b$. An integer $d$ is called a **common divisor** of $a$ and $b$ if $d \mid a$ and $d \mid b$. The **greatest common divisor** of integers $a$ and $b$ is a positive integer $d$ such that $d$ is a common divisor of $a$ and $b$ and if $d'$ is any other common divisor of $a$ and $b$, then $d' \mid d$. We write $d = \gcd(a, b)$; for example, $\gcd(24, 36) = 12$ and $\gcd(120, 102) = 6$. We say that two integers $a$ and $b$ are **relatively prime** if $\gcd(a, b) = 1$.

**Theorem 8.1.2** *Let $a$ and $b$ be nonzero integers. Then there exist integers $r$ and $s$ such that*

$$\gcd(a, b) = ar + bs.$$

*Furthermore, the greatest common divisor of $a$ and $b$ is unique.*

*Proof.* Let

$$S = \{am + bn : m, n \in \mathbb{Z} \text{ and } am + bn > 0\}.$$

Clearly, the set $S$ is nonempty; hence, by the Well-Ordering Principle $S$ must have a smallest member, say $d = ar + bs$. We claim that $d = \gcd(a, b)$. Write $a = dq + r'$ where $0 \leq r' < d$. If $r' > 0$, then

$$\begin{aligned} r' &= a - dq \\ &= a - (ar + bs)q \\ &= a - arq - bsq \\ &= a(1 - rq) + b(-sq), \end{aligned}$$

which is in $S$. But this would contradict the fact that $d$ is the smallest member of $S$. Hence, $r' = 0$ and $d$ divides $a$. A similar argument shows that $d$ divides $b$. Therefore, $d$ is a common divisor of $a$ and $b$.

Suppose that $d'$ is another common divisor of $a$ and $b$, and we want to show that $d' \mid d$. If we let $a = d'h$ and $b = d'k$, then

$$d = ar + bs = d'hr + d'ks = d'(hr + ks).$$

So $d'$ must divide $d$. Hence, $d$ must be the unique greatest common divisor of $a$ and $b$. ■

**Corollary 8.1.3** *Let $a$ and $b$ be two integers that are relatively prime. Then there exist integers $r$ and $s$ such that $ar + bs = 1$.*

### 8.1.1 The Euclidean Algorithm

Among other things, Theorem 8.1.2 allows us to compute the greatest common divisor of two integers.

**Example 8.1.4** Let us compute the greatest common divisor of 945 and 2415. First observe that

$$\begin{aligned} 2415 &= 945 \cdot 2 + 525 \\ 945 &= 525 \cdot 1 + 420 \\ 525 &= 420 \cdot 1 + 105 \\ 420 &= 105 \cdot 4 + 0. \end{aligned}$$

Reversing our steps, 105 divides 420, 105 divides 525, 105 divides 945, and 105 divides 2415. Hence, 105 divides both 945 and 2415. If $d$ were another common divisor of 945 and 2415, then $d$ would also have to divide 105. Therefore, $\gcd(945, 2415) = 105$.

If we work backward through the above sequence of equations, we can also obtain numbers $r$ and $s$ such that $945r + 2415s = 105$. Observe that

$$\begin{aligned} 105 &= 525 + (-1) \cdot 420 \\ &= 525 + (-1) \cdot [945 + (-1) \cdot 525] \\ &= 2 \cdot 525 + (-1) \cdot 945 \\ &= 2 \cdot [2415 + (-2) \cdot 945] + (-1) \cdot 945 \\ &= 2 \cdot 2415 + (-5) \cdot 945. \end{aligned}$$

So $r = -5$ and $s = 2$. Notice that $r$ and $s$ are not unique, since $r = 41$ and $s = -16$ would also work. □

To compute $\gcd(a, b) = d$, we are using repeated divisions to obtain a decreasing sequence of positive integers $r_1 > r_2 > \cdots > r_n = d$; that is,

$$\begin{aligned} b &= aq_1 + r_1 \\ a &= r_1q_2 + r_2 \\ r_1 &= r_2q_3 + r_3 \\ &\vdots \\ r_{n-2} &= r_{n-1}q_n + r_n \\ r_{n-1} &= r_nq_{n+1}. \end{aligned}$$

To find $r$ and $s$ such that $ar + bs = d$, we begin with this last equation and substitute results obtained from the previous equations:

$$\begin{aligned} d &= r_n \\ &= r_{n-2} - r_{n-1}q_n \\ &= r_{n-2} - q_n(r_{n-3} - q_{n-1}r_{n-2}) \\ &= -q_n r_{n-3} + (1 + q_n q_{n-1})r_{n-2} \\ &\vdots \\ &= ra + sb. \end{aligned}$$

The algorithm that we have just used to find the greatest common divisor $d$ of two integers $a$ and $b$ and to write $d$ as the linear combination of $a$ and $b$ is known as the **Euclidean algorithm**.

## 8.2 Polynomial Rings

Throughout this chapter we shall assume that $R$ is a commutative ring with identity. Any expression of the form

$$f(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

where $a_i \in R$ and $a_n \neq 0$, is called a **polynomial over** $R$ with **indeterminate** $x$. The elements $a_0, a_1, \ldots, a_n$ are called the **coefficients** of $f$. The coefficient $a_n$ is called the **leading coefficient**. A polynomial is called **monic** if the leading coefficient is 1. If $n$ is the largest nonnegative number for which $a_n \neq 0$, we say that the **degree** of $f$ is $n$ and write $\deg f(x) = n$. If no such $n$ exists—that is, if $f = 0$ is the zero polynomial—then the degree of $f$ is defined to be $-\infty$. We will denote the set of all polynomials with coefficients in a ring $R$ by $R[x]$. Two polynomials are equal exactly when their corresponding coefficients are equal; that is, if we let

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ q(x) &= b_0 + b_1 x + \cdots + b_m x^m, \end{aligned}$$

then $p(x) = q(x)$ if and only if $a_i = b_i$ for all $i \geq 0$.

To show that the set of all polynomials forms a ring, we must first define addition and multiplication. We define the sum of two polynomials as follows. Let

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ q(x) &= b_0 + b_1 x + \cdots + b_m x^m. \end{aligned}$$

Then the sum of $p(x)$ and $q(x)$ is

$$p(x) + q(x) = c_0 + c_1 x + \cdots + c_k x^k,$$

where $c_i = a_i + b_i$ for each $i$. We define the product of $p(x)$ and $q(x)$ to be

$$p(x)q(x) = c_0 + c_1 x + \cdots + c_{m+n} x^{m+n},$$

where

$$c_i = \sum_{k=0}^{i} a_k b_{i-k} = a_0 b_i + a_1 b_{i-1} + \cdots + a_{i-1} b_1 + a_i b_0$$

for each $i$. Notice that in each case some of the coefficients may be zero.

**Example 8.2.1** Suppose that

$$p(x) = 3 + 0x + 0x^2 + 2x^3 + 0x^4$$

and

$$q(x) = 2 + 0x - x^2 + 0x^3 + 4x^4$$

are polynomials in $\mathbb{Z}[x]$. If the coefficient of some term in a polynomial is zero, then we usually just omit that term. In this case we would write $p(x) = 3+2x^3$ and $q(x) = 2 - x^2 + 4x^4$. The sum of these two polynomials is

$$p(x) + q(x) = 5 - x^2 + 2x^3 + 4x^4.$$

The product,

$$p(x)q(x) = (3 + 2x^3)(2 - x^2 + 4x^4) = 6 - 3x^2 + 4x^3 + 12x^4 - 2x^5 + 8x^7,$$

can be calculated either by determining the $c_i$s in the definition or by simply multiplying polynomials in the same way as we have always done. □

**Example 8.2.2** Let

$$p(x) = 3 + 3x^3 \qquad \text{and} \qquad q(x) = 4 + 4x^2 + 4x^4$$

be polynomials in $\mathbb{Z}_{12}[x]$. The sum of $p(x)$ and $q(x)$ is $7 + 4x^2 + 3x^3 + 4x^4$. The product of the two polynomials is the zero polynomial. This example tells us that we can not expect $R[x]$ to be an integral domain if $R$ is not an integral domain. □

**Theorem 8.2.3** *Let $R$ be a commutative ring with identity. Then $R[x]$ is a commutative ring with identity.*

*Proof.* Our first task is to show that $R[x]$ is an abelian group under polynomial addition. The zero polynomial, $f(x) = 0$, is the additive identity. Given a polynomial $p(x) = \sum_{i=0}^{n} a_i x^i$, the inverse of $p(x)$ is easily verified to be $-p(x) = \sum_{i=0}^{n} (-a_i)x^i = -\sum_{i=0}^{n} a_i x^i$. Commutativity and associativity follow immediately from the definition of polynomial addition and from the fact that addition in $R$ is both commutative and associative.

To show that polynomial multiplication is associative, let

$$\begin{aligned} p(x) &= \sum_{i=0}^{m} a_i x^i, \\ q(x) &= \sum_{i=0}^{n} b_i x^i, \\ r(x) &= \sum_{i=0}^{p} c_i x^i. \end{aligned}$$

Then

$$\begin{aligned} [p(x)q(x)]r(x) &= \left[ \left( \sum_{i=0}^{m} a_i x^i \right) \left( \sum_{i=0}^{n} b_i x^i \right) \right] \left( \sum_{i=0}^{p} c_i x^i \right) \\ &= \left[ \sum_{i=0}^{m+n} \left( \sum_{j=0}^{i} a_j b_{i-j} \right) x^i \right] \left( \sum_{i=0}^{p} c_i x^i \right) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=0}^{m+n+p} \left[ \sum_{j=0}^{i} \left( \sum_{k=0}^{j} a_k b_{j-k} \right) c_{i-j} \right] x^i \\
&= \sum_{i=0}^{m+n+p} \left( \sum_{j+k+l=i} a_j b_k c_l \right) x^i \\
&= \sum_{i=0}^{m+n+p} \left[ \sum_{j=0}^{i} a_j \left( \sum_{k=0}^{i-j} b_k c_{i-j-k} \right) \right] x^i \\
&= \left( \sum_{i=0}^{m} a_i x^i \right) \left[ \sum_{i=0}^{n+p} \left( \sum_{j=0}^{i} b_j c_{i-j} \right) x^i \right] \\
&= \left( \sum_{i=0}^{m} a_i x^i \right) \left[ \left( \sum_{i=0}^{n} b_i x^i \right) \left( \sum_{i=0}^{p} c_i x^i \right) \right] \\
&= p(x)[q(x)r(x)]
\end{aligned}
$$

The commutativity and distribution properties of polynomial multiplication are proved in a similar manner. We shall leave the proofs of these properties as an exercise. ■

**Proposition 8.2.4** *Let $p(x)$ and $q(x)$ be polynomials in $R[x]$, where $R$ is an integral domain. Then $\deg p(x) + \deg q(x) = \deg(p(x)q(x))$. Furthermore, $R[x]$ is an integral domain.*

*Proof.* Suppose that we have two nonzero polynomials

$$p(x) = a_m x^m + \cdots + a_1 x + a_0$$

and

$$q(x) = b_n x^n + \cdots + b_1 x + b_0$$

with $a_m \neq 0$ and $b_n \neq 0$. The degrees of $p(x)$ and $q(x)$ are $m$ and $n$, respectively. The leading term of $p(x)q(x)$ is $a_m b_n x^{m+n}$, which cannot be zero since $R$ is an integral domain; hence, the degree of $p(x)q(x)$ is $m + n$, and $p(x)q(x) \neq 0$. Since $p(x) \neq 0$ and $q(x) \neq 0$ imply that $p(x)q(x) \neq 0$, we know that $R[x]$ must also be an integral domain. ■

We also want to consider polynomials in two or more variables, such as $x^2 - 3xy + 2y^3$. Let $R$ be a ring and suppose that we are given two indeterminates $x$ and $y$. Certainly we can form the ring $(R[x])[y]$. It is straightforward but perhaps tedious to show that $(R[x])[y] \cong R([y])[x]$. We shall identify these two rings by this isomorphism and simply write $R[x, y]$. The ring $R[x, y]$ is called the **ring of polynomials in two indeterminates $x$ and $y$ with coefficients in $R$**. We can define the **ring of polynomials in $n$ indeterminates with coefficients in $R$** similarly. We shall denote this ring by $R[x_1, x_2, \ldots, x_n]$.

**Theorem 8.2.5** *Let $R$ be a commutative ring with identity and $\alpha \in R$. Then we have a ring homomorphism $\phi_\alpha : R[x] \to R$ defined by*

$$\phi_\alpha(p(x)) = p(\alpha) = a_n \alpha^n + \cdots + a_1 \alpha + a_0,$$

*where $p(x) = a_n x^n + \cdots + a_1 x + a_0$.*

*Proof.* Let $p(x) = \sum_{i=0}^{n} a_i x^i$ and $q(x) = \sum_{i=0}^{m} b_i x^i$. It is easy to show that $\phi_\alpha(p(x)+q(x)) = \phi_\alpha(p(x))+\phi_\alpha(q(x))$. To show that multiplication is preserved

under the map $\phi_\alpha$, observe that

$$\begin{aligned}\phi_\alpha(p(x))\phi_\alpha(q(x)) &= p(\alpha)q(\alpha)\\ &= \left(\sum_{i=0}^{n} a_i\alpha^i\right)\left(\sum_{i=0}^{m} b_i\alpha^i\right)\\ &= \sum_{i=0}^{m+n}\left(\sum_{k=0}^{i} a_k b_{i-k}\right)\alpha^i\\ &= \phi_\alpha(p(x)q(x)).\end{aligned}$$

■

The map $\phi_\alpha : R[x] \to R$ is called the **evaluation homomorphism** at $\alpha$.

**Example 8.2.6** The polynomial function for two different polynomials can be the same. This is captured by the kernel of the evaluation homomorphism. Consider the polynomials $x, x^3 \in \mathbb{Z}_3[x]$. As polynomials they are different, as they differ in the coefficients of $x$ and $x^3$. However, as functions on $\mathbb{Z}_3$ they are the same. Their difference $x^3 - x$ is in the kernel of the evaluation homomorphism $\phi_\alpha$ for each $\alpha \in \mathbb{Z}_3$. □

## 8.3 The Division Algorithm

Recall that the division algorithm for integers (Theorem 8.1.1) says that if $a$ and $b$ are integers with $b > 0$, then there exist unique integers $q$ and $r$ such that $a = bq + r$, where $0 \leq r < b$. The algorithm by which $q$ and $r$ are found is just long division. A similar theorem exists for polynomials. The division algorithm for polynomials has several important consequences. Since its proof is very similar to the corresponding proof for integers, it is worthwhile to review Theorem 8.1.1 at this point.

**Theorem 8.3.1 Division Algorithm.** *Let $f(x)$ and $g(x)$ be polynomials in $F[x]$, where $F$ is a field and $g(x)$ is a nonzero polynomial. Then there exist unique polynomials $q(x), r(x) \in F[x]$ such that*

$$f(x) = g(x)q(x) + r(x),$$

*where either $\deg r(x) < \deg g(x)$ or $r(x)$ is the zero polynomial.*

*Proof.* We will first consider the existence of $q(x)$ and $r(x)$. If $f(x)$ is the zero polynomial, then

$$0 = 0 \cdot g(x) + 0;$$

hence, both $q$ and $r$ must also be the zero polynomial. Now suppose that $f(x)$ is not the zero polynomial and that $\deg f(x) = n$ and $\deg g(x) = m$. If $m > n$, then we can let $q(x) = 0$ and $r(x) = f(x)$. Hence, we may assume that $m \leq n$ and proceed by induction on $n$. If

$$\begin{aligned}f(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0\\ g(x) &= b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0\end{aligned}$$

the polynomial

$$f'(x) = f(x) - \frac{a_n}{b_m}x^{n-m}g(x)$$

has degree less than $n$ or is the zero polynomial. By induction, there exist polynomials $q'(x)$ and $r(x)$ such that

$$f'(x) = q'(x)g(x) + r(x),$$

where $r(x) = 0$ or the degree of $r(x)$ is less than the degree of $g(x)$. Now let

$$q(x) = q'(x) + \frac{a_n}{b_m}x^{n-m}.$$

Then

$$f(x) = g(x)q(x) + r(x),$$

with $r(x)$ the zero polynomial or $\deg r(x) < \deg g(x)$.

To show that $q(x)$ and $r(x)$ are unique, suppose that there exist two other polynomials $q_1(x)$ and $r_1(x)$ such that $f(x) = g(x)q_1(x) + r_1(x)$ with $\deg r_1(x) < \deg g(x)$ or $r_1(x) = 0$, so that

$$f(x) = g(x)q(x) + r(x) = g(x)q_1(x) + r_1(x),$$

and

$$g(x)[q(x) - q_1(x)] = r_1(x) - r(x).$$

If $q(x) - q_1(x)$ is not the zero polynomial, then

$$\deg(g(x)[q(x) - q_1(x)]) = \deg(r_1(x) - r(x)) \geq \deg g(x).$$

However, the degrees of both $r(x)$ and $r_1(x)$ are strictly less than the degree of $g(x)$; therefore, $r(x) = r_1(x)$ and $q(x) = q_1(x)$. ■

**Example 8.3.2** The division algorithm merely formalizes long division of polynomials, a task we have been familiar with since high school. For example, suppose that we divide $x^3 - x^2 + 2x - 3$ by $x - 2$.

$$\begin{array}{rrrrrrrrrr}
 & & & x^2 & + & x & + & 4 & & \\
\cline{4-10}
x & - & 2 & \multicolumn{7}{l}{\big| \; x^3 - x^2 + 2x - 3} \\
 & & & x^3 & - & 2x^2 & & & & \\
\cline{4-10}
 & & & & & x^2 & + & 2x & - & 3 \\
 & & & & & x^2 & - & 2x & & \\
\cline{5-10}
 & & & & & & & 4x & - & 3 \\
 & & & & & & & 4x & - & 8 \\
\cline{7-10}
 & & & & & & & & & 5
\end{array}$$

Hence, $x^3 - x^2 + 2x - 3 = (x - 2)(x^2 + x + 4) + 5$. □

Let $p(x)$ be a polynomial in $F[x]$ and $\alpha \in F$. We say that $\alpha$ is a **zero** or **root** of $p(x)$ if $p(x)$ is in the kernel of the evaluation homomorphism $\phi_\alpha$. All we are really saying here is that $\alpha$ is a zero of $p(x)$ if $p(\alpha) = 0$.

**Corollary 8.3.3** *Let $F$ be a field. An element $\alpha \in F$ is a zero of $p(x) \in F[x]$ if and only if $x - \alpha$ is a factor of $p(x)$ in $F[x]$.*

*Proof.* Suppose that $\alpha \in F$ and $p(\alpha) = 0$. By the division algorithm, there exist polynomials $q(x)$ and $r(x)$ such that

$$p(x) = (x - \alpha)q(x) + r(x)$$

and the degree of $r(x)$ must be less than the degree of $x - \alpha$. Since the degree of $r(x)$ is less than 1, $r(x) = a$ for $a \in F$; therefore,

$$p(x) = (x - \alpha)q(x) + a.$$

But

$$0 = p(\alpha) = 0 \cdot q(\alpha) + a = a;$$

consequently, $p(x) = (x - \alpha)q(x)$, and $x - \alpha$ is a factor of $p(x)$.

Conversely, suppose that $x - \alpha$ is a factor of $p(x)$; say $p(x) = (x - \alpha)q(x)$. Then $p(\alpha) = 0 \cdot q(\alpha) = 0$. ■

**Corollary 8.3.4** *Let $F$ be a field. A nonzero polynomial $p(x)$ of degree $n$ in $F[x]$ can have at most $n$ distinct zeros in $F$.*

*Proof.* We will use induction on the degree of $p(x)$. If $\deg p(x) = 0$, then $p(x)$ is a constant polynomial and has no zeros. Let $\deg p(x) = 1$. Then $p(x) = ax + b$ for some $a$ and $b$ in $F$. If $\alpha_1$ and $\alpha_2$ are zeros of $p(x)$, then $a\alpha_1 + b = a\alpha_2 + b$ or $\alpha_1 = \alpha_2$.

Now assume that $\deg p(x) > 1$. If $p(x)$ does not have a zero in $F$, then we are done. On the other hand, if $\alpha$ is a zero of $p(x)$, then $p(x) = (x - \alpha)q(x)$ for some $q(x) \in F[x]$ by Corollary 8.3.3. The degree of $q(x)$ is $n - 1$ by Proposition 8.2.4. Let $\beta$ be some other zero of $p(x)$ that is distinct from $\alpha$. Then $p(\beta) = (\beta - \alpha)q(\beta) = 0$. Since $\alpha \neq \beta$ and $F$ is a field, $q(\beta) = 0$. By our induction hypothesis, $q(x)$ can have at most $n - 1$ zeros in $F$ that are distinct from $\alpha$. Therefore, $p(x)$ has at most $n$ distinct zeros in $F$. ■

Let $F$ be a field. A monic polynomial $d(x)$ is a **greatest common divisor** of polynomials $p(x), q(x) \in F[x]$ if $d(x)$ evenly divides both $p(x)$ and $q(x)$; and, if for any other polynomial $d'(x)$ dividing both $p(x)$ and $q(x)$, $d'(x) \mid d(x)$. We write $d(x) = \gcd(p(x), q(x))$. Two polynomials $p(x)$ and $q(x)$ are **relatively prime** if $\gcd(p(x), q(x)) = 1$.

**Proposition 8.3.5** *Let $F$ be a field and suppose that $d(x)$ is a greatest common divisor of two polynomials $p(x)$ and $q(x)$ in $F[x]$. Then there exist polynomials $r(x)$ and $s(x)$ such that*

$$d(x) = r(x)p(x) + s(x)q(x).$$

*Furthermore, the greatest common divisor of two polynomials is unique.*

*Proof.* Let $d(x)$ be the monic polynomial of smallest degree in the set

$$S = \{f(x)p(x) + g(x)q(x) : f(x), g(x) \in F[x]\}.$$

We can write $d(x) = r(x)p(x) + s(x)q(x)$ for two polynomials $r(x)$ and $s(x)$ in $F[x]$. We need to show that $d(x)$ divides both $p(x)$ and $q(x)$. We shall first show that $d(x)$ divides $p(x)$. By the division algorithm, there exist polynomials $a(x)$ and $b(x)$ such that $p(x) = a(x)d(x) + b(x)$, where $b(x)$ is either the zero polynomial or $\deg b(x) < \deg d(x)$. Therefore,

$$\begin{aligned} b(x) &= p(x) - a(x)d(x) \\ &= p(x) - a(x)(r(x)p(x) + s(x)q(x)) \\ &= p(x) - a(x)r(x)p(x) - a(x)s(x)q(x) \\ &= p(x)(1 - a(x)r(x)) + q(x)(-a(x)s(x)) \end{aligned}$$

is a linear combination of $p(x)$ and $q(x)$ and therefore must be in $S$. However, $b(x)$ must be the zero polynomial since $d(x)$ was chosen to be of smallest degree; consequently, $d(x)$ divides $p(x)$. A symmetric argument shows that $d(x)$ must also divide $q(x)$; hence, $d(x)$ is a common divisor of $p(x)$ and $q(x)$.

To show that $d(x)$ is a greatest common divisor of $p(x)$ and $q(x)$, suppose that $d'(x)$ is another common divisor of $p(x)$ and $q(x)$. We will show that $d'(x) \mid d(x)$. Since $d'(x)$ is a common divisor of $p(x)$ and $q(x)$, there exist polynomials $u(x)$ and $v(x)$ such that $p(x) = u(x)d'(x)$ and $q(x) = v(x)d'(x)$. Therefore,

$$d(x) = r(x)p(x) + s(x)q(x)$$

$$= r(x)u(x)d'(x) + s(x)v(x)d'(x)$$
$$= d'(x)[r(x)u(x) + s(x)v(x)].$$

Since $d'(x) \mid d(x)$, $d(x)$ is a greatest common divisor of $p(x)$ and $q(x)$.

Finally, we must show that the greatest common divisor of $p(x)$ and $q(x)$ is unique. Suppose that $d'(x)$ is another greatest common divisor of $p(x)$ and $q(x)$. We have just shown that there exist polynomials $u(x)$ and $v(x)$ in $F[x]$ such that $d(x) = d'(x)[r(x)u(x) + s(x)v(x)]$. Since

$$\deg d(x) = \deg d'(x) + \deg[r(x)u(x) + s(x)v(x)]$$

and $d(x)$ and $d'(x)$ are both greatest common divisors, $\deg d(x) = \deg d'(x)$. Since $d(x)$ and $d'(x)$ are both monic polynomials of the same degree, it must be the case that $d(x) = d'(x)$. ■

Notice the similarity between the proof of Proposition 8.3.5 and the proof of Theorem 8.1.2.

## 8.4 Ideals in $F[x]$

A nonconstant polynomial $f(x) \in F[x]$ is **irreducible** over a field $F$ if $f(x)$ cannot be expressed as a product of two polynomials $g(x)$ and $h(x)$ in $F[x]$, where the degrees of $g(x)$ and $h(x)$ are both smaller than the degree of $f(x)$. If a polynomial is not irreducible, it is **reducible**. Irreducible polynomials function as the "prime numbers" of polynomial rings.

**Corollary 8.4.1** *Let $F$ be a field and $p(x) \in F[x]$ of degree at most 3. Then $p(x)$ is irreducible over $F$ if and only if it has no zero in $F$.*

*Proof.* In a factorization of $p(x)$ into polynomials of smaller degree than $p(x)$ at least one polynomial has to have degree 1. Now, Corollary 8.3.3 yields the claim. ■

The former statement is wrong for polynomials of higher degree. For example, the polynomial $x^4 - 25 \in \mathbb{Q}[x]$ has no zeros in $\mathbb{Q}$ but it can be written as $(x^2 - 5)(x^2 + 5)$.

**Example 8.4.2** The polynomial $x^2 - 2 \in \mathbb{Q}[x]$ is irreducible since it cannot be factored any further over the rational numbers. Similarly, $x^2 + 1$ is irreducible over the real numbers. □

**Example 8.4.3** The polynomial $p(x) = x^3 + x^2 + 2 \in \mathbb{Z}_3[x]$ is irreducible over $\mathbb{Z}_3$. Suppose that this polynomial was reducible over $\mathbb{Z}_3$. By the division algorithm there would have to be a factor of the form $x - a$, where $a$ is some element in $\mathbb{Z}_3$. Hence, it would have to be true that $p(a) = 0$. However,

$$p(0) = 2$$
$$p(1) = 1$$
$$p(2) = 2.$$

Therefore, $p(x)$ has no zeros in $\mathbb{Z}_3$ and must be irreducible. □

**Example 8.4.4** Let $F$ be some field and let $p_1(x_1, \ldots, x_n)$, $p_2(x_1, \ldots, x_n)$, $\ldots, p_m(x_1, \ldots, x_n)$ be polynomials in the polynomial ring $F[x_1, \ldots, x_n]$. Consider the system of polynomial equations

$$p_1(x_1, \ldots, x_n) = 0$$
$$p_2(x_1, \ldots, x_n) = 0$$

$$\vdots$$
$$p_m(x_1, \ldots, x_n) = 0$$

Let $(x_1^*, \ldots, x_n^*)$ be a solution to this system. That means that the $n$ polynomials are in the kernel of the evaluation homomorphism $\phi_{(x_1^*, \ldots, x_n^*)}$. In particular, a point $z$ in $F^n$ is a solution to the system exactly if these polynomials are in the kernel of its evaluation homomorphism $\phi_z$. Recall that the kernel of a ring homomorphism is an ideal; this ideal is the set of all polynomials

$$h_1(x_1, \ldots, x_n) \cdot p_1(x_1, \ldots, x_n) + \cdots + h_m(x_1, \ldots, x_n) \cdot p_m(x_1, \ldots, x_n)$$

where $h_1(x_1, \ldots, x_n), h_2(x_1, \ldots, x_n), \ldots, h_m(x_1, \ldots, x_n) \in F[x_1, \ldots, x_n]$. Note that this ideal is independent of the specific solution but (implicitly) encodes the set of all solutions to the system. □

Let $F$ be a field. Recall that a principal ideal in $F[x]$ is an ideal $\langle p(x) \rangle$ generated by some polynomial $p(x)$; that is,

$$\langle p(x) \rangle = \{p(x)q(x) : q(x) \in F[x]\}.$$

**Example 8.4.5** The polynomial $x^2$ in $F[x]$ generates the ideal $\langle x^2 \rangle$ consisting of all polynomials with no constant term or term of degree 1. □

**Corollary 8.4.6** *Let $F$ be a field and suppose that $d(x)$ is the greatest common divisor of two polynomials $p(x)$ and $q(x)$ in $F[x]$. Then the smallest ideal of $F[x]$ containing $p(x)$ and $q(x)$ is $\langle d(x) \rangle$.*

*Proof.* Let $I$ be the smallest ideal containing $p(x)$ and $q(x)$. By Proposition 8.3.5, we can represent $d(x) = r(x)p(x) + s(x)q(x)$ with $r(x), s(x) \in F[x]$. Hence, we get $d(x) \in I$ which implies $\langle d(x) \rangle \subseteq I$. From the definition of the greatest common divisor, we get that $d(x)$ divides $p(x)$ and $q(x)$, so $p(x), q(x) \in \langle d(x) \rangle$. Now, the minimality of $I$ yields $\langle d(x) \rangle = I$. ■

**Theorem 8.4.7** *If $F$ is a field, then every ideal in $F[x]$ is a principal ideal.*

*Proof.* Let $I$ be an ideal of $F[x]$. If $I$ is the zero ideal, the theorem is easily true. Suppose that $I$ is a nontrivial ideal in $F[x]$, and let $p(x) \in I$ be a nonzero element of minimal degree. If $\deg p(x) = 0$, then $p(x)$ is a nonzero constant and 1 must be in $I$. Since 1 generates all of $F[x]$, $\langle 1 \rangle = I = F[x]$ and $I$ is again a principal ideal.

Now assume that $\deg p(x) \geq 1$ and let $f(x)$ be any element in $I$. By the division algorithm there exist $q(x)$ and $r(x)$ in $F[x]$ such that $f(x) = p(x)q(x) + r(x)$ and $\deg r(x) < \deg p(x)$. Since $f(x), p(x) \in I$ and $I$ is an ideal, $r(x) = f(x) - p(x)q(x)$ is also in $I$. However, since we chose $p(x)$ to be of minimal degree, $r(x)$ must be the zero polynomial. Since we can write any element $f(x)$ in $I$ as $p(x)q(x)$ for some $q(x) \in F[x]$, it must be the case that $I = \langle p(x) \rangle$. ■

**Example 8.4.8** It is not the case that every ideal in the ring $F[x, y]$ is a principal ideal. Consider the ideal of $F[x, y]$ generated by the polynomials $x$ and $y$. This is the ideal of $F[x, y]$ consisting of all polynomials with no constant term. Since both $x$ and $y$ are in the ideal, no single polynomial can generate the entire ideal. □

**Theorem 8.4.9** *Let $F$ be a field and suppose that $p(x) \in F[x]$. Then the ideal generated by $p(x)$ is maximal if and only if $p(x)$ is irreducible.*

*Proof.* Suppose that $p(x)$ generates a maximal ideal of $F[x]$. Then $\langle p(x) \rangle$ is also a prime ideal of $F[x]$. Since a maximal ideal must be properly contained

inside $F[x]$, $p(x)$ cannot be a constant polynomial. Let us assume that $p(x)$ factors into two polynomials of lesser degree, say $p(x) = f(x)g(x)$. Since $\langle p(x) \rangle$ is a prime ideal one of these factors, say $f(x)$, is in $\langle p(x) \rangle$ and therefore be a multiple of $p(x)$. But this would imply that $\langle p(x) \rangle \subset \langle f(x) \rangle$, which is impossible since $\langle p(x) \rangle$ is maximal.

Conversely, suppose that $p(x)$ is irreducible over $F[x]$. Let $I$ be an ideal in $F[x]$ containing $\langle p(x) \rangle$. By Theorem 8.4.7, $I$ is a principal ideal; hence, $I = \langle f(x) \rangle$ for some $f(x) \in F[x]$. Since $p(x) \in I$, it must be the case that $p(x) = f(x)g(x)$ for some $g(x) \in F[x]$. However, $p(x)$ is irreducible; hence, either $f(x)$ or $g(x)$ is a constant polynomial. If $f(x)$ is constant, then $I = F[x]$ and we are done. If $g(x)$ is constant, then $f(x)$ is a constant multiple of $I$ and $I = \langle p(x) \rangle$. Thus, there are no proper ideals of $F[x]$ that properly contain $\langle p(x) \rangle$. ■

### 8.4.1 Historical Note

Throughout history, the solution of polynomial equations has been a challenging problem. The Babylonians knew how to solve the equation $ax^2 + bx + c = 0$. Omar Khayyam (1048–1131) devised methods of solving cubic equations through the use of geometric constructions and conic sections. The algebraic solution of the general cubic equation $ax^3 + bx^2 + cx + d = 0$ was not discovered until the sixteenth century. An Italian mathematician, Luca Pacioli (ca. 1445–1509), wrote in *Summa de Arithmetica* that the solution of the cubic was impossible. This was taken as a challenge by the rest of the mathematical community.

Scipione del Ferro (1465–1526), of the University of Bologna, solved the "depressed cubic,"

$$ax^3 + cx + d = 0.$$

He kept his solution an absolute secret. This may seem surprising today, when mathematicians are usually very eager to publish their results, but in the days of the Italian Renaissance secrecy was customary. Academic appointments were not easy to secure and depended on the ability to prevail in public contests. Such challenges could be issued at any time. Consequently, any major new discovery was a valuable weapon in such a contest. If an opponent presented a list of problems to be solved, del Ferro could in turn present a list of depressed cubics. He kept the secret of his discovery throughout his life, passing it on only on his deathbed to his student Antonio Fior (ca. 1506–?).

Although Fior was not the equal of his teacher, he immediately issued a challenge to Niccolo Fontana (1499–1557). Fontana was known as Tartaglia (the Stammerer). As a youth he had suffered a blow from the sword of a French soldier during an attack on his village. He survived the savage wound, but his speech was permanently impaired. Tartaglia sent Fior a list of 30 various mathematical problems; Fior countered by sending Tartaglia a list of 30 depressed cubics. Tartaglia would either solve all 30 of the problems or absolutely fail. After much effort Tartaglia finally succeeded in solving the depressed cubic and defeated Fior, who faded into obscurity.

At this point another mathematician, Gerolamo Cardano (1501–1576), entered the story. Cardano wrote to Tartaglia, begging him for the solution to the depressed cubic. Tartaglia refused several of his requests, then finally revealed the solution to Cardano after the latter swore an oath not to publish the secret or to pass it on to anyone else. Using the knowledge that he had obtained from Tartaglia, Cardano eventually solved the general cubic

$$ax^3 + bx^2 + cx + d = 0.$$

Cardano shared the secret with his student, Ludovico Ferrari (1522–1565), who solved the general quartic equation,

$$ax^4 + bx^3 + cx^2 + dx + e = 0.$$

In 1543, Cardano and Ferrari examined del Ferro's papers and discovered that he had also solved the depressed cubic. Cardano felt that this relieved him of his obligation to Tartaglia, so he proceeded to publish the solutions in *Ars Magna* (1545), in which he gave credit to del Ferro for solving the special case of the cubic. This resulted in a bitter dispute between Cardano and Tartaglia, who published the story of the oath a year later.

## 8.5 The Chinese Remainder Theorem and Software Design

The Chinese Remainder Theorem is a result from elementary number theory about the solution of systems of simultaneous congruences. The Chinese mathematician Sun-tsï wrote about the theorem in the first century A.D. This theorem has some interesting consequences in the design of software for parallel processors.

**Lemma 8.5.1** *Let $m$ and $n$ be positive integers such that* $\gcd(m, n) = 1$. *Then for $a, b \in \mathbb{Z}$ the system*

$$\begin{aligned} x &\equiv a \pmod{m} \\ x &\equiv b \pmod{n} \end{aligned}$$

*has a solution. If $x_1$ and $x_2$ are two solutions of the system, then $x_1 \equiv x_2 \pmod{mn}$.*

*Proof.* The equation $x \equiv a \pmod{m}$ has a solution since $a + km$ satisfies the equation for all $k \in \mathbb{Z}$. We must show that there exists an integer $k_1$ such that

$$a + k_1 m \equiv b \pmod{n}.$$

This is equivalent to showing that

$$k_1 m \equiv (b - a) \pmod{n}$$

has a solution for $k_1$. Since $m$ and $n$ are relatively prime, there exist integers $s$ and $t$ such that $ms + nt = 1$. Consequently,

$$(b - a)ms = (b - a) - (b - a)nt,$$

or

$$[(b - a)s]m \equiv (b - a) \pmod{n}.$$

Now let $k_1 = (b - a)s$.

To show that any two solutions are congruent modulo $mn$, let $c_1$ and $c_2$ be two solutions of the system. That is,

$$\begin{aligned} c_i &\equiv a \pmod{m} \\ c_i &\equiv b \pmod{n} \end{aligned}$$

for $i = 1, 2$. Then

$$\begin{aligned} c_2 &\equiv c_1 \pmod{m} \\ c_2 &\equiv c_1 \pmod{n}. \end{aligned}$$

Therefore, both $m$ and $n$ divide $c_1 - c_2$. Consequently, $c_2 \equiv c_1 \pmod{mn}$. ■

**Remark 8.5.2** Using the terminology of Lemma 8.5.1, there are integers $w, v \in \mathbb{Z}$ fulfilling $wm + vn = 1$. Then a solution of the system is given by $wmb + vna$ as one can check by taking congruences.

**Example 8.5.3** Let us solve the system

$$\begin{aligned} x &\equiv 3 \pmod 4 \\ x &\equiv 4 \pmod 5. \end{aligned}$$

Using the Euclidean algorithm, we can find integers $s$ and $t$ such that $4s+5t = 1$. Two such integers are $s = 4$ and $t = -3$. Consequently,

$$x = a + k_1 m = 3 + 4k_1 = 3 + 4[(5 - 4)4] = 19.$$

□

By induction, we get the version with several difference congruences. The actual proof is left as exercise.

**Theorem 8.5.4 Chinese Remainder Theorem.** *Let $n_1, n_2, \ldots, n_k$ be positive integers such that $\gcd(n_i, n_j) = 1$ for $i \neq j$. Then for any integers $a_1, \ldots, a_k$, the system*

$$\begin{aligned} x &\equiv a_1 \pmod{n_1} \\ x &\equiv a_2 \pmod{n_2} \\ &\vdots \\ x &\equiv a_k \pmod{n_k} \end{aligned}$$

*has a solution. Furthermore, any two solutions of the system are congruent modulo $n_1 n_2 \cdots n_k$.*

**Example 8.5.5** Let us solve the system

$$\begin{aligned} x &\equiv 3 \pmod 4 \\ x &\equiv 4 \pmod 5 \\ x &\equiv 1 \pmod 9 \\ x &\equiv 5 \pmod 7. \end{aligned}$$

From Example 8.5.3 we know that 19 is a solution of the first two congruences and any other solution of the system is congruent to 19 (mod 20). Hence, we can reduce the system to a system of three congruences:

$$\begin{aligned} x &\equiv 19 \pmod{20} \\ x &\equiv 1 \pmod 9 \\ x &\equiv 5 \pmod 7. \end{aligned}$$

Solving the next two equations, we can reduce the system to

$$\begin{aligned} x &\equiv 19 \pmod{180} \\ x &\equiv 5 \pmod 7. \end{aligned}$$

Solving this last system, we find that 19 is a solution for the system that is unique up to modulo 1260. □

One interesting application of the Chinese Remainder Theorem in the design of computer software is that the theorem allows us to break up a calculation involving large integers into several less formidable calculations. A

computer will handle integer calculations only up to a certain size due to the size of its processor chip, which is usually a 32 or 64-bit processor chip. For example, the largest integer available on a computer with a 64-bit processor chip is

$$2^{63} - 1 = 9{,}223{,}372{,}036{,}854{,}775{,}807.$$

Larger processors such as 128 or 256-bit have been proposed or are under development. There is even talk of a 512-bit processor chip. The largest integer that such a chip could store with be $2^{511} - 1$, which would be a 154 digit number. However, we would need to deal with much larger numbers to break sophisticated encryption schemes.

Special software is required for calculations involving larger integers which cannot be added directly by the machine. By using the Chinese Remainder Theorem we can break down large integer additions and multiplications into calculations that the computer can handle directly. This is especially useful on parallel processing computers which have the ability to run several programs concurrently.

Most computers have a single central processing unit (CPU) containing one processor chip and can only add two numbers at a time. To add a list of ten numbers, the CPU must do nine additions in sequence. However, a parallel processing computer has more than one CPU. A computer with 10 CPUs, for example, can perform 10 different additions at the same time. If we can take a large integer and break it down into parts, sending each part to a different CPU, then by performing several additions or multiplications simultaneously on those parts, we can work with an integer that the computer would not be able to handle as a whole.

**Example 8.5.6** Suppose that we wish to multiply 2134 by 1531. We will use the integers 95, 97, 98, and 99 because they are relatively prime. We can break down each integer into four parts:

$$\begin{aligned} 2134 &\equiv 44 \pmod{95} \\ 2134 &\equiv 0 \pmod{97} \\ 2134 &\equiv 76 \pmod{98} \\ 2134 &\equiv 55 \pmod{99} \end{aligned}$$

and

$$\begin{aligned} 1531 &\equiv 11 \pmod{95} \\ 1531 &\equiv 76 \pmod{97} \\ 1531 &\equiv 61 \pmod{98} \\ 1531 &\equiv 46 \pmod{99}. \end{aligned}$$

Multiplying the corresponding equations, we obtain

$$\begin{aligned} 2134 \cdot 1531 &\equiv 44 \cdot 11 \equiv 9 \pmod{95} \\ 2134 \cdot 1531 &\equiv 0 \cdot 76 \equiv 0 \pmod{97} \\ 2134 \cdot 1531 &\equiv 76 \cdot 61 \equiv 30 \pmod{98} \\ 2134 \cdot 1531 &\equiv 55 \cdot 46 \equiv 55 \pmod{99}. \end{aligned}$$

Each of these four computations can be sent to a different processor if our computer has several CPUs. By the above calculation, we know that $2134 \cdot 1531$ is a solution of the system

$$x \equiv 9 \pmod{95}$$

$$\begin{aligned} x &\equiv 0 \pmod{97} \\ x &\equiv 30 \pmod{98} \\ x &\equiv 55 \pmod{99}. \end{aligned}$$

The Chinese Remainder Theorem tells us that solutions are unique up to modulo $95 \cdot 97 \cdot 98 \cdot 99 = 89{,}403{,}930$. Solving this system of congruences for $x$ tells us that $2134 \cdot 1531 = 3{,}267{,}154$.

The conversion of the computation into the four subcomputations will take some computing time. In addition, solving the system of congruences can also take considerable time. However, if we have many computations to be performed on a particular set of numbers, it makes sense to transform the problem as we have done above and to perform the necessary calculations simultaneously. □

## 8.6 Additional insights

### 8.6.1

**Lemma 8.6.1** *Let $p(x) \in \mathbb{Q}[x]$. Then*

$$p(x) = \frac{r}{s}(a_0 + a_1 x + \cdots + a_n x^n),$$

*where $r, s, a_0, \ldots, a_n$ are integers, the $a_i$'s are relatively prime, and $r$ and $s$ are relatively prime.*

*Proof.* Suppose that

$$p(x) = \frac{b_0}{c_0} + \frac{b_1}{c_1}x + \cdots + \frac{b_n}{c_n}x^n,$$

where the $b_i$'s and the $c_i$'s are integers. We can rewrite $p(x)$ as

$$p(x) = \frac{1}{c_0 \cdots c_n}(d_0 + d_1 x + \cdots + d_n x^n),$$

where $d_0, \ldots, d_n$ are integers. Let $d$ be the greatest common divisor of $d_0, \ldots, d_n$. Then

$$p(x) = \frac{d}{c_0 \cdots c_n}(a_0 + a_1 x + \cdots + a_n x^n),$$

where $d_i = da_i$ and the $a_i$'s are relatively prime. Reducing $d/(c_0 \cdots c_n)$ to its lowest terms, we can write

$$p(x) = \frac{r}{s}(a_0 + a_1 x + \cdots + a_n x^n),$$

where $\gcd(r, s) = 1$. ■

**Theorem 8.6.2 Gauss's Lemma.** *Let $p(x) \in \mathbb{Z}[x]$ be a monic polynomial such that $p(x)$ factors into a product of two polynomials $\alpha(x)$ and $\beta(x)$ in $\mathbb{Q}[x]$, where the degrees of both $\alpha(x)$ and $\beta(x)$ are less than the degree of $p(x)$. Then $p(x) = a(x)b(x)$, where $a(x)$ and $b(x)$ are monic polynomials in $\mathbb{Z}[x]$ with $\deg \alpha(x) = \deg a(x)$ and $\deg \beta(x) = \deg b(x)$.*

*Proof.* By Lemma 8.6.1, we can assume that

$$\alpha(x) = \frac{c_1}{d_1}(a_0 + a_1 x + \cdots + a_m x^m) = \frac{c_1}{d_1}\alpha_1(x)$$

$$\beta(x) = \frac{c_2}{d_2}(b_0 + b_1x + \cdots + b_nx^n) = \frac{c_2}{d_2}\beta_1(x),$$

where the $a_i$'s are relatively prime and the $b_i$'s are relatively prime. Consequently,

$$p(x) = \alpha(x)\beta(x) = \frac{c_1c_2}{d_1d_2}\alpha_1(x)\beta_1(x) = \frac{c}{d}\alpha_1(x)\beta_1(x),$$

where $c/d$ is the product of $c_1/d_1$ and $c_2/d_2$ expressed in lowest terms. Hence, $dp(x) = c\alpha_1(x)\beta_1(x)$.

If $d = 1$, then $ca_mb_n = 1$ since $p(x)$ is a monic polynomial. Hence, either $c = 1$ or $c = -1$. If $c = 1$, then either $a_m = b_n = 1$ or $a_m = b_n = -1$. In the first case $p(x) = \alpha_1(x)\beta_1(x)$, where $\alpha_1(x)$ and $\beta_1(x)$ are monic polynomials with $\deg\alpha(x) = \deg\alpha_1(x)$ and $\deg\beta(x) = \deg\beta_1(x)$. In the second case $a(x) = -\alpha_1(x)$ and $b(x) = -\beta_1(x)$ are the correct monic polynomials since $p(x) = (-\alpha_1(x))(-\beta_1(x)) = a(x)b(x)$. The case in which $c = -1$ can be handled similarly.

Now suppose that $d \neq 1$. Since $\gcd(c, d) = 1$, there exists a prime $p$ such that $p \mid d$ and $p \nmid c$. Also, since the coefficients of $\alpha_1(x)$ are relatively prime, there exists a coefficient $a_i$ such that $p \nmid a_i$. Similarly, there exists a coefficient $b_j$ of $\beta_1(x)$ such that $p \nmid b_j$. Let $\alpha_1'(x)$ and $\beta_1'(x)$ be the polynomials in $\mathbb{Z}_p[x]$ obtained by reducing the coefficients of $\alpha_1(x)$ and $\beta_1(x)$ modulo $p$. Since $p \mid d$, $\alpha_1'(x)\beta_1'(x) = 0$ in $\mathbb{Z}_p[x]$. However, this is impossible since neither $\alpha_1'(x)$ nor $\beta_1'(x)$ is the zero polynomial and $\mathbb{Z}_p[x]$ is an integral domain. Therefore, $d = 1$ and the theorem is proven. ■

**Corollary 8.6.3** *Let $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a polynomial with coefficients in $\mathbb{Z}$ and $a_0 \neq 0$. If $p(x)$ has a zero in $\mathbb{Q}$, then $p(x)$ also has a zero $\alpha$ in $\mathbb{Z}$. Furthermore, $\alpha$ divides $a_0$.*

*Proof.* Let $p(x)$ have a zero $a \in \mathbb{Q}$. Then $p(x)$ must have a linear factor $x - a$. By Gauss's Lemma, $p(x)$ has a factorization with a linear factor in $\mathbb{Z}[x]$. Hence, for some $\alpha \in \mathbb{Z}$

$$p(x) = (x - \alpha)(x^{n-1} + \cdots - a_0/\alpha).$$

Thus $a_0/\alpha \in \mathbb{Z}$ and so $\alpha \mid a_0$. ■

**Example 8.6.4** Let $p(x) = x^4 - 2x^3 + x + 1$. We shall show that $p(x)$ is irreducible over $\mathbb{Q}[x]$. Assume that $p(x)$ is reducible. Then either $p(x)$ has a linear factor, say $p(x) = (x - \alpha)q(x)$, where $q(x)$ is a polynomial of degree three, or $p(x)$ has two quadratic factors.

If $p(x)$ has a linear factor in $\mathbb{Q}[x]$, then it has a zero in $\mathbb{Z}$. By Corollary 8.6.3, any zero must divide 1 and therefore must be $\pm 1$; however, $p(1) = 1$ and $p(-1) = 3$. Consequently, we have eliminated the possibility that $p(x)$ has any linear factors.

Therefore, if $p(x)$ is reducible it must factor into two quadratic polynomials, say

$$\begin{aligned} p(x) &= (x^2 + ax + b)(x^2 + cx + d) \\ &= x^4 + (a + c)x^3 + (ac + b + d)x^2 + (ad + bc)x + bd, \end{aligned}$$

where each factor is in $\mathbb{Z}[x]$ by Gauss's Lemma. Hence,

$$\begin{aligned} a + c &= -2 \\ ac + b + d &= 0 \\ ad + bc &= 1 \\ bd &= 1. \end{aligned}$$

Since $bd = 1$, either $b = d = 1$ or $b = d = -1$. In either case $b = d$ and so

$$ad + bc = b(a + c) = 1.$$

Since $a + c = -2$, we know that $-2b = 1$. This is impossible since $b$ is an integer. Therefore, $p(x)$ must be irreducible over $\mathbb{Q}$. □

**Theorem 8.6.5 Eisenstein's Criterion.** *Let $p$ be a prime and suppose that*

$$f(x) = a_n x^n + \cdots + a_0 \in \mathbb{Z}[x].$$

*If $p \mid a_i$ for $i = 0, 1, \ldots, n-1$, but $p \nmid a_n$ and $p^2 \nmid a_0$, then $f(x)$ is irreducible over $\mathbb{Q}$.*

*Proof.* By Gauss's Lemma, we need only show that $f(x)$ does not factor into polynomials of lower degree in $\mathbb{Z}[x]$. Let

$$f(x) = (b_r x^r + \cdots + b_0)(c_s x^s + \cdots + c_0)$$

be a factorization in $\mathbb{Z}[x]$, with $b_r$ and $c_s$ not equal to zero and $r, s < n$. Since $p^2$ does not divide $a_0 = b_0 c_0$, either $b_0$ or $c_0$ is not divisible by $p$. Suppose that $p \nmid b_0$ and $p \mid c_0$. Since $p \nmid a_n$ and $a_n = b_r c_s$, neither $b_r$ nor $c_s$ is divisible by $p$. Let $m$ be the smallest value of $k$ such that $p \nmid c_k$. Then

$$a_m = b_0 c_m + b_1 c_{m-1} + \cdots + b_m c_0$$

is not divisible by $p$, since each term on the right-hand side of the equation is divisible by $p$ except for $b_0 c_m$. Therefore, $m = n$ since $a_i$ is divisible by $p$ for $m < n$. Hence, $f(x)$ cannot be factored into polynomials of lower degree and therefore must be irreducible. ■

**Example 8.6.6** The polynomial

$$f(x) = 16x^5 - 9x^4 + 3x^2 + 6x - 21$$

is easily seen to be irreducible over $\mathbb{Q}$ by Eisenstein's Criterion if we let $p = 3$. □

Eisenstein's Criterion is more useful in constructing irreducible polynomials of a certain degree over $\mathbb{Q}$ than in determining the irreducibility of an arbitrary polynomial in $\mathbb{Q}[x]$: given an arbitrary polynomial, it is not very likely that we can apply Eisenstein's Criterion. The real value of Theorem 8.6.5 is that we now have an easy method of generating irreducible polynomials of any degree.

## 8.7 Core Exercises

**1.** Use the division algorithm to find $q(x)$ and $r(x)$ such that $a(x) = q(x)b(x) + r(x)$ with $\deg r(x) < \deg b(x)$ for each of the following pairs of polynomials.

(a) $a(x) = 5x^3 + 6x^2 - 3x + 4$ and $b(x) = x - 2$ in $\mathbb{Z}_7[x]$

(b) $a(x) = 6x^4 - 2x^3 + x^2 - 3x + 1$ and $b(x) = x^2 + x - 2$ in $\mathbb{Z}_7[x]$

(c) $a(x) = 4x^5 - x^3 + x^2 + 4$ and $b(x) = x^3 - 2$ in $\mathbb{Z}_5[x]$

(d) $a(x) = x^5 + x^3 - x^2 - x$ and $b(x) = x^3 + x$ in $\mathbb{Z}_2[x]$

**2.** Find the greatest common divisor of each of the following pairs $p(x)$ and $q(x)$ of polynomials. If $d(x) = \gcd(p(x), q(x))$, find two polynomials $a(x)$ and $b(x)$ such that $a(x)p(x) + b(x)q(x) = d(x)$.

(a) $p(x) = x^3 - 6x^2 + 14x - 15$ and $q(x) = x^3 - 8x^2 + 21x - 18$, where $p(x), q(x) \in \mathbb{Q}[x]$

(b) $p(x) = x^3+x^2-x+1$ and $q(x) = x^3+x-1$, where $p(x), q(x) \in \mathbb{Z}_2[x]$

(c) $p(x) = x^3+x^2-4x+4$ and $q(x) = x^3+3x-2$, where $p(x), q(x) \in \mathbb{Z}_5[x]$

(d) $p(x) = x^3 - 2x + 4$ and $q(x) = 4x^3 + x + 3$, where $p(x), q(x) \in \mathbb{Q}[x]$

**3.** Find all of the irreducible polynomials of degrees 2 and 3 in $\mathbb{Z}_2[x]$.

**4.** Show that the division algorithm does not hold for $\mathbb{Z}[x]$. Why does it fail?

**5.** **Homomorphisms between $R$ and $R[x]$.** Let $R$ be an integral domain.

(a) Give two different ring homomorphisms from $R$ to $R[x]$ and two different ring homomorphisms from $R[x]$ to $R$.

(b) Show the following: if the characteristic of $R$ is finite, there is no ring isomorphism between $R$ and $R[x]$.

**6.** Let $F$ be a field. Denote the set of $(n \times n)$-matrices with coefficients in $F$ by $M_n(F)$.

(a) Check that $M_n(F)$ forms a ring.

(b) For a fixed matrix $A \in M_n(F)$, define the map $\psi_A : F[x] \to M_n(F)$ by

$$\psi_A(p(x)) = \sum_{i=0}^{n} c_i A^i$$

where $p(x) = \sum_{i=0}^{n} c_i x^i = c_0+c_1x+c_2x^2+\cdots+c_nx^n$ for some $c_i \in F$ with $i \in \{0, 1, \ldots, n\}$. Show that $\psi_A$ is a ring homomorphism.

(c) Define $q(B) := \psi_B(q(x))$ for $q(x) \in F[x]$ and $B \in M_n(F)$. Now fix a matrix $A \in M_n(F)$ and consider the set

$$H = \{g(x) \mid g(x) \in F[x], g(A) = 0\}.$$

Show that there is a polynomial $m(x) \in F[x]$ such that $m(x)$ divides each polynomial in $H$; that is for each polynomial $h(x) \in H$ there is a polynomial $k(x) \in F[x]$ such that $h(x) = m(x)k(x)$.

(d) What is the meaning of the zeros of the polynomial $m(x)$ from (c)?

**7.** Prove Theorem 8.5.4.

**8.** Solve each of the following systems of congruences.

(a)
$$\begin{aligned} x &\equiv 2 \pmod 5 \\ x &\equiv 6 \pmod{11} \end{aligned}$$

(b)
$$\begin{aligned} x &\equiv 3 \pmod 7 \\ x &\equiv 0 \pmod 8 \\ x &\equiv 5 \pmod{15} \end{aligned}$$

(c)
$$\begin{aligned} x &\equiv 2 \pmod 4 \\ x &\equiv 4 \pmod 7 \\ x &\equiv 7 \pmod 9 \\ x &\equiv 5 \pmod{11} \end{aligned}$$

(d)
$$\begin{aligned} x &\equiv 3 \pmod 5 \\ x &\equiv 0 \pmod 8 \\ x &\equiv 1 \pmod{11} \\ x &\equiv 5 \pmod{13} \end{aligned}$$

**9.** ***Advanced*** **The Chinese Remainder Theorem for Rings.** Let $R$ be a ring and $I$ and $J$ be ideals in $R$ such that $I + J = R$.

(a) Show that for any $r$ and $s$ in $R$, the system of equations

$$\begin{aligned} x &\equiv r \pmod I \\ x &\equiv s \pmod J \end{aligned}$$

has a solution.

(b) In addition, prove that any two solutions of the system are congruent modulo $I \cap J$.

(c) Let $I$ and $J$ be ideals in a ring $R$ such that $I + J = R$. Show that there exists a ring isomorphism

$$R/(I \cap J) \cong R/I \times R/J.$$

## 8.8 Additional Exercises

**1.** List all of the polynomials of degree 3 or less in $\mathbb{Z}_2[x]$.

**2.** Compute each of the following.

(a) $(5x^2 + 3x - 4) + (4x^2 - x + 9)$ in $\mathbb{Z}_{12}[x]$

(b) $(5x^2 + 3x - 4)(4x^2 - x + 9)$ in $\mathbb{Z}_{12}[x]$

(c) $(7x^3 + 3x^2 - x) + (6x^2 - 8x + 4)$ in $\mathbb{Z}_9[x]$

(d) $(3x^2 + 2x - 4) + (4x^2 + 2)$ in $\mathbb{Z}_5[x]$

(e) $(3x^2 + 2x - 4)(4x^2 + 2)$ in $\mathbb{Z}_5[x]$

(f) $(5x^2 + 3x - 2)^2$ in $\mathbb{Z}_{12}[x]$

**3.** Find all of the zeros for each of the following polynomials.

(a) $5x^3 + 4x^2 - x + 9$ in $\mathbb{Z}_{12}[x]$

(b) $3x^3 - 4x^2 - x + 4$ in $\mathbb{Z}_5[x]$

(c) $5x^4 + 2x^2 - 3$ in $\mathbb{Z}_7[x]$

(d) $x^3 + x + 1$ in $\mathbb{Z}_2[x]$

**4.** Find all of the units in $\mathbb{Z}[x]$.

**5.** Find a unit $p(x)$ in $\mathbb{Z}_4[x]$ such that $\deg p(x) > 1$.

**6.** Which of the following polynomials are irreducible over $\mathbb{Q}[x]$?

(a) $x^4 - 2x^3 + 2x^2 + x + 4$ (c) $3x^5 - 4x^3 - 6x^2 + 6$

(b) $x^4 - 5x^3 + 3x - 2$ (d) $5x^5 - 6x^4 - 3x^2 + 9x - 15$

**7.** Give two different factorizations of $x^2 + x + 8$ in $\mathbb{Z}_{10}[x]$.

**8.** Prove or disprove: There exists a polynomial $p(x)$ in $\mathbb{Z}_6[x]$ of degree $n$ with more than $n$ distinct zeros.

**9.** If $F$ is a field, show that $F[x_1, \ldots, x_n]$ is an integral domain.

**10.** Prove or disprove: $x^p + a$ is irreducible for any $a \in \mathbb{Z}_p$, where $p$ is prime.

**11.** Let $f(x)$ be irreducible in $F[x]$, where $F$ is a field. If $f(x) \mid p(x)q(x)$, prove that either $f(x) \mid p(x)$ or $f(x) \mid q(x)$.

**12.** Suppose that $R$ and $S$ are isomorphic rings. Prove that $R[x] \cong S[x]$.

**13.** Let $F$ be a field and $a \in F$. If $p(x) \in F[x]$, show that $p(a)$ is the remainder obtained when $p(x)$ is divided by $x - a$.

**14. The Rational Root Theorem.**Let

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x],$$

where $a_n \neq 0$. Prove that if $p(r/s) = 0$, where $\gcd(r, s) = 1$, then $r \mid a_0$ and $s \mid a_n$.

**15.** Let $\mathbb{Q}^*$ be the multiplicative group of positive rational numbers. Prove that $\mathbb{Q}^*$ is isomorphic to $(\mathbb{Z}[x], +)$.

**16. Cyclotomic Polynomials.**The polynomial

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

is called the **cyclotomic polynomial.** Show that $\Phi_p(x)$ is irreducible over $\mathbb{Q}$ for any prime $p$.

**17.** If $F$ is a field, show that there are infinitely many irreducible polynomials in $F[x]$.

**18.** Let $R$ be a commutative ring with identity. Prove that multiplication is commutative in $R[x]$.

**19.** Let $R$ be a commutative ring with identity. Prove that multiplication is distributive in $R[x]$.

**20.** Show that $x^p - x$ has $p$ distinct zeros in $\mathbb{Z}_p$, for any prime $p$. Conclude that

$$x^p - x = x(x - 1)(x - 2) \cdots (x - (p - 1)).$$

**21.** Let $F$ be a field and $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ be in $F[x]$. Define $f'(x) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}$ to be the **derivative** of $f(x)$.

(a) Prove that

$$(f + g)'(x) = f'(x) + g'(x).$$

Conclude that we can define a homomorphism of abelian groups $D : F[x] \to F[x]$ by $D(f(x)) = f'(x)$.

(b) Calculate the kernel of $D$ if $\operatorname{char} F = 0$.

(c) Calculate the kernel of $D$ if $\operatorname{char} F = p$.

(d) Prove that

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

(e) Suppose that we can factor a polynomial $f(x) \in F[x]$ into linear factors, say

$$f(x) = a(x - a_1)(x - a_2) \cdots (x - a_n).$$

Prove that $f(x)$ has no repeated factors if and only if $f(x)$ and $f'(x)$ are relatively prime.

**22.** Let $F$ be a field. Show that $F[x]$ is never a field.

**23.** Let $R$ be an integral domain. Prove that $R[x_1, \ldots, x_n]$ is an integral domain.

**24.** Let $R$ be a commutative ring with identity. Show that $R[x]$ has a subring $R'$ isomorphic to $R$.

**25.** Let $p(x)$ and $q(x)$ be polynomials in $R[x]$, where $R$ is a commutative ring with identity. Prove that $\deg(p(x) + q(x)) \leq \max(\deg p(x), \deg q(x))$.

**26.** Use the method of parallel computation outlined in the text to calculate 2234+4121 by dividing the calculation into four separate additions modulo 95, 97, 98, and 99.

**27.** Explain why the method of parallel computation outlined in the text fails for 2134·1531 if we attempt to break the calculation down into two smaller calculations modulo 98 and 99.

## 8.9 Material

## 8.10 Hints to Selected Exercises

### 8.7 · Core Exercises

**8.7.4.** The integers $\mathbb{Z}$ do not form a field.

### 8.8 · Additional Exercises

**8.8.2.** (a) $9x^2 + 2x + 5$; (b) $8x^4 + 7x^3 + 2x^2 + 7x$.

**8.8.3.** (a) No zeros in $\mathbb{Z}_{12}$; (c) 3, 4.

**8.8.5.** Look at $(2x + 1)$.

**8.8.6.** (a) Reducible; (c) irreducible.

**8.8.7.** One factorization is $x^2 + x + 8 = (x + 2)(x + 9)$.

**8.8.10.** False.

**8.8.12.** Let $\phi : R \to S$ be an isomorphism. Define $\overline{\phi} : R[x] \to S[x]$ by $\overline{\phi}(a_0 + a_1x + \cdots + a_nx^n) = \phi(a_0) + \phi(a_1)x + \cdots + \phi(a_n)x^n$.

**8.8.16. Cyclotomic Polynomials.**

The polynomial

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

is called the **cyclotomic polynomial.** Show that $\Phi_p(x)$ is irreducible over $\mathbb{Q}$ for any prime $p$.

**8.8.22.** Find a nontrivial proper ideal in $F[x]$.

# Chapter 9

# Vector Spaces and Field Extensions

Basic learning goals

1. Basic terminology and working with vector spaces and subspaces.
2. Basic terminology and working with linear (in)dependence and bases.
3. Constructions of extension fields.
4. Working with algebraic and transcendental elements.
5. Applying linear algebra to study extension fields.
6. Terminology of algebraic closure and splitting fields.

It is natural to ask whether or not some field $F$ is contained in a larger field. We think of the rational numbers, which reside inside the real numbers, while in turn, the real numbers live inside the complex numbers. We can also study the fields between $\mathbb{Q}$ and $\mathbb{R}$ and inquire as to the nature of these fields.

More specifically if we are given a field $F$ and a polynomial $p(x) \in F[x]$, we can ask whether or not we can find a field $E$ containing $F$ such that $p(x)$ factors into linear factors over $E[x]$. For example, if we consider the polynomial

$$p(x) = x^4 - 5x^2 + 6$$

in $\mathbb{Q}[x]$, then $p(x)$ factors as $(x^2 - 2)(x^2 - 3)$. However, both of these factors are irreducible in $\mathbb{Q}[x]$. If we wish to find a zero of $p(x)$, we must go to a larger field. Certainly the field of real numbers will work, since

$$p(x) = (x - \sqrt{2})(x + \sqrt{2})(x - \sqrt{3})(x + \sqrt{3}).$$

It is possible to find a smaller field in which $p(x)$ has a zero, namely

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

We wish to be able to compute and study such fields for arbitrary polynomials over a field $F$.

We start by recalling the mathematical structures called vector spaces. As with groups and rings, they are given by a simple list of axioms.

## 9.1 Vector Spaces

A **vector space** $V$ over a field $F$ is an abelian group with a **scalar multiplication** $\alpha \cdot v$ or $\alpha v$ defined for all $\alpha \in F$ and all $v \in V$ satisfying the following axioms.

- $\alpha(\beta v) = (\alpha\beta)v$;
- $(\alpha + \beta)v = \alpha v + \beta v$;
- $\alpha(u + v) = \alpha u + \alpha v$;
- $1v = v$;

where $\alpha, \beta \in F$ and $u, v \in V$.

The elements of $V$ are called **vectors**; the elements of $F$ are called **scalars**. It is important to notice that in most cases two vectors cannot be multiplied. In general, it is only possible to multiply a vector with a scalar. To differentiate between the scalar zero and the vector zero, we will usually write them as 0 and **0**, respectively.

Let us examine several examples of vector spaces. Some of them will be quite familiar; others will seem less so.

Note that scalar multiplication is a special form of a group action of the multiplicative group $(F \setminus \{0\}, \cdot)$ on $V$.

**Example 9.1.1** The $n$-tuples of real numbers, denoted by $\mathbb{R}^n$, form a vector space over $\mathbb{R}$. Given vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ and $\alpha$ in $\mathbb{R}$, we can define vector addition by

$$u + v = (u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)$$

and scalar multiplication by

$$\alpha u = \alpha(u_1, \ldots, u_n) = (\alpha u_1, \ldots, \alpha u_n).$$

□

**Example 9.1.2** If $F$ is a field, then $F[x]$ is a vector space over $F$. The vectors in $F[x]$ are simply polynomials, and vector addition is just polynomial addition. If $\alpha \in F$ and $p(x) \in F[x]$, then scalar multiplication is defined by $\alpha p(x)$. □

**Example 9.1.3** The set of all continuous real-valued functions on a closed interval $[a, b]$ is a vector space over $\mathbb{R}$. If $f(x)$ and $g(x)$ are continuous on $[a, b]$, then $(f + g)(x)$ is defined to be $f(x) + g(x)$. Scalar multiplication is defined by $(\alpha f)(x) = \alpha f(x)$ for $\alpha \in \mathbb{R}$. For example, if $f(x) = \sin x$ and $g(x) = x^2$, then $(2f + 5g)(x) = 2\sin x + 5x^2$. □

**Example 9.1.4** Let $V = \mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. Then $V$ is a vector space over $\mathbb{Q}$. If $u = a + b\sqrt{2}$ and $v = c + d\sqrt{2}$, then $u + v = (a + c) + (b + d)\sqrt{2}$ is again in $V$. Also, for $\alpha \in \mathbb{Q}$, $\alpha v$ is in $V$. We will leave it as an exercise to verify that all of the vector space axioms hold for $V$. □

**Proposition 9.1.5** *Let $V$ be a vector space over $F$. Then each of the following statements is true.*

1. $0v = \mathbf{0}$ *for all* $v \in V$.
2. $\alpha\mathbf{0} = \mathbf{0}$ *for all* $\alpha \in F$.
3. *If* $\alpha v = \mathbf{0}$, *then either* $\alpha = 0$ *or* $v = \mathbf{0}$.
4. $(-1)v = -v$ *for all* $v \in V$.

5. $-(\alpha v) = (-\alpha)v = \alpha(-v)$ *for all* $\alpha \in F$ *and all* $v \in V$.

*Proof.* To prove (1), observe that

$$0v = (0 + 0)v = 0v + 0v;$$

consequently, $\mathbf{0} + 0v = 0v + 0v$. Since $V$ is an abelian group, $\mathbf{0} = 0v$.

The proof of (2) is almost identical to the proof of (1). For (3), we are done if $\alpha = 0$. Suppose that $\alpha \neq 0$. Multiplying both sides of $\alpha v = \mathbf{0}$ by $1/\alpha$, we have $v = \mathbf{0}$.

To show (4), observe that

$$v + (-1)v = 1v + (-1)v = (1 - 1)v = 0v = \mathbf{0},$$

and so $-v = (-1)v$. We will leave the proof of (5) as an exercise. ■

## 9.2 Subspaces

Just as groups have subgroups and rings have subrings, vector spaces also have substructures. Let $V$ be a vector space over a field $F$, and $W$ a subset of $V$. Then $W$ is a **subspace** of $V$ if it is closed under vector addition and scalar multiplication; that is, if $u, v \in W$ and $\alpha \in F$, it will always be the case that $u + v$ and $\alpha v$ are also in $W$.

**Example 9.2.1** Let $W$ be the subspace of $\mathbb{R}^3$ defined by $W = \{(x_1, 2x_1 + x_2, x_1 - x_2) : x_1, x_2 \in \mathbb{R}\}$. We claim that $W$ is a subspace of $\mathbb{R}^3$. Since

$$\begin{aligned}\alpha(x_1, 2x_1 + x_2, x_1 - x_2) &= (\alpha x_1, \alpha(2x_1 + x_2), \alpha(x_1 - x_2)) \\ &= (\alpha x_1, 2(\alpha x_1) + \alpha x_2, \alpha x_1 - \alpha x_2),\end{aligned}$$

$W$ is closed under scalar multiplication. To show that $W$ is closed under vector addition, let $u = (x_1, 2x_1 + x_2, x_1 - x_2)$ and $v = (y_1, 2y_1 + y_2, y_1 - y_2)$ be vectors in $W$. Then

$$u + v = (x_1 + y_1, 2(x_1 + y_1) + (x_2 + y_2), (x_1 + y_1) - (x_2 + y_2)).$$

□

**Example 9.2.2** Let $W$ be the subset of polynomials of $F[x]$ with no odd-power terms. If $p(x)$ and $q(x)$ have no odd-power terms, then neither will $p(x) + q(x)$. Also, $\alpha p(x) \in W$ for $\alpha \in F$ and $p(x) \in W$. □

Let $V$ be any vector space over a field $F$ and suppose that $v_1, v_2, \ldots, v_n$ are vectors in $V$ and $\alpha_1, \alpha_2, \ldots, \alpha_n$ are scalars in $F$. Any vector $w$ in $V$ of the form

$$w = \sum_{i=1}^{n} \alpha_i v_i = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

is called a **linear combination** of the vectors $v_1, v_2, \ldots, v_n$. The **span** of vectors $v_1, v_2, \ldots, v_n$ is the set of vectors obtained from all possible linear combinations of $v_1, v_2, \ldots, v_n$. If $W$ is the span set of $v_1, v_2, \ldots, v_n$, then we say that $W$ is **spanned** by $v_1, v_2, \ldots, v_n$.

**Proposition 9.2.3** *Let* $S = \{v_1, v_2, \ldots, v_n\}$ *be vectors in a vector space* $V$. *Then the span of* $S$ *is a subspace of* $V$.

*Proof.* Let $u$ and $v$ be in $S$. We can write both of these vectors as linear

combinations of the $v_i$'s:

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$
$$v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n.$$

Then

$$u + v = (\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + \cdots + (\alpha_n + \beta_n)v_n$$

is a linear combination of the $v_i$'s. For $\alpha \in F$,

$$\alpha u = (\alpha\alpha_1)v_1 + (\alpha\alpha_2)v_2 + \cdots + (\alpha\alpha_n)v_n$$

is in the span of $S$. ■

## 9.3 Linear Independence

Let $S = \{v_1, v_2, \ldots, v_n\}$ be a set of vectors in a vector space $V$. If there exist scalars $\alpha_1, \alpha_2 \ldots \alpha_n \in F$ such that not all of the $\alpha_i$'s are zero and

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \mathbf{0},$$

then $S$ is said to be **linearly dependent**. If the set $S$ is not linearly dependent, then it is said to be **linearly independent**. More specifically, $S$ is a linearly independent set if

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \mathbf{0}$$

implies that

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

for any set of scalars $\{\alpha_1, \alpha_2 \ldots \alpha_n\}$.

**Proposition 9.3.1** *Let $\{v_1, v_2, \ldots, v_n\}$ be a set of linearly independent vectors in a vector space. Suppose that*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n.$$

*Then $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$.*

*Proof.* If

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n,$$

then

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_n - \beta_n)v_n = \mathbf{0}.$$

Since $v_1, \ldots, v_n$ are linearly independent, $\alpha_i - \beta_i = 0$ for $i = 1, \ldots, n$. ■

The definition of linear dependence makes more sense if we consider the following proposition.

**Proposition 9.3.2** *A set $\{v_1, v_2, \ldots, v_n\}$ of vectors in a vector space $V$ is linearly dependent if and only if one of the $v_i$'s is a linear combination of the rest.*

*Proof.* Suppose that $\{v_1, v_2, \ldots, v_n\}$ is a set of linearly dependent vectors. Then there exist scalars $\alpha_1, \ldots, \alpha_n$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \mathbf{0},$$

with at least one of the $\alpha_i$'s not equal to zero. Suppose that $\alpha_k \neq 0$. Then

$$v_k = -\frac{\alpha_1}{\alpha_k} v_1 - \cdots - \frac{\alpha_{k-1}}{\alpha_k} v_{k-1} - \frac{\alpha_{k+1}}{\alpha_k} v_{k+1} - \cdots - \frac{\alpha_n}{\alpha_k} v_n.$$

Conversely, suppose that

$$v_k = \beta_1 v_1 + \cdots + \beta_{k-1} v_{k-1} + \beta_{k+1} v_{k+1} + \cdots + \beta_n v_n.$$

Then

$$\beta_1 v_1 + \cdots + \beta_{k-1} v_{k-1} - v_k + \beta_{k+1} v_{k+1} + \cdots + \beta_n v_n = \mathbf{0}.$$

■

The following proposition is a consequence of the fact that any system of homogeneous linear equations with more unknowns than equations will have a nontrivial solution. We leave the details of the proof for the end-of-chapter exercises.

**Proposition 9.3.3** *Suppose that a vector space $V$ is spanned by $n$ vectors. If $m > n$, then any set of $m$ vectors in $V$ must be linearly dependent.*

A set $\{e_1, e_2, \ldots, e_n\}$ of vectors in a vector space $V$ is called a **basis** for $V$ if $\{e_1, e_2, \ldots, e_n\}$ is a linearly independent set that spans $V$.

**Example 9.3.4** The vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$ form a basis for $\mathbb{R}^3$. The set certainly spans $\mathbb{R}^3$, since any arbitrary vector $(x_1, x_2, x_3)$ in $\mathbb{R}^3$ can be written as $x_1e_1 + x_2e_2 + x_3e_3$. Also, none of the vectors $e_1, e_2, e_3$ can be written as a linear combination of the other two; hence, they are linearly independent. The vectors $e_1, e_2, e_3$ are not the only basis of $\mathbb{R}^3$: the set $\{(3, 2, 1), (3, 2, 0), (1, 1, 1)\}$ is also a basis for $\mathbb{R}^3$. □

**Example 9.3.5** Let $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. The sets $\{1, \sqrt{2}\}$ and $\{1 + \sqrt{2}, 1 - \sqrt{2}\}$ are both bases of $\mathbb{Q}(\sqrt{2})$. □

From the last two examples it should be clear that a given vector space has several bases. In fact, there are an infinite number of bases for both of these examples. *In general, there is no unique basis for a vector space.* However, every basis of $\mathbb{R}^3$ consists of exactly three vectors, and every basis of $\mathbb{Q}(\sqrt{2})$ consists of exactly two vectors. This is a consequence of the next proposition.

**Proposition 9.3.6** *Let $\{e_1, e_2, \ldots, e_m\}$ and $\{f_1, f_2, \ldots, f_n\}$ be two bases for a vector space $V$. Then $m = n$.*

*Proof.* Since $\{e_1, e_2, \ldots, e_m\}$ is a basis, it is a linearly independent set. By Proposition 9.3.3, $n \leq m$. Similarly, $\{f_1, f_2, \ldots, f_n\}$ is a linearly independent set, and the last proposition implies that $m \leq n$. Consequently, $m = n$. ■

If $\{e_1, e_2, \ldots, e_n\}$ is a basis for a vector space $V$, then we say that the **dimension** of $V$ is $n$ and we write $\dim V = n$. We will leave the proof of the following theorem as an exercise.

**Theorem 9.3.7** *Let $V$ be a vector space of dimension $n$.*

1. *If $S = \{v_1, \ldots, v_n\}$ is a set of linearly independent vectors for $V$, then $S$ is a basis for $V$.*
2. *If $S = \{v_1, \ldots, v_n\}$ spans $V$, then $S$ is a basis for $V$.*
3. *If $S = \{v_1, \ldots, v_k\}$ is a set of linearly independent vectors for $V$ with $k < n$, then there exist vectors $v_{k+1}, \ldots, v_n$ such that*

$$\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$$

*is a basis for $V$.*

## 9.4 Extension Fields

A field $F$ is a **subfield** of a field $E$, if $F$ is a subring of $E$ and it is also a field. Then the field $E$ is an **extension field** of $F$. The field $F$ is called the **base field**. We write $F \subset E$.

If $E$ is a field extension of $F$ and $S$ is a subset of $E$, then we denote the smallest field containing $F$ and $S$ by $F(S)$. If $S$ is finite, so it is given by elements $\alpha_1, \ldots, \alpha_n$ in $E$, we denote the smallest field containing $F$ and $\alpha_1, \ldots, \alpha_n$ by $F(\alpha_1, \ldots, \alpha_n)$. We say that $F(S)$ arises from the **adjunction** of $S$ to $F$. If $E = F(\alpha)$ for some $\alpha \in E$, then $E$ is a **simple extension** of $F$.

Let $E$ be a field extension of a field $F$. We can regard $E$ as a vector space over $F$ and so bring the machinery of linear algebra to bear on the problems that we will encounter in our study of fields. With this point of view, the elements in the field $E$ are vectors and the elements in the field $F$ are scalars. We can think of addition in $E$ as adding vectors. When we multiply an element in $E$ by an element of $F$, we are multiplying a vector by a scalar.

**Example 9.4.1** For example, let

$$F = \mathbb{Q}(\sqrt{2}\,) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

and let $E = \mathbb{Q}(\sqrt{2} + \sqrt{3}\,)$ be the smallest field containing both $\mathbb{Q}$ and $\sqrt{2} + \sqrt{3}$. Both $E$ and $F$ are extension fields of the rational numbers. We claim that $E$ is an extension field of $F$. To see this, we need only show that $\sqrt{2}$ is in $E$. Since $\sqrt{2} + \sqrt{3}$ is in $E$, $1/(\sqrt{2} + \sqrt{3}\,) = \sqrt{3} - \sqrt{2}$ must also be in $E$. Taking linear combinations of $\sqrt{2} + \sqrt{3}$ and $\sqrt{3} - \sqrt{2}$, we find that $\sqrt{2}$ and $\sqrt{3}$ must both be in $E$. □

**Example 9.4.2** Let $p(x) = x^2 + x + 1 \in \mathbb{Z}_2[x]$. Since neither 0 nor 1 is a root of this polynomial, we know that $p(x)$ is irreducible over $\mathbb{Z}_2$. We will construct a field extension of $\mathbb{Z}_2$ containing an element $\alpha$ such that $p(\alpha) = 0$. By Theorem 8.4.9, the ideal $\langle p(x) \rangle$ generated by $p(x)$ is maximal; hence, $\mathbb{Z}_2[x]/\langle p(x) \rangle$ is a field. Let $f(x) + \langle p(x) \rangle$ be an arbitrary element of $\mathbb{Z}_2[x]/\langle p(x) \rangle$. By the division algorithm,

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

where the degree of $r(x)$ is less than the degree of $x^2 + x + 1$. Therefore,

$$f(x) + \langle x^2 + x + 1 \rangle = r(x) + \langle x^2 + x + 1 \rangle.$$

The only possibilities for $r(x)$ are then 0, 1, $x$, and $1 + x$. Consequently, $E = \mathbb{Z}_2[x]/\langle x^2 + x + 1 \rangle$ is a field with four elements and must be a field extension of $\mathbb{Z}_2$, containing a zero $\alpha$ of $p(x)$. The field $\mathbb{Z}_2(\alpha)$ consists of elements

$$\begin{aligned} 0 + 0\alpha &= 0 \\ 1 + 0\alpha &= 1 \\ 0 + 1\alpha &= \alpha \\ 1 + 1\alpha &= 1 + \alpha. \end{aligned}$$

Notice that $\alpha^2 + \alpha + 1 = 0$; hence, if we compute $(1 + \alpha)^2$,

$$(1 + \alpha)(1 + \alpha) = 1 + \alpha + \alpha + (\alpha)^2 = \alpha.$$

Other calculations are accomplished in a similar manner. We summarize these computations in the following tables, which tell us how to add and multiply elements in $E$. □

| $+$ | $0$ | $1$ | $\alpha$ | $1+\alpha$ |
|---|---|---|---|---|
| $0$ | $0$ | $1$ | $\alpha$ | $1+\alpha$ |
| $1$ | $1$ | $0$ | $1+\alpha$ | $\alpha$ |
| $\alpha$ | $\alpha$ | $1+\alpha$ | $0$ | $1$ |
| $1+\alpha$ | $1+\alpha$ | $\alpha$ | $1$ | $0$ |

**Figure 9.4.3** Addition Table for $\mathbb{Z}_2(\alpha)$

| $\cdot$ | $0$ | $1$ | $\alpha$ | $1+\alpha$ |
|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $0$ |
| $1$ | $0$ | $1$ | $\alpha$ | $1+\alpha$ |
| $\alpha$ | $0$ | $\alpha$ | $1+\alpha$ | $1$ |
| $1+\alpha$ | $0$ | $1+\alpha$ | $1$ | $\alpha$ |

**Figure 9.4.4** Multiplication Table for $\mathbb{Z}_2(\alpha)$

The following theorem, due to Kronecker, is so important and so basic to our understanding of fields that it is often known as the Fundamental Theorem of Field Theory.

**Theorem 9.4.5** *Let $F$ be a field and let $p(x)$ be a nonconstant polynomial in $F[x]$. Then there exists an extension field $E$ of $F$ and an element $\alpha \in E$ such that $p(\alpha) = 0$.*

*Proof.* To prove this theorem, we will employ the method that we used to construct Example 9.4.2. By picking an irreducible factor of $p(x)$ instead, we can assume that $p(x)$ is an irreducible polynomial. We wish to find an extension field $E$ of $F$ containing an element $\alpha$ such that $p(\alpha) = 0$. The ideal $\langle p(x)\rangle$ generated by $p(x)$ is a maximal ideal in $F[x]$ by Theorem 8.4.9; hence, $F[x]/\langle p(x)\rangle$ is a field. We claim that $E = F[x]/\langle p(x)\rangle$ is the desired field.

We first show that $E$ is a field extension of $F$. We can define a homomorphism of commutative rings by the map $\psi : F \to F[x]/\langle p(x)\rangle$, where $\psi(a) = a + \langle p(x)\rangle$ for $a \in F$. It is easy to check that $\psi$ is indeed a ring homomorphism. Observe that

$$\psi(a) + \psi(b) = (a + \langle p(x)\rangle) + (b + \langle p(x)\rangle) = (a + b) + \langle p(x)\rangle = \psi(a + b)$$

and

$$\psi(a)\psi(b) = (a + \langle p(x)\rangle)(b + \langle p(x)\rangle) = ab + \langle p(x)\rangle = \psi(ab).$$

To prove that $\psi$ is one-to-one, assume that

$$a + \langle p(x)\rangle = \psi(a) = \psi(b) = b + \langle p(x)\rangle.$$

Then $a - b$ is a multiple of $p(x)$, since it lives in the ideal $\langle p(x)\rangle$. Since $p(x)$ is a nonconstant polynomial, the only possibility is that $a - b = 0$. Consequently, $a = b$ and $\psi$ is injective. Since $\psi$ is one-to-one, we can identify $F$ with the subfield $\{a + \langle p(x)\rangle : a \in F\}$ of $E$ and view $E$ as an extension field of $F$.

It remains for us to prove that $p(x)$ has a zero $\alpha \in E$. Set $\alpha = x + \langle p(x)\rangle$. Then $\alpha$ is in $E$. If $p(x) = a_0 + a_1x + \cdots + a_nx^n$, then

$$\begin{aligned} p(\alpha) &= a_0 + a_1(x + \langle p(x)\rangle) + \cdots + a_n(x + \langle p(x)\rangle)^n \\ &= a_0 + (a_1x + \langle p(x)\rangle) + \cdots + (a_nx^n + \langle p(x)\rangle) \\ &= a_0 + a_1x + \cdots + a_nx^n + \langle p(x)\rangle \\ &= 0 + \langle p(x)\rangle. \end{aligned}$$

Therefore, we have found an element $\alpha \in E = F[x]/\langle p(x)\rangle$ such that $\alpha$ is a zero of $p(x)$. ∎

**Example 9.4.6** Let $p(x) = x^5 + x^4 + 1 \in \mathbb{Z}_2[x]$. Then $p(x)$ has irreducible factors $x^2 + x + 1$ and $x^3 + x + 1$. For a field extension $E$ of $\mathbb{Z}_2$ such that $p(x)$ has a root in $E$, we can let $E$ be either $\mathbb{Z}_2[x]/\langle x^2 + x + 1\rangle$ or $\mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle$. We will leave it as an exercise to show that $\mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle$ is a field with $2^3 = 8$ elements. □

**Example 9.4.7** Recall how to compute multiplicative inverses in $\mathbb{Z}_p$ of an element $k \in \{1, 2, \ldots, p-1\}$ for a prime number $p$. For this, one can compute two integers $r, s \in \mathbb{Z}$ such that $rp + sk = 1$. Then $rp + sk \equiv sk \equiv 1 \pmod{p}$ and the congruence class of $s$ is the multiplicative inverse of $k$. Note that the congruence is actually given by the isomorphism $\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z}$.

Now, let $F$ be a field and $q(x) \in F[x]$ be an irreducible polynomial in $F[x]$. Then $F[x]/\langle q(x)\rangle$ forms a field as well. Let $\ell(x)$ be a polynomial in $F[x]$ with $\deg \ell(x) < \deg q(x)$. This gives rise to an element $\ell(x) + \langle q(x)\rangle$ in $F[x]/\langle q(x)\rangle$. Since $q(x)$ is irreducible and $\ell(x)$ has smaller degree, the two polynomials are relatively prime. Analogously to the above, we know that there exist polynomials $v(x), w(x) \in F[x]$ such that $w(x)\ell(x) + v(x)q(x) = 1$. Then the element $w(x) + \langle q(x)\rangle \in F[x]/\langle q(x)\rangle$ is the multiplicative inverse of $\ell(x) + \langle q(x)\rangle$:

$$\begin{aligned}(\ell(x) + \langle q(x)\rangle)\,(w(x) + \langle q(x)\rangle) &= \ell(x)w(x) + \langle q(x)\rangle \\ &= (\ell(x)w(x) + v(x)q(x)) + \langle q(x)\rangle = 1 + \langle q(x)\rangle.\end{aligned}$$

To illustrate this, we apply it to Example 9.4.2. We determine the multiplicative inverse of $x+\langle x^2+x+1\rangle$ in $\mathbb{Z}_2[x]/\langle x^2+x+1\rangle$. Note that $(x+1)x+(x^2+x+1) = 1$. Hence, the multiplicative inverse is $x+1+\langle x^2+x+1\rangle$. One can also see this in the multiplication table Figure 9.4.4 as there $\alpha(1+\alpha) = (1+\alpha)\alpha = 1$. □

### 9.4.1 Algebraic Elements

An element $\alpha$ in an extension field $E$ over $F$ is **algebraic** over $F$ if $f(\alpha) = 0$ for some nonzero polynomial $f(x) \in F[x]$. An element in $E$ that is not algebraic over $F$ is **transcendental** over $F$. An extension field $E$ of a field $F$ is an **algebraic extension** of $F$ if every element in $E$ is algebraic over $F$.

**Example 9.4.8** Both $\sqrt{2}$ and $i$ are algebraic over $\mathbb{Q}$ since they are zeros of the polynomials $x^2 - 2$ and $x^2 + 1$, respectively. Clearly $\pi$ and $e$ are algebraic over the real numbers; however, it is a nontrivial fact that they are transcendental over $\mathbb{Q}$. Numbers in $\mathbb{R}$ that are algebraic over $\mathbb{Q}$ are in fact quite rare. Almost all real numbers are transcendental over $\mathbb{Q}$.[1](In many cases we do not know whether or not a particular number is transcendental; for example, it is still not known whether $\pi + e$ is transcendental or algebraic.) □

A complex number that is algebraic over $\mathbb{Q}$ is an **algebraic number**. A **transcendental number** is an element of $\mathbb{C}$ that is transcendental over $\mathbb{Q}$.

Note that algebraic numbers are essentially those, which we can actually encode, at least by its polynomial that has it as a zero.

**Example 9.4.9** We will show that $\sqrt{2+\sqrt{3}}$ is algebraic over $\mathbb{Q}$. If $\alpha = \sqrt{2+\sqrt{3}}$, then $\alpha^2 = 2 + \sqrt{3}$. Hence, $\alpha^2 - 2 = \sqrt{3}$ and $(\alpha^2 - 2)^2 = 3$. Since $\alpha^4 - 4\alpha^2 + 1 = 0$, it must be true that $\alpha$ is a zero of the polynomial $x^4 - 4x^2 + 1 \in \mathbb{Q}[x]$. □

[1]The probability that a real number chosen at random from the interval $[0, 1]$ will be transcendental over the rational numbers is one.

It is very easy to give an example of an extension field $E$ over a field $F$, where $E$ contains an element transcendental over $F$. The following theorem characterizes transcendental extensions.

**Theorem 9.4.10** *Let $E$ be an extension field of $F$ and $\alpha \in E$. Then $\alpha$ is transcendental over $F$ if and only if $F(\alpha)$ is isomorphic to*

$$F(x) = \left\{ \frac{p(x)}{q(x)} \mid p(x), q(x) \in F[x], q(x) \neq 0 \right\},$$

*the field of rational functions.*

*Proof.* Let $\phi_\alpha : F[x] \to E$ be the evaluation homomorphism for $\alpha$. Then $\alpha$ is transcendental over $F$ if and only if $\phi_\alpha(p(x)) = p(\alpha) \neq 0$ for all nonconstant polynomials $p(x) \in F[x]$. This is true if and only if $\ker \phi_\alpha = \{0\}$; that is, it is true exactly when $\phi_\alpha$ is one-to-one. Hence, $E$ must contain a copy of $F[x]$. The smallest field containing $F[x]$ is the field of fractions $F(x)$, so $E$ must contain a copy of this field. [2] ■

We have a more interesting situation in the case of algebraic extensions.

**Theorem 9.4.11** *Let $E$ be an extension field of a field $F$ and $\alpha \in E$ with $\alpha$ algebraic over $F$. Then there is a unique irreducible monic polynomial $p(x) \in F[x]$ of smallest degree such that $p(\alpha) = 0$. If $f(x)$ is another polynomial in $F[x]$ such that $f(\alpha) = 0$, then $p(x)$ divides $f(x)$.*

Let $E$ be an extension field of $F$ and $\alpha \in E$ be algebraic over $F$. The unique monic polynomial $p(x)$ of the last theorem is called the **minimal polynomial** for $\alpha$ over $F$. The degree of $p(x)$ is the **degree of $\alpha$ over $F$**.

**Example 9.4.12** Let $f(x) = x^2 - 2$ and $g(x) = x^4 - 4x^2 + 1$. These polynomials are the minimal polynomials of $\sqrt{2}$ and $\sqrt{2+\sqrt{3}}$, respectively. □

**Proposition 9.4.13** *Let $E$ be a field extension of $F$ and $\alpha \in E$ be algebraic over $F$. Then $F(\alpha) \cong F[x]/\langle p(x) \rangle$, where $p(x)$ is the minimal polynomial of $\alpha$ over $F$.*

*Proof.* Let $\phi_\alpha : F[x] \to E$ be the evaluation homomorphism. The kernel of this map is $\langle p(x) \rangle$, where $p(x)$ is the minimal polynomial of $\alpha$. By the First Isomorphism Theorem for rings, the image of $\phi_\alpha$ in $E$ is isomorphic to $F(\alpha)$ since it contains both $F$ and $\alpha$. ■

**Theorem 9.4.14** *Let $E = F(\alpha)$ be a simple extension of $F$, where $\alpha \in E$ is algebraic over $F$. Suppose that the degree of $\alpha$ over $F$ is $n$. Then every element $\beta \in E$ can be expressed uniquely in the form*

$$\beta = b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1}$$

*for $b_i \in F$.*

*Proof.* Since $\phi_\alpha(F[x]) \cong F(\alpha)$, every element in $E = F(\alpha)$ must be of the form $\phi_\alpha(f(x)) = f(\alpha)$, where $f(\alpha)$ is a polynomial in $\alpha$ with coefficients in $F$. Let

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

be the minimal polynomial of $\alpha$. Then $p(\alpha) = 0$; hence,

$$\alpha^n = -a_{n-1}\alpha^{n-1} - \cdots - a_0.$$

[2] The last sentence of this proof omits several details but the formal argument would go beyond the scope of this course. It involves the construction of the field of fractions.

Similarly,

$$\begin{aligned}\alpha^{n+1} &= \alpha\alpha^n \\ &= -a_{n-1}\alpha^n - a_{n-2}\alpha^{n-1} - \cdots - a_0\alpha \\ &= -a_{n-1}(-a_{n-1}\alpha^{n-1} - \cdots - a_0) - a_{n-2}\alpha^{n-1} - \cdots - a_0\alpha.\end{aligned}$$

Continuing in this manner, we can express every monomial $\alpha^m$, $m \geq n$, as a linear combination of powers of $\alpha$ that are less than $n$. Hence, any $\beta \in F(\alpha)$ can be written as

$$\beta = b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1}.$$

To show uniqueness, suppose that

$$\beta = b_0 + b_1\alpha + \cdots + b_{n-1}\alpha^{n-1} = c_0 + c_1\alpha + \cdots + c_{n-1}\alpha^{n-1}$$

for $b_i$ and $c_i$ in $F$. Then

$$g(x) = (b_0 - c_0) + (b_1 - c_1)x + \cdots + (b_{n-1} - c_{n-1})x^{n-1}$$

is in $F[x]$ and $g(\alpha) = 0$. Since the degree of $g(x)$ is less than the degree of $p(x)$, the irreducible polynomial of $\alpha$, $g(x)$ must be the zero polynomial. Consequently,

$$b_0 - c_0 = b_1 - c_1 = \cdots = b_{n-1} - c_{n-1} = 0,$$

or $b_i = c_i$ for $i = 0, 1, \ldots, n-1$. Therefore, we have shown uniqueness. ■

**Example 9.4.15** Since $x^2+1$ is irreducible over $\mathbb{R}$, $\langle x^2+1\rangle$ is a maximal ideal in $\mathbb{R}[x]$. So $E = \mathbb{R}[x]/\langle x^2+1\rangle$ is a field extension of $\mathbb{R}$ that contains a root of $x^2+1$. Let $\alpha = x + \langle x^2+1\rangle$. We can identify $E$ with the complex numbers. By Proposition 9.4.13, $E$ is isomorphic to $\mathbb{R}(\alpha) = \{a + b\alpha : a, b \in \mathbb{R}\}$. We know that $\alpha^2 = -1$ in $E$, since

$$\begin{aligned}\alpha^2 + 1 &= (x + \langle x^2+1\rangle)^2 + (1 + \langle x^2+1\rangle) \\ &= (x^2+1) + \langle x^2+1\rangle \\ &= 0.\end{aligned}$$

Hence, we have an isomorphism of $\mathbb{R}(\alpha)$ with $\mathbb{C}$ defined by the map that takes $a + b\alpha$ to $a + bi$. □

The view of a field extension $E$ over a field $F$ as a vector space is especially fruitful if it is a finite dimensional vector space over $F$. In particular, Theorem 9.4.14 states that $E = F(\alpha)$ is finite dimensional vector space over $F$ with basis $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$.

If an extension field $E$ of a field $F$ is a finite dimensional vector space over $F$ of dimension $n$, then we say that $E$ is a **finite extension of degree** $n$ **over** $F$. We write

$$[E : F] = n.$$

to indicate the dimension of $E$ over $F$.

**Theorem 9.4.16** *Every finite extension field $E$ of a field $F$ is an algebraic extension.*

*Proof.* Let $\alpha \in E$. Since $[E : F] = n$, the elements

$$1, \alpha, \ldots, \alpha^n$$

cannot be linearly independent. Hence, there exist $a_i \in F$, not all zero, such

that

$$a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0 = 0.$$

Therefore,

$$p(x) = a_n x^n + \cdots + a_0 \in F[x]$$

is a nonzero polynomial with $p(\alpha) = 0$. ■

**Remark 9.4.17** Theorem 9.4.16 says that every finite extension of a field $F$ is an algebraic extension. The converse is false, however. We will leave it as an exercise to show that the set of all elements in $\mathbb{R}$ that are algebraic over $\mathbb{Q}$ forms an infinite field extension of $\mathbb{Q}$.

The next theorem is a counting theorem, similar to Lagrange's Theorem in group theory. Theorem 9.4.18 will prove to be an extremely useful tool in our investigation of finite field extensions.

**Theorem 9.4.18** *If $E$ is a finite extension of $F$ and $K$ is a finite extension of $E$, then $K$ is a finite extension of $F$ and*

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

The following corollary is easily proved using mathematical induction.

**Corollary 9.4.19** *If $F_i$ is a field for $i = 1, \ldots, k$ and $F_{i+1}$ is a finite extension of $F_i$, then $F_k$ is a finite extension of $F_1$ and*

$$[F_k : F_1] = [F_k : F_{k-1}] \cdots [F_2 : F_1].$$

**Example 9.4.20** Let us determine an extension field of $\mathbb{Q}$ containing $\sqrt{3}+\sqrt{5}$. It is easy to determine that the minimal polynomial of $\sqrt{3}+\sqrt{5}$ is $x^4-16x^2+4$. It follows that

$$[\mathbb{Q}(\sqrt{3}+\sqrt{5}\,) : \mathbb{Q}] = 4.$$

We know that $\{1, \sqrt{3}\,\}$ is a basis for $\mathbb{Q}(\sqrt{3}\,)$ over $\mathbb{Q}$. Hence, $\sqrt{3}+\sqrt{5}$ cannot be in $\mathbb{Q}(\sqrt{3}\,)$. It follows that $\sqrt{5}$ cannot be in $\mathbb{Q}(\sqrt{3}\,)$ either. Therefore, $\{1, \sqrt{5}\,\}$ is a basis for $\mathbb{Q}(\sqrt{3}, \sqrt{5}\,) = (\mathbb{Q}(\sqrt{3}\,))(\sqrt{5}\,)$ over $\mathbb{Q}(\sqrt{3}\,)$ and $\{1, \sqrt{3}, \sqrt{5}, \sqrt{3}\sqrt{5} = \sqrt{15}\,\}$ is a basis for $\mathbb{Q}(\sqrt{3}, \sqrt{5}\,) = \mathbb{Q}(\sqrt{3}+\sqrt{5}\,)$ over $\mathbb{Q}$. This example shows that it is possible that some extension $F(\alpha_1, \ldots, \alpha_n)$ is actually a simple extension of $F$ even though $n > 1$. □

**Example 9.4.21** Let us compute a basis for $\mathbb{Q}(\sqrt[3]{5}, \sqrt{5}\,i)$, where $\sqrt{5}$ is the positive square root of 5 and $\sqrt[3]{5}$ is the real cube root of 5. We know that $\sqrt{5}\,i \notin \mathbb{Q}(\sqrt[3]{5}\,)$, so

$$[\mathbb{Q}(\sqrt[3]{5}, \sqrt{5}\,i) : \mathbb{Q}(\sqrt[3]{5}\,)] = 2.$$

It is easy to determine that $\{1, \sqrt{5}i\,\}$ is a basis for $\mathbb{Q}(\sqrt[3]{5}, \sqrt{5}\,i)$ over $\mathbb{Q}(\sqrt[3]{5}\,)$. We also know that $\{1, \sqrt[3]{5}, (\sqrt[3]{5}\,)^2\}$ is a basis for $\mathbb{Q}(\sqrt[3]{5}\,)$ over $\mathbb{Q}$. Hence, a basis for $\mathbb{Q}(\sqrt[3]{5}, \sqrt{5}\,i)$ over $\mathbb{Q}$ is

$$\{1, \sqrt{5}\,i, \sqrt[3]{5}, (\sqrt[3]{5}\,)^2, (\sqrt[6]{5}\,)^5 i, (\sqrt[6]{5}\,)^7 i\},$$

where $(\sqrt[6]{5}\,)^7 i = 5\sqrt[6]{5}\,i$ so that we could also use $\sqrt[6]{5}\,i$. □

## 9.4.2 Algebraic Closure

We include a short discussion of algebraic closures. More details are not part of the core topics and they are provided in Subsection 9.5.2.

Let $E$ be a field extension of a field $F$. We define the **algebraic closure** of a field $F$ in $E$ to be the field consisting of all elements in $E$ that are algebraic over $F$. A field $F$ is **algebraically closed** if every nonconstant polynomial

in $F[x]$ has a root in $F$. The smallest algebraically closed extension field of a field is its **algebraic closure**. We list two important properties.

**Theorem 9.4.22**

1. *A field $F$ is algebraically closed if and only if every nonconstant polynomial in $F[x]$ factors into linear factors over $F[x]$.*
2. *Every field $F$ has a unique algebraic closure.*

Finally, we state the fundamental theorem about the complex numbers $\mathbb{C}$.

**Theorem 9.4.23 Fundamental Theorem of Algebra.** *The field of complex numbers is algebraically closed.*

### 9.4.3 Splitting Fields

Let $F$ be a field and $p(x)$ be a nonconstant polynomial in $F[x]$. We already know that we can find a field extension of $F$ that contains a root of $p(x)$. However, we would like to know whether an extension $E$ of $F$ containing all of the roots of $p(x)$ exists. In other words, can we find a field extension of $F$ such that $p(x)$ factors into a product of linear polynomials? What is the "smallest" extension containing all the roots of $p(x)$?

Let $F$ be a field and $p(x) = a_0 + a_1x + \cdots + a_nx^n$ be a nonconstant polynomial in $F[x]$. An extension field $E$ of $F$ is a **splitting field** of $p(x)$ if there exist elements $\alpha_1, \ldots, \alpha_n$ in $E$ such that $E = F(\alpha_1, \ldots, \alpha_n)$ and

$$p(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n).$$

A polynomial $p(x) \in F[x]$ **splits** in $E$ if it is the product of linear factors in $E[x]$.

We summarize the crucial statement about the existence of splitting fields; further details can be found in Subsection 9.5.3.

**Theorem 9.4.24** *Let $p(x)$ be a polynomial in $F[x]$. Then there exists a splitting field $K$ of $p(x)$ that is unique up to isomorphism.*

**Example 9.4.25** Let $p(x) = x^4 + 2x^2 - 8$ be in $\mathbb{Q}[x]$. Then $p(x)$ has irreducible factors $x^2 - 2$ and $x^2 + 4$. Therefore, the field $\mathbb{Q}(\sqrt{2}, i)$ is a splitting field for $p(x)$. □

**Example 9.4.26** Let $p(x) = x^3 - 3$ be in $\mathbb{Q}[x]$. Then $p(x)$ has a root in the field $\mathbb{Q}(\sqrt[3]{3}\,)$. However, this field is not a splitting field for $p(x)$ since the complex cube roots of 3,

$$\frac{-\sqrt[3]{3} \pm (\sqrt[6]{3}\,)^5 i}{2},$$

are not in $\mathbb{Q}(\sqrt[3]{3}\,)$. □

## 9.5 Additional insights

### 9.5.1 More on Finite Extensions

**Corollary 9.5.1** *Let $E$ be an extension field of $F$. If $\alpha \in E$ is algebraic over $F$ with minimal polynomial $p(x)$ and $\beta \in F(\alpha)$ with minimal polynomial $q(x)$, then* $\deg q(x)$ *divides* $\deg p(x)$.

*Proof.* We know that $\deg p(x) = [F(\alpha) : F]$ and $\deg q(x) = [F(\beta) : F]$. Since

$F \subset F(\beta) \subset F(\alpha)$,

$$[F(\alpha) : F] = [F(\alpha) : F(\beta)][F(\beta) : F].$$

■

**Theorem 9.5.2** *Let $E$ be a field extension of $F$. Then the following statements are equivalent.*

1. *$E$ is a finite extension of $F$.*
2. *There exists a finite number of algebraic elements $\alpha_1, \ldots, \alpha_n \in E$ such that $E = F(\alpha_1, \ldots, \alpha_n)$.*
3. *There exists a sequence of fields*

$$E = F(\alpha_1, \ldots, \alpha_n) \supset F(\alpha_1, \ldots, \alpha_{n-1}) \supset \cdots \supset F(\alpha_1) \supset F,$$

*where each field $F(\alpha_1, \ldots, \alpha_i)$ is algebraic over $F(\alpha_1, \ldots, \alpha_{i-1})$.*

*Proof.* (1) ⇒ (2). Let $E$ be a finite algebraic extension of $F$. Then $E$ is a finite dimensional vector space over $F$ and there exists a basis consisting of elements $\alpha_1, \ldots, \alpha_n$ in $E$ such that $E = F(\alpha_1, \ldots, \alpha_n)$. Each $\alpha_i$ is algebraic over $F$ by Theorem 9.4.16.

(2) ⇒ (3). Suppose that $E = F(\alpha_1, \ldots, \alpha_n)$, where every $\alpha_i$ is algebraic over $F$. Then

$$E = F(\alpha_1, \ldots, \alpha_n) \supset F(\alpha_1, \ldots, \alpha_{n-1}) \supset \cdots \supset F(\alpha_1) \supset F,$$

where each field $F(\alpha_1, \ldots, \alpha_i)$ is algebraic over $F(\alpha_1, \ldots, \alpha_{i-1})$.

(3) ⇒ (1). Let

$$E = F(\alpha_1, \ldots, \alpha_n) \supset F(\alpha_1, \ldots, \alpha_{n-1}) \supset \cdots \supset F(\alpha_1) \supset F,$$

where each field $F(\alpha_1, \ldots, \alpha_i)$ is algebraic over $F(\alpha_1, \ldots, \alpha_{i-1})$. Since

$$F(\alpha_1, \ldots, \alpha_i) = F(\alpha_1, \ldots, \alpha_{i-1})(\alpha_i)$$

is simple extension and $\alpha_i$ is algebraic over $F(\alpha_1, \ldots, \alpha_{i-1})$, it follows that

$$[F(\alpha_1, \ldots, \alpha_i) : F(\alpha_1, \ldots, \alpha_{i-1})]$$

is finite for each $i$. Therefore, $[E : F]$ is finite. ■

### 9.5.2 Algebraic Closure

Given a field $F$, the question arises as to whether or not we can find a field $E$ such that every polynomial $p(x)$ has a root in $E$. This leads us to the following theorem.

**Theorem 9.5.3** *Let $E$ be an extension field of $F$. The set of elements in $E$ that are algebraic over $F$ form a field.*

*Proof.* Let $\alpha, \beta \in E$ be algebraic over $F$. Then $F(\alpha, \beta)$ is a finite extension of $F$. Since every element of $F(\alpha, \beta)$ is algebraic over $F$, $\alpha \pm \beta$, $\alpha\beta$, and $\alpha/\beta$ ($\beta \neq 0$) are all algebraic over $F$. Consequently, the set of elements in $E$ that are algebraic over $F$ form a field. ■

**Corollary 9.5.4** *The set of all algebraic numbers forms a field; that is, the set of all complex numbers that are algebraic over $\mathbb{Q}$ makes up a field.*

Let $E$ be a field extension of a field $F$. We define the **algebraic closure** of a field $F$ in $E$ to be the field consisting of all elements in $E$ that are algebraic over $F$. A field $F$ is **algebraically closed** if every nonconstant polynomial in $F[x]$ has a root in $F$.

**Theorem 9.5.5** *A field $F$ is algebraically closed if and only if every nonconstant polynomial in $F[x]$ factors into linear factors over $F[x]$.*

*Proof.* Let $F$ be an algebraically closed field. If $p(x) \in F[x]$ is a nonconstant polynomial, then $p(x)$ has a zero in $F$, say $\alpha$. Therefore, $x-\alpha$ must be a factor of $p(x)$ and so $p(x) = (x-\alpha)q_1(x)$, where $\deg q_1(x) = \deg p(x) - 1$. Continue this process with $q_1(x)$ to find a factorization

$$p(x) = (x-\alpha)(x-\beta)q_2(x),$$

where $\deg q_2(x) = \deg p(x) - 2$. The process must eventually stop since the degree of $p(x)$ is finite.

Conversely, suppose that every nonconstant polynomial $p(x)$ in $F[x]$ factors into linear factors. Let $ax-b$ be such a factor. Then $p(b/a) = 0$. Consequently, $F$ is algebraically closed. ■

**Corollary 9.5.6** *An algebraically closed field $F$ has no proper algebraic extension $E$.*

*Proof.* Let $E$ be an algebraic extension of $F$; then $F \subset E$. For $\alpha \in E$, the minimal polynomial of $\alpha$ is $x - \alpha$. Therefore, $\alpha \in F$ and $F = E$. ■

**Theorem 9.5.7** *Every field $F$ has a unique algebraic closure.*

It is a nontrivial fact that every field has a unique algebraic closure. The proof is not extremely difficult, but requires some rather sophisticated set theory. We refer the reader to [3], [4], or [8] for a proof of this result.

We now state the Fundamental Theorem of Algebra, first proven by Gauss at the age of 22 in his doctoral thesis. This theorem states that every polynomial with coefficients in the complex numbers has a root in the complex numbers. The proof of this theorem goes beyond the scope of these notes.

**Theorem 9.5.8 Fundamental Theorem of Algebra.** *The field of complex numbers is algebraically closed.*

### 9.5.3 Splitting Fields

**Theorem 9.5.9** *Let $p(x) \in F[x]$ be a nonconstant polynomial. Then there exists a splitting field $E$ for $p(x)$.*

*Proof.* We will use mathematical induction on the degree of $p(x)$. If $\deg p(x) = 1$, then $p(x)$ is a linear polynomial and $E = F$. Assume that the theorem is true for all polynomials of degree $k$ with $1 \leq k < n$ and let $\deg p(x) = n$. We can assume that $p(x)$ is irreducible; otherwise, by our induction hypothesis, we are done. By Theorem 9.4.5, there exists a field $K$ such that $p(x)$ has a zero $\alpha_1$ in $K$. Hence, $p(x) = (x-\alpha_1)q(x)$, where $q(x) \in K[x]$. Since $\deg q(x) = n-1$, there exists a splitting field $E \supset K$ of $q(x)$ that contains the zeros $\alpha_2, \ldots, \alpha_n$ of $p(x)$ by our induction hypothesis. Consequently,

$$E = K(\alpha_2, \ldots, \alpha_n) = F(\alpha_1, \ldots, \alpha_n)$$

is a splitting field of $p(x)$. ■

The question of uniqueness now arises for splitting fields. This question is answered in the affirmative. Given two splitting fields $K$ and $L$ of a polynomial

$p(x) \in F[x]$, there exists a field isomorphism $\phi : K \to L$ that preserves $F$. In order to prove this result, we must first prove a lemma.

**Lemma 9.5.10** *Let $\phi : E \to F$ be an isomorphism of fields. Let $K$ be an extension field of $E$ and $\alpha \in K$ be algebraic over $E$ with minimal polynomial $p(x)$. Suppose that $L$ is an extension field of $F$ such that $\beta$ is root of the polynomial in $F[x]$ obtained from $p(x)$ under the image of $\phi$. Then $\phi$ extends to a unique isomorphism $\overline{\phi} : E(\alpha) \to F(\beta)$ such that $\overline{\phi}(\alpha) = \beta$ and $\overline{\phi}$ agrees with $\phi$ on $E$.*

*Proof.* If $p(x)$ has degree $n$, then by Theorem 9.4.14 we can write any element in $E(\alpha)$ as a linear combination of $1, \alpha, \ldots, \alpha^{n-1}$. Therefore, the isomorphism that we are seeking must be

$$\overline{\phi}(a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}) = \phi(a_0) + \phi(a_1)\beta + \cdots + \phi(a_{n-1})\beta^{n-1},$$

where

$$a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}$$

is an element in $E(\alpha)$. The fact that $\overline{\phi}$ is an isomorphism could be checked by direct computation; however, it is easier to observe that $\overline{\phi}$ is a composition of maps that we already know to be isomorphisms.

We can extend $\phi$ to be an isomorphism from $E[x]$ to $F[x]$, which we will also denote by $\phi$, by letting

$$\phi(a_0 + a_1x + \cdots + a_nx^n) = \phi(a_0) + \phi(a_1)x + \cdots + \phi(a_n)x^n.$$

This extension agrees with the original isomorphism $\phi : E \to F$, since constant polynomials get mapped to constant polynomials. By assumption, $\phi(p(x)) = q(x)$; hence, $\phi$ maps $\langle p(x)\rangle$ onto $\langle q(x)\rangle$. Consequently, we have an isomorphism $\psi : E[x]/\langle p(x)\rangle \to F[x]/\langle q(x)\rangle$. By Proposition 9.4.13, we have isomorphisms $\sigma : E[x]/\langle p(x)\rangle \to E(\alpha)$ and $\tau : F[x]/\langle q(x)\rangle \to F(\beta)$, defined by evaluation at $\alpha$ and $\beta$, respectively. Therefore, $\overline{\phi} = \tau\psi\sigma^{-1}$ is the required isomorphism (see Figure 9.5.11).

$$\begin{array}{ccc} E[x]/\langle p(x)\rangle & \xrightarrow{\psi} & F[x]/\langle q(x)\rangle \\ \downarrow{\scriptstyle \sigma} & & \downarrow{\scriptstyle \tau} \\ E(\alpha) & \xrightarrow{\overline{\phi}} & F(\beta) \\ \downarrow & & \downarrow \\ E & \xrightarrow{\phi} & F \end{array}$$

**Figure 9.5.11**

We leave the proof of uniqueness as a exercise. ■

**Theorem 9.5.12** *Let $\phi : E \to F$ be an isomorphism of fields and let $p(x)$ be a nonconstant polynomial in $E[x]$ and $q(x)$ the corresponding polynomial in $F[x]$ under the isomorphism. If $K$ is a splitting field of $p(x)$ and $L$ is a splitting field of $q(x)$, then $\phi$ extends to an isomorphism $\psi : K \to L$.*

*Proof.* We will use mathematical induction on the degree of $p(x)$. We can assume that $p(x)$ is irreducible over $E$. Therefore, $q(x)$ is also irreducible over $F$. If $\deg p(x) = 1$, then by the definition of a splitting field, $K = E$ and $L = F$ and there is nothing to prove.

Assume that the theorem holds for all polynomials of degree less than $n$.

Since $K$ is a splitting field of $p(x)$, all of the roots of $p(x)$ are in $K$. Choose one of these roots, say $\alpha$, such that $E \subset E(\alpha) \subset K$. Similarly, we can find a root $\beta$ of $q(x)$ in $L$ such that $F \subset F(\beta) \subset L$. By Lemma 9.5.10, there exists an isomorphism $\overline{\phi} : E(\alpha) \to F(\beta)$ such that $\overline{\phi}(\alpha) = \beta$ and $\overline{\phi}$ agrees with $\phi$ on $E$ (see Figure 9.5.13).

$$\begin{array}{ccc} K & \xrightarrow{\psi} & L \\ \downarrow{\scriptstyle \sigma} & & \downarrow{\scriptstyle \tau} \\ E(\alpha) & \xrightarrow{\overline{\phi}} & F(\beta) \\ \downarrow & & \downarrow \\ E & \xrightarrow{\phi} & F \end{array}$$

**Figure 9.5.13**

Now write $p(x) = (x-\alpha)f(x)$ and $q(x) = (x-\beta)g(x)$, where the degrees of $f(x)$ and $g(x)$ are less than the degrees of $p(x)$ and $q(x)$, respectively. The field extension $K$ is a splitting field for $f(x)$ over $E(\alpha)$, and $L$ is a splitting field for $g(x)$ over $F(\beta)$. By our induction hypothesis there exists an isomorphism $\psi : K \to L$ such that $\psi$ agrees with $\overline{\phi}$ on $E(\alpha)$. Hence, there exists an isomorphism $\psi : K \to L$ such that $\psi$ agrees with $\phi$ on $E$. ■

**Corollary 9.5.14** *Let $p(x)$ be a polynomial in $F[x]$. Then there exists a splitting field $K$ of $p(x)$ that is unique up to isomorphism.*

## 9.6 Core Exercises

**1.** Prove Theorem 9.4.11.

**2.** Prove Theorem 9.4.18

**3.** Consider the field $E = \mathbb{Q}(\sqrt{11}, \sqrt{13}\,)$.

(a) Prove that $E$ is a vector space over $\mathbb{Q}$. Find a basis. What is $[E : \mathbb{Q}]$?

(b) Let $\phi$: $E \to E$ be the map given by $\phi(z) = (\sqrt{11} + \sqrt{13})z$. Show that $\phi$ is a linear map but not a ring homomorphism. Determine a matrix representation of $\phi$ with respect to the basis from (a).

**4.**

(a) Prove or disprove: $\mathbb{Q}(\sqrt{2}\,) \cong \mathbb{Q}(\sqrt{3}\,)$.

(b) Prove that the fields $\mathbb{Q}(\sqrt[4]{3}\,)$ and $\mathbb{Q}(\sqrt[4]{3}\,i)$ are isomorphic but not equal.

(c) Show that there is a bijective linear map between $\mathbb{Q}(\sqrt{2}\,)$ and $\mathbb{Q}(i\,)$ as vector spaces over $\mathbb{Q}$ but that they are not isomorphic as fields.

**5.**

(a) Show that $\mathbb{Q}[x]/\langle x^4 - 1\rangle$ is not a field.

(b) Determine the splitting field of $x^4 - 1 \in \mathbb{Q}[x]$.

(c) Determine the splitting field of $x^3 - 11 \in \mathbb{Q}[x]$.

(d) Show that $\mathbb{Q}[x]/\langle x^3 - 11\rangle$ is not isomorphic to the splitting field of $x^3 - 11$.

(e) Determine the multiplicative inverse of $x^2+x+\langle x^3-11\rangle$ in $\mathbb{Q}[x]/\langle x^3-11\rangle$.

**6.** Let $F$ be a field.

(a) Show that $F[x]$ is a vector space over $F$. Vector addition is polynomial addition, and scalar multiplication is defined by $\alpha p(x)$ for $\alpha \in F$.

(b) Show that the set $P_n$ of all polynomials of degree less than $n$ form a subspace of the vector space $F[x]$. Find a basis for $P_n$ and determine the dimension of $P_n$.

(c) Show that the derivative is a linear map but not a ring homomorphism.

**7.** Consider the field extension $\mathbb{Q}(\sqrt[4]{3}, i)$ over $\mathbb{Q}$.

(a) Find a basis for the field extension $\mathbb{Q}(\sqrt[4]{3}, i)$ over $\mathbb{Q}$. Conclude that $[\mathbb{Q}(\sqrt[4]{3}, i) : \mathbb{Q}] = 8$.

(b) Find all subfields $F$ of $\mathbb{Q}(\sqrt[4]{3}, i)$ such that $[F : \mathbb{Q}] = 2$.

(c) Find all subfields $F$ of $\mathbb{Q}(\sqrt[4]{3}, i)$ such that $[F : \mathbb{Q}] = 4$.

***Note*** For this exercise, it is not so important to actually construct all subfields but to familiarize with how one can construct these subfields.

**8.**

(a) Prove or disprove: $\pi$ is algebraic over $\mathbb{Q}(\pi^3)$.

(b) If every irreducible polynomial $p(x)$ in $F[x]$ is linear, show that $F$ is an algebraically closed field.

(c) Let $\alpha, \beta$ be transcendental over $\mathbb{Q}$. Prove that either $\alpha\beta$ or $\alpha + \beta$ is also transcendental.

**9.** ***Advanced*** **Automorphism acts as permutation.**Let $E$ be an algebraic extension of a field $F$, and let $\sigma$ be an automorphism of $E$ (that is an isomorphism from $E$ to itself) leaving $F$ fixed. Let $\alpha \in E$. Show that $\sigma$ induces a permutation of the set of all zeros of the minimal polynomial of $\alpha$ that are in $E$.

**10.** ***Background*** **Linear Transformations.**Let $V$ and $W$ be vector spaces over a field $F$, of dimensions $m$ and $n$, respectively. If $T : V \to W$ is a map satisfying

$$T(u + v) = T(u) + T(v)$$
$$T(\alpha v) = \alpha T(v)$$

for all $\alpha \in F$ and all $u, v \in V$, then $T$ is called a **linear transformation** from $V$ into $W$.

(a) Prove that the **kernel** of $T$, $\ker(T) = \{v \in V : T(v) = \mathbf{0}\}$, is a subspace of $V$. The kernel of $T$ is sometimes called the **null space** of $T$.

(b) Prove that the **range** or **range space** of $T$, $R(V) = \{w \in W : T(v) = w \text{ for some } v \in V\}$, is a subspace of $W$.

(c) Show that $T : V \to W$ is injective if and only if $\ker(T) = \{\mathbf{0}\}$.

(d) Let $\{v_1, \ldots, v_k\}$ be a basis for the null space of $T$. We can extend this basis to be a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_m\}$ of $V$. Why? Prove that $\{T(v_{k+1}), \ldots, T(v_m)\}$ is a basis for the range of $T$. Conclude that the range of $T$ has dimension $m - k$.

(e) Let $\dim V = \dim W$. Show that a linear transformation $T : V \to W$ is injective if and only if it is surjective.

## 9.7 Additional Exercises

**1.** Let $F$ be a field and denote the set of $n$-tuples of $F$ by $F^n$. Given vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ in $F^n$ and $\alpha$ in $F$, define vector addition by

$$u + v = (u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n)$$

and scalar multiplication by

$$\alpha u = \alpha(u_1, \ldots, u_n) = (\alpha u_1, \ldots, \alpha u_n).$$

Prove that $F^n$ is a vector space of dimension $n$ under these operations.

**2.** Which of the following sets are subspaces of $\mathbb{R}^3$? If the set is indeed a subspace, find a basis for the subspace and compute its dimension.

(a) $\{(x_1, x_2, x_3) : 3x_1 - 2x_2 + x_3 = 0\}$

(b) $\{(x_1, x_2, x_3) : 3x_1 + 4x_3 = 0, 2x_1 - x_2 + x_3 = 0\}$

(c) $\{(x_1, x_2, x_3) : x_1 - 2x_2 + 2x_3 = 2\}$

(d) $\{(x_1, x_2, x_3) : 3x_1 - 2x_2^2 = 0\}$

**3.** Show that the set of all possible solutions $(x, y, z) \in \mathbb{R}^3$ of the equations

$$\begin{aligned} Ax + By + Cz &= 0 \\ Dx + Ey + Cz &= 0 \end{aligned}$$

form a subspace of $\mathbb{R}^3$.

**4.** Let $W$ be the subset of continuous functions on $[0, 1]$ such that $f(0) = 0$. Prove that $W$ is a subspace of $C[0, 1]$.

**5.** Let $V$ be a vector space over $F$. Prove that $-(\alpha v) = (-\alpha)v = \alpha(-v)$ for all $\alpha \in F$ and all $v \in V$.

**6.** Let $V$ be a vector space of dimension $n$. Prove each of the following statements.

(a) If $S = \{v_1, \ldots, v_n\}$ is a set of linearly independent vectors for $V$, then $S$ is a basis for $V$.

(b) If $S = \{v_1, \ldots, v_n\}$ spans $V$, then $S$ is a basis for $V$.

(c) If $S = \{v_1, \ldots, v_k\}$ is a set of linearly independent vectors for $V$ with $k < n$, then there exist vectors $v_{k+1}, \ldots, v_n$ such that

$$\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$$

is a basis for $V$.

**7.** Prove that any set of vectors containing $\mathbf{0}$ is linearly dependent.

**8.** Let $V$ be a vector space. Show that $\{\mathbf{0}\}$ is a subspace of $V$ of dimension zero.

**9.** If a vector space $V$ is spanned by $n$ vectors, show that any set of $m$ vectors in $V$ must be linearly dependent for $m > n$.

**10.** Let $V$ and $W$ be finite dimensional vector spaces of dimension $n$ over a field $F$. Suppose that $T : V \to W$ is a vector space isomorphism. If $\{v_1, \ldots, v_n\}$ is a basis of $V$, show that $\{T(v_1), \ldots, T(v_n)\}$ is a basis of $W$. Conclude that any vector space over a field $F$ of dimension $n$ is isomorphic to $F^n$.

**11.** **Direct Sums.**Let $U$ and $V$ be subspaces of a vector space $W$. The sum of $U$ and $V$, denoted $U + V$, is defined to be the set of all vectors of the form $u + v$, where $u \in U$ and $v \in V$.

(a) Prove that $U + V$ and $U \cap V$ are subspaces of $W$.

(b) If $U + V = W$ and $U \cap V = \mathbf{0}$, then $W$ is said to be the **direct sum.** In this case, we write $W = U \oplus V$. Show that every element $w \in W$ can be written uniquely as $w = u + v$, where $u \in U$ and $v \in V$.

(c) Let $U$ be a subspace of dimension $k$ of a vector space $W$ of dimension $n$. Prove that there exists a subspace $V$ of dimension $n - k$ such that $W = U \oplus V$. Is the subspace $V$ unique?

(d) If $U$ and $V$ are arbitrary subspaces of a vector space $W$, show that

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$$

**12.** **Dual Spaces.**Let $V$ and $W$ be finite dimensional vector spaces over a field $F$.

(a) Show that the set of all linear transformations from $V$ into $W$, denoted by $\mathrm{Hom}(V, W)$, is a vector space over $F$, where we define vector addition as follows:

$$(S + T)(v) = S(v) + T(v)$$
$$(\alpha S)(v) = \alpha S(v),$$

where $S, T \in \mathrm{Hom}(V, W)$, $\alpha \in F$, and $v \in V$.

(b) Let $V$ be an $F$-vector space. Define the **dual space** of $V$ to be $V^* = \mathrm{Hom}(V, F)$. Elements in the dual space of $V$ are called **linear functionals.** Let $v_1, \ldots, v_n$ be an ordered basis for $V$. If $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ is any vector in $V$, define a linear functional $\phi_i : V \to F$ by $\phi_i(v) = \alpha_i$. Show that the $\phi_i$'s form a basis for $V^*$. This basis is called the **dual basis** of $v_1, \ldots, v_n$ (or simply the dual basis if the context makes the meaning clear).

(c) Consider the basis $\{(3, 1), (2, -2)\}$ for $\mathbb{R}^2$. What is the dual basis for $(\mathbb{R}^2)^*$?

(d) Let $V$ be a vector space of dimension $n$ over a field $F$ and let $V^{**}$ be the dual space of $V^*$. Show that each element $v \in V$ gives rise to an element $\lambda_v$ in $V^{**}$ and that the map $v \mapsto \lambda_v$ is an isomorphism of $V$ with $V^{**}$.

**13.** Prove that $\mathbb{Q}(\sqrt{2}\,)$ is a vector space.

**14.** Prove that the complex numbers are a vector space of dimension 2 over $\mathbb{R}$.

**15.** Show that each of the following numbers is algebraic over $\mathbb{Q}$ by finding the minimal polynomial of the number over $\mathbb{Q}$.

(a) $\sqrt{1/3+\sqrt{7}}$

(b) $\sqrt{3}+\sqrt[3]{5}$

(c) $\sqrt{3}+\sqrt{2}\,i$

(d) $\cos\theta + i\sin\theta$ for $\theta = 2\pi/n$ with $n \in \mathbb{N}$

(e) $\sqrt{\sqrt[3]{2}-i}$

**16.** Find a basis for each of the following field extensions. What is the degree of each extension?

(a) $\mathbb{Q}(\sqrt{3},\sqrt{6}\,)$ over $\mathbb{Q}$

(b) $\mathbb{Q}(\sqrt[3]{2},\sqrt[3]{3}\,)$ over $\mathbb{Q}$

(c) $\mathbb{Q}(\sqrt{2},i)$ over $\mathbb{Q}$

(d) $\mathbb{Q}(\sqrt{3},\sqrt{5},\sqrt{7}\,)$ over $\mathbb{Q}$

(e) $\mathbb{Q}(\sqrt{2},\sqrt[3]{2}\,)$ over $\mathbb{Q}$

(f) $\mathbb{Q}(\sqrt{8}\,)$ over $\mathbb{Q}(\sqrt{2}\,)$

(g) $\mathbb{Q}(i,\sqrt{2}+i,\sqrt{3}+i)$ over $\mathbb{Q}$

(h) $\mathbb{Q}(\sqrt{2}+\sqrt{5}\,)$ over $\mathbb{Q}(\sqrt{5}\,)$

(i) $\mathbb{Q}(\sqrt{2},\sqrt{6}+\sqrt{10}\,)$ over $\mathbb{Q}(\sqrt{3}+\sqrt{5}\,)$

**17.** Find the splitting field for each of the following polynomials.

(a) $x^4 - 10x^2 + 21$ over $\mathbb{Q}$

(b) $x^4 + 1$ over $\mathbb{Q}$

(c) $x^3 + 2x + 2$ over $\mathbb{Z}_3$

(d) $x^3 - 3$ over $\mathbb{Q}$

**18.** Show that $\mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle$ is a field with eight elements. Construct a multiplication table for the multiplicative group of the field.

**19.** Prove that $\mathbb{Q}(\sqrt{3},\sqrt[4]{3},\sqrt[8]{3},\ldots)$ is an algebraic extension of $\mathbb{Q}$ but not a finite extension.

**20.** Let $p(x)$ be a nonconstant polynomial of degree $n$ in $F[x]$. Prove that there exists a splitting field $E$ for $p(x)$ such that $[E : F] \leq n!$.

**21.** Let $K$ be an algebraic extension of $E$, and $E$ an algebraic extension of $F$. Prove that $K$ is algebraic over $F$. [*Caution*: Do not assume that the extensions are finite.]

**22.** Prove or disprove: $\mathbb{Z}[x]/\langle x^3 - 2\rangle$ is a field.

**23.** Let $F$ be a field of characteristic $p$. Prove that $p(x) = x^p - a$ either is irreducible over $F$ or splits in $F$.

**24.** Let $E$ be the algebraic closure of a field $F$. Prove that every polynomial $p(x)$ in $F[x]$ splits in $E$.

**25.** Show that the set of all elements in $\mathbb{R}$ that are algebraic over $\mathbb{Q}$ form a field extension of $\mathbb{Q}$ that is not finite.

**26.** Show that $\mathbb{Q}(\sqrt{3}, \sqrt{7}) = \mathbb{Q}(\sqrt{3} + \sqrt{7})$. Extend your proof to show that $\mathbb{Q}(\sqrt{a}, \sqrt{b}) = \mathbb{Q}(\sqrt{a} + \sqrt{b})$, where $a \neq b$ and neither $a$ nor $b$ is a perfect square.

**27.** Let $E$ be a finite extension of a field $F$. If $[E : F] = 2$, show that $E$ is a splitting field of $F$ for some polynomial $f(x) \in F[x]$.

**28.** Let $E$ be an extension field of $F$ and $\alpha \in E$ be transcendental over $F$. Prove that every element in $F(\alpha)$ that is not in $F$ is also transcendental over $F$.

## 9.8 Material

1. Vector Spaces[1]
2. Field Definition (Quick)[2]
3. Field Definition (Expanded)[3]
4. Field Examples (Infinite Fields)[4]

## 9.9 Hints to Selected Exercises

### 9.6 · Core Exercises

**9.6.1.** Consider the kernel of the evaluation homomorphism. It has a specific structure as each ideal in $F[x]$ is a principal ideal.

Furthermore, you may want to use that a product is not zero in an integral domain if the factors are not zero.

**9.6.2.** Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for $E$ as a vector space over $F$ and $\{\beta_1, \ldots, \beta_m\}$ be a basis for $K$ as a vector space over $E$. Show that $\{\alpha_i \beta_j\}$ is a basis for $K$ over $F$.

**9.6.8.** What is the minimal polynomial of $\pi$ over $\mathbb{Q}(\pi^3)$?

**9.6.10.** ***Background*** **Linear Transformations.**

(a) Let $u, v \in \ker(T)$ and $\alpha \in F$. Then

$$T(u + v) = T(u) + T(v) = 0$$
$$T(\alpha v) = \alpha T(v) = \alpha 0 = 0.$$

Hence, $u + v, \alpha v \in \ker(T)$, and $\ker(T)$ is a subspace of $V$.

(c) The statement that $T(u) = T(v)$ is equivalent to $T(u - v) = T(u) - T(v) = 0$, which is true if and only if $u - v = 0$ or $u = v$.

### 9.7 · Additional Exercises

**9.7.2.** (a) Subspace of dimension 2 with basis $\{(1, 0, -3), (0, 1, 2)\}$; (d) not a subspace

**9.7.5.** Since $0 = \alpha 0 = \alpha(-v + v) = \alpha(-v) + \alpha v$, it follows that $-\alpha v = \alpha(-v)$.

**9.7.7.** Let $v_0 = 0, v_1, \ldots, v_n \in V$ and $\alpha_0 \neq 0, \alpha_1, \ldots, \alpha_n \in F$. Then $\alpha_0 v_0 + \cdots + \alpha_n v_n = 0$.

---

[1] www.socratica.com/lesson/vector-spaces
[2] www.socratica.com/lesson/field-definition-quick
[3] www.socratica.com/lesson/field-definition
[4] www.socratica.com/lesson/field-examples-infinite-fields

**9.7.11. Direct Sums.**

(a) Let $u, u' \in U$ and $v, v' \in V$. Then

$$(u+v)+(u'+v') = (u+u')+(v+v') \in U+V$$
$$\alpha(u+v) = \alpha u + \alpha v \in U+V.$$

**9.7.15.** (a) $x^4 - (2/3)x^2 - 62/9$; (c) $x^4 - 2x^2 + 25$.

**9.7.16.** (a) $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$; (c) $\{1, i, \sqrt{2}, \sqrt{2}\,i\}$; (e) $\{1, 2^{1/6}, 2^{1/3}, 2^{1/2}, 2^{2/3}, 2^{5/6}\}$.

**9.7.17.** (a) $\mathbb{Q}(\sqrt{3}, \sqrt{7})$.

**9.7.18.** Use the fact that the elements of $\mathbb{Z}_2[x]/\langle x^3+x+1\rangle$ are 0, 1, $\alpha$, $1+\alpha$, $\alpha^2$, $1+\alpha^2$, $\alpha+\alpha^2$, $1+\alpha+\alpha^2$ and the fact that $\alpha^3+\alpha+1=0$.

**9.7.21.** Suppose that $E$ is algebraic over $F$ and $K$ is algebraic over $E$. Let $\alpha \in K$. It suffices to show that $\alpha$ is algebraic over some finite extension of $F$. Since $\alpha$ is algebraic over $E$, it must be the zero of some polynomial $p(x) = \beta_0 + \beta_1 x + \cdots + \beta_n x^n$ in $E[x]$. Hence $\alpha$ is algebraic over $F(\beta_0, \ldots, \beta_n)$.

**9.7.26.** Since $\{1, \sqrt{3}, \sqrt{7}, \sqrt{21}\}$ is a basis for $\mathbb{Q}(\sqrt{3}, \sqrt{7})$ over $\mathbb{Q}$, $\mathbb{Q}(\sqrt{3}, \sqrt{7}) \supset \mathbb{Q}(\sqrt{3}+\sqrt{7})$. Since $[\mathbb{Q}(\sqrt{3}, \sqrt{7}) : \mathbb{Q}] = 4$, $[\mathbb{Q}(\sqrt{3}+\sqrt{7}) : \mathbb{Q}] = 2$ or 4. Since the degree of the minimal polynomial of $\sqrt{3}+\sqrt{7}$ is 4, $\mathbb{Q}(\sqrt{3}, \sqrt{7}) = \mathbb{Q}(\sqrt{3}+\sqrt{7})$.

**9.7.28.** Let $\beta \in F(\alpha)$ not in $F$. Then $\beta = p(\alpha)/q(\alpha)$, where $p$ and $q$ are polynomials in $\alpha$ with $q(\alpha) \neq 0$ and coefficients in $F$. If $\beta$ is algebraic over $F$, then there exists a polynomial $f(x) \in F[x]$ such that $f(\beta) = 0$. Let $f(x) = a_0 + a_1 x + \cdots + a_n x^n$. Then

$$0 = f(\beta) = f\left(\frac{p(\alpha)}{q(\alpha)}\right) = a_0 + a_1\left(\frac{p(\alpha)}{q(\alpha)}\right) + \cdots + a_n\left(\frac{p(\alpha)}{q(\alpha)}\right)^n.$$

Now multiply both sides by $q(\alpha)^n$ to show that there is a polynomial in $F[x]$ that has $\alpha$ as a zero.

# Chapter 10

# Finite Fields and Geometric Constructions

Basic learning goals

1. Existence and basic properties of finite fields.
2. Formalization of geometric constructions and constructible numbers.
3. Impossibility of certain constructions.

We finish with two topics, finite fields and geometric constructions, that use several of the tools developed so far.

Finite fields appear in many applications of algebra, including coding theory and cryptography. We already know a family of finite fields, $\mathbb{Z}_p$, where $p$ is prime. In this chapter we will show that a unique finite field of order $p^n$ exists for every prime $p$, where $n$ is a positive integer. Finite fields are also called Galois fields in honor of Évariste Galois, who was one of the first mathematicians to investigate them.

In ancient Greece, three classic problems were posed. These problems are geometric in nature and involve straightedge-and-compass constructions from what is now high school geometry; that is, we are allowed to use only a straightedge and compass to solve them. The problems are: trisecting an angle, squaring the circle and doubling the cube. After puzzling mathematicians for over two thousand years, each of these constructions was finally shown to be impossible. We will use the theory of fields to provide a proof that the solutions do not exist. It is quite remarkable that the long-sought solution to each of these three geometric problems came from abstract algebra.

## 10.1 Structure of a Finite Field

Recall that a field $F$ has **characteristic** $p$ if $p$ is the smallest positive integer such that for every nonzero element $\alpha$ in $F$, we have $p\alpha = 0$. If no such integer exists, then $F$ has characteristic 0. From Theorem 7.2.7 we know that $p$ must be prime if it is not 0. Suppose that $F$ is a finite field with $n$ elements. Then, by Lagrange's Theorem, $n\alpha = 0$ for all $\alpha$ in $F$. Consequently, characteristic of $F$ is not zero but must be $p$, where $p$ is a prime dividing $n$. This discussion is summarized in the following proposition.

**Proposition 10.1.1** *If $F$ is a finite field, then the characteristic of $F$ is $p$, where $p$ is prime.*

Throughout this chapter we will assume that $p$ is a prime number unless otherwise stated.

**Proposition 10.1.2** *If $F$ is a finite field of characteristic $p$, then the order of $F$ is $p^n$ for some $n \in \mathbb{N}$.*

*Proof.* Let $\phi : \mathbb{Z} \to F$ be the ring homomorphism defined by $\phi(n) = n \cdot 1$. Since the characteristic of $F$ is $p$, the kernel of $\phi$ must be $p\mathbb{Z}$ and the image of $\phi$ must be a subfield of $F$ isomorphic to $\mathbb{Z}_p$. We will denote this subfield by $K$. Since $F$ is a finite field, it must be a finite extension of $K$ and, therefore, an algebraic extension of $K$. Suppose that $[F : K] = n$ is the dimension of $F$, where $F$ is a $K$ vector space. There must exist elements $\alpha_1, \ldots, \alpha_n \in F$ such that any element $\alpha$ in $F$ can be written uniquely in the form

$$\alpha = a_1\alpha_1 + \cdots + a_n\alpha_n,$$

where the $a_i$'s are in $K$. Since there are $p$ elements in $K$, there are $p^n$ possible linear combinations of the $\alpha_i$'s. Therefore, the order of $F$ must be $p^n$. ■

**Lemma 10.1.3 Freshman's Dream.** *Let $p$ be prime and $D$ be an integral domain of characteristic $p$. Then*

$$a^{p^n} + b^{p^n} = (a + b)^{p^n}$$

*for all positive integers $n$.*

*Proof.* We will prove this lemma using mathematical induction on $n$. We can use the binomial formula to verify the case for $n = 1$; that is,

$$(a + b)^p = \sum_{k=0}^{p} \binom{p}{k} a^k b^{p-k}.$$

If $0 < k < p$, then

$$\binom{p}{k} = \frac{p!}{k!(p-k)!}$$

must be divisible by $p$, since $p$ cannot divide $k!(p - k)!$. Note that $D$ is an integral domain of characteristic $p$, so all but the first and last terms in the sum must be zero. Therefore, $(a + b)^p = a^p + b^p$.

Now suppose that the result holds for all $k$, where $1 \leq k \leq n$. By the induction hypothesis,

$$(a + b)^{p^{n+1}} = ((a + b)^p)^{p^n} = (a^p + b^p)^{p^n} = (a^p)^{p^n} + (b^p)^{p^n} = a^{p^{n+1}} + b^{p^{n+1}}.$$

Therefore, the lemma is true for $n + 1$ and the proof is complete. ■

The next statement is an amazing insight in the theory of fields. Its proof would need a bit more tools, hence it is deferred to Subsection 10.3.1.

**Theorem 10.1.4** *For every prime $p$ and every positive integer $n$, there exists a finite field $F$ with $p^n$ elements. Furthermore, any field of order $p^n$ is isomorphic to the splitting field of $x^{p^n} - x$ over $\mathbb{Z}_p$.*

The unique finite field with $p^n$ elements is called the **Galois field** of order $p^n$. We will denote this field by $\mathrm{GF}(p^n)$.

**Theorem 10.1.5** *Every subfield of the Galois field $\mathrm{GF}(p^n)$ has $p^m$ elements, where $m$ divides $n$. Conversely, if $m \mid n$ for $m > 0$, then there exists a unique*

*subfield of* $\mathrm{GF}(p^n)$ *isomorphic to* $\mathrm{GF}(p^m)$.

*Proof.* Let $F$ be a subfield of $E = \mathrm{GF}(p^n)$. Then $F$ must be a field extension of $K$ that contains $p^m$ elements, where $K$ is isomorphic to $\mathbb{Z}_p$. Then $m \mid n$, since $[E : K] = [E : F][F : K]$.

To prove the converse, suppose that $m \mid n$ for some $m > 0$. Then $p^m - 1$ divides $p^n - 1$. Consequently, $x^{p^m-1} - 1$ divides $x^{p^n-1} - 1$. Therefore, $x^{p^m} - x$ must divide $x^{p^n} - x$, and every zero of $x^{p^m} - x$ is also a zero of $x^{p^n} - x$. Thus, $\mathrm{GF}(p^n)$ contains, as a subfield, a splitting field of $x^{p^m} - x$, which must be isomorphic to $\mathrm{GF}(p^m)$. ■

**Example 10.1.6** The subfields of $\mathrm{GF}(p^{24})$ ordered by inclusion are given in Figure 10.1.7. □

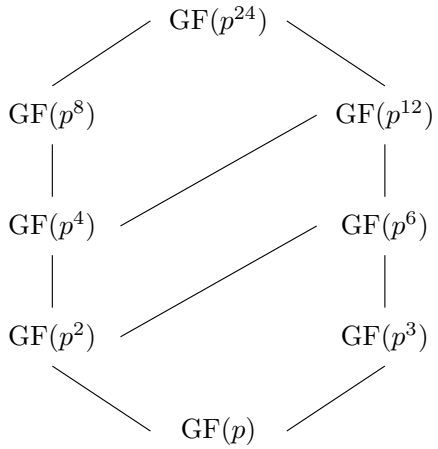



**Figure 10.1.7** Subfields of $\mathrm{GF}(p^{24})$

With each field $F$ we have a multiplicative group of nonzero elements of $F$ which we will denote by $F^*$. The multiplicative group of any finite field is cyclic. This result follows from the more general result of the next theorem; its proof is posed as an exercise.

**Theorem 10.1.8** *If $G$ is a finite subgroup of $F^*$, the multiplicative group of nonzero elements of a field $F$, then $G$ is cyclic.*

**Corollary 10.1.9** *The multiplicative group of all nonzero elements of a finite field is cyclic.*

**Corollary 10.1.10** *Every finite extension $E$ of a finite field $F$ is a simple extension of $F$.*

*Proof.* Let $\alpha$ be a generator for the cyclic group $E^*$ of nonzero elements of $E$. Then $E = F(\alpha)$. ■

**Example 10.1.11** The finite field $\mathrm{GF}(2^4)$ is isomorphic to the field $\mathbb{Z}_2[x]/\langle 1 + x + x^4 \rangle$. Therefore, the elements of $\mathrm{GF}(2^4)$ can be taken to be

$$\{a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3 : a_i \in \mathbb{Z}_2 \text{ and } 1 + \alpha + \alpha^4 = 0\}.$$

Remembering that $1 + \alpha + \alpha^4 = 0$, we add and multiply elements of $\mathrm{GF}(2^4)$ exactly as we add and multiply polynomials. The multiplicative group of $\mathrm{GF}(2^4)$ is isomorphic to $\mathbb{Z}_{15}$ with generator $\alpha$:

$$\begin{array}{lll} \alpha^1 = \alpha & \alpha^6 = \alpha^2 + \alpha^3 & \alpha^{11} = \alpha + \alpha^2 + \alpha^3 \\ \alpha^2 = \alpha^2 & \alpha^7 = 1 + \alpha + \alpha^3 & \alpha^{12} = 1 + \alpha + \alpha^2 + \alpha^3 \end{array}$$

$$\begin{array}{lll} \alpha^3 = \alpha^3 & \alpha^8 = 1 + \alpha^2 & \alpha^{13} = 1 + \alpha^2 + \alpha^3 \\ \alpha^4 = 1 + \alpha & \alpha^9 = \alpha + \alpha^3 & \alpha^{14} = 1 + \alpha^3 \\ \alpha^5 = \alpha + \alpha^2 & \alpha^{10} = 1 + \alpha + \alpha^2 & \alpha^{15} = 1. \end{array}$$

□

## 10.2 Geometric Constructions

Finally, we will consider three classic problems from ancient Greece. The problems can be stated as follows.

1. Given an arbitrary angle, can one trisect the angle into three equal subangles using only a straightedge and compass?
2. Given an arbitrary circle, can one construct a square with the same area using only a straightedge and compass?
3. Given a cube, can one construct the edge of another cube having twice the volume of the original? Again, we are only allowed to use a straightedge and compass to do the construction.

First, let us determine more specifically what we mean by a straightedge and compass, and also examine the nature of these problems in a bit more depth. To begin with, *a straightedge is not a ruler*. We cannot measure arbitrary lengths with a straightedge. It is merely a tool for drawing a line through two points. The statement that the trisection of an arbitrary angle is impossible means that there is at least one angle that is impossible to trisect with a straightedge-and-compass construction. Certainly it is possible to trisect an angle in special cases. We can construct a 30° angle; hence, it is possible to trisect a 90° angle. However, we will show that it is impossible to construct a 20° angle. Therefore, we cannot trisect a 60° angle.

### 10.2.1 Constructible Numbers

A real number $\alpha$ is **constructible** if we can construct a line segment of length $|\alpha|$ in a finite number of steps from a segment of unit length by using a straightedge and compass.

**Theorem 10.2.1** *The set of all constructible real numbers forms a subfield $F$ of the field of real numbers.*

*Proof.* Let $\alpha$ and $\beta$ be constructible numbers. We must show that $\alpha+\beta$, $\alpha-\beta$, $\alpha\beta$, and $\alpha/\beta$ ($\beta \neq 0$) are also constructible numbers. We can assume that both $\alpha$ and $\beta$ are positive with $\alpha > \beta$. It is quite obvious how to construct $\alpha + \beta$ and $\alpha - \beta$. To find a line segment with length $\alpha\beta$, we assume that $\beta > 1$ and construct the triangle in Figure 10.2.2 such that triangles $\triangle ABC$ and $\triangle ADE$ are similar. Since $\alpha/1 = x/\beta$, the line segment $x$ has length $\alpha\beta$. A similar construction can be made if $\beta < 1$. We will leave it as an exercise to show that the same triangle can be used to construct $\alpha/\beta$ for $\beta \neq 0$. ■

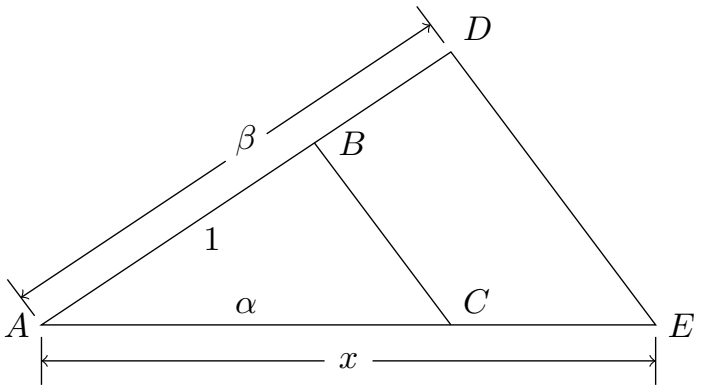



**Figure 10.2.2** Construction of products

**Lemma 10.2.3** *If $\alpha$ is a constructible number, then $\sqrt{\alpha}$ is a constructible number.*

*Proof.* In Figure 10.2.4 the triangles $\triangle ABD$, $\triangle BCD$, and $\triangle ABC$ are similar; hence, $1/x = x/\alpha$, or $x^2 = \alpha$. ■

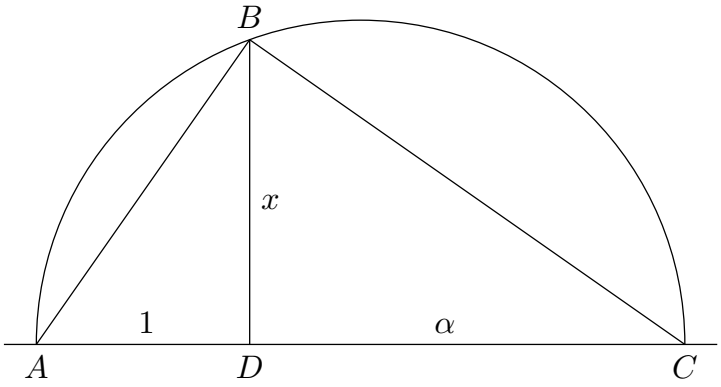



**Figure 10.2.4** Construction of roots

By Theorem 10.2.1, we can locate in the plane any point $P = (p, q)$ that has rational coordinates $p$ and $q$. We need to know what other points can be constructed with a compass and straightedge from points with rational coordinates.

**Lemma 10.2.5** *Let $F$ be a subfield of $\mathbb{R}$.*

1. *If a line contains two points in $F$, then it has the equation $ax+by+c = 0$, where $a$, $b$, and $c$ are in $F$.*

2. *If a circle has a center at a point with coordinates in $F$ and a radius that is also in $F$, then it has the equation $x^2 + y^2 + dx + ey + f = 0$, where $d$, $e$, and $f$ are in $F$.*

*Proof.* Let $(x_1, y_1)$ and $(x_2, y_2)$ be points on a line whose coordinates are in $F$. If $x_1 = x_2$, then the equation of the line through the two points is $x - x_1 = 0$, which has the form $ax + by + c = 0$. If $x_1 \neq x_2$, then the equation of the line through the two points is given by

$$y - y_1 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1),$$

which can also be put into the proper form.

To prove the second part of the lemma, suppose that $(x_1, y_1)$ is the center

of a circle of radius $r$. Then the circle has the equation

$$(x - x_1)^2 + (y - y_1)^2 - r^2 = 0.$$

This equation can easily be put into the appropriate form. ■

Starting with a field of constructible numbers $F$, we have three possible ways of constructing additional points in $\mathbb{R}$ with a compass and straightedge.

1. To find possible new points in $\mathbb{R}$, we can take the intersection of two lines, each of which passes through two known points with coordinates in $F$.
2. The intersection of a line that passes through two points that have coordinates in $F$ and a circle whose center has coordinates in $F$ with radius of a length in $F$ will give new points in $\mathbb{R}$.
3. We can obtain new points in $\mathbb{R}$ by intersecting two circles whose centers have coordinates in $F$ and whose radii are of lengths in $F$.

The first case gives no new points in $\mathbb{R}$, since the solution of two equations of the form $ax + by + c = 0$ having coefficients in $F$ will always be in $F$. The third case can be reduced to the second case. Let

$$x^2 + y^2 + d_1x + e_1y + f_1 = 0$$
$$x^2 + y^2 + d_2x + e_2y + f_2 = 0$$

be the equations of two circles, where $d_i$, $e_i$, and $f_i$ are in $F$ for $i = 1, 2$. These circles have the same intersection as the circle

$$x^2 + y^2 + d_1x + e_1x + f_1 = 0$$

and the line

$$(d_1 - d_2)x + b(e_2 - e_1)y + (f_2 - f_1) = 0.$$

The last equation is that of the chord passing through the intersection points of the two circles. Hence, the intersection of two circles can be reduced to the case of an intersection of a line with a circle.

Considering the case of the intersection of a line and a circle, we must determine the nature of the solutions of the equations

$$ax + by + c = 0$$
$$x^2 + y^2 + dx + ey + f = 0.$$

If we eliminate $y$ from these equations, we obtain an equation of the form $Ax^2 + Bx + C = 0$, where $A$, $B$, and $C$ are in $F$. The $x$ coordinate of the intersection points is given by

$$x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

and is in $F(\sqrt{\alpha}\,)$, where $\alpha = B^2 - 4AC > 0$. We have proven the following lemma.

**Lemma 10.2.6** *Let $F$ be a field of constructible numbers. Then the points determined by the intersections of lines and circles in $F$ lie in the field $F(\sqrt{\alpha}\,)$ for some $\alpha$ in $F$.*

**Theorem 10.2.7** *A real number $\alpha$ is a constructible number if and only if there exists a sequence of fields*

$$\mathbb{Q} = F_0 \subset F_1 \subset \cdots \subset F_k$$

*such that* $F_i = F_{i-1}(\sqrt{\alpha_i})$ *with* $\alpha_i \in F_i$ *and* $\alpha \in F_k$. *In particular, there exists an integer* $k > 0$ *such that* $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 2^k$.

*Proof.* The existence of the $F_i$'s and the $\alpha_i$'s is a direct consequence of Lemma 10.2.6 and of the fact that

$$[F_k : \mathbb{Q}] = [F_k : F_{k-1}][F_{k-1} : F_{k-2}] \cdots [F_1 : \mathbb{Q}] = 2^k.$$

■

**Corollary 10.2.8** *The field of all constructible numbers is an algebraic extension of* $\mathbb{Q}$.

As we can see by the field of constructible numbers, not every algebraic extension of a field is a finite extension.

### 10.2.2 Doubling the Cube and Squaring the Circle

We are now ready to investigate the classical problems of doubling the cube and squaring the circle. We can use the field of constructible numbers to show exactly when a particular geometric construction can be accomplished.

**Doubling the cube is impossible.** Given the edge of the cube, it is impossible to construct with a straightedge and compass the edge of the cube that has twice the volume of the original cube. Let the original cube have an edge of length 1 and, therefore, a volume of 1. If we could construct a cube having a volume of 2, then this new cube would have an edge of length $\sqrt[3]{2}$. However, $\sqrt[3]{2}$ is a zero of the irreducible polynomial $x^3 - 2$ over $\mathbb{Q}$; hence,

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

This is impossible, since 3 is not a power of 2.

**Squaring the circle.** Suppose that we have a circle of radius 1. The area of the circle is $\pi$; therefore, we must be able to construct a square with side $\sqrt{\pi}$. This is impossible since $\pi$ and consequently $\sqrt{\pi}$ are both transcendental. Therefore, using a straightedge and compass, it is not possible to construct a square with the same area as the circle.

### 10.2.3 Trisecting an Angle

*Trisecting an arbitrary angle is impossible.* We will show that it is impossible to construct a 20° angle. Consequently, a 60° angle cannot be trisected. We first need to calculate the triple-angle formula for the cosine:

$$\begin{aligned}
\cos 3\theta &= \cos(2\theta + \theta) \\
&= \cos 2\theta \cos\theta - \sin 2\theta \sin\theta \\
&= (2\cos^2\theta - 1)\cos\theta - 2\sin^2\theta \cos\theta \\
&= (2\cos^2\theta - 1)\cos\theta - 2(1 - \cos^2\theta)\cos\theta \\
&= 4\cos^3\theta - 3\cos\theta.
\end{aligned}$$

The angle $\theta$ can be constructed if and only if $\alpha = \cos\theta$ is constructible. Let $\theta = 20°$. Then $\cos 3\theta = \cos 60° = 1/2$. By the triple-angle formula for the cosine,

$$4\alpha^3 - 3\alpha = \frac{1}{2}.$$

Therefore, $\alpha$ is a zero of $8x^3 - 6x - 1$. This polynomial has degree 3 and no root in $\mathbb{Q}$. Hence, it is irreducible over $\mathbb{Q}[x]$. Thus, $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3$. Consequently, $\alpha$ cannot be a constructible number.

### 10.2.4 Historical Note

Algebraic number theory uses the tools of algebra to solve problems in number theory. Modern algebraic number theory began with Pierre de Fermat (1601–1665). Certainly we can find many positive integers that satisfy the equation $x^2 + y^2 = z^2$; Fermat conjectured that the equation $x^n + y^n = z^n$ has no positive integer solutions for $n \geq 3$. He stated in the margin of his copy of the Latin translation of Diophantus' *Arithmetica* that he had found a marvelous proof of this theorem, but that the margin of the book was too narrow to contain it. Building on work of other mathematicians, it was Andrew Wiles who finally succeeded in proving Fermat's Last Theorem in the 1990s. Wiles's achievement was reported on the front page of the *New York Times*.

Attempts to prove Fermat's Last Theorem have led to important contributions to algebraic number theory by such notable mathematicians as Leonhard Euler (1707–1783). Significant advances in the understanding of Fermat's Last Theorem were made by Ernst Kummer (1810–1893). Kummer's student, Leopold Kronecker (1823–1891), became one of the leading algebraists of the nineteenth century. Kronecker's theory of ideals and his study of algebraic number theory added much to the understanding of fields.

David Hilbert (1862–1943) and Hermann Minkowski (1864–1909) were among the mathematicians who led the way in this subject at the beginning of the twentieth century. Hilbert and Minkowski were both mathematicians at Göttingen University in Germany. Göttingen was truly one the most important centers of mathematical research during the last two centuries. The large number of exceptional mathematicians who studied there included Gauss, Dirichlet, Riemann, Dedekind, Noether, and Weyl.

André Weil answered questions in number theory using algebraic geometry, a field of mathematics that studies geometry by studying commutative rings. From about 1955 to 1970, Alexander Grothendieck dominated the field of algebraic geometry. Pierre Deligne, a student of Grothendieck, solved several of Weil's number-theoretic conjectures. One of the most recent contributions to algebra and number theory is Gerd Faltings' proof of the Mordell conjecture. This conjecture of Mordell, now known as Faltings' theorem, essentially says that certain polynomials $p(x, y)$ in $\mathbb{Z}[x, y]$ have only a finite number of integral solutions.

## 10.3 Additional insights

### 10.3.1 Separable polynomials and derivatives

Let $F$ be a field. A polynomial $f(x) \in F[x]$ of degree $n$ is **separable** if it has $n$ distinct roots in the splitting field of $f(x)$; that is, $f(x)$ is separable when it factors into distinct linear factors over the splitting field of $f$. An extension $E$ of $F$ is a **separable extension** of $F$ if every element in $E$ is the root of a separable polynomial in $F[x]$.

**Example 10.3.1** The polynomial $x^2 - 2$ is separable over $\mathbb{Q}$ since it factors as $(x - \sqrt{2}\,)(x + \sqrt{2}\,)$. In fact, $\mathbb{Q}(\sqrt{2}\,)$ is a separable extension of $\mathbb{Q}$. Let $\alpha = a + b\sqrt{2}$ be any element in $\mathbb{Q}(\sqrt{2}\,)$. If $b = 0$, then $\alpha$ is a root of $x - a$. If

$b \neq 0$, then $\alpha$ is the root of the separable polynomial

$$x^2 - 2ax + a^2 - 2b^2 = (x - (a + b\sqrt{2}\,))(x - (a - b\sqrt{2}\,)).$$

□

Fortunately, we have an easy test to determine the separability of any polynomial. Let

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

be any polynomial in $F[x]$. Define the **derivative** of $f(x)$ to be

$$f'(x) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}.$$

**Lemma 10.3.2** *Let $F$ be a field and $f(x) \in F[x]$. Then $f(x)$ is separable if and only if $f(x)$ and $f'(x)$ are relatively prime.*

*Proof.* Let $f(x)$ be separable. Then $f(x)$ factors over some extension field of $F$ as $f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n)$, where $\alpha_i \neq \alpha_j$ for $i \neq j$. Taking the derivative of $f(x)$, we see that

$$\begin{aligned} f'(x) &= (x - \alpha_2)\cdots(x - \alpha_n) \\ &\quad + (x - \alpha_1)(x - \alpha_3)\cdots(x - \alpha_n) \\ &\quad + \cdots + (x - \alpha_1)\cdots(x - \alpha_{n-1}). \end{aligned}$$

Hence, $f(x)$ and $f'(x)$ can have no common factors.

To prove the converse, we will show that the contrapositive of the statement is true. Suppose that $f(x) = (x - \alpha)^k g(x)$, where $k > 1$. Differentiating, we have

$$f'(x) = k(x - \alpha)^{k-1} g(x) + (x - \alpha)^k g'(x).$$

Therefore, $f(x)$ and $f'(x)$ have a common factor. ■

**Theorem 10.3.3** *For every prime $p$ and every positive integer $n$, there exists a finite field $F$ with $p^n$ elements. Furthermore, any field of order $p^n$ is isomorphic to the splitting field of $x^{p^n} - x$ over $\mathbb{Z}_p$.*

*Proof.* Let $f(x) = x^{p^n} - x$ and let $F$ be the splitting field of $f(x)$. Then by Lemma 10.3.2, $f(x)$ has $p^n$ distinct zeros in $F$, since $f'(x) = p^n x^{p^n - 1} - 1 = -1$ is relatively prime to $f(x)$. We claim that the roots of $f(x)$ form a subfield of $F$. Certainly 0 and 1 are zeros of $f(x)$. If $\alpha$ and $\beta$ are zeros of $f(x)$, then $\alpha + \beta$ and $\alpha\beta$ are also zeros of $f(x)$, since $\alpha^{p^n} + \beta^{p^n} = (\alpha + \beta)^{p^n}$ and $\alpha^{p^n}\beta^{p^n} = (\alpha\beta)^{p^n}$. We also need to show that the additive inverse and the multiplicative inverse of each root of $f(x)$ are roots of $f(x)$. For any zero $\alpha$ of $f(x)$, we know that $-\alpha$ is also a zero of $f(x)$, since

$$f(-\alpha) = (-\alpha)^{p^n} - (-\alpha) = -\alpha^{p^n} + \alpha = -(\alpha^{p^n} - \alpha) = 0,$$

provided $p$ is odd. If $p = 2$, then

$$f(-\alpha) = (-\alpha)^{2^n} - (-\alpha) = \alpha + \alpha = 0.$$

If $\alpha \neq 0$, then $(\alpha^{-1})^{p^n} = (\alpha^{p^n})^{-1} = \alpha^{-1}$. Since the zeros of $f(x)$ form a subfield of $F$ and $f(x)$ splits in this subfield, the subfield must be all of $F$.

Let $E$ be any other field of order $p^n$. To show that $E$ is isomorphic to $F$, we must show that every element in $E$ is a root of $f(x)$. Certainly 0 is a root of $f(x)$. Let $\alpha$ be a nonzero element of $E$. The order of the multiplicative group of nonzero elements of $E$ is $p^n - 1$; hence, $\alpha^{p^n - 1} = 1$ or $\alpha^{p^n} - \alpha = 0$. Since $E$ contains $p^n$ elements, $E$ must be a splitting field of $f(x)$; however, by Corollary 9.5.14, the splitting field of any polynomial is unique up to isomorphism. ■

## 10.4 Core Exercises

**1.** Prove or disprove: $\mathbb{Q}^*$ is cyclic.

**2.** Calculate $[\mathrm{GF}(p^m) : \mathrm{GF}(p^n)]$, where $n \mid m$.

**3.** Let $\alpha$ be a zero of $x^3 + x^2 + 1$ over $\mathbb{Z}_2$. Construct a finite field of order 8. Show that $x^3 + x^2 + 1$ splits in $\mathbb{Z}_2(\alpha)$.

**4.** Prove that the cosine of one degree ($\cos 1°$) is algebraic over $\mathbb{Q}$ but not constructible.

**5.** What are the subfields of $\mathrm{GF}(p^{30})$ and how are they ordered by inclusion?

**6.** Construct a finite field of order 27.

**7.** Prove Theorem 10.1.8.

**8.**

(a) Show that the power set $\mathcal{P}(X)$ forms a vector space over $\mathbb{Z}_2$ where sum is symmetric difference.

(b) Part (a) together with Exercise 7.6.9 shows that $\mathcal{P}(X)$ is a ring and a vector space at the same time. Such a structure is also called **(associative) algebra**. Give another example of an associative algebra.

**9.** Let $G = (V, E)$ be a simple undirected graph.

(a) Show that the set of functions

$$\mathcal{E} = \{f \mid f : E \to \mathbb{Z}_2\}$$

forms a vector space over $\mathbb{Z}_2$.

(b) The incidence vector of a subset $F \subseteq E$ is the function $\chi_F$ in $\mathcal{E}$ with $\chi_F(e) = 0$ for $e \notin F$ and $\chi_F(e) = 1$ for $e \in F$. An induced cycle is a cycle that is also an induced subgraph.

Let $C$ be the subspace of $\mathcal{E}$ generated by the incidence vectors of cycles. Then there is a basis of $C$ consisting only of incidence vectors of induced cycles.

**10.** Prove that the **Frobenius map** $\Phi : \mathrm{GF}(p^n) \to \mathrm{GF}(p^n)$ given by $\Phi : \alpha \mapsto \alpha^p$ is an automorphism (that is an isomorphism from $E$ to itself). Show that its order is $n$; that is, $\Phi^n$ is the identity map on $\mathrm{GF}(p^n)$.

**11.** ***Advanced*** Read The Fast Fourier Transform in a Finite Field[1].

## 10.5 Additional Exercises

**1.** Show that the regular 9-gon is not constructible with a straightedge and compass, but that the regular 20-gon is constructible.

**2.** Can a cube be constructed with three times the volume of a given cube?

**3.** Prove that if $\alpha$ and $\beta$ are constructible numbers such that $\beta \neq 0$, then so is $\alpha/\beta$.

**4.** Calculate each of the following.

[1] www.ams.org/journals/mcom/1971-25-114/S0025-5718-1971-0301966-0/S0025-5718-1971-0301966-0.pdf

(a) $[\mathrm{GF}(3^6) : \mathrm{GF}(3^3)]$ (c) $[\mathrm{GF}(625) : \mathrm{GF}(25)]$

(b) $[\mathrm{GF}(128) : \mathrm{GF}(16)]$ (d) $[\mathrm{GF}(p^{12}) : \mathrm{GF}(p^2)]$

**5.** Factor each of the following polynomials in $\mathbb{Z}_2[x]$.

(a) $x^5 - 1$ (c) $x^9 - 1$

(b) $x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$ (d) $x^4 + x^3 + x^2 + x + 1$

**6.** Prove or disprove: $\mathbb{Z}_2[x]/\langle x^3 + x + 1\rangle \cong \mathbb{Z}_2[x]/\langle x^3 + x^2 + 1\rangle$.

**7.** Prove or disprove: There exists a finite field that is algebraically closed.

**8.** Let $p$ be prime. Prove that the field of rational functions $\mathbb{Z}_p(x)$ is an infinite field of characteristic $p$.

**9.** Let $D$ be an integral domain of characteristic $p$. Prove that $(a-b)^{p^n} = a^{p^n} - b^{p^n}$ for all $a, b \in D$.

**10.** Show that every element in a finite field can be written as the sum of two squares.

**11.** Let $E$ and $F$ be subfields of a finite field $K$. If $E$ is isomorphic to $F$, show that $E = F$.

**12.** Let $E$ be an extension of a finite field $F$, where $F$ has $q$ elements. Let $\alpha \in E$ be algebraic over $F$ of degree $n$. Prove that $F(\alpha)$ has $q^n$ elements.

**13.** Show that for every $n$ there exists an irreducible polynomial of degree $n$ in $\mathbb{Z}_p[x]$.

**14.** Show that every element in $\mathrm{GF}(p^n)$ can be written in the form $a^p$ for some unique $a \in \mathrm{GF}(p^n)$.

**15.** Let $E$ and $F$ be subfields of $\mathrm{GF}(p^n)$. If $|E| = p^r$ and $|F| = p^s$, what is the order of $E \cap F$?

**16.** **Wilson's Theorem.**Let $p$ be prime. Prove that $(p-1)! \equiv -1 \pmod{p}$.

**17.** Let $F \subset E \subset K$ be fields. If $K$ is a separable extension of $F$, show that $K$ is also separable extension of $E$.

## 10.6 Material

1. Application of quaternions: Quaternions[1]
2. Beyond the scope of the course: Modules[2]

## 10.7 Hints to Selected Exercises

### 10.4 · Core Exercises

**10.4.3.** There are eight elements in $\mathbb{Z}_2(\alpha)$. Exhibit two more zeros of $x^3+x^2+1$ other than $\alpha$ in these eight elements.

**10.4.7.** The group $G$ is abelian as it is a subgroup of the multiplicative group of a field; therefore, we can apply the Fundamental Theorem of Finite Abelian Groups. You may also want to use that the polynomial $x^r - 1$ has at most $r$ distinct roots.

### 10.5 · Additional Exercises

**10.5.2.** False.

[1] eater.net/quaternions/video/intro
[2] www.socratica.com/lesson/modules

**10.5.4.** Make sure that you have a field extension.

**10.5.5.** (a) $x^5 - 1 = (x+1)(x^4 + x^3 + x^2 + x + 1)$; (c) $x^9 - 1 = (x+1)(x^2 + x + 1)(x^6 + x^3 + 1)$.

**10.5.6.** True.

**10.5.7.** False.

**10.5.12.** Since $\alpha$ is algebraic over $F$ of degree $n$, we can write any element $\beta \in F(\alpha)$ uniquely as $\beta = a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1}$ with $a_i \in F$. There are $q^n$ possible $n$-tuples $(a_0, a_1, \ldots, a_{n-1})$.

**10.5.16. Wilson's Theorem.**

Factor $x^{p-1} - 1$ over $\mathbb{Z}_p$.

**10.5.17.** If $p(x) \in F[x]$, then $p(x) \in E[x]$.

# Appendix A

# GNU Free Documentation License

Version 1.3, 3 November 2008



**0. PREAMBLE.** The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

**1. APPLICABILITY AND DEFINITIONS.** This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers

are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

**2. VERBATIM COPYING.** You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

**3. COPYING IN QUANTITY.** If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

**4. MODIFICATIONS.** You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if

there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties — for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

**5. COMBINING DOCUMENTS.** You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

**6. COLLECTIONS OF DOCUMENTS.** You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

**7. AGGREGATION WITH INDEPENDENT WORKS.** A compilation of the Document or its derivatives with other separate and independent

documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

**8. TRANSLATION.** Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

**9. TERMINATION.** You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

**10. FUTURE REVISIONS OF THIS LICENSE.** The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See www.gnu.org/copyleft.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or

any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

**11. RELICENSING.** "Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

**ADDENDUM: How to use this License for your documents.** To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
Copyright (C)  YEAR  YOUR NAME.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.3
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with… Texts." line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

# Appendix B

# Comments to Core Exercises

## 1 · Abstract Algebra: Getting Started
## 1.4 · Core Exercises

**1.4.2.** One can use various different representations; similar to Figure 1.2.6.

**1.4.3.** There are at least two ways to find such an element: by checking all possible equivalence classes or by using the extended Euclidean algorithm.

(a) 3; (b) 38

**1.4.4.** What it means to be 'well defined' in this case is made explicit here; try to get a feeling for what that means.

A direct way to approach this exercise is by considering four arbitrary numbers $a, b, c, d \in \mathbb{Z}$ with $a \equiv c \pmod{n}$ and $b \equiv d \pmod{n}$. Then, show that $a + b \equiv c + d \pmod{n}$ and $a \cdot b \equiv c \cdot d \pmod{n}$. This follows by applying the definition of congruence and divisibility.

**1.4.5.** Nearly every choice of two matrices does the job.

**1.4.6.** An interesting example from Linear Algebra is cross-product. But there are many other examples.

For the second part, there is no desired answer; just think about different ways to 'put parantheses'.

**1.4.7. Riddle for Word Problem.**

This exercise is just to get familiar with words and rewriting rules.

(a) The shortest word which it can be reduced to is $a$. This can be seeing as $a$ appears an even number of times in all words in $R$. In particular, one cannot get the empty word.

(b) Yes, one can get the empty word. One possible sequence is

$$cbaabcabc^{-1} - cbcabc^{-1} - cbca - cbcbaaa - caa - bc^{-1}caa - baa - \epsilon$$

(c) This is the inverse which comes up later for products of elements in arbitrary groups.

## 2 · Groups
## 2.4 · Core Exercises

**2.4.1.** Use one of the possibilities to check if something is a subgroup: the definition, Proposition 2.2.7 or Proposition 2.2.8. Furthermore, one has to use that determinants are multiplicative and that the inverse of a matrix with integer entries has again integer entries if its determinant is 1.

**2.4.2.** This follows directly by checking the three properties of a group.

**2.4.3.** Similarly to Proposition 2.1.11, one has to rely merely on the group axioms. A general approach for showing uniqueness is by assuming that there are two elements and then showing that they are the same.

**2.4.4.** Exercise 1.4.5 shows that matrices are not commutative. So one can use these two matrices to do this exemplary calculation. The purpose of this exercise is to familiarize with this property of inverses. Furthermore, it is the basis for part (c) of Exercise 1.4.7

**2.4.5.** This is a standard exercise that can be found online / in other textbooks. To mention a few helpful observations:

(a) Show by induction that $g^m g^n = g^{m+1} g^{n-1} = \cdots = g^{m+n}$

(b) Then use $g^m = g^{m+n} g^{-n}$.

(c) Show by induction $g^m \dots g^m = g^{2m} \dots g^m = g^{mn}$.

(d) Combine the insights $(gh)^n = (h^{-1}g^{-1})^{-n} = ((gh)^{-1})^{-n}$.

**2.4.6.** This follows similarly to Exercise 2.4.1 by using a characterization of subgroups.

**2.4.7.** The crucial idea is to take all possible combinations of elements in $T$, as well as their inverses.
(a) $\{b, d\}$; (b) $\{\mathrm{id}, \rho_1, \rho_2\}$; (c) $G$

**2.4.8. Programming problem.**

Such a program shall be an implementation of the group axioms. Note how to check the quantified statements: all-quantifiers are usually resolved by loops over the respective set; the identity elements and the inverses first have to be identified by checking their properties.

## 3 · Cyclic Groups and Permutation Groups
## 3.5 · Core Exercises

**3.5.1.** This is a standard exercise which can be found elsewhere. Recall what commutative means and the form of the elements of a cyclic group as powers of a generator.

**3.5.2.** This is a standard exercise which can be found elsewhere.

(a) Show two directions!

(b) What does it mean to divide something? Then recall exponentiation!

(c) In the other direction, if $n$ does not divide $k$, there is a non-zero remainder.

**3.5.3.** This is a standard exercise which can be found elsewhere.
To show this exercise, recall what composition of two permutations means. Use that these elements not apearing in a cycle are fixed points of that cycle.

**3.5.4.** This is a standard exercise which can be found elsewhere.

(a) Show that taking $(g, h)$ to the power $\mathrm{lcm}(r, s)$ yields the identity element of $G \times H$ which is the pair of the two identities of $G$ and $H$. Furthermore, observe that the order of $(g, h)$ actually has to be a multiple of $r$ and $s$.

(b) Follows by induction from (a).

(c) The product of cycles can be considered as an element of the symmetric groups corresponding to the elements contained in the cycles. Now, the

claim follows from (b). Alternatively, one can show it directly analogously to (a) and (b).

**3.5.5.** Check that the product $(a_1, a_2, \ldots, a_n)(a_1, a_n, a_{n-1}, \ldots, a_2)$ yields the identity permutation; you might have to use induction for this.

**3.5.6.** The discussion about integers is just to get creative.

(a) 7; you might use brute force here.

(b) 4; there is a smarter way by applying Exercise 3.5.4 and Exercise 3.5.5 one the cycle structure of the permutation.

(c) 3; brute force works but one can diagonalize the matrices as they are symmetric so one can actually take logarithm over integers.

**3.5.7. Dihedral groups.**

(a) $rsr^2srs^3 = rsrssrsrs = s$ This has order 2.

(b) $r^3$ has order 4, it is the product of three cycles of length 4. $s$ has two fixed points and is a product of 5 transpositions.

(c) To do this, go through all elements and list the elements of the cyclic subgroup generated by it. The reflections are always of order two and their cyclic subgroup contains only the identity additional to them. The cyclic subgroups generated by rotations are all contained in the cyclic group generated by $r$.

**3.5.8. Programming problem.**

To check that this is a multiplicative inverse, check if $z$ is coprime with $k$. Then you can either check brute force by going through all equivalence classes or use the Extended Euclidean Algorithm on $z$ and $k$; note that there is a representation of 1 as combination of $z$ and $k$. This yields the inverse.

# 4 · Homomorphisms

## 4.5 · Core Exercises

**4.5.1.** (1) Suppose that $e$ and $e'$ are the identities of $G_1$ and $G_2$, respectively; then

$$e'\phi(e) = \phi(e) = \phi(ee) = \phi(e)\phi(e).$$

By cancellation, $\phi(e) = e'$.

(2) This statement follows from the fact that

$$\phi(g^{-1})\phi(g) = \phi(g^{-1}g) = \phi(e) = e'.$$

(3) This follows by induction from the homomorphism property $\phi(g^n) = \phi(g^{n-1}g) = \phi(g^{n-1})\phi(g) = \phi(g)^{n-1}\phi(g)$.

(4) The set $\phi(H_1)$ is nonempty since the identity of $G_2$ is in $\phi(H_1)$. Suppose that $H_1$ is a subgroup of $G_1$ and let $x$ and $y$ be in $\phi(H_1)$. There exist elements $a, b \in H_1$ such that $\phi(a) = x$ and $\phi(b) = y$. Since

$$xy^{-1} = \phi(a)[\phi(b)]^{-1} = \phi(ab^{-1}) \in \phi(H_1),$$

$\phi(H_1)$ is a subgroup of $G_2$ by Proposition 2.2.8.

**4.5.2.** (1) The image has six elements. Additional to the example above, one

gets

$$\psi((1)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\psi((12)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad .$$

The remaining three matrices are obtained analogously.

(2) There are essentially two ways to approach this. Either one can check the homomorphism property for all pairs explicitly or one shows it more abstractly. The first approach goes through all pairs of elements in $S_3$ like

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \psi((12)(123)) \stackrel{?}{=} \psi((1)(23)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

and indeed

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad .$$

Note that one has to check actually $6 \cdot 6 = 36$ pairs then. For the second approach, let $\alpha, \beta \in S_3$ and $j \in \{1, 2, 3\}$. Then the $(i, j)$-entry of the matrix $\psi(\alpha) \cdot \psi(\beta)$ is

$$\sum_{k=1}^{3} \psi(\alpha)_{ik} \cdot \psi(\beta)_{kj} \quad .$$

Properly checking this expression with the definition of $\psi$ yields the claim.

(3) The kernel is $\{id\}$, it is an isomorphism.

(4) Indeed, there is such an isomorphism from $S_n$ to the group of **permutation matrices** for all positive integers $n$.

Even or odd.

**4.5.3.** (a) is a homomorphism with kernel $\{1\}$.

(b) Computing the product

$$\begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$$

directly shows that it is a homomorphism from $(\mathbb{R}, +)$ to the multiplicative group $GL_2(\mathbb{R})$. The kernel is only 0.

(c) is not a homomorphism.

(d) Compare with the definition of the determinant and Example 4.1.7.

(e) It is a homomorphism. The kernel is all matrices whose $(2, 2)$-entry is 0.

**4.5.4.** A simple approach to list all subgroups is by checking the generated subgroup for each subset of a group. Then to check which ones are normal, one can use the definition of being normal or the criteria. For example, we consider the subgroup of $S_3$ generated by $(12)$. This is just $H = \{(1), (12)\}$. Furthermore, the subgroup generated by $(123)$ is $\{(1), (123), (132)\}$. The check for being normal is done in Example 6.1.2.

**4.5.5.** (a) One can check that one of them is cyclic while the other is not. Or one can check that $\mathbb{Z}_8$ has an element of order 8 while each element in $\mathbb{Z}_2 \times \mathbb{Z}_4$ has order at most 4.

(b) Use Theorem 4.2.10 and the solution is analogous to Example 4.2.11.

**4.5.6.** This follows essentially from the definition of a group homomorphism and the definition of function composition. The purpose of this exercise is to prepare for the insight that a group action (introduced in the next chapter) is a group homomorphism.

**4.5.7.** We denote the map by $\phi$. Each element in $\mathcal{F}(S)$ can be written as a word (or product) of elements in $S$ and their inverses. So let $v, w \in \mathcal{F}(S)$ where we write $v = v_1^{\epsilon_1} \dots v_m^{\epsilon_m}$ and $w = w_1^{\nu_1} \dots w_n^{\nu_n}$ with the $v_i$ and $w_i$ elements in $S$ and the $\epsilon_i$ and $\nu_i$ in $\{-1, 1\}$. Then their product is the concatenation $vw = v_1 \dots v_m w_1 \dots w_n$. We get

$$\phi(vw) = g_{v_1}^{\epsilon_1} \dots g_{v_m}^{\epsilon_m} g_{w_1}^{\nu_1} \dots g_{w_n}^{\nu_n} = \phi(v)\phi(w).$$

**4.5.8. Riddle with nails and a picture.**

(a) Words of length 0: $\epsilon$ Words of length 2: $aa^{-1}, a^{-1}a, bb^{-1}, b^{-1}b$ Words of length 4: $\binom{4}{2}$ choices of two minus and two plus signs in the expression $a^{\sigma_1}a^{\sigma_2}a^{\sigma_3}a^{\sigma_4}$ and analogously for $b$. So that are $2 \cdot \binom{4}{2} = 12$ possibilities. One pair of $a$ and $a^{-1}$ and one pair of $b$ and $b^{-1}$. They can appear as $aa^{-1}bb^{-1}, abb^{-1}a^{-1}, bb^{-1}aa^{-1}$, and with the inverses swapped. This yields $2 \cdot 2 \cdot 4 = 16$ possibilities. Summarizing, we get $1 + 4 + 12 + 16 = 33$. Note that there is a nice interpretation in terms of 'putting correct parentheses', see also Associahedron[1].

(b) Analogously to (a).

(c) That is a standard riddle. Play around with it or search for the solution.

(d) Model the loops of the string by letters $a, b, c$; going left or right around a nail corresponds to a letter or its inverse. The elements, which are in the kernel of all three homomorphisms $\psi_a, \psi_b, \psi_c$ correspond to the required ways of hanging the picture. This idea is not restricted to two or three nails but generalizes analogously.

# 5 · Cosets and Group actions
## 5.5 · Core Exercises

**5.5.1.** (a) The left cosets are $\langle 8\rangle$, $1+\langle 8\rangle$, $2+\langle 8\rangle$, $3+\langle 8\rangle$, $4+\langle 8\rangle$, $5+\langle 8\rangle$, $6+\langle 8\rangle$, and $7+\langle 8\rangle$. These are also the right cosets since the group is commutative. The index is 8.

(b) We have $U(8) = \{1, 3, 5, 7\}$ and $\langle 3\rangle = \{1, 3\}$. The cosets are $\{1, 3\}$ and $\{5, 7\}$. As the group is commutative, the left and right cosets are the same. The index is 2.

(c) The left cosets are

$$\begin{aligned} D_4 &= \{(1), (1\,3), (2\,4), (1\,4\,3\,2), (1\,2\,3\,4), (1\,2)(3\,4), (1\,4)(2\,3), (1\,3)(2\,4)\}, \\ (1\,2)D_4 &= \{(1\,2), (1\,3\,2), (2\,4\,1), (1\,4\,3), (2\,3\,4), (3\,4), (1\,4\,2\,3), (1\,3\,2\,4)\}, \\ (1\,4)D_4 &= \{(1\,4), (1\,3\,4), (1\,4\,2), (4\,3\,2), (1\,2\,3), (1\,2\,4\,3), (2\,3), (1\,3\,4\,2)\}. \end{aligned}$$

Therefore, the index is 3. The three right cosets are determined analogously.

**5.5.2.** One can pick examples, e.g., from Exercise 5.5.1.

**5.5.3.** (a) The cosets are $3\mathbb{Z}$, $1 + 3\mathbb{Z}$, $2 + 3\mathbb{Z}$.

(b) The orbits are $3\mathbb{Z}$ (the orbit containing 0), $1 + 3\mathbb{Z}$ (the orbit containing 1), $2 + 3\mathbb{Z}$ (the orbit containing 2).

(c)/(d) The results of (a) and (b) are the same. Let $G$ be an arbitrary group and $H$ be a subgroup. Then the orbits of the action of $H$ on $G$ by left multiplication are the left cosets. This follows by writing out the definitions of cosets and orbits.

**5.5.4.** (a) There is a single orbit containing all cosets. The stabilizer subgroup of the coset $gS_3$ for $g \in S_4$ is the subgroup $gS_3g^{-1}$ of $S_4$.

(b) / (c) The neutral element of $\mathrm{Sym}(\{gH \mid g \in H\})$ is the function which maps each coset to itself. So we want to find those elements $k$ in $G$ such that $kgH = gH$ for all $g \in G$. These are the elements in $\bigcap_{g\in G} gHg^{-1}$. For $G = S_4$ and $H = S_3$ this yields only the identity (1).

**5.5.5.** (a) $X_{(1)} = \{1,2,3\}$, $X_{(1\,2)} = \{3\}$, $X_{(1\,3)} = \{2\}$, $X_{(2\,3)} = \{1\}$, $X_{(1\,2\,3)} = X_{(1\,3\,2)} = \emptyset$. $G_1 = \{(1), (2\,3)\}$, $G_2 = \{(1), (1\,3)\}$, $G_3 = \{(1), (1\,2)\}$. There is only a single orbit, namely $X$ itself. Indeed, $|G_1| = |G_2| = |G_3| = 2$ and $\frac{|G|}{|G_1|} = \frac{6}{2} = 3 = |X|$.

(b) Analogously to (a).

**5.5.6.** (a) Let $(x, y) \in \mathbb{R}^2$ be a point in the plane. Rotating by $\theta$ yields $A \begin{pmatrix} x \\ y \end{pmatrix}$ where

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

In particular, this group action boils down to a special case of the action of the multiplicative group of rotation matrices. One can also check explicitly the two properties of a group action.

(b) This is the circle around the origin through $P$.

(c) The stabilizer group of $P$ is $2\pi\mathbb{Z}$.

**5.5.7.** This is a standard exercise which can be found elsewhere.

**5.5.8.** This follows by checking the three properties of an equivalence relation. The explicit proof is standard and can be found elsewhere.

**5.5.10. Connection to Linear Algebra.**

(a) This is just matrix equivalence. A canonical form is given by the diagonal matrices with rank-many ones on the diagonal.

(b) Row-echelon form.

# 6 · Quotients and Motivation for Rings
## 6.6 · Core Exercises

**6.6.1.** (1) $\Rightarrow$ (2). For a given $g \in G$ and $n \in N$, there exists an $n'$ in $N$ such that $gn = n'g$. Therefore, $gng^{-1} = n' \in N$ or $gNg^{-1} \subset N$.

(2) $\Rightarrow$ (3). Let $g \in G$. Since $gNg^{-1} \subset N$, we need only show $N \subset gNg^{-1}$. For $n \in N$, $g^{-1}ng = g^{-1}n(g^{-1})^{-1} \in N$. Hence, $g^{-1}ng = n'$ for some $n' \in N$. Therefore, $n = gn'g^{-1}$ is in $gNg^{-1}$.

(3) $\Rightarrow$ (1). Suppose that $gNg^{-1} = N$ for all $g \in G$. Then for any $n \in N$ there exists an $n' \in N$ such that $gng^{-1} = n'$. Consequently, $gn = n'g$ or $gN \subset Ng$. Similarly, $Ng \subset gN$.

**6.6.2.** See Example 6.1.4 for the factor group of the normal subgroup $N = \{(1), (1\,2\,3), (1\,3\,2)\}$ of $S_3$. The factor group is isomorphic to $\mathbb{Z}_2$.

**6.6.3.** (c) Let $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ be an element of $T$. Then its inverse is $\begin{pmatrix} \frac{1}{a} & \frac{-b}{ac} \\ 0 & \frac{1}{c} \end{pmatrix}$ Now, for an element $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ in $U$ we obtain

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \cdot \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{a} & \frac{-b}{ac} \\ 0 & \frac{1}{c} \end{pmatrix} = \begin{pmatrix} 1 & \frac{ax}{c} \\ 0 & 1 \end{pmatrix}.$$

This is in $U$, so $U$ is normal.

(d) Let $t = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ and $t' = \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix}$ be two elements of $T$. Then

$$\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix} = \begin{pmatrix} aa' & ab' + bc' \\ 0 & cc' \end{pmatrix}$$

and

$$\begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix} \cdot \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} aa' & a'b + b'c \\ 0 & cc' \end{pmatrix}$$

Hence, there is a matrix $u \in U$ with $tt'u = t't$. Therefore, the cosets $tt'U$ and $t'tU$ are the same.

(e) No, conjugation by an element in $GL_2(\mathbb{R})$ can destroy the triangular structure.

**6.6.4.** At first, recall that the preimage of a subgroup under a group homomorphism is a subgroup. Secondly, each element in $H$ is a coset of $N$ in $G$. They all have the same cardinality $|N|$ and they are pairwise disjoint. Hence, the preimage of $H$ has order $|H| \cdot |N|$.

**6.6.5.** To show that this relation is an equivalence relation, one can check the three properties (reflexivity, symmetry, transitivity). The equivalence classes are just the cosets in $G/\ker(\phi)$ as one can see by writing the definitions out.

**6.6.6.** (b) $bbbabaaabaaab \sim abaaabaaab \sim abaaabab \sim ababab \sim \epsilon$

(c) Note that $ab = (1\,4\,3\,5\,2)$ which has order 5. Observe that $b$, $b^2$ and $aba$ are 3-cycles. This allows to show that there are no other relations among the generators except for those given by the presentation. Therefore, the kernel is the normal closure of the subgroup generated by the relations.

(d) For example, let $\phi$ be defined by

$$a \mapsto (1\,2) \quad b \mapsto (2\,3) \quad c \mapsto (3\,4).$$

This gives rise to a well-defined group homomorphism because

$$\begin{gathered} \phi(a^2) = (1\,2)(1\,2) = (1) \quad \phi(b^2) = (1) \quad \phi(c^2) = (1) \\ \phi((ac)^2) = (1\,2)(3\,4)(1\,2)(3\,4)\phi(a)^2 = (1) \\ \phi((ab)^3) = ((1\,2)(2\,3))^3 = (1\,2\,3)^3 = (1) \\ \phi((bc)^3) = (2\,3\,4)^3 = (1) \end{gathered}$$

Since this homomorphism is an isomorphism, one can check the word by checking its image under $\phi$.

**6.6.7.** This exercise is mainly for getting a feeling for the interplay between additive and multiplicative structure to prepare for the following chapters on rings. We give some insights using forward reference to notions which are introduced in the following chapters. A pair $(a, b) \in \mathbb{Z}$ is a solution to the equation if and only if $a + \sqrt{5}b$ is a unit of the ring $\mathbb{Z}[\sqrt{5}]\{u + \sqrt{5}v \mid u, v \in \mathbb{Z}\}$. Note that for a solution $(a, b)$ also the pairs $(-a, b), (a, -b), (-a, -b)$ are solutions. By checking the squares $1^2, 2^2, 3^2, \ldots, 9^2$, one sees that $(9, 4)$ is the smallest positive solution. All other solutions arise as powers of this in the ring $\mathbb{Z}[\sqrt{5}]$.

# 7 · Rings

## 7.6 · Core Exercises

**7.6.1.** A heuristic approach is to check first if the sets are closed under addition and multiplication. Particular attention should be paid to elements which look

'special'. For checking if a ring is a field, a heuristic approach is to check if all elements have multiplicative inverses, so in particular those that look 'special'.

(a) $7\mathbb{Z}$ is a ring but not a field;
(b) $\mathbb{Z}_{18}$ is a ring but not a field;
(c) $\mathbb{Q}(\sqrt{2}\,)$ is a field;
(d) $\mathbb{Q}(\sqrt{2}, \sqrt{3}\,)$ is a field.
(f) $R$ is not a ring.

**7.6.2.** (a) $0\mathbb{Z}_{18} = \{0\}$, $9\mathbb{Z}_{18} = \{0, 9\}$, $6\mathbb{Z}_{18} = \{0, 6, 12\}$, $3\mathbb{Z}_{18} = \{0, 3, 6, 9, 12, 15\}$, $2\mathbb{Z}_{18} = \{0, 2, 4, 6, 8, 10, 12, 14, 16\}$. As one can see from the inclusions, $2\mathbb{Z}_{18}$ and $3\mathbb{Z}_{18}$ are maximal, therefore also prime. All the other ideals are not prime since they are either not proper or their factorring has zero divisors.

Analogous to (a).

(c) There are no nontrivial ideals.

(d) All ideals are of the form $\mathbb{M}_2(n\mathbb{Z})$ for some integer $n$. To get an intuition for the ideal, observe that the ring is not commutative. Therefore, an ideal generated by one element is already 'quite big'. One can check this, by looking at two-sided products of a generator with arbitrary matrices.

(e) There are no nontrivial ideals.

**7.6.3.** This exercise is to recall properties of Cayley tables of groups and to see the interplay of addition and multiplication in a field. Note that we characterize finite fields in later chapters. So this field is actually GF(4). Furthermore, the additive and multiplicative groups are commutative. By the Fundamental Theorem of Finite Abelian Groups, the additive group is either $\mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$ and the multiplicative group is isomorphic to $\mathbb{Z}_3$.

**7.6.4.** We need to check that it is a group homomorphism for addition and it also preserves the multiplicative structure. Addition is immediate and multiplication follows from matrix multiplication. The kernel of $\phi$ is just $\{0\} \subset \mathbb{C}$. Hence, the first isomorphism theorem shows that $\phi$ is actually an isomorphism on its image.

**7.6.5.** These are standard properties whose proofs can be found elsewhere.

**7.6.6.** The zero ideal $\{0\}$ is a principal ideal since $\langle 0\rangle = \{0\}$. If $I$ is any nonzero ideal in $\mathbb{Z}$, then $I$ must contain some positive integer $m$. There exists a least positive integer $n$ in $I$ by the Principle of Well-Ordering. Now let $a$ be any element in $I$. Using the division algorithm, we know that there exist integers $q$ and $r$ such that

$$a = nq + r$$

where $0 \leq r < n$. This equation tells us that $r = a - nq \in I$, but $r$ must be 0 since $n$ is the least positive element in $I$. Therefore, $a = nq$ and $I = \langle n\rangle$.

**7.6.7.** (a) This is just the addition and multiplication table of $\mathbb{Z}_6$.

(b) This is just the addition and multiplication table of $\mathbb{Z}_3$.

**7.6.8.** Since a ring homomorphism is also a group homomorphism of the additive group of the ring, we need $\phi(0+6\mathbb{Z}) = 0+15\mathbb{Z}$. Now, we consider the image of the unity $1 + 6\mathbb{Z}$. So let

$$\phi(1 + 6\mathbb{Z}) = z + 15\mathbb{Z}.$$

Then

$$6z + 15\mathbb{Z} = \phi(6 \cdot (1 + 6\mathbb{Z})) = \phi((1 + 6\mathbb{Z}) + \cdots + (1 + 6\mathbb{Z})) = \phi(6\mathbb{Z}) = 0 + 15\mathbb{Z}.$$

Therefore, $15 \mid 6z$ implying $5 \mid z$. Indeed this leads to three valid options with $z = 0$, $z = 5$ and $z = 10$.

**7.6.9.** (a) The zero element is the empty set. The ring properties follow from the properties of symmetric difference and intersection.

(b) The unit is $X$. The characteristic of the ring is 2.

(c) The isomorphism is given by mapping a set $Y \subset X$ to its characteristic vector given by

$$\chi_Y(x) = \begin{cases} 1 & x \in Y \\ 0 & x \notin Y \end{cases}.$$

That this map is actually an isomorphism follows by directly checking the properties of an isomorphism.

**7.6.10.** (a) Zero element: maps all elements to the neutral element of $H$. The inverse of an element $f \in \mathrm{End}(H)$ is given by the map with $h \mapsto -f(h)$ for all $h \in H$. The other properties follow from the properties of $H$ and associativity of function composition.

(b) $\mathbb{R}^2$ is an abelian group under component-wise addition. The group homomorphisms on itself are given by $GL_2(\mathbb{R})$.

# 8 · Polynomials and Remainders

## 8.7 · Core Exercises

**8.7.1.** (a) $5x^3 + 6x^2 - 3x + 4 = (5x^2 + 2x + 1)(x - 2) + 6$

(b) $6x^4 - 2x^3 + x^2 - 3x + 1 = (6x^2 - x)(x^2 + x - 2) + 2x + 1$

(c) $4x^5 - x^3 + x^2 + 4 = (4x^2 + 4)(x^3 + 3) + 4x^2 + 2$.

**8.7.2.** The procedure is exemplarily demonstrated on (b). Note that, since $q(x)$ is irreducible, the gcd of the two polynomials is 1. We will see that also from the algorithm. The Euclidean algorithm yields

$$\begin{aligned} p(x) &= 1 \cdot q(x) + x^2 \\ q(x) &= x \cdot x^2 + x + 1 \\ x^2 &= (x+1) \cdot (x+1) + 1 \end{aligned}$$

Therefore, we obtain (where we use $-1 = 1$ in $\mathbb{Z}_2$)

$$1 = x^2 - (x+1) \cdot (x+1) = x^2 + (x+1) \cdot (q(x) + x \cdot x^2) =$$
$$p(x) + q(x) + (x+1) \cdot (q(x) + x \cdot (p(x) + q(x)) = (x^2 + x + 1) \cdot p(x) + x^2 \cdot q(x)$$

The results for the others are: (a) gcd is $x - 3$, (c) gcd is 1, (d) gcd is 1.

**8.7.3.** A possible procedure is to list all polynomials and then check which of them are irreducible. There 4 polynomials of degree 2 and 8 polynomials of degree 3; note that the leading coefficient must not be zero, otherwise the polynomial has a smaller degree. Furthermore, note that the polynomials have degree smaller than 4, so we can use the simple irreducibility criterion stated above.

**8.7.4.** Consider the polynomials $p(x) = 2x + 1$ and $q(x) = 3x + 1$ in $\mathbb{Z}[x]$. There are no polynomials $s(x), r(x) \in \mathbb{Z}[x]$ with $q(x) = s(x)p(x) + r(x)$ with $\deg r(x) < \deg p(x)$. This is because the leading coefficients of $p(x)$ and $q(x)$ have no multiplicative inverses.

**8.7.5. Homomorphisms between $R$ and $R[x]$.**

(a) The ring homomorphisms from $R$ to $R[x]$: mapping an element of $R$ to its copy in $R[x]$ or mapping everything to 0. The ring homomorphisms from $R[x]$ to $R$: any evaluation homomorphism or again mapping everything to 0.

(b) The image of $x$ in such an isomorphism could not have a finite additive order; this contradicts the finite characteristic.

**8.7.6.** (a), (b) can be checked right from the definitions.

(c) $H$ is the kernel of the ring homomorphism $\psi_A$, hence an ideal of $F[x]$. Since every ideal of $F[x]$ is a principal ideal, there is such a polynomial as required, namely its generator.

(d) The zeros are the eigenvalues of $A$ and this polynomial is called the minimal polynomial of the matrix.

**8.7.7.** We will use mathematical induction on the number of equations in the system. If there are $k = 2$ equations, then the theorem is true by Lemma 8.5.1. Now suppose that the result is true for a system of $k$ equations or less and that we wish to find a solution of

$$\begin{aligned} x &\equiv a_1 \pmod{n_1} \\ x &\equiv a_2 \pmod{n_2} \\ &\vdots \\ x &\equiv a_{k+1} \pmod{n_{k+1}}. \end{aligned}$$

Considering the first $k$ equations, there exists a solution that is unique modulo $n_1 \cdots n_k$, say $a$. Since $n_1 \cdots n_k$ and $n_{k+1}$ are relatively prime, the system

$$\begin{aligned} x &\equiv a \pmod{n_1 \cdots n_k} \\ x &\equiv a_{k+1} \pmod{n_{k+1}} \end{aligned}$$

has a solution that is unique modulo $n_1 \ldots n_{k+1}$ by the lemma.

**8.7.8.** (a) $x \equiv 17 \pmod{55}$

(b) $x \equiv 80 \pmod{840}$

(c) $x \equiv 214 \pmod{2772}$.

**8.7.9.** ***Advanced*** **The Chinese Remainder Theorem for Rings.**

The proof can be found in any advanced algebra book.

# 9 · Vector Spaces and Field Extensions
# 9.6 · Core Exercises

**9.6.1.** Let $\phi_\alpha : F[x] \to E$ be the evaluation homomorphism. The kernel of $\phi_\alpha$ is a principal ideal generated by some $p(x) \in F[x]$ with $\deg p(x) \geq 1$. We know that such a polynomial exists, since each ideal in $F[x]$ is a principal ideal and $\alpha$ is algebraic. The ideal $\langle p(x) \rangle$ consists exactly of those elements of $F[x]$ having $\alpha$ as a zero. If $f(\alpha) = 0$ and $f(x)$ is not the zero polynomial, then $f(x) \in \langle p(x) \rangle$ and $p(x)$ divides $f(x)$. So $p(x)$ is a polynomial of minimal degree having $\alpha$ as a zero. Any other polynomial of the same degree having $\alpha$ as a zero must have the form $\beta p(x)$ for some $\beta \in F$.

Suppose now that $p(x) = r(x)s(x)$ is a factorization of $p(x)$ into polynomials of lower degree. Since $p(\alpha) = 0$, $r(\alpha)s(\alpha) = 0$; consequently, either $r(\alpha) = 0$ or $s(\alpha) = 0$, which contradicts the fact that $p$ is of minimal degree. Therefore, $p(x)$ must be irreducible.

**9.6.2.** Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for $E$ as a vector space over $F$ and $\{\beta_1, \ldots, \beta_m\}$ be a basis for $K$ as a vector space over $E$. We claim that $\{\alpha_i \beta_j\}$ is a basis for $K$ over $F$. We will first show that these vectors span $K$. Let $u \in K$. Then $u = \sum_{j=1}^{m} b_j \beta_j$ and $b_j = \sum_{i=1}^{n} a_{ij} \alpha_i$, where $b_j \in E$ and $a_{ij} \in F$. Then

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

So the $mn$ vectors $\alpha_i \beta_j$ must span $K$ over $F$.

We must show that $\{\alpha_i\beta_j\}$ are linearly independent. Recall that a set of vectors $v_1, v_2, \ldots, v_n$ in a vector space $V$ are linearly independent if

$$c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0$$

implies that

$$c_1 = c_2 = \cdots = c_n = 0.$$

Let

$$u = \sum_{i,j} c_{ij}(\alpha_i\beta_j) = 0$$

for $c_{ij} \in F$. We need to prove that all of the $c_{ij}$'s are zero. We can rewrite $u$ as

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

where $\sum_i c_{ij}\alpha_i \in E$. Since the $\beta_j$'s are linearly independent over $E$, it must be the case that

$$\sum_{i=1}^{n} c_{ij}\alpha_i = 0$$

for all $j$. However, the $\alpha_j$ are also linearly independent over $F$. Therefore, $c_{ij} = 0$ for all $i$ and $j$, which completes the proof.

**9.6.3.** (a) The vector space axioms follow mostly from the field axioms of $E$. $\mathbb{Q}(\sqrt{11}, \sqrt{13})$ has basis $\{1, \sqrt{11}, \sqrt{13}, \sqrt{143}\}$ over $\mathbb{Q}$. In particular, it has dimension 4 as vector space over $\mathbb{Q}$. Therefore, the degree is 4.

(b) Linearity follows from the properties of multiplication in the field $E$. While $\phi$ is a group homomorphism of the additive group, it does not behave well with the multiplication operation. It does not fulfill $\phi(1)\cdot\phi(1) = \phi(1\cdot 1) = \phi(1)$.

**9.6.4.** (a) There are two possible interpretations. The two are isomorphic as vector spaces over $\mathbb{Q}$. However, as fields they are not isomorphic. There is no element in $\mathbb{Q}(\sqrt{3})$ whose square is 2.

(b) They are both isomorphic to $\mathbb{Q}[x]/\langle x^4 - 3\rangle$. But only one is contained in $\mathbb{R}$.

**9.6.5.** (a) It is not a field since it has zero divisors.

(b) $\mathbb{Q}[i]$.

(e) We need to find $q(x), s(x) \in \mathbb{Q}[x]$ such that $q(x)\cdot(x^2 + x) + s(x)\cdot(x^3 - 11) = 1$. This can be done with the Euclidean algorithm. Then the inverse is $q(x) + \langle x^3 - 11\rangle$.

**9.6.6.** (a) and (b) are basic linear algebra.

We have

$$\frac{d}{dx}(\lambda_1 p(x) + \lambda_2 q(x)) = \lambda_1 \frac{d}{dx}p(x) + \lambda_2 \frac{d}{dx}q(x)$$

but

$$\frac{d}{dx}(x^2) \neq \left(\frac{d}{dx}x\right)^2.$$

**9.6.7.** (a) A basis is $\{1, \sqrt[4]{3}, \sqrt[4]{9}, \sqrt[4]{27}, i, \sqrt[4]{3}i, \sqrt[4]{9}i, \sqrt[4]{27}i\}$. Therefore the dimension and, hence, the degree is 8.

Since we have a basis of the field $\mathbb{Q}(\sqrt[4]{3}, i)$, we have an explicit expression for each element as linear combination of the basis elements with scalars from $\mathbb{Q}$. It is instructive to check which elements one obtains by taking combinations of powers of such a combination. For example, one finds $4(\sqrt{3}+i)^{-1} = \sqrt{3} - i$.

This implies $\mathbb{Q}(\sqrt{3}, i) = \mathbb{Q}(\sqrt{3} + i)$; that field has degree 4. As a next step, check what one can get from $\sqrt[4]{3} + i$.

**9.6.8.** (a) It is algebraic with minimal polynomial $x^3 - \pi^3$.

(b) Since all irreducible polynomials are linear and all nonconstant polynomials factor into irreducible polynomials, we get that all nonconstant polynomials split into linear factors. Hence, $F$ is algebraically closed.

(c) If both are not transcendantal then $\mathbb{Q}(\alpha\beta, \alpha + \beta)$ is a finite (algebraic) extension. But $(x - \alpha)(x - \beta) = x^2 - (\alpha + \beta) + \alpha\beta$ has $\alpha$ and $\beta$ as roots, so they are algebraic over $\mathbb{Q}(\alpha\beta, \alpha + \beta)$. But then they are also algebraic over $\mathbb{Q}$.

# 10 · Finite Fields and Geometric Constructions
# 10.4 · Core Exercises

**10.4.1.** It is not cyclic. Assume there is a generator $\frac{a}{b} \in \mathbb{Q}^*$. Let $p$ be a prime number which is coprime to $a$ and $b$. Then

$$p \notin \{\left(\frac{a}{b}\right)^k \mid k \in \mathbb{Z}\},$$

a contradiction.

**10.4.2.** The answer is $\frac{m}{n}$. To get an intuition for the solution, consider the case $n = 1$. This is exactly the situation in Proposition 10.1.2, where we use a basis of $\mathrm{GF}(p^n)$ with $n$ elements.

**10.4.3.** The polynomial $q(x) = x^3 + x^2 + 1 \in \mathbb{Z}_2[x]$ is irreducible, so $E = \mathbb{Z}_2[x]/\langle q(x)\rangle$ is a field. The element $x + \langle q(x)\rangle$ in $E$ is a root of the polynomial $z^3 + z^2 + 1$ in $E[z]$ and we can consider $E$ as an extension field of $\mathbb{Z}_2$. Now, let $\gamma$ be the coset $x + \langle q(x)\rangle$.

By construction, we know that $\gamma \in E$ is a root of $z^3 + z^2 + 1$. We know that neither 0 nor 1 is a root. To find the other roots, we can check the other five elements of $E$. We find that also $\gamma^2$ is a root. We can also use polynomial division to find the remaining root.

$$(z^3 + z^2 + 1) : (z + \gamma) = z^2 + (\gamma + 1)z + \gamma^2 + \gamma,$$

using $\gamma^3 + \gamma^2 + 1 = 0$ Furthermore,

$$(z^2 + (\gamma + 1)z + \gamma^2 + \gamma) : (z + \gamma^2) = z + \gamma^2 + \gamma + 1$$

using $\gamma^4 + \gamma^3 + \gamma = \gamma(\gamma^3 + \gamma^2 + 1) = 0$. So the three roots are $\gamma, \gamma^2, \gamma^2 + \gamma + 1$. As they are all contained in $E$, it is a splitting field of the polynomial.

**10.4.4.** To show that it is algebraic, one can use a formula similar to the triple-angle formula above whose root is $\cos 1°$. It is however not constructible since otherwise also a 20° angle would be constructible.

**10.4.5.** This is analogous to Figure 10.1.7.

**10.4.6.** Find an irreducible polynomial $p(x)$ in $\mathbb{Z}_3[x]$ of degree 3 and show that $\mathbb{Z}_3[x]/\langle p(x)\rangle$ has 27 elements.

**10.4.7.** Let $G$ be a finite subgroup of $F^*$ of order $n$. By the Fundamental Theorem of Finite Abelian Groups

$$G \cong \mathbb{Z}_{p_1^{e_1}} \times \cdots \times \mathbb{Z}_{p_k^{e_k}},$$

where $n = p_1^{e_1} \cdots p_k^{e_k}$ and the $p_1, \ldots, p_k$ are (not necessarily distinct) primes. Let $m$ be the least common multiple of $p_1^{e_1}, \ldots, p_k^{e_k}$. Then $G$ contains an element of order $m$. Since every $\alpha$ in $G$ satisfies $x^r - 1$ for some $r$ dividing $m$, $\alpha$ must also be a root of $x^m - 1$. Since $x^m - 1$ has at most $m$ roots in $F$,

$n \leq m$. On the other hand, we know that $m \leq |G|$; therefore, $m = n$. Thus, $G$ contains an element of order $n$ and must be cyclic.

# Appendix C

# Notation

The following table defines the notation used in this book. Page numbers or references refer to the first appearance of each symbol.

| Symbol | Description | Page |
|---|---|---|
| $(a_{ij})$ | matrix | 2 |
| $a \equiv b \pmod{n}$ | $a$ is congruent to $b$ modulo $n$ | 3 |
| $\mathbb{Z}_n$ | the integers modulo $n$ | 4 |
| $U(n)$ | group of units in $\mathbb{Z}_n$ | 15 |
| $\mathbb{M}_n(\mathbb{R})$ | the $n \times n$ matrices with entries in $\mathbb{R}$ | 15 |
| $\det A$ | the determinant of $A$ | 15 |
| $GL_n(\mathbb{R})$ | the general linear group | 15 |
| $Q_8$ | the group of quaternions | 16 |
| $\mathbb{C}^*$ | the multiplicative group of complex numbers | 16 |
| $\lvert G \rvert$ | the order of a group | 16 |
| $\mathbb{R}^*$ | the multiplicative group of real numbers | 18 |
| $\mathbb{Q}^*$ | the multiplicative group of rational numbers | 18 |
| $\mathbb{T}$ | the circle group | 18 |
| $SL_n(\mathbb{R})$ | the special linear group | 18 |
| $\operatorname{cis}\theta$ | $\cos\theta + i\sin\theta$ | 23 |
| $\mathbb{T}$ | the circle group | 24 |
| $\langle a \rangle$ | cyclic group generated by $a$ | 25 |
| $\langle a \rangle$ | cyclic group generated by $a$ | 29 |
| $\lvert a \rvert$ | the order of an element $a$ | 29 |
| $S_n$ | the symmetric group on $n$ letters | 32 |
| $(a_1, a_2, \ldots, a_k)$ | cycle of length $k$ | 35 |
| $\mathbb{T}$ | the circle group | 38 |
| $A_n$ | the alternating group on $n$ letters | 40 |
| $D_n$ | the dihedral group | 42 |
| $\ker\phi$ | kernel of $\phi$ | 48 |
| $G \cong H$ | $G$ is isomorphic to a group $H$ | 49 |
| $\mathcal{L}_H$ | the set of left cosets of a subgroup $H$ in a group $G$ | 61 |
| $\mathcal{R}_H$ | the set of right cosets of a subgroup $H$ in a group $G$ | 61 |
| $[G : H]$ | index of a subgroup $H$ in a group $G$ | 61 |

(Continued on next page)

| Symbol | Description | Page |
|---|---|---|
| $a \nmid b$ | $a$ does not divide $b$ | 62 |
| $\mathcal{O}_x$ | orbit of $x$ | 64 |
| $X_g$ | fixed point set of $g$ | 64 |
| $G_x$ | isotropy subgroup of $x$ | 65 |
| $N(H)$ | normalizer of s subgroup $H$ | 71 |
| $G/N$ | factor group of $G$ mod $N$ | 83 |
| $\mathbb{H}$ | the ring of quaternions | 101 |
| $\mathbb{Z}[i]$ | the Gaussian integers | 103 |
| $\operatorname{char} R$ | characteristic of a ring $R$ | 103 |
| $\mathbb{Z}_{(p)}$ | ring of integers localized at $p$ | 113 |
| $a \mid b$ | $a$ divides $b$ | 116 |
| $\gcd(a, b)$ | greatest common divisor of $a$ and $b$ | 116 |
| $\deg f(x)$ | degree of a polynomial | 118 |
| $R[x]$ | ring of polynomials over a ring $R$ | 118 |
| $R[x_1, x_2, \ldots, x_n]$ | ring of polynomials in $n$ indeterminants | 120 |
| $\phi_\alpha$ | evaluation homomorphism at $\alpha$ | 120 |
| $\dim V$ | dimension of a vector space $V$ | 141 |
| $F(S)$ | smallest field containing $F$ and $S$ | 142 |
| $F(\alpha_1, \ldots, \alpha_n)$ | smallest field containing $F$ and $\alpha_1, \ldots, \alpha_n$ | 142 |
| $[E : F]$ | dimension of a field extension of $E$ over $F$ | 146 |
| $U \oplus V$ | direct sum of vector spaces $U$ and $V$ | 155 |
| $\operatorname{Hom}(V, W)$ | set of all linear transformations from $U$ into $V$ | 155 |
| $V^*$ | dual of a vector space $V$ | 155 |
| $\mathrm{GF}(p^n)$ | Galois field of order $p^n$ | 160 |
| $F^*$ | multiplicative group of a field $F$ | 161 |

# Index

## Colophon

This book was authored and produced with PreTeXt[1].

[1] pretextbook.org