





# **Distance and Transformation: A Study of Topological and Algebraic Structures.**

Sarthak Dattatray Dhobale  
AMMOC - An International Math Circle



## Contents

|                                                                                    |     |
|------------------------------------------------------------------------------------|-----|
| Chapter 1. Preface                                                                 | 7   |
| Chapter 2. Topological Structures                                                  | 9   |
| 1. Metric Spaces                                                                   | 9   |
| 2. Normed spaces                                                                   | 16  |
| 3. Inner Product Spaces                                                            | 20  |
| 4. Convergence in the Metric Spaces                                                | 23  |
| 5. Some Point-Set Topology of Metric Spaces                                        | 25  |
| 6. Dense Sets and Closures                                                         | 32  |
| 7. Cauchy sequences and complete metric spaces                                     | 34  |
| 8. Compactness                                                                     | 41  |
| 9. History of Compactness                                                          | 45  |
| 10. Future Works                                                                   | 47  |
| Chapter 3. Introduction to groups                                                  | 49  |
| 1. Groups - Definitions and Elementary Properties                                  | 49  |
| 2. Quaternionic Group                                                              | 54  |
| 3. General Linear Groups                                                           | 56  |
| 4. Dihedral group                                                                  | 58  |
| 5. Cyclic Groups                                                                   | 60  |
| 6. Permutation Groups                                                              | 61  |
| 7. Homomorphism, Isomorphism, Automorphism                                         | 64  |
| 8. The Special orthogonal and unitary group                                        | 70  |
| 9. Direct Product, Polyhedra, and Platonic solids                                  | 70  |
| 10. Group Actions                                                                  | 73  |
| Chapter 4. Study of sub-groups                                                     | 79  |
| 1. Structure, sub-structure and smallest sub-structure - a philosophical interlude | 79  |
| 2. Derived or Commutator subgroup                                                  | 80  |
| 3. Centralizers and Normalizers, Stabilizers and Kernels                           | 81  |
| 4. Cyclic Groups, and Cyclic Subgroups                                             | 83  |
| 5. Subgroup generated by subset of a group                                         | 87  |
| Chapter 5. Group homomorphism and quotient structure                               | 91  |
| 1. Normal Subgroups and Quotient Groups                                            | 91  |
| 2. Size of the Cosets, Lagrange's Theorem, and Consequences                        | 94  |
| 3. The isomorphism theorems                                                        | 100 |
| Chapter 6. Group action                                                            | 105 |
| 1. Groups acting on sets                                                           | 105 |
| 2. Important Consequences of group action                                          | 110 |
| 3. Sylow's theorems                                                                | 112 |
| Chapter 7. Introduction to linear transformation                                   | 117 |
| 1. Linear Transformation, Null space, and Range                                    | 117 |

---

|                                                        |     |
|--------------------------------------------------------|-----|
| 2. Algebras over a Field                               | 120 |
| 3. Linear Independence and Dimension                   | 121 |
| 4. Linear Transformations                              | 127 |
| 5. The Endomorphism Algebra of a Vector Space          | 132 |
| 6. The Matrix Representation of a Linear Map           | 133 |
| Chapter 8. Rings and Fields                            | 139 |
| 1. Main definitions and theorems                       | 139 |
| 2. Problems and solutions                              | 144 |
| Appendix A. Sets Relations, Functions, and Partitions  | 165 |
| 1. Sets, Functions, Equivalence Relation and Partition | 165 |
| 2. Mapping between sets                                | 166 |
| Appendix. Bibliography                                 | 169 |

## CHAPTER 1

### Preface

This project is the culmination of my one year of devoted study of algebraic and topological structures. The genesis of this project lies in my failure to qualify in the third round of India's national mathematical Olympiad. I contemplated and talked to my mentor, [Yaashaa Golovanov](#), that instead of continuing to prepare for the IMO, for one last time, I would invest all my energy and time to study pure mathematics. This is how [Abel's theorem in Problems and Solutions based on the lectures of Professor V.I. Arnold](#), [Elementary Real and Complex Analysis by Georgi Shilov](#), [Analysis by Terence Tao](#), and some other minor topics became foundational directed reading projects under the guidance of my mentor. I reaped the benefits of doing these projects as prestigious contests such as [Caltech Math Meet](#), [Princeton University Mathematics Competition \(PUMAC\)](#), and [Berkeley Math Tournament](#) were centered around **Linear and Abstract Algebra**, **Measure Theory and Fractals**, and **Origametry** respectively. I led my team to remarkable victories in these contests, including [third place internationally](#) in Caltech Math Meet'24. My handwritten solutions to the problem of Caltech Math Meet are available [here](#). Solutions to subjective papers of BMT and PUMAC are available on my [website](#). These were jointly typed by my team members and myself and are placed on the website as a record for reference.

In this document, the section 5 of the chapter 2 is joint work in collaboration with [Tisya Rawat](#) is a sophomore at Caltech, majoring in mathematics, and we have been classmates at AMMOC since 2021. The last chapter 8 – on Rings and Fields – is a joint work with [Rama Lakshmi Chunduri](#), a current student at the University of California, San Diego.

None of this would have been possible without endless support from my mentor, Yaashaa Golovanov, who entered my life when my family was facing financial hardship, and I was struggling to find a resource to learn mathematics. He admitted me into his program (AMMOC) towards the end of 2020, and since then, he has been blessing me with his knowledge and mentorship without charging any fee. Learning such advanced mathematics at this age is a very rare privilege that I have received from my mentor. Furthermore, the amount of extraordinary effort that he has poured –into supervising this and [other research projects including proofs of various equivalences of Zorn's lemma, their applications, and group theoretical aspects of Rubik's cube](#) and proofreading this manuscript—is unbelievable. I am forever and infinitely grateful to him for his kindness and mentorship.



## CHAPTER 2

### Topological Structures

**Philosophical viewpoint:** “Anything that depends upon the fundamental least upper bound axioms and thus on its various equivalent versions – the Bolzano-Weierstrass theorem, the Monotone convergence theorem, the Cauchy completeness of  $\mathbb{R}$ , the Nested interval property and Archimedean principle (that there is no smallest positive real number) – is known as Mathematical Analysis and rest everything else is Algebra.” *Excerpt from the first lecture on Analysis by my mentor – [Yaashaa Golovanov](#).*

#### 1. Metric Spaces

**History:** The idea of an abstract space with metric properties listed in definition 1.1 was enunciated by René Maurice Fréchet. However, the term metric space was coined by Felix Hausdorff, one of the founders of modern topology.

Fréchet’s work laid the foundation for convergence, and thus continuity, in spaces beyond the realm of Euclidean spaces.

One of the high points of this section is my counter-example to a problem and solution to another very difficult problem that was incorrectly written in [Tom Körner’s](#) advanced analysis book, [A Companion to Analysis](#), which surveys the results of mathematical analysis at pace, however, takes the reader to remarkably deep digressions. Please read [this email](#) that was written to me by Professor Korner appreciating my arguments.



FIGURE 1. René Maurice Fréchet



FIGURE 2. Felix Hausdorff

**Definition 1.1 (Metric Spaces).** A metric space is a non-empty set  $X$  together with a distance function  $d : X \times X \rightarrow [0, \infty)$  which satisfies the following axioms

- (1)  $d(x, x) = 0$  for all  $x \in X$ .
- (2)  $d(x, y) > 0$  whenever  $x, y$  are distinct elements in  $X$ .
- (3)  $d(x, y) = d(y, x)$  for all  $x, y \in X$ . This property is known as the **symmetry** of the metric function  $d$ .
- (4)  $d(x, z) \leq d(x, y) + d(y, z)$  for all  $x, y, z \in X$ . This property is known as the **triangle inequality** of the distance function  $d$ .

Next we discuss a very good problem that we have taken from [Tom Körner's](#) formidable analysis text: [A Companion to Analysis](#) published as a title in the series "Graduate Studies in Mathematics, AMS".

**Example 1.1 (Construction of new metric using the old).** Let  $(M, d)$  be a metric space. Define  $d_2(x, y) := d(x, y)^2$ . Show that  $(M, d_2)$  is a metric space. Show further that, a sequence in  $(M, d)$  is convergent if and only if it is convergent in  $(M, d_2)$ .

**Remark 1.1 (An interesting story).** Initially, I had to work around to prove the claim. In fact, I wrote a proof to support the claim made in this question, and then the proof was read by [Tisya Rawat](#) during our analysis class of our mentor Yaashaa Golovanov, and she found a gap in the proof. This event forced me to think for a counter example furnished below. I wrote an email to Professor Tom W. Korner, one of the masters of modern analysis, highlighting the mistake in one of the problems and the solution to another problem. Prof. Korner replied that my arguments were really well founded.

**Solution 1.1 (Counterexample).** Consider the set  $M = \{1, 2, 3\}$  and  $d : M \times M \rightarrow M$  is defined as  $d(1, 1) = d(2, 2) = d(3, 3) = 0$ ,  $d(1, 2) = d(2, 1) = 3$ ,  $d(1, 3) = d(3, 1) = 4$ ,  $d(2, 3) = d(3, 2) = 6$ . It is easy to verify that all axioms of metric hold true. Note that,  $d_2(2, 1) + d_2(1, 3) = 3^2 + 4^2 = 25 < 36 = d_2(2, 3)$ . This is a violation of the triangle inequality.

**Example 1.2 (A more general metric inspired by the previous problem).** Let  $\alpha > 0$ . Define a new metric on  $M$  as  $d_\alpha := d(x, y)^\alpha$ . Determine values of  $\alpha$  for which  $(M, d_\alpha)$  is always a metric spaces. Moreover, for those values of  $\alpha$ , is it always true that a sequence in  $(M, d)$  is convergent if and only if it is convergent in  $(M, d_\alpha)$ ?

**Solution 1.2.** Let  $d(x, y) = a$ ,  $d(y, z) = b$  and  $d(x, z) = c$ . We have to find  $\alpha$  such that  $a^\alpha + b^\alpha > c^\alpha$ . It is sufficient to prove that  $a^\alpha + b^\alpha > (a + b)^\alpha$ . Divide both sides by  $b^\alpha$  and take  $x = \frac{a}{b}$  to get ,

$$x^\alpha + 1 - (x + 1)^\alpha > 0.$$

We have to find  $\alpha$  such that  $f(x) = x^\alpha + 1 - (x + 1)^\alpha$  is greater than zero for all  $x > 0$ .

We claim that  $0 < \alpha \leq 1$  works. Since for  $\alpha = 1$ , the given claim is true. So, we only prove for  $0 < \alpha < 1$ . Notice that  $f(0) = 0$  and  $f'(x) = \alpha x^{\alpha-1} - \alpha(1+x)^{\alpha-1}$ . As  $0 < \alpha < 1$ ,  $1 - \alpha > 0$ , we have

$$\begin{aligned} (1+x)^{1-\alpha} &> x^{1-\alpha} \\ x^{\alpha-1} &> (1+x)^{\alpha-1} \end{aligned}$$

Thus,  $f'(x) > 0$  for  $x > 0$  i.e.,  $f(x)$  is increasing with  $f(0) = 0$ . Thus,  $f(x) > 0$  for all  $x > 0$  when  $0 < \alpha \leq 1$ .

**Example 1.3.** Show that the real line with the metric  $d(x, y) := |x - y|$  is indeed a metric space.

**Solution 1.3.** Axioms (a), (b), (c) are clearly true. We now have to prove that

$$|x - z| \leq |x - y| + |y - z|.$$

Let  $a = x - y$  and  $b = y - z$ . By the triangle inequality,

$$|a + b| \leq |a| + |b|,$$

we have

$$|x - z| = |x - y + y - z| \leq |x - y| + |y - z|$$

In the next example, we construct concrete examples of functions which are not metric on the underlying set.

**Example 1.4.** Let  $X$  be a set, and let  $d : X \times X \rightarrow [0, \infty)$  be a function

- (1) Give an example of a pair  $(X, d)$  which obeys axioms (2,3,4) of Definition 1.1, but not (1).
- (2) Give an example of a pair  $(X, d)$  which obeys axioms (1,3,4) of Definition 1.1, but not (2).
- (3) Give an example of a pair  $(X, d)$  which obeys axioms (1,2,3) of Definition 1.1, but not (4). (Hint: try examples where  $X$  is a finite set.)

**Solution 1.4.** (1) Consider  $X = \mathbb{R}$ , and  $d$  defined by  $d(x, x) = 1$  and  $d(x, y) = 2$  for all  $x \neq y \in \mathbb{R}$ .  
 (2) Consider  $X = \mathbb{R}$ , and  $d$  defined by  $d(x, y) = 0$  for all  $x, y \in \mathbb{R}$ .  
 (3) Consider the finite set  $X := \{1, 2, 3\}$  and the metric  $d$  defined by  $d(1, 2) = d(2, 1) = d(2, 3) = d(3, 2) := 1.5$ , and  $d(1, 3) = d(3, 1) := 5$ , and  $d(x, x) = 0$  for all  $x \in X$ . Note that  $d(1, 2) + d(2, 3) = 3 < d(1, 3) = 5$ . This is a violation of triangle inequality.

**Example 1.5** (Terence Tao - Analysis II, Exercise 1.1.3). Show that the pair  $(Y, d|_{Y \times Y})$  defined in Example 12.1.5 is indeed a metric space.

**Solution 1.5.** Since  $d|_{Y \times Y}(x, y) := d(x, y)$ , then the application  $d|_{Y \times Y}$  obeys all four statements (1)–(4) of the definition of the metric space. Thus,  $(Y, d|_{Y \times Y})$  is indeed a metric space.

**Example 1.6** (Terence Tao - Analysis II, Exercise 1.1.3). Let  $n \geq 1$ , and let  $a_1, a_2, \dots, a_n$  and  $b_1, b_2, \dots, b_n$  be real numbers. Verify the identity

$$\left( \sum_{i=1}^n a_i b_i \right)^2 + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i b_j - a_j b_i)^2 = \sum_{i=1}^n a_i^2 \sum_{j=1}^n b_j^2,$$

and conclude the **Cauchy-Schwarz inequality**.

$$\left| \sum_{i=1}^n a_i b_i \right| \leq \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \left( \sum_{i=1}^n b_i^2 \right)^{1/2}$$

Then use the Cauchy-Schwarz inequality to prove the **triangle inequality**.

$$\left( \sum_{i=1}^n (a_i + b_i)^2 \right)^{1/2} \leq \left( \sum_{i=1}^n a_i^2 \right)^{1/2} + \left( \sum_{j=1}^n b_j^2 \right)^{1/2}.$$

**Solution 1.6.** We prove this using **Lagrange identity**:<sup>1</sup>

$$\left( \sum_{k=1}^n a_k b_k \right)^2 + \sum_{i=1}^n \sum_{j=i+1}^n (a_i b_j - a_j b_i)^2 = \left( \sum_{i=1}^n a_i^2 \right) \left( \sum_{j=1}^n b_j^2 \right).$$

This identity implies that

$$\left( \sum_{k=1}^n a_k b_k \right)^2 \leq \left( \sum_{i=1}^n a_i^2 \right) \left( \sum_{j=1}^n b_j^2 \right) \implies \left( \sum_{k=1}^n a_k b_k \right) \leq \left( \sum_{i=1}^n a_i^2 \right)^{\frac{1}{2}} \left( \sum_{j=1}^n b_j^2 \right)^{\frac{1}{2}}.$$

<sup>1</sup>Prof. M. Berry once formulated the following two principles (a) The Arnold Principle. If a notion bears a personal name, then this name is not the name of the discoverer. (b) The Berry Principle. The Arnold Principle is applicable to itself.

PROOF. First, We will prove the identity Consider

$$\begin{aligned} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i b_j - a_j b_i)^2 &= \frac{1}{2} \sum_{i=1}^n ((a_i b_1 - a_1 b_i)^2 + (a_i b_2 - a_2 b_i)^2 + \dots (a_i b_n - a_n b_i)^2) \\ &= \frac{1}{2} \sum_{i=1}^n (a_i^2 (b_1^2 + b_2^2 + \dots b_n^2) + b_i^2 (a_1^2 + \dots a_n^2) - 2a_i b_i (a_1 b_1 + \dots a_n b_n)) \end{aligned}$$

We can further simplify this as

$$\begin{aligned} \frac{1}{2} ((a_1^2 + a_2^2 + \dots a_n^2)(b_1^2 + b_2^2 + \dots b_n^2) + (b_1^2 + b_2^2 + \dots b_n^2)(a_1^2 + a_2^2 + \dots a_n^2) - 2(a_1 b_1 + \dots a_n b_n)^2) \\ \frac{1}{2} (2(a_1^2 + a_2^2 + \dots a_n^2)(b_1^2 + b_2^2 + \dots b_n^2) - 2(a_1 b_1 + \dots a_n b_n)^2) \\ (a_1^2 + a_2^2 + \dots a_n^2)(b_1^2 + b_2^2 + \dots b_n^2) - (a_1 b_1 + \dots a_n b_n)^2 \end{aligned}$$

Thus, We have

$$\begin{aligned} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i b_j - a_j b_i)^2 &= (a_1^2 + a_2^2 + \dots a_n^2)(b_1^2 + b_2^2 + \dots b_n^2) - (a_1 b_1 + \dots a_n b_n)^2 \\ \left( \sum_{i=1}^n a_i b_i \right)^2 + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (a_i b_j - a_j b_i)^2 &= \sum_{i=1}^n a_i^2 \sum_{j=1}^n b_j^2 \end{aligned}$$

By the identity its clear that

$$\begin{aligned} \left( \sum_{i=1}^n a_i b_i \right)^2 &\leq \sum_{i=1}^n a_i^2 \sum_{j=1}^n b_j^2 \\ \left( \sum_{i=1}^n a_i b_i \right)^2 &\leq \sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2 \\ \left| \sum_{i=1}^n a_i b_i \right| &\leq \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \left( \sum_{i=1}^n b_i^2 \right)^{1/2}. \end{aligned}$$

Thus, we are done with the proof of Cauchy Schwartz. We will use it to prove the final inequality. By cauchy schwartz we have,

$$\begin{aligned} \sum_{i=1}^n a_i b_i &\leq \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \left( \sum_{i=1}^n b_i^2 \right)^{1/2} \\ \sum_{i=1}^n 2a_i b_i &\leq 2 \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \left( \sum_{i=1}^n b_i^2 \right)^{1/2} \end{aligned}$$

Adding  $(a_1^2 + a_2^2 + \dots + a_n^2) + (b_1^2 + b_2^2 + \dots + b_n^2)$  on both sides, we get:

$$\sum_{i=1}^n a_i^2 + \sum_{i=1}^n b_i^2 + \sum_{i=1}^n 2a_i b_i \leq \left( \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \right)^2 + \left( \left( \sum_{i=1}^n b_i^2 \right)^{1/2} \right)^2 + 2 \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \left( \sum_{i=1}^n b_i^2 \right)^{1/2}$$

The RHS is basically

$$\left( \left( \sum_{i=1}^n a_i^2 \right)^{1/2} + \left( \sum_{i=1}^n b_i^2 \right)^{1/2} \right)^2$$

The LHS is

$$(a_1 + b_1)^2 + (a_2 + b_2)^2 + \cdots + (a_n + b_n)^2 = \sum_{i=1}^n (a_i + b_i)^2$$

So, we have

$$\sum_{i=1}^n (a_i + b_i)^2 \leq \left( \left( \sum_{i=1}^n a_i^2 \right)^{1/2} + \left( \sum_{i=1}^n b_i^2 \right)^{1/2} \right)^2$$

Taking the square root on both sides we get that

$$\left( \sum_{i=1}^n (a_i + b_i)^2 \right)^{1/2} \leq \left( \sum_{i=1}^n a_i^2 \right)^{1/2} + \left( \sum_{i=1}^n b_i^2 \right)^{1/2}$$

□

**Example 1.7 (Euclidean Spaces).** Let  $n \geq 1$  be a natural number, and let  $\mathbb{R}^n$  be the space of  $n$ -tuples of real numbers:

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\}.$$

We define the **Euclidean Metric**, also known as  $l^2$  metric as  $d_{l^2} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  by

$$d_{l^2}((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) := \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

We claim that  $(\mathbb{R}^n, d_{l^2})$  is indeed a metric space.

**Solution 1.7.** (1) For all  $x = (x_1, x_2, x_3, \dots, x_n) \in \mathbb{R}^n$ , we have  $d_{l^2}(x, x) = \sqrt{\sum_{i=1}^n (x_i - x_i)^2} = 0$  is obviously true.

(2) **Positivity:** Let  $x = (x_1, x_2, x_3, \dots, x_n), y = (y_1, y_2, y_3, \dots, y_n) \in \mathbb{R}^n$ . If  $x \neq y$ , then there exists at least one  $i : 1 \leq i \leq n$  and  $x_i \neq y_i$ , and for this particular  $i$ ,  $(x_i - y_i)^2 > 0$ , and hence  $d_{l^2}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} > 0$ . Therefore, the given metric satisfies the “positivity”.

(3) **Symmetry:** Let  $x = (x_1, x_2, x_3, \dots, x_n), y = (y_1, y_2, y_3, \dots, y_n) \in \mathbb{R}^n$ . Then symmetry is induced from the fact that for any two real numbers  $a$  and  $b$ , we have  $(a - b)^2 = (b - a)^2$ . Therefore,

$$d_{l^2}(y, x) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} = d_{l^2}(x, y)$$

as expected.

(4) **Triangle inequality:** Let  $x = (x_1, x_2, x_3, \dots, x_n), y = (y_1, y_2, y_3, \dots, y_n), z = (z_1, z_2, z_3, \dots, z_n) \in \mathbb{R}^n$ . We have

$$\begin{aligned} d_{l^2}(x, z) &:= \left( \sum_{i=1}^n (x_i - z_i)^2 \right)^{1/2} \\ &= \left( \sum_{i=1}^n (x_i - y_i + y_i - z_i)^2 \right)^{1/2}. \end{aligned}$$

Now set with  $a_i := x_i - y_i$  and  $b_i := y_i - z_i$ . Now, with the help of 1.6, we get

$$\leq \left( \sum_{i=1}^n a_i^2 \right)^{1/2} + \left( \sum_{i=1}^n b_i^2 \right)^{1/2}$$

Again, switching back to the original variables  $x_i, y_i, z_i$

$$\begin{aligned} &\leq \left( \sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2} + \left( \sum_{i=1}^n (y_i - z_i)^2 \right)^{1/2} \\ &\leq d_{l^2}(x, y) + d_{l^2}(y, z). \end{aligned}$$

which proves the triangle inequality for **Euclidean distance** and hence  $d_{l^2}$  - the Euclidean distance is indeed a metric on  $\mathbb{R}^n$ .

Thus,  $(\mathbb{R}^n, d_{l^2})$  is indeed a metric space.

**Example 1.8 (Taxicab metric).** Let  $n \geq 1$ , and let  $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\}$ . Define  $d_{l^1}((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) := \sum_{i=1}^n |x_i - y_i|$ . We claim that  $(\mathbb{R}^n, d_{l^1})$  is indeed a metric space, where  $d_{l^1}$  is known as **Taxicab metric**.

**Solution 1.8.** (1) For all  $x = (x_1, x_2, x_3, \dots, x_n)$ , we have  $d_{l^1}(x, x) = \sum_{i=1}^n |x_i - x_i| = 0$  is obvious due to the definition of  $|x|$  on  $\mathbb{R}$ .

(2) **Positivity:** Let  $x = (x_1, x_2, x_3, \dots, x_n), y = (y_1, y_2, y_3, \dots, y_n) \in \mathbb{R}^n$ . If  $x \neq y$ , then there exists at least one  $i : 1 \leq i \leq n$  and  $x_i \neq y_i$ , and for this particular  $i$ ,  $|(x_i - y_i)| > 0$ , and therefore,  $d_{l^1}(x, y) = \sum_{i=1}^n |x_i - y_i| > 0$ .

(3) **Symmetry:** Let  $x = (x_1, x_2, x_3, \dots, x_n), y = (y_1, y_2, y_3, \dots, y_n) \in \mathbb{R}^n$ . Now, by definition of  $|x| := \max\{x, -x\}$  on  $\mathbb{R}$ , we have  $|x_i - y_i| = |y_i - x_i|$ . Using this, we conclude that

$$d_{l^1}(y, x) = \sum_{i=1}^n |y_i - x_i| = \sum_{i=1}^n |x_i - y_i| = d_{l^1}(x, y).$$

(4) **Triangle inequality:** For real numbers  $x, y, z$  we have  $|x - z| \leq |x - y| + |y - z|$ . Thus, for all  $x, y, z \in \mathbb{R}^n$ , we have

$$d_{l^1}(x, z) := \sum_{i=1}^n |x_i - z_i| \leq \sum_{i=1}^n (|x_i - y_i| + |y_i - z_i|) =: d_{l^1}(x, y) + d_{l^1}(y, z).$$

Therefore, the **Taxicab metric** is indeed a metric on  $\mathbb{R}^n$ .

**Example 1.9 (Sup norm metric).** Let  $n \geq 1$ , and let  $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \in \mathbb{R}\}$ . Define  $d_{l^\infty}((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) := \sup\{|x_i - y_i| : 1 \leq i \leq n\}$ . We claim that  $(\mathbb{R}^n, d_{l^\infty})$  is indeed a metric space.

**Solution 1.9.** (1) For all  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ , we clearly have  $d_{l^\infty}(x, x) = \sup\{|x_i - x_i| : 1 \leq i \leq n\} = 0$  because each of  $|x_i - x_i| = 0$ , and therefore, the supremum of the set is 0.

(2) **Positivity:** Let  $x = (x_1, x_2, x_3, \dots, x_n), y = (y_1, y_2, y_3, \dots, y_n) \in \mathbb{R}^n$ . If  $x \neq y$ , then there exists at least one  $i : 1 \leq i \leq n$  and  $x_i \neq y_i$ , and for this particular  $i$ ,  $|(x_i - y_i)| > 0$ , and  $d_{l^\infty}(x, y) = \sup\{|x_j - y_j| : 1 \leq j \leq n\} \geq |x_i - y_i| > 0$ . This proves that the given metric does satisfies the positivity property.

(3) **Symmetry:** For all  $x, y \in \mathbb{R}^n$ , we have

$$d_{l^\infty}(x, y) = \sup\{|x_i - y_i| : 1 \leq i \leq n\} = \sup\{|y_i - x_i| : 1 \leq i \leq n\} = d_{l^\infty}(y, x)$$

is obviously true.

(4) **Triangle inequality.** Let be  $x, y, z \in \mathbb{R}^n$ . We have  $|x_i - z_i| \leq |x_i - y_i| + |y_i - z_i|$  for all  $1 \leq i \leq n$ . Supremum of the set of elements is always greater than or equal to each element of the set. Therefore,  $|x_i - y_i| \leq d_{l^\infty}(x, y)$  and  $|y_i - z_i| \leq d_{l^\infty}(y, z)$  for all  $1 \leq i \leq n$ . Thus, we have  $|x_i - z_i| \leq |x_i - y_i| + |y_i - z_i| \leq d_{l^\infty}(x, y) + d_{l^\infty}(y, z)$  for all  $1 \leq i \leq n$ . This precisely means that  $d_{l^\infty}(x, y) + d_{l^\infty}(y, z)$  acts as an upper bound of the set  $\{|x_i - z_i| : 1 \leq i \leq n\}$ . Since supremum is the least of all upper bound, therefore,

$$d_{l^\infty}(x, z) := \sup\{|x_i - z_i| : 1 \leq i \leq n\} \leq d_{l^\infty}(x, y) + d_{l^\infty}(y, z)$$

stands true.

**Theorem 1.1.** We claim the following algebraic relation between **Euclidean metric**, **Sup norm metric**, and **Taxicab metric**. For all  $x, y \in \mathbb{R}^n$ ,

$$\frac{1}{\sqrt{n}} d_{l^2}(x, y) \leq d_{l^\infty}(x, y) \leq d_{l^1}(x, y).$$

**PROOF.** Supremum of a finite set is one of the elements of the given set itself. This is the cornerstone of this proof. So in the given setup there exist some  $1 \leq k \leq n$  such that  $d_{l^\infty}(x, y) := \sup\{|x_i - y_i| : 1 \leq$

$i \leq n\} = |x_k - y_k|$  and this  $k$  depends upon the chosen  $x, y \in \mathbb{R}^n$ . We will refer this  $k$  in the subsequent part of the proof.

- **First inequality:** First note that since each term  $(x_i - y_i)^2$  is less than or equal to  $(x_k - y_k)^2$ , therefore, in the subsequent proof, we will replace, for each  $i$ , the term  $(x_i - y_i)^2$  by  $(x_k - y_k)^2$ .

$$\begin{aligned} \frac{1}{\sqrt{n}} d_{l^2}(x, y) &:= \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \\ &\leq \sqrt{\frac{1}{n} \sum_{i=1}^n (x_k - y_k)^2} \\ &\leq \sqrt{\frac{n}{n} (x_k - y_k)^2} \\ &= |x_k - y_k| =: d_{l^\infty}(x, y) \end{aligned}$$

as expected.

- **Second inequality:**

$$\begin{aligned} d_{l^\infty}(x, y) &:= |x_k - y_k| = \sqrt{(x_k - y_k)^2} \\ &\leq \sqrt{(x_k - y_k)^2 + \sum_{1 \leq i \leq n; i \neq k} (x_i - y_i)^2} \\ &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \\ &:= d_{l^2}(x, y) \end{aligned}$$

□

**Definition 1.2 (Diameter, Oscillation and Bounded function).** Let  $(X, d)$  be a metric space and let  $E \subseteq X$ . Define **diameter**  $\varnothing$  of set  $E$  to be the number  $\varnothing(E) = \sup_{x, y \in E} d(x, y)$ . We say a set is **bounded** if  $\varnothing(E) < \infty$ . More specifically, a set in a metric space is bounded if it can be contained or enclosed in a ball  $B(x_0, r)$  of finite positive radius. If  $f : S \rightarrow (X, d)$  is a mapping and  $A$  is a non-empty subset of  $S$ , we define the **oscillation**  $\Omega(f, A)$  of  $f$  on  $A$  to be the diameter of  $f(A)$  :  $\Omega(f, A) = \sup\{d(f(a), f(a')) | a, a' \in A\}$ . Function  $f$  is **bounded** if  $\Omega(f, S) = \varnothing(f(S)) < \infty$ . A subset  $X$  of the real line is said to be bounded if we have  $X \subset [-M, M]$  for some real number  $M > 0$ .

Now, we define **uniform metric** on the set  $B(S, X)$  of bounded function defined from the set  $S$  to another metric space  $(X, d)$ .

**Proposition 1.1 (Uniform metric).** Let  $S$  be a set and  $(X, d)$  be a metric space. Let  $B(S, X)$  be the set of all bounded functions defined from  $S$  to  $(X, d)$ . Let  $f, g \in B(S, X)$ . Define  $d_\infty(f, g) := \sup\{d(f(s), g(s)) | f, g \in B(S, X)\}$ . Then  $d_\infty$  is a metric on  $B(S, X)$ . This metric is known as **uniform metric**.

PROOF. First we prove that  $d_\infty(f, g)$  is indeed finite value for each  $f, g \in B(S, X)$ . Let  $s_0$  be arbitrarily fixed.

$$d(f(s), g(s)) \leq d(f(s), f(s_0)) + d(f(s_0), g(s_0)) + d(g(s_0), g(s)) \leq \Omega(f, S) + d(f(s_0), g(s_0)) + \Omega(g, S)$$

Now we take supremum and arrive at

$$d_\infty(f, g) \leq \Omega(f, S) + d(f(s_0), g(s_0)) + \Omega(g, S)$$

. Since  $f, g$  are bounded mapping so  $\varnothing(f(S)) = \Omega(f, S)$  and  $\varnothing(g(S)) = \Omega(g, S)$  are bounded. Moreover,  $d$  is a metric on  $X$ , so  $d(f(s_0), g(s_0))$  is also finite. Hence,  $d_\infty$  is a finite positive value for each  $f, g \in B(S, X)$ .

It is immediately obvious that  $d_\infty$  satisfies positive definiteness and symmetry. We prove that  $d_\infty$  satisfies triangle inequality. Let  $f, g, h \in B(S, X)$ .

$$d(f(s), h(s)) \leq d(f(s), g(s)) + d(g(s), h(s)) \leq d_\infty(f(s), g(s)) + d_\infty(g(s), h(s))$$

We take supremum and arrive at

$$d_\infty(f(s), h(s)) \leq d_\infty(f(s), g(s)) + d_\infty(g(s), h(s)).$$

□

**Example 1.10 (Discrete metric space).** Show that metric defined on set  $X$  as

$$d_{disc}(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

indeed is a metric on  $X$ .

**Solution 1.10.** (1) For all  $x \in X$ ,  $d_{disc}(x, x) := 0$  by definition of discrete metric,  $d_{disc}$

(2) **Positivity:** For all  $x \neq y \in X$ , we have  $d_{disc}(x, y) := 1 > 0$  again by the definition of discrete metric,  $d_{disc}$ .

(3) **Symmetry:** For all  $x, y \in X$ , we have either  $x = y$  or  $x \neq y$ , and in either case, by the definition of discrete metric, it is obvious that this metric is symmetric.

(4) **Triangle inequality.** Let be  $x, y, z \in X$ , and let's consider  $d_{disc}(x, z)$ .

- If  $x = z$ , then  $d_{disc}(x, z) = 0$ . And since  $d_{disc}$  is a non-negative application, we clearly have  $0 =: d_{disc}(x, z) \leq d_{disc}(x, y) + d_{disc}(y, z)$  for all  $y \in X$ .
- If  $x \neq z$ , then we cannot have both  $x = y$  and  $y = z$  (it would be a clear contradiction with  $x \neq z$ ). Thus, at least one of the propositions " $x \neq y$ ", " $y \neq z$ " is true. Another way to say that is  $d_{disc}(x, y) + d_{disc}(y, z) \geq 1$ . But since  $d_{disc}(x, z) := 1$ , we have actually  $d_{disc}(x, y) + d_{disc}(y, z) \geq d_{disc}(x, z)$ , as expected.

## 2. Normed spaces

In the field of mathematical analysis, we often consider metric over real and complex vector spaces. Since vector spaces are algebraic structure, **our goal is to construct metric on these vector spaces that relates to their algebraic structure**. In particular, we will consider those metric on real ( $\mathbb{R}$ ) and complex ( $\mathbb{C}$ ) vector spaces which are **translation-invariant**. By this, we mean  $d(x + v, y + v) = d(x, y)$  for all  $x, y \in V$ . In such construction, it is a natural requirement that if we scale the vectors by  $\lambda$ , then distance respects scaling. That is,  $d(\lambda x, \lambda y) = |\lambda|d(x, y)$

**Definition 2.1 (Norm).** Suppose  $X$  is a real or complex vector space. A norm  $\|x\|$  on  $X$  is a real-valued function which satisfies following properties.

(N1) **Zero vector**  $|x| \geq 0$  for all  $x \in X$ . The equality holds iff  $x = 0$ .

(N2) **Scaling factor or homogeneity:**  $\|\lambda x\| = |\lambda|\|x\|$  for all  $\lambda \in \mathbb{F} = \mathbb{R}, \mathbb{C}$  and all  $x \in X$ .

(N3) **Triangle inequality:**  $\|x + y\| \leq \|x\| + \|y\|$  for all  $x, y \in X$ .

Next theorem proves three *fundamental* inequalities in mathematical analysis. We will use these three inequalities heavily in the examples of various normed linear spaces.

**Theorem 2.1 (Three classical inequalities).**

(1) **Young's inequality** Suppose that  $x, y \geq 0$  and let  $a, b, c$  be arbitrary vectors in  $\mathbb{R}^k$ .

If  $p, q > 1$  are such that  $\frac{1}{p} + \frac{1}{q} = 1$ , then

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

(2) Prove that if  $x, y \geq 0$  and  $a, b, c$  are arbitrary vectors in  $\mathbb{R}^k$ : if  $p, q > 1$  are such that  $\frac{1}{p} + \frac{1}{q} = 1$ , then

$$\sum_{i=1}^k |a_i| \cdot |b_i| \leq \left[ \sum_{i=1}^k |a_i|^p \right]^{\frac{1}{p}} \left[ \sum_{i=1}^k |b_i|^q \right]^{\frac{1}{q}}$$

**Hölder's inequality**

(3) **Minkowski's inequality** For each  $p \geq 1$  define a metric  $d_p$  as

$$d_p(x, y) = \left[ \sum_{i=1}^k |x_i - y_i|^p \right]^{\frac{1}{p}}$$

Prove that if  $x, y \geq 0$  and  $a, b, c$  are arbitrary vectors in  $\mathbb{R}^k$ : if  $p > 1$ , then

$$d_p(a, b) \leq d_p(a, c) + d_p(c, b)$$

PROOF. (1) **Young's inequality**

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

Let  $y \geq 0$  be fixed and consider a function  $f$  such that:

$$f(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy, \quad x \geq 0$$

We need to show that  $f(x) \geq 0$  to prove the inequality.

Since  $f'(x) = x^{p-1} - y$ , this function decreases for  $x^{p-1} < y$  and increases for  $x^{p-1} > y$ . It attains its Minimum value at  $x_0 = y^{\frac{1}{p-1}}$  ( the point at which  $f'(x) = 0$  ). We prove that  $f(x_0) = 0$

$$f(x_0) = \frac{y^{\frac{p}{p-1}}}{p} + \frac{y^q}{q} - y^{\frac{1}{p-1}} y$$

As,  $\frac{1}{p} + \frac{1}{q} = 1 \implies q = \frac{p}{p-1}$ . So,

$$f(x_0) = y^q \left( \frac{1}{p} + \frac{1}{q} \right) - y^{\frac{1}{p-1}+1} = y^q - y^{\frac{p}{p-1}} = 0$$

at the minimum value of  $x_0$  so  $f(x) \geq 0$ .  $\square$

(2) **Hölder's inequality** The result is obvious when at least one of the  $a$  or  $b$  is zero. Now, for the case when  $a, b$  are both nonzero, we need to show that  $\sum_{i=1}^k x_i y_i \leq 1$ , where

$$x_i = |a_i| \cdot \left[ \sum_{i=1}^k |a_i|^p \right]^{\frac{-1}{p}}$$

$$y_i = |b_i| \cdot \left[ \sum_{i=1}^k |b_i|^q \right]^{\frac{-1}{q}}$$

Applying Young's inequality to  $x_i$  and  $y_i$ , we get  $x_i y_i \leq \frac{x_i^p}{p} + \frac{y_i^q}{q}$  for each  $i$ . We sum this inequality for  $1 \leq i \leq k$ . So, we get

$$\begin{aligned} \sum_{i=1}^k |a_i| |b_i| \left[ \sum_{i=1}^k |a_i|^p \right]^{\frac{-1}{p}} \left[ \sum_{i=1}^k |b_i|^q \right]^{\frac{-1}{q}} &\leq \sum_{i=1}^k \frac{|a_i|^p}{p \left[ \sum_{i=1}^k |a_i|^p \right]} + \sum_{i=1}^k \frac{|b_i|^q}{q \left[ \sum_{i=1}^k |b_i|^q \right]} \\ &\leq \frac{1}{\left[ \sum_{i=1}^k |a_i|^p \right]} \left( \sum_{i=1}^k \frac{|a_i|^p}{p} \right) + \frac{1}{\left[ \sum_{i=1}^k |b_i|^q \right]} \left( \sum_{i=1}^k \frac{|b_i|^q}{q} \right) = \frac{1}{p} + \frac{1}{q} = 1 \end{aligned}$$

This proves **Hölder's inequality** :

$$\sum_{i=1}^k |a_i| \cdot |b_i| \leq \left[ \sum_{i=1}^k |a_i|^p \right]^{\frac{1}{p}} \left[ \sum_{i=1}^k |b_i|^q \right]^{\frac{1}{q}}$$

(3) **Minkowski's inequality** To begin with, we use the triangle inequality on  $\mathbb{R}$  to prove that

$$d_p(a, b)^p \leq \sum_{i=1}^k |a_i - c_i| |a_i - b_i|^{p-1} + \sum_{i=1}^k |c_i - b_i| |a_i - b_i|^{p-1}$$

Now, we will prove the above inequality

$$\begin{aligned} d_p(a, b)^p &= \sum_{i=1}^k |a_i - b_i|^p = \sum_{i=1}^k |a_i - b_i|^{p-1} |a_i - b_i| \leq \sum_{i=1}^k |a_i - b_i|^{p-1} (|a_i - c_i| + |c_i - b_i|) \\ &= \sum_{i=1}^k |a_i - c_i| |a_i - b_i|^{p-1} + \sum_{i=1}^k |c_i - b_i| |a_i - b_i|^{p-1} \end{aligned}$$

Letting  $q = \frac{p}{p-1}$  we have  $\frac{1}{p} + \frac{1}{q} = 1$  so Holder's Inequality gives

$$\begin{aligned} \sum_{i=1}^k |a_i - c_i| |a_i - b_i|^{p-1} &\leq \left[ \sum_{i=1}^k |a_i - c_i|^p \right]^{\frac{1}{p}} \cdot \left[ \sum_{i=1}^k |a_i - b_i|^{q(p-1)} \right]^{\frac{1}{q}} \\ \sum_{i=1}^k |c_i - b_i| |a_i - b_i|^{p-1} &\leq \left[ \sum_{i=1}^k |c_i - b_i|^p \right]^{\frac{1}{p}} \cdot \left[ \sum_{i=1}^k |a_i - b_i|^{q(p-1)} \right]^{\frac{1}{q}} \end{aligned}$$

Adding the above two inequalities and noting  $q(p-1) = p$  and  $\frac{1}{q} = \frac{p-1}{p}$ , we get

$$\sum_{i=1}^k |a_i - b_i|^p \leq \left[ \sum_{i=1}^k |a_i - c_i|^p \right]^{\frac{1}{p}} \cdot \left[ \sum_{i=1}^k |a_i - b_i|^p \right]^{\frac{p-1}{p}} + \left[ \sum_{i=1}^k |c_i - b_i|^p \right]^{\frac{1}{p}} \cdot \left[ \sum_{i=1}^k |a_i - b_i|^p \right]^{\frac{p-1}{p}}$$

So by the definition of  $d_p$ , we get

$$d_p(a, b)^p \leq d_p(a, c) d_p(a, b)^{p-1} + d_p(c, b) d_p(a, b)^{p-1}$$

Dividing by  $d_p(a, b)^{p-1}$ , we get the Minkowski's inequality whenever  $d_p(a, b) \neq 0$  and the inequality holds trivially whenever  $d_p(a, b) = 0$ . □

**Example 2.1 (Classical normed space).** We give some classical examples of normed vector spaces, which are vector spaces equipped with a norm that defines the "length" or "size" of vectors in that space:

(1) **Euclidean Space**  $\mathbb{R}^n$  with Euclidean Norm defined as:

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

where  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ .

We do verify that Euclidean norm is indeed a norm.

- **Positive Definiteness:**  $\|x\|_2 \geq 0$  with equality if and only if  $x = 0$ .
- **Homogeneity:**  $\|\alpha x\|_2 = |\alpha| \|x\|_2$ .

• **Triangle Inequality:**  $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$  (This follows from the Cauchy-Schwarz inequality and the Minkowski inequality in  $\mathbb{R}^n$ ).

- (2) **The vector space of all continuous Functions on  $[a, b]$  with Supremum Norm** defined as:

$$\|f\|_{\infty} = \sup_{x \in [a, b]} |f(x)|$$

We verify that *Supremum norm* is indeed a norm.

- **Positive Definiteness:**  $\|f\|_{\infty} \geq 0$  and  $\|f\|_{\infty} = 0$  only if  $f(x) = 0$  for all  $x \in [a, b]$ .
- **Homogeneity:** For any scalar  $\alpha$ ,  $\|\alpha f\|_{\infty} = |\alpha| \|f\|_{\infty}$ .
- **Triangle Inequality**  $\|f+g\|_{\infty} \leq \|f\|_{\infty} + \|g\|_{\infty}$  because for all  $x \in [a, b]$ ,  $|(f+g)(x)| \leq |f(x)| + |g(x)|$ .

- (3) **Sequence Space  $\ell^p$**

Consider the space of all sequences  $x = (x_n)$  such that:

$$\sum_{n=1}^{\infty} |x_n|^p < \infty$$

for  $1 \leq p < \infty$ . Define the  $p$ -norm:

$$\|x\|_p = \left( \sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}$$

We verify that it is indeed a norm.

- **Positive Definiteness:**  $\|x\|_p \geq 0$  with equality if and only if  $x_n = 0$  for all  $n$ .
- **Homogeneity:**  $\|\alpha x\|_p = |\alpha| \|x\|_p$ .
- **Triangle Inequality:** This is Minkowski's inequality for sequences:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

- (4) **Space of Square-Summable Sequences  $\ell^2$ .**

This is a specific case of  $\ell^p$  where  $p = 2$ . Therefore, proofs follows from the previous one. That is, sequences  $x = (x_n)$  where:

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty$$

and the  $\ell^2$ -norm is given by:

$$\|x\|_2 = \sqrt{\sum_{n=1}^{\infty} |x_n|^2}$$

This norm satisfies all the properties of a norm, including the triangle inequality, which can be proven using the Cauchy-Schwarz inequality in this infinite-dimensional setting.

The norm provides a way to measure the size of vectors or functions that is consistent with our intuitive notion of distance and size. At the same time, norm also allows for the application of analytical tools like convergence and completeness.

**Definition 2.2 (Convex set and convex function).** A subset  $A$  of a real or complex vector space is said to be **convex** if for each  $x, y \in A$ ,  $tx + (1-t)y \in A$  for all  $t \in [0, 1]$ . A real valued function on convex subset  $A$  of a real or complex vector space is said to be **convex** if for all  $x, y \in C$  and for all  $t \in [0, 1]$  we have  $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ .

**Proposition 2.1.** If  $\|\cdot\|$  is a norm on a real or complex vector space  $V$ , then  $\|\cdot\|$  is a convex function on  $V$ .

PROOF. Let  $x, y \in V$  and  $0 \leq t \leq 1$ . Then,

$$\|(1-t)x + ty\| \leq \|(1-t)x\| + \|ty\| = (1-t)\|x\| + t\|y\|.$$

This proves the claim. From this it is obvious that unit open and closed balls (defined with respect to norm as a metric) are convex subsets of a normed linear space  $V$ .  $\square$

We know that every norm indeed induces a metric. In the next theorem, we discuss a **criterion for a metric to induce a norm**.

**Theorem 2.2.** Let  $d$  be a metric on a real or complex vector space  $V$ . Then the following are equivalent–

- (1)  $d$  is a translation invariant metric and for all scalars  $\lambda$  and for  $x, y \in V$   $d(\lambda x, \lambda y) = |\lambda|d(x, y)$ .
- (2) there exists a norm  $\|\cdot\|$  on  $V$  such that  $d(x, y) := \|x - y\|$

PROOF. If  $d$  is a translation-invariant metric with the desired scaling properties, and we set  $\|x\| = d(x, 0)$ , then  $\|x\| = 0$  if and only if  $x = 0$ . Now, by multiple application of translation-invariant property we get:

$$\|x + y\| = d(x + y, 0) = d(x, -y) \leq d(x, 0) + d(0, -y) = d(x, 0) + d(y, 0) = \|x\| + \|y\|,$$

. Also,

$$\|\lambda x\| = d(\lambda x, 0) = d(\lambda x, \lambda 0) = |\lambda|d(x, 0) = |\lambda|\|x\|.$$

Thus we proved that (1) implies (2).

Conversely, assume that  $\|\cdot\|$  is a norm on  $V$ . We set  $d(x, y) = \|x - y\|$  and show that  $d$  is indeed a metric on  $V$ :

$$\begin{aligned} d(x, y) &= \|x - y\| = \|y - x\| = d(y, x), \\ d(x, y) &= 0 \text{ if and only if } \|x - y\| = 0, \text{ iff } x - y = 0, \text{ iff } x = y, \end{aligned}$$

and

$$d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z),$$

thus, triangle inequality holds true. Further,

$$d(x + a, y + a) = \|(x + a) - (y + a)\| = \|x - y\| = d(x, y)$$

thus,  $d$  is translation-invariant, and

$$d(\lambda x, \lambda y) = \|\lambda x - \lambda y\| = \|\lambda(x - y)\| = |\lambda|\|x - y\| = |\lambda|d(x, y).$$

Hence, (2) implies (1). □

### 3. Inner Product Spaces

In 1885, Hermann Schwarz gave another proof of Cauchy's inequality. Schwarz's proof is quite different from Cauchy's, and extends to a more general and more abstract setting. This provides some important examples of normed spaces and this forms the genesis of this new definition. Let  $V$  be a real vector space. We say  $\langle \cdot, \cdot \rangle$  is an inner product on  $V \times V \rightarrow \mathbb{R}$  if

- (1) **bilinearity**  $\langle \alpha v + \beta v', u \rangle = \alpha \langle v, u \rangle + \beta \langle v', u \rangle$  and  $\langle v, \alpha u + \beta u' \rangle = \alpha \langle v, u \rangle + \beta \langle v, u' \rangle$  for all  $u, u', v, v' \in V$  and  $\alpha, \beta \in \mathbb{R}$ .
- (2) **Symmetry**  $\langle x, y \rangle = \langle y, x \rangle$  for all  $x, y$  in  $V$ .
- (3) **positive definiteness**  $\langle x, x \rangle > 0$  for all  $x \in V - \{0\}$ .

A function that satisfies (1) and (2) is known as *symmetric bilinear form*. The inner product on  $\mathbb{R}^d$  given by  $\sum_{i=1}^n a_i b_i$ , where  $u = (a_1, \dots, a_n)$  and  $v = (b_1, \dots, b_n)$  is as *usual inner product* on  $\mathbb{R}^n$ .

Let  $V$  be a vector space over the complex field  $\mathbb{C}$ . An inner product on  $V$  is a function  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$  (often written as  $\langle u, v \rangle$  for  $u, v \in V$ ) that satisfies the following properties for all vectors  $u, v, w \in V$  and all scalars  $a, b \in \mathbb{C}$ :

- (1) **Linearity in the First Argument:**

$$\langle au + bv, w \rangle = a \langle u, w \rangle + b \langle v, w \rangle$$

- (2) **Sesquilinearity in the Second Argument:**

$$\langle v, \alpha u + \beta u' \rangle = \bar{\alpha} \langle v, u \rangle + \bar{\beta} \langle v, u' \rangle$$

where  $\bar{\cdot}$  denotes complex conjugation.

- (3) **Hermitian condition**  $\langle u, v \rangle = \overline{\langle v, u \rangle}$  where  $\bar{\cdot}$  denotes complex conjugation.
- (4) **Positive definiteness**

$$\langle v, v \rangle \geq 0 \quad \text{and} \quad \langle v, v \rangle = 0 \iff v = 0$$

A function that satisfies (1) and (2) is known as *Hermitian bilinear form*. The *usual inner product* on  $\mathbb{C}^d$ . The main difference between real and complex vector space is requirement of conjugate symmetry in complex inner product spaces.

Define  $\|x\| = \sqrt{\langle x, x \rangle}$ . We prove that it is indeed a norm and we call it *a norm induced by the given inner product* on a vector space. It is easy to verify the first two axioms of norm. However, to prove it satisfies the triangle inequality we need the fundamental result **Cauchy - Schwartz inequality** in inner product spaces.

**Theorem 3.1 (Cauchy-Schwartz inequality).** Let  $V$  be an inner product space. Then

$$|\langle u, v \rangle|^2 \leq \langle u, u \rangle \langle v, v \rangle$$

. In terms of norm induced by this inner product,

$$|\langle u, v \rangle|^2 \leq \|u\|^2 \|v\|^2$$

The equality holds iff  $u, v$  are collinear vectors.

PROOF. This depends upon the quadratic nature of the inner product and by this we mean that  $\|v\|^2 = \langle v, v \rangle = \sum_{i=1}^n v_i^2$ . Cauchy-Schwartz inequality is trivially true if at least one of the vectors  $x$ , or  $y$  is zero vector. Also, we prove it for complex vector space and for real it would follow from our proof.

Let us assume  $\lambda \in \mathbb{C}$ . Then,

$$0 \leq \|x + \lambda y\|^2 = \langle x + \lambda y, x + \lambda y \rangle = \langle x, x \rangle + \bar{\lambda} \langle x, y \rangle + \lambda \langle y, x \rangle + |\lambda|^2 \langle y, y \rangle.$$

Put

$$\lambda = -\frac{\langle x, y \rangle}{\langle y, y \rangle} \cdot \frac{\|x\|}{\|y\|}.$$

It follows that

$$0 \leq \|x\|^2 - 2 \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} \cdot \frac{\|x\|}{\|y\|} + \frac{\|x\|^2}{\|y\|^2} \|y\|^2 = 2 \left( \|x\|^2 - |\langle x, y \rangle| \cdot \frac{\|x\|}{\|y\|} \right),$$

so that  $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$ .

If  $x = 0$  or  $y = 0$ , then equality holds, and in that case,  $x$  and  $y$  are linearly dependent. Otherwise, if equality holds, then  $\|x + \lambda y\| = 0$ , so that  $x + \lambda y = 0$ , and  $x$  and  $y$  are linearly dependent. Conversely, if  $x$  and  $y$  are linearly dependent, then  $x = \alpha y$  for some scalar  $\alpha$ , and so

$$|\langle x, y \rangle| = |\alpha| \|y\|^2 = \|x\| \cdot \|y\|.$$

Note that if  $u, v \in \mathbb{R}^n$  or  $\mathbb{C}^n$  and we consider the usual inner product  $\langle u, v \rangle = \sum_{i=1}^n u_i v_i$  then Cauchy-Schwartz just becomes Cauchy's inequality. □

**Proposition 3.1 (Inner product spaces are normed spaces).** Let  $(V, \langle \cdot, \cdot \rangle)$  be a (real or complex) inner product space and set  $\|v\| = \langle v, v \rangle^{1/2}$ . Then  $(V, \|\cdot\|)$  is a normed space.

PROOF. We use Definition 2.1. Axioms N1 and N2 are straightforward verifications. Now for the axiom N3, we first need a lemma

**Lemma 3.1.**

$$\operatorname{Re}(\langle v_1, v_2 \rangle) \leq |\langle v_1, v_2 \rangle|$$

Where,  $\operatorname{Re}(\langle v_1, v_2 \rangle)$  is the real part of the inner product and  $|\langle v_1, v_2 \rangle|$  is the absolute value of the inner product.

PROOF. (1) **Case 1: Real Inner Product Space**

In a real inner product space, the inner product  $\langle v_1, v_2 \rangle$  is always a real number. Therefore,

$$\operatorname{Re}(\langle v_1, v_2 \rangle) = \langle v_1, v_2 \rangle.$$

As we know that the modulus of a real number is simply its absolute value and hence,

$$|\langle v_1, v_2 \rangle| = \max(\langle v_1, v_2 \rangle, -\langle v_1, v_2 \rangle).$$

Thus,

$$\operatorname{Re}(\langle v_1, v_2 \rangle) = \langle v_1, v_2 \rangle \leq |\langle v_1, v_2 \rangle|,$$

which is obviously true.

(2) **Case 2: Complex Inner Product Space**

In a complex inner product space, the inner product  $\langle v_1, v_2 \rangle$  is generally a complex number. Let:

$$\langle v_1, v_2 \rangle = a + bi, \quad \text{where } a = \operatorname{Re}(\langle v_1, v_2 \rangle) \text{ and } b = \operatorname{Im}(\langle v_1, v_2 \rangle).$$

The modulus of  $\langle v_1, v_2 \rangle$  is given by:

$$|\langle v_1, v_2 \rangle| = \sqrt{a^2 + b^2}.$$

Now, observe that:

$$\operatorname{Re}(\langle v_1, v_2 \rangle) = a.$$

Clearly, the modulus satisfies:

$$|\langle v_1, v_2 \rangle| = \sqrt{a^2 + b^2} \geq a,$$

because  $b^2 \geq 0$ . Therefore:

$$\operatorname{Re}(\langle v_1, v_2 \rangle) \leq |\langle v_1, v_2 \rangle|.$$

Hence we conclude that in both real and complex inner product spaces, the inequality

$$\operatorname{Re}(\langle v_1, v_2 \rangle) \leq |\langle v_1, v_2 \rangle|$$

always holds true. □

Now using this Lemma and Cauchy-Schwartz inequality, we get

$$\|v_1 + v_2\|^2 = \|v_1\|^2 + 2\operatorname{Re}(\langle v_1, v_2 \rangle) + \|v_2\|^2 \leq \|v_1\|^2 + 2|\langle v_1, v_2 \rangle| + \|v_2\|^2 \leq \|v_1\|^2 + 2\|v_1\|\|v_2\| + \|v_2\|^2 = (\|v_1\| + \|v_2\|)^2.$$

□

**Proposition 3.2 (Polarization and the Parallelogram Law).** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and  $\|\cdot\|$  the corresponding norm. Then

(i) **(Polarization)**

$$4\langle v_1, v_2 \rangle = \|v_1 + v_2\|^2 - \|v_1 - v_2\|^2, \quad \text{for } V \text{ real, while}$$

$$4\langle v_1, v_2 \rangle = \|v_1 + v_2\|^2 - \|v_1 - v_2\|^2 + i\|v_1 + iv_2\|^2 - i\|v_1 - iv_2\|^2, \quad \text{if } E \text{ is complex.}$$

(i) **(Parallelogram law)**

$$2\|v_1\|^2 + 2\|v_2\|^2 = \|v_1 - v_2\|^2 + \|v_1 + v_2\|^2.$$

PROOF. (i) We prove first for the case of complex vector space and to do so, we manipulate right hand side as follows.

$$\begin{aligned} \|v_1 + v_2\|^2 - \|v_1 - v_2\|^2 + i\|v_1 + iv_2\|^2 - i\|v_1 - iv_2\|^2 &= \|v_1\|^2 + 2\Re\langle v_1, v_2 \rangle + \|v_2\|^2 \\ &\quad - \|v_1\|^2 - 2\Re\langle v_1, v_2 \rangle - \|v_2\|^2 \\ &\quad + i\|v_1\|^2 + 2i\Re\langle v_1, iv_2 \rangle + i\|v_2\|^2 \\ &\quad - i\|v_1\|^2 - 2i\Re\langle v_1, iv_2 \rangle - i\|v_2\|^2 \\ &= 4\Re(\langle v_1, v_2 \rangle) + 4i\Im(\langle v_1, v_2 \rangle) \\ &= 4\langle v_1, v_2 \rangle. \end{aligned}$$

The real case is proved in a similar way.

(ii) For this too, we again manipulate the right hand side:

$$\begin{aligned}\|e_1 + e_2\|^2 + \|e_1 - e_2\|^2 &= \|e_1\|^2 + 2\Re\langle e_1, e_2 \rangle + \|e_2\|^2 \\ &\quad + \|e_1\|^2 - 2\Re\langle e_1, e_2 \rangle + \|e_2\|^2 \\ &= 2\|e_1\|^2 + 2\|e_2\|^2.\end{aligned}$$

□

*So, we have proved that every inner product space forms a norm linear space. However, not every norm comes from inner product space. To understand whether a norm is induced by an inner product or not it is necessary to check if the given norm satisfies the parallelogram law. For example, consider the norm*

$$\|x\| = \sum_{i=1}^n |x^i|$$

is not induced by any inner product since this norm fails to satisfy the parallelogram law.

**Normed Spaces are Metric Spaces.** We have seen that inner product spaces are normed spaces. Now, we show that normed spaces are metric spaces.

Thus, we have the following hierarchy of generality:

**Inner Product Spaces  $\subset$  Normed Spaces  $\subset$  Metric Spaces  $\subset$  Topological Spaces.**



FIGURE 3. Hermann Schwarz

#### 4. Convergence in the Metric Spaces

We briefly discuss the concept of convergence and limit here, and we will have an elaborate discussion on this topic in the subsequent sections. Given that convergence would be defined with respect to a metric, here the metric could be on an abstract set or be induced by structures like inner products and norms on linear spaces.

**Definition 4.1** (Convergence and limit). Let  $(X, d)$  be a metric space. Let  $(x_n)_{n=0}^\infty$  be a sequence of points in  $X$ . Let  $L \in X$ . We say that  $(x_n)_{n=0}^\infty$  converges to  $L$ , **with respect to metric  $d$**  if for every  $\epsilon > 0 \exists N_\epsilon \geq 1 : d(x_n, L) \leq \epsilon$  for all  $n \geq N_\epsilon$ . Note that here we used  $\epsilon$  as a subscript in  $N_\epsilon$  to emphasize that  $N$  is dependent upon  $\epsilon$ . This is equivalent to say that  $\lim_{n \rightarrow \infty} d(x_n, L) = 0$ .

**Example 4.1** (Terence Tao - Analysis II, Exercise 1.1.15). Let  $X = \{(a_n)_{n=0}^\infty : \sum_{n=0}^\infty |a_n| < \infty\}$ . be the space of absolutely convergent sequences. Define  $l^1, l^\infty$  metrics on this space by

$$d_{l^1}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = \sum_{n=0}^\infty |a_n - b_n|;$$

$$d_{l^\infty}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = \sup_{n \in \mathbb{N}} |a_n - b_n|;$$

Show that these are both metric on  $X$ , but show that there exist sequences  $x^1, x^2, \dots$  of elements of  $X$  which are convergent with respect to  $d_{l^\infty}$  metric but not with respect to the  $d_{l^1}$ . Conversely, show that any sequence which converges in  $d_{l^1}$  automatically converges in  $d_{l^\infty}$

**Solution 4.1.** Note that

$$d_{l^1}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = \sum_{n=0}^\infty |a_n - b_n| \leq \sum_{n=0}^\infty |a_n| + \sum_{n=0}^\infty |b_n|.$$

By the very definition of space  $X$ , each of these sums is bounded and more specifically each of these two sums have a unique positive real number as their values. Hence  $d_{l^1}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = \sum_{n=0}^\infty |a_n - b_n|$  is indeed a function from  $X$  to  $[0, \infty)$ . Now, we prove that it satisfies the axiom of being a metric. Positive definiteness and symmetry properties are very obvious by the definition of  $|\cdot|$  function on  $\mathbb{R}$ . The triangle inequality follows from the following fact -

$$|a_n - c_n| = |a_n - b_n + b_n - c_n| \leq |a_n - b_n| + |b_n - c_n|$$

Therefore,

$$\sum_{n=1}^\infty |a_n - c_n| = \sum_{n=1}^\infty |a_n - b_n + b_n - c_n| \leq \sum_{n=1}^\infty |a_n - b_n| + \sum_{n=1}^\infty |b_n - c_n|$$

Therefore, the metric satisfies the triangle inequality. That is,

$$d_{l^1}((a_n)_{n=0}^\infty, (c_n)_{n=0}^\infty) \leq d_{l^1}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) + d_{l^1}((b_n)_{n=0}^\infty, (c_n)_{n=0}^\infty)$$

Now for the second metric

$$d_{l^\infty}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = \sup_{n \in \mathbb{N}} |a_n - b_n|;$$

we prove that it is indeed a function with values in  $[0, \infty)$ . To do so, note that

$$\sup_{n \in \mathbb{N}} |a_n - b_n| \leq \sup_{n \in \mathbb{N}} |a_n| + \sup_{n \in \mathbb{N}} |b_n| \leq \sum_{n=1}^\infty |a_n| + \sum_{n=1}^\infty |b_n|$$

and extreme right of this inequality is indeed a finite and unique positive real number. Hence, the given metric  $d_{l^\infty}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = \sup_{n \in \mathbb{N}} |a_n - b_n|$  takes values in  $[0, \infty)$ . Positive definiteness and symmetry are very easy to verify. We only prove the triangle inequality. Note that

$$|a_n - c_n| = |a_n - b_n + b_n - c_n| \leq |a_n - b_n| + |b_n - c_n|$$

So, we take supremum over the set  $n \in \mathbb{N}$  and get

$$\sup_{n \in \mathbb{N}} |a_n - c_n| = \sup_{n \in \mathbb{N}} |a_n - b_n + b_n - c_n| \leq \sup_{n \in \mathbb{N}} |a_n - b_n| + \sup_{n \in \mathbb{N}} |b_n - c_n|$$

Therefore, the metric satisfies the triangle inequality. That is,

$$d_{l^\infty}((a_n)_{n=0}^\infty, (c_n)_{n=0}^\infty) \leq d_{l^\infty}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) + d_{l^\infty}((b_n)_{n=0}^\infty, (c_n)_{n=0}^\infty)$$

**Example 4.2** (Terence Tao - Analysis II, Exercise 1.1.16). Let  $(x_n)_{n=1}^{\infty}$  and  $(y_n)_{n=1}^{\infty}$  be two sequences in a metric space  $(X, d)$ . Suppose that  $(x_n)_{n=1}^{\infty}$  converges to a point  $x \in X$  and  $(y_n)_{n=1}^{\infty}$  converges to a point  $y \in X$ . Show that  $\lim_{n \rightarrow \infty} d(x_n, y_n) = d(x, y)$ .

**Solution 4.2.** Note that in metric space  $(X, d)$  if  $x, y, z, u$  are points then we can apply the **generalized triangle inequality** to arrive at  $d(x, y) \leq d(x, z) + d(z, u) + d(u, y) \implies d(x, y) - d(z, u) \leq d(x, z) + d(y, u)$ . Also, we can reformulate it as  $d(z, u) \leq d(z, x) + d(x, y) + d(y, u) \implies d(z, u) - d(x, y) \leq d(z, x) + d(y, u)$ . Combining these two inequalities, we conclude

$$|d(x, y) - d(z, u)| \leq d(x, z) + d(y, u). \text{ This is known as quadrilateral inequality}$$

. In order to prove that  $\lim_{n \rightarrow \infty} d(x_n, y_n) = d(x, y)$  it is enough to show that for each  $\epsilon > 0$ , there exists a  $N \geq 1$ , depending upon this  $\epsilon$ , such that for  $n \geq N$ , we have  $|d(x_n, y_n) - d(x, y)| \leq \epsilon$ . Since  $(x_n)_{n=1}^{\infty}$  converges to  $x$  therefore for a given  $\epsilon > 0$ , there exists an  $N_1$  such that for all  $n \geq N_1$  we have

$$d(x_n, x) \leq \frac{\epsilon}{2}.$$

Also,  $(y_n)_{n=1}^{\infty}$  converges to  $y$  therefore for a given  $\epsilon > 0$ , there exists an  $N_2$  such that for all  $n \geq N_2$  we have

$$d(y_n, y) \leq \frac{\epsilon}{2}.$$

Now, set  $N := \max N_1, N_2$ . So for  $n \geq N$  we have

$$d(x_n, x) \leq \frac{\epsilon}{2} \text{ and}$$

$$d(y_n, y) \leq \frac{\epsilon}{2}.$$

Add these two and we get  $d(x_n, x) + d(y_n, y) \leq \epsilon$  for all  $n \geq N$ . But by **quadrilateral inequality**,

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) \leq \epsilon$$

for all  $n \geq N$ . This is what we had set to prove.

## 5. Some Point-Set Topology of Metric Spaces

This particular section is joint work in collaboration with Tisya Rawat (a sophomore at Caltech. She is majoring in Mathematics and we have been classmates at AMMOC since 2021).

**Definition 5.1 (Balls).** Let  $(X, d)$  be a metric space, let  $x_0$  be a point in  $X$ , and let  $r > 0$ . We define the **ball**  $B_{(X, d)}(x_0, r)$  in  $X$ , centered at  $x_0$ , and with radius  $r$ , in the metric  $d$ , to be the set

$$B_{(X, d)}(x_0, r) := \{x \in X : d(x, x_0) < r\}.$$

**Definition 5.2 (Bounded sets in metric spaces).** Let  $(X, d)$  be a metric space and let  $E \subseteq X$ . We say that  $E$  is bounded if, for any arbitrarily chosen point  $a \in X$ , the distance  $d(a, e) < \lambda$ , for all  $e \in E$ , is bounded above by some positive number  $\lambda$ . This is equivalent to say that  $E$  is a bounded set if it can be contained in some ball of certain finite positive radius.

**Definition 5.3 (Interior, Exterior, Boundary).** Let  $(X, d)$  be a metric space, let  $E$  be a subset of  $X$ , and let  $x_0$  be a point in  $X$ . We say that  $x_0$  is an **interior point** of  $E$  if there exists a radius  $r > 0$  such that  $B(x_0, r) \subseteq E$ . We say that  $x_0$  is an **exterior point** of  $E$  if there exists a  $r > 0$  such that  $B(x_0, r) \cap E = \emptyset \iff B(x_0) \subset E^c$ . We say that  $x_0$  is a **boundary point** of  $E$  if it is neither an interior point nor an exterior point of  $E$ .

**Remark 5.1.** So, by definition 5.3, we can see that the interior point lies inside the set  $E$ . That is  $\text{int}(E) \subseteq E$ . Also, we can see that the exterior point of  $E$  lies outside the set  $E$ . Also if  $x_0$  is a boundary point  $\iff$  for each  $r > 0$ , we have  $B(x_0, r) \cap E \neq \emptyset$  and  $B(x_0, r) \cap E^c \neq \emptyset$

**Definition 5.4 (Limit points and closure).** Let  $(X, d)$  be a metric space, let  $E$  be a subset of  $X$ , and let  $x_0$  be a point in  $X$ . We say that  $x_0$  is an **adherent or limit point** of  $E$  if for every radius  $r > 0$ , the ball  $B(x_0, r)$  has a non-empty intersection with  $E$ . The set of all adherent points of  $E$  is called the **closure** of  $E$  and is denoted  $\bar{E}$ .

**Definition 5.5 (Limit point of a sequence).** Let  $(x_n)_{n=1}^{\infty}$  be a sequence of points in a metric space  $(X, d)$ . We can construct a set  $E$  using the points of this sequence and then we can define the limit point of this sequence  $(x_n)_{n=1}^{\infty}$  as the limit point of the set  $E$  in the sense of Definition 5.4

**Proposition 5.1 (Characterization of limit/adherent point).** Let  $(X, d)$  be a metric space and  $E$  be a subset of  $X$ . Let  $x_0$  be a point in  $X$ . Then the following statements are logically equivalent.

- (a)  $x_0$  is an adherent point of  $E$ .
- (b)  $x_0$  is either an interior point or a boundary point of  $E$ .
- (c) There exists a sequence  $(x_n)_{n=1}^{\infty}$  in  $E$  which converges to  $x_0$  with respect to the metric  $d$ .

**Remark 5.2 (On the Axiom of Choice).** For some of the implications one will need the axiom of choice and one author of this project namely, Sarthak Dhobale has done a very detailed project on Axiom of Choice, its several equivalent formulation and its various implications for example

- Existence of basis for a vector space
- Existence of maximal subgroup of a finitely generated group
- Existence of Maximal Ideal
- Hahn Banach Theorem

The project is available at his [webpage](#).

PROOF. (1)  $\implies$  (2): Let  $x_0$  is an adherent point of  $E$  and we prove (2). First note that for  $x_0$  to be an exterior point there must exists a radius  $r > 0$  such that  $B(x_0, r) \cap E = \emptyset$ . However, by definition 5.4 of adherent point, for each  $r > 0$ , the ball  $B(x_0, r) \cap E \neq \emptyset$ . Hence, no such positive  $r$  exists for which  $x_0$  can be an exterior point. Therefore, every adherent/limit point is either an interior point or a boundary point.

For (2)  $\implies$  (1). Let  $x_0$  be either an interior point or a boundary point of  $E$ . So,  $x_0$  is not an exterior point of  $E$ . This means that for every radius  $r > 0$ , the ball  $B(x_0, r)$  has a non-empty intersection with  $E$ , which is precisely the definition of adherent point.

To prove (1)  $\implies$  (3): Let us assume that  $x_0$  is an adherent or limit point in the sense of the definition 5.4. So, for each given  $\epsilon > 0$ , the ball  $B(x_0, \epsilon)$  has non-empty intersection with  $E$ . In particular, for each  $n$  if we set  $\epsilon_n := \frac{1}{n}$ , then for each of  $\epsilon_n$ , there exists an  $x_n$  such that  $x_n \in B(x_0, \epsilon_n) \cap E$  and therefore, for these  $x_n$ , which we chose using axiom of choice, we have  $d(x_n, x_0) < \epsilon_n = \frac{1}{n}$  and hence the sequence  $(x_n)_{n=1}^{\infty}$  converges to  $x_0$  and each element of this sequence comes from  $E$ .

To prove (3)  $\implies$  1. Let us assume that there exists a sequence  $(x_n)_{n=1}^{\infty}$  of points in  $E$  that converges to  $x_0$ . Then infinitely many points of that sequence lie in ball  $B(x_0, \epsilon)$  of any positive radius  $\epsilon$ . This precisely is the condition for the  $x_0$  to be adherent point. This concludes the proof of the Proposition 5.1  $\square$

Now we state an immediate corollary of this proposition in context of sequences.

**Corollary 5.1.** Let  $(x_n)_{n=1}^{\infty}$  be a sequence of elements in metric space  $(X, d)$ . Set  $E = \{x_n | n \in \mathbb{N}\}$ . Then, in the sense of Proposition 5.1, a number  $L$  will known as a limit point of sequence  $(x_n)_{n=1}^{\infty}$  if and only if there exists a subsequence of the sequence  $(x_n)_{n=1}^{\infty}$  which converges to  $L$ .

**Corollary 5.2.** Let  $(X, d)$  be a metric space, and let  $E$  be a subset of  $X$ . Then  $E = \text{int}(E) \cup \partial E = X \setminus \text{ext}(E)$ .

PROOF. Note that if  $x_0 \in X$  is an exterior point of  $E$ , then there exists an  $r > 0$  such that  $B(x_0, r) \cap E = \emptyset$ . In particular,  $x_0 \notin E$ . Every point of  $E$  is either its interior or boundary point.  $\square$

**Corollary 5.3 (Supremum and infimum of a bounded set of real line).** Let  $E$  be a non empty subset of the real number bounded from above. Let  $\sup E \notin E$ . Then  $\sup E$  is a limit point of  $E$ . Further, let  $F$  be a non-empty subset of real number bounded from below such that  $\inf F \notin F$ , then  $\inf F$  is a limit point of  $F$ .

PROOF. Since supremum  $\sup E$  is the **smallest of all upper bounds**, therefore, for each integer  $n > 0$ ,  $\sup E - \frac{1}{n}$  is not an upper bound for the set  $E$ . This implies that for each  $n \geq 1$ , there exists an element  $x_n \in E$  such that  $\sup E - \frac{1}{n} < x_n \leq \sup E$ . This sequence  $(x_n)_{n=1}^{\infty}$  lies entirely in  $E$  and converges to  $\sup E$ . Therefore, by Proposition 5.1, we conclude that  $\sup E$  is indeed an adherent/limit point of  $E$ . Similarly, we can conclude the result for infimum of the set  $E$ .  $\square$

**Theorem 5.1.** If the distance between any two points of a sequence  $(x_n)_{n=1}^{\infty}$  is bounded below by a positive constant  $C$ , that is,  $d(x_i, x_j) > C$  for all  $i, j \in \mathbb{N}$ , then this sequence  $(x_n)_{n=1}^{\infty}$  has no limit point.

PROOF. Seeking a contradiction, let us assume that this sequence has a limit point  $L$ . Set  $\epsilon = \frac{C}{2}$ . Now for each  $i, j \in \mathbb{N}$ . For this setup, we can eventually find  $i, j \in \mathbb{N}$  such that  $d(x_i, L) < \frac{C}{2}$  and  $d(L, x_j) < \frac{C}{2}$ . Adding these two inequations we get  $d(x_i, x_j) < d(x_i, L) + d(L, x_j) < \frac{C}{2} + \frac{C}{2} = C$ , contradicting the given hypothesis that for all  $i, j \in \mathbb{N}$  we have  $d(x_i, x_j) > C$ .  $\square$

**Theorem 5.2 (Nested interval theorem).** Let  $Q$  be any system of nested closed intervals  $[a, b]$ . There exists at least one point of  $\mathbb{R}$  belonging to every interval of  $Q$ . More exactly, the numbers

$$\xi = \sup\{a : [a, b] \in Q\}$$

and

$$\eta = \inf\{b : [a, b] \in Q\}$$

exist ( $\xi \leq \eta$ ) and the interval  $[\xi, \eta]$  is the intersection of all the intervals of  $Q$ .

PROOF. Let  $Q$  be a system of nested closed intervals. Define the following sets:

-  $A = \{a \mid [a, b] \in Q\}$ , i.e., the set of all left endpoints of intervals in  $Q$ . -  $B = \{b \mid [a, b] \in Q\}$ , i.e., the set of all right endpoints of intervals in  $Q$ .

Since  $Q$  is a system of nested closed intervals, for any two intervals  $[a_1, b_1]$  and  $[a_2, b_2]$  in  $Q$ , one must be contained within the other. Without loss of generality, assume  $[a_1, b_1] \subseteq [a_2, b_2]$ . Thus, we have:

$$a_2 \leq a_1 \leq b_1 \leq b_2.$$

Therefore, for any arbitrary  $a \in A$  and  $b \in B$ , it follows that:

$$a \leq b.$$

By the lemma, the supremum of  $A$  exists and is denoted by  $\xi$ , and the infimum of  $B$  exists and is denoted by  $\eta$  and  $\xi \leq \eta$ .

By the definitions of  $\xi$  and  $\eta$ , for each interval  $[a, b] \in Q$ , it holds that:

$$a \leq \xi \leq \eta \leq b.$$

This implies:

$$[\xi, \eta] \subseteq [a, b]$$

for every interval  $[a, b] \in Q$ . Consequently:

$$[\xi, \eta] \subseteq \bigcap_{[a, b] \in Q} [a, b].$$

Now, we need to prove that:

$$\bigcap_{[a, b] \in Q} [a, b] \subseteq [\xi, \eta].$$

To do this, suppose there exists a point  $x$  such that  $x \notin [\xi, \eta]$ . We need to show that  $x \notin [a, b]$  for any interval  $[a, b] \in Q$ .

There are two cases to consider:

1. Case 1:  $x < \xi$

Since  $\xi$  is the supremum of  $A$ , there must be some left hand point  $a$  such that  $x < a < \xi$ . Therefore,  $x < a$ , which implies:

$$x \notin [a, b].$$

2. Case 2:  $x > \eta$

Since  $\eta$  is the infimum of  $B$ , there must be some right hand point  $b$  such that  $x > b > \eta$ . Therefore,  $x > b$ , which implies:

$$x \notin [a, b].$$

In both cases,  $x \notin [a, b]$ . Thus,  $x \notin \bigcap_{[a,b] \in Q} [a, b]$  whenever  $x \notin [\xi, \eta]$ .

Hence:

$$\bigcap_{[a,b] \in Q} [a, b] \subseteq [\xi, \eta].$$

Combining both inclusions, we conclude:

$$\bigcap_{[a,b] \in Q} [a, b] = [\xi, \eta].$$

□

**Theorem 5.3 (Bolzano-Weierstrass).** Every infinite set of points in a closed interval  $[a, b] \subset \mathbb{R}$  has at least one limit point

**History:** This theorem was first proved by Bolzano in 1817 as a lemma in his proof of the intermediate value theorem. Around a half century later, this result became one of the principal results in the development of modern analysis and was again proved by Karl Weierstrass, often known as 'father of modern analysis'. Since then, it is known as Bolzano-Weierstrass theorem.

**PROOF.** Let  $(x_n)_{n \in \mathbb{N}}$  be a sequence of real numbers with  $|x_n| \leq L$  for all  $n \in \mathbb{N}$ .

**Step 1 - Search of the limit point**

- Set  $a_0 = -L$  and  $b_0 = L$ . Note that  $|a_0 - b_0| = 2L$ .
- Divide the interval  $[a_0, b_0]$  into two halves. Since infinitely points of sequence lie in the the interval  $[-L, L]$ , therefore, both half can't contain only finitely many points of  $s_n$ . That is, one half must contain infinitely many of  $s_n$ . Choose one such half and name it  $[a_1, b_1]$ . Note that  $|b_1 - a_1| = \frac{1}{2}|b_0 - a_0| = L$ , and that  $a_1 \in \left\{a_0, \frac{a_0+b_0}{2}\right\}$ . So  $a_1 \geq a_0$ . There are infinitely many  $s_n$  in  $[a_1, b_1]$ . Select one, say  $s_{k_1}$ .
- Divide the interval  $[a_1, b_1]$  into two halves. At least one half must contain infinitely many  $s_n$ . Pick one such half and call it  $[a_2, b_2]$ . Note that  $|b_2 - a_2| = \frac{1}{2}|b_1 - a_1| = \frac{L}{2}$ , and that  $a_2 \in \left\{a_1, \frac{a_1+b_1}{2}\right\}$ . So  $a_2 \geq a_1$ . There are infinitely many  $s_n$  in  $[a_2, b_2]$ . Select one with  $s_{k_2}$ .

In this way we generate

$$[a_0, b_0] \supseteq [a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots,$$

$$|b_{n+1} - a_{n+1}| = \frac{1}{2}|b_n - a_n| = \frac{L}{2^{n+1}}.$$

Note that  $|b_n - a_n| = \frac{L}{2^n} \rightarrow 0$  as  $n \rightarrow \infty$ . Thus, there is a unique limit  $s \in \bigcap_{n=1}^{\infty} [a_n, b_n]$  such that  $s_n \rightarrow s$  as  $n \rightarrow \infty$ .

**Step 2 - to conjecture the limit**

The sequence  $a_1, a_2, a_3, \dots$  is monotone increasing and bounded above by  $b_0 = L$ , so it converges. Call the limit  $s$ .



FIGURE 4. Bernard Bolzano, one of the earliest mathematicians to begin instilling rigor into mathematical analysis



FIGURE 5. Karl Weierstrass, often known as "father of modern analysis"

**Step 3: To prove that  $s_n \rightarrow s$ .**

Let  $\epsilon > 0$ . Since  $(a_n)_{n \in \mathbb{N}}$  converges to  $s$ , there exists  $N_1 \in \mathbb{N}$  such that  $|a_n - s| < \frac{\epsilon}{2}$  whenever  $n \geq N_1$ .

Since  $s_n$  lies in the interval  $[a_n, b_n]$ , and the length of the interval  $[a_n, b_n] = |b_n - a_n| = \frac{L}{2^n}$ , the distance from  $s_n$  to  $s$ , at most  $\frac{L}{2^n}$ , which converges to zero as  $n \rightarrow \infty$ , we can find an  $N_2 \in \mathbb{N}$  such that  $\frac{L}{2^n} < \frac{\epsilon}{2}$  whenever  $n \geq N_2$ .

Choose  $N = \max\{N_1, N_2\}$ . Then  $n \geq N \implies |s_n - s| \leq |s_n - a_n| + |a_n - s| < \epsilon$ . ■

□

**Definition 5.6 (Open and closed sets).** Let  $(X, d)$  be metric space, and let  $E \subset X$ . We say that  $E$  is closed in  $X$  if it contains all of its **boundary points**, that  $\partial E \subseteq E$ . We say that  $E$  is **open** in  $X$  if it does not contains any of its boundary points. That is  $\partial E \cap E = \emptyset$ . If  $E$  contains some of its boundary points and doesn't contains others, then we say  $E$  is neither open nor closed.

**Remark 5.3 (Algebraic closure vs topological closure).** An algebraic structure (group, rings, fields etc.) is closed under one or more arithmetic operations of additions, subtraction, multiplication, and division if performing these operations on elements of underlying structure gives back again an element within the same structure. In metric structure the closure is under "limits". The limit of each convergent sequence must be inside the set.

**Proposition 5.2 (Basic properties of open and closed sets).** Let  $(X, d)$  be a metric space

- (1) Let  $E$  be a subset of  $X$ . Then  $E$  is open if and only if  $E = \text{int}(E)$ . In other words,  $E$  is open if and only if for every  $x \in E$ , there exists an  $r > 0$  such that  $B(x, r) \subseteq E$ .
- (2) Let  $E$  be a subset of  $X$ . Then  $E$  is closed if and only if  $E$  contains all its adherent points. In other words,  $E$  is closed if and only if for every convergent sequence  $(x_n)_{n=m}^{\infty}$  in  $E$ , the limit  $\lim_{n \rightarrow \infty} x_n$  of that sequence also lies in  $E$ .

- (3) For any  $x_0 \in X$  and  $r > 0$ , the ball  $B(x_0, r)$  is an open set. The set  $\{x \in X : d(x, x_0) \leq r\}$  is a closed set. (**This set is sometimes called the closed ball of radius  $r$  centered at  $x_0$ .**)
- (4) Any singleton set  $\{x_0\}$ , where  $x_0 \in X$ , is automatically closed.
- (5) If  $E$  is a subset of  $X$ , then  $E$  is open if and only if the complement  $X/E := \{x \in X : x \notin E\}$  is closed. **This basically can be stated as "a set is open if and only if none of its points are adherent points of its complement".**
- (6) If  $E_1, E_2, \dots, E_n$  are a finite collection of open sets in  $X$ , then  $E_1 \cap E_2 \cap \dots \cap E_n$  is also open. If  $F_1, \dots, F_n$  is a finite collection of closed sets in  $X$ , then  $F_1 \cup F_2 \cup \dots \cup F_n$  is also closed.
- (7) If  $\{E_\alpha\}_{\alpha \in I}$  is a collection of open sets in  $X$  (where the index set  $I$  could be **finite, countable, or uncountable**), then the union  $\bigcup_{\alpha \in I} E_\alpha := \{x \in X : x \in E_\alpha \text{ for some } \alpha \in I\}$  is also open. If  $\{F_\alpha\}_{\alpha \in I}$  is a collection of closed sets in  $X$ , then the intersection  $\bigcap_{\alpha \in I} F_\alpha := \{x \in X : x \in F_\alpha \text{ for all } \alpha \in I\}$  is also a closed set.
- (8) If  $E$  is any subset of  $X$ , then  $\text{int}(E)$  is the largest open set which is contained in  $E$ ; in other words,  $\text{int}(E)$  is open, and given any other open set  $V \subseteq E$ , we have  $V \subseteq \text{int}(E)$ . Similarly,  $\bar{E}$  is the smallest closed set which contains  $E$ ; in other words,  $\bar{E}$  is closed, and given any other closed set  $K \supseteq E$ , we must have  $K \supseteq \bar{E}$ .

PROOF. (1) Suppose  $E$  is open. So by Definition 5.6,  $E$  does not contain any of its boundary point, that is,  $\partial E \cap E = \emptyset$ . **Our goal is to prove that  $E = \text{int}(E)$**  and to achieve it, we use set theoretical inclusion. First note that due to Remark 5.1, we know that  $\text{int}(E) \subseteq E$ . So we set out to prove  $E \subseteq \text{int}(E)$ . Let  $x_0 \in E$ . We need to prove that  $x_0 \in \text{int}(E)$ .  $E$  is open  $\iff$  it does not contain any of its boundary points, see Definition 5.3. Also, since  $x_0 \in E$ , therefore, by remark 5.1,  $x_0$  can not be an exterior point either, that is, there is no  $r > 0$  for which  $B(x_0, r)$  has non-empty intersection with  $E$ . This means for each  $r > 0$ , we have  $B(x_0, r) \neq \emptyset$ . So, either  $x_0$  is either a boundary point or an interior point. However, since  $E$  is open so it does not contain any of its boundary points. Hence,  $x_0$  must be an interior point. **Next, our goal is to prove that  $E = \text{int}(E)$**  implies that  $E$  is open in the sense of the Definition 5.6. So, we need to prove that  $E = \text{int}E$  implies that  $E$  does not contain any of its boundary points. Let  $x_0$  be the boundary point of  $E$ . So, in the sense of Definition 5.6 and Remark 5.1,  $x_0 \notin E$ .

- (2) We need to show that  **$E$  is closed  $\iff E$  contains all of its adherent points**. To prove in the forward direction, let us assume that  $E$  is closed and we need to show that it contains all of its adherent points. Let  $x_0$  be an adherent point of  $E$ . By Proposition 5.1, we know that  $x_0$  is either an interior or boundary point. If  $x_0$  is an interior point, then by the Remark 5.1,  $x_0 \in E$ . If  $x_0$  is a boundary point then given that  $E$  is closed and which by Definition 5.6 implies that  $E$  contains all its boundary points and hence  $x_0 \in E$ .

To prove in backward direction  $E$  contains all of its adherent point implies that  $E$  is closed, which by definition means that  $E$  contains all of its boundary point. Since, by Proposition 5.1 we know that boundary points are adherent point and that we are given that  $E$  contains all of its adherent points. Therefore,  $E$  contains all its boundary point. Hence,  $E$  is closed. The last equivalence follows directly from what we proved and characterization of adherent points as proved in Proposition 5.1.

- (3) By Definition 5.6 of open sets, we need to prove that each point of  $B(x_0, r)$  is its interior point. Let  $x \in B(x_0, r)$ . We prove that there exists a  $r' > 0$  such that  $B(x, r') \subseteq B(x_0, r)$ . We claim that  $r' = r - d(x_0, x)$ , which is definitely positive, works. That is,  $B(x, r') \subseteq B(x_0, r)$ . Let  $u \in B(x, r')$ . So,  $d(u, x) < r'$ . Now,  $d(x_0, u) \leq d(x_0, x) + d(x, u) < d(x_0, x) + r' = r$ . **Now, to prove that  $\bar{B}(x_0, r) = \{x \in X | d(x_0, x) \leq r\}$  is indeed a closed ball.** We need to show that this set contains all of its boundary point. Let  $u \in X$  be its boundary point. We need to prove that  $u \in \bar{B}(x_0, r)$ . Suppose that  $u \notin \bar{B}(x_0, r)$ . This means that  $d(u, x_0) > r$ . Define  $r_0 = \frac{r - d(u, x_0)}{2}$ . Now consider a ball of radius  $r_0$  and centered at  $u$ . Since boundary points are adherent points, and therefore, in view of their characterization in Proposition 5.1, there exists a sequence of elements

- in  $\bar{B}(x_0, r)$  which converges to  $u$  and therefore, infinitely many points of this sequence will lie inside the ball  $B(u, r_0)$  which lies completely outside the ball  $\bar{B}(x_0, r)$ , a contradiction.
- (4) We have to prove that singleton set  $\{x_0\}$  is a closed set. To do so we refer to part (2) of this proposition itself, namely Proposition 5.2, according to which, "a set is closed iff every convergent sequence of elements in the given set converges inside the given set itself." Now, for a set that has just one element, any sequence of elements in this set will be the constant sequence defined by the element given in this singleton set. The limit of constant sequence is the term of the sequence itself, and hence it lies in the given singleton set. Therefore, singletons are closed.
- (5) Let us assume that  $E$  is open and then we aim to prove that  $E^c = \{x \in X | x \notin E\}$  is closed in  $X$ . Let  $x_n$  be a sequence of points in  $E^c$ . Let  $x_n$  converges with respect to metric  $d$  and let  $L$  be the limit of this sequence. We claim that  $L \in E^c$ . Seeking a contradiction, assume that  $L \notin E^c$ . Then  $L \in E$ . But  $E$  is an open set and therefore, eventually infinitely many terms of this sequence  $x_n$  will lie in  $E$  which contradicts the fact that  $x_n$  is a sequence of points in  $E^c$ . Now assume that  $E^c$  is closed and we aim to prove that  $E$  is open. We need to prove that each point  $x \in E$  is an interior point of  $E$ . For this we need to find at least one  $r > 0$  such that  $B(x, r) \subseteq E$ . Seeking a contradiction, let us assume that no such  $r > 0$  exists. This means for each  $r > 0$ , the ball  $B(x, r)$  does not completely lie within  $E$ . Which basically means that for each  $r > 0$ ,  $B(x, r) \cap E^c \neq \emptyset$ . But this, by definition 5.4, means that  $x$  is an adherent point of  $E^c$ . However, we had assumed that  $E^c$  is closed, therefore,  $x \in E^c$ . But this contradicts the facts that  $x \in E$ .
- (6) Let  $E_1, \dots, E_n$  be open sets. We have to show that their intersection is an open set. Let  $x \in \bigcap_{i=1}^n E_i$ . Thus, for each  $1 \leq i \leq n$ ,  $x \in E_i$ , and since each  $E_i$  is open therefore, there exists a  $r_i > 0$  such that  $B(x, r_i) \subseteq E_i$ . Set  $r := \min(r_1, \dots, r_n)$ . We have  $B(x, r) \subseteq B(x, r_i) \subseteq E_i$  for all  $1 \leq i \leq n$ . That is  $B(x, r) \subseteq \bigcap_{i=1}^n E_i$ . Thus,  $E_1 \cap \dots \cap E_n$  is an open set.
- Also, let  $F_1, \dots, F_n$  be closed sets. We need to prove that  $\bigcup_{i=1}^n F_i$  is a closed set. By part (5) of this proposition, we need to show that its complement  $X \setminus \bigcup_{i=1}^n F_i$  is open. However, note that  $X \setminus \bigcup_{i=1}^n F_i = \bigcap_{i=1}^n X \setminus F_i$ . By part (e), each of  $X \setminus F_i$  is open and we have proven that finite intersection of open sets is an open sets. This implies that,  $X \setminus \bigcup_{i=1}^n F_i$  is open. However, note that  $X \setminus \bigcup_{i=1}^n F_i = \bigcap_{i=1}^n X \setminus F_i$  is an open set. Hence, the finite union of closed sets is a closed set.
- (7) **To prove that arbitrary union of an open set is an open set.** Let  $\{E_\alpha\}_{\alpha \in I}$  be a collection of open sets in  $X$  (where the index set  $I$  could be **finite, countable, or uncountable**), then the union  $\bigcup_{\alpha \in I} E_\alpha := \{x \in X : x \in E_\alpha \text{ for some } \alpha \in I\}$  is also an open set. Let  $x \in \bigcup_{\alpha \in I} E_\alpha$ . Then there exists some  $\alpha_1 \in I$  such that  $x \in E_{\alpha_1}$  and since  $E_{\alpha_1}$  is open in  $X$ , therefore, there exists an  $r_{\alpha_1} > 0$  such that  $B(x, r_{\alpha_1}) \subseteq E_{\alpha_1} \subseteq \bigcup_{\alpha \in I} E_\alpha$ . Therefore,  $x$  is an interior point of  $\bigcup_{\alpha \in I} E_\alpha$ . Hence, the arbitrary union of an open set is an open set. **Now, a set is open iff its complement in  $X$  is closed.** Therefore, the complement of  $X \setminus \bigcup_{\alpha \in I} E_\alpha = \bigcap_{\alpha \in I} X \setminus E_\alpha$  is closed and note that for each  $\alpha \in I$ ,  $X \setminus E_\alpha$  is closed. This proves that the **arbitrary intersection of the closed sets is a closed set in  $X$ .**
- (8) To prove that  $\text{int}E$  is open we will use part (7) wherein we established that arbitrary union of an open set is an open set. Note that for each  $x \in \text{int}E$  there exists at least one  $r_x > 0 : B(x, r_x) \subseteq E$ , where  $x$  is used as a subscript to emphasize the fact that  $r_x$  is dependent upon  $x$ . Now

$$\text{int}E = \bigcup_{x \in E} B(x, r_x)$$

is the union of a family of open sets, each of which is contained in  $E$  and hence  $\text{int}E$  is an open set contained in  $E$ . Now, let us assume that  $D$  is another open set contained in  $E$ . Since  $D$  is open therefore, for each  $y \in D$ , there exist an  $r_y > 0$  such that  $B(y, r_y) \subseteq D \subset E$  and hence each point of  $D$  is an interior point of  $E$ . Therefore,  $D$  is contained in  $\text{int}(E)$ . To prove that  $\bar{E}$  is the smallest closed set containing  $E$ . Let us assume that  $K$  is a closed set that contains  $E$ . Therefore, every convergent sequence of elements in  $E$  has limit in  $K$  because  $K$  is a closed set. So,  $K$  contains limit of sequences of elements in  $E$ . Hence  $K$  contains all adherent points of  $E$  which essentially means that  $\bar{E} \subseteq K$ .

□

**Corollary 5.4.** Prove that every finite set is closed.

PROOF. Note that we proved in previous proposition that singleton set is closed as well as finite union of closed sets is closed. Therefore, a finite set is closed reason being it is union of finite number of closed set. □

## 6. Dense Sets and Closures

**Definition 6.1 (Dense sets).** Let  $(M, d)$  be a metric space. Let  $A \subset M$ . Let  $B$  be another subset of  $M$ . We say  $A$  is (everywhere) dense relative to  $B$  if every point of  $B$  is either a point of  $A$  or a limit point of  $A$ . In other words, we say  $A$  is dense relative to  $B$  iff  $\forall r > 0$  and for all  $b \in B$ , we have  $B(b, r) \cap A \neq \emptyset$

The property of being dense is **transitive** in the following sense.

**Proposition 6.1.** Let  $(M, d)$  be a metric space. Let  $A, B, C$  be subsets of  $M$ . Let  $A$  be dense relative to  $B$ , let  $B$  be dense relative to  $C$ . Prove that  $A$  is dense relative to  $C$ .

PROOF. To prove that  $A$  is dense relative to  $C$ , we use Definition 6.1 and prove that for every  $r > 0$ ,  $B(c, r) \cap A \neq \emptyset$ . Let  $r > 0$  be given and let  $c_0 \in C$  be any arbitrary element of  $C$ . Since  $B$  is dense relative to  $C$ , therefore,  $B(c_0, r) \cap B \neq \emptyset$ . This further means that there exists at least one point  $b_0 \in B$  such that  $b_0 \in B(c_0, r)$ . Since  $B(c_0, r)$  is open, and  $b_0 \in B(c_0, r)$  so there exists  $r_1 > 0$  such that  $B(b_0, r_1) \subseteq B(c_0, r)$ . Since  $b_0 \in B$  and  $A$  is dense in  $B$  so there exists a point  $a_0 \in A$  such that  $a_0 \in B(b_0, r_1) \subseteq B(c_0, r)$ . Since  $c_0 \in C$  and  $r > 0$  were arbitrarily chosen therefore, we proved that every ball of positive radius centered at any point of  $C$  has non empty intersection with  $A$ . Hence, by Definition 6.1 we conclude that  $A$  is dense in  $C$ . □

Using previous proposition we prove a very useful result.

**Corollary 6.1 (Closure of any set is a closet set).** Let  $(M, d)$  be a metric space. Let  $A \subset M$ . Prove that  $\bar{A}$  is a closed set. That is  $\bar{\bar{A}} = \bar{A}$ .

PROOF. We will prove it using equality of sets via inclusions;  $\bar{A} \subset \bar{\bar{A}}$  and  $\bar{\bar{A}} \subset \bar{A}$ . The first inclusion obviously holds true. For second inclusion we appeal to the previous proposition. By definition of closure, every set is dense in its own closer.  $A$  is dense in  $\bar{A}$  and  $\bar{A}$  is dense in  $\bar{\bar{A}}$ . Therefore, by Proposition 6.1,  $A$  is dense in  $\bar{\bar{A}}$ . That is every point of  $\bar{\bar{A}}$  is either a point of  $A$  or the limit point of  $A$  and hence each point of  $\bar{\bar{A}}$  lies in the closure of  $A$ . That is  $\bar{\bar{A}} \subset \bar{A}$  and this implies that  $\bar{\bar{A}} = \bar{A}$ . That is, closure of any set is always closed. □

**Remark 6.1.** We begin with any set  $A$  and consider limit of all sequences of elements in  $A$ , including those constant sequences. This makes  $A$  to be a closed set. Limits of limits are limits, that is,  $\bar{\bar{A}} = \bar{A}$ . Hence, we don't add any more points to obtain closure of  $\bar{A}$ . This kind of operation has special name **Idempotent**

**Lemma 6.1.** Let  $E$  be normed linear space. The set  $U = \{x \in E : \|x\| < 1\}$  and  $V = \{x \in E : \|x\| \leq 1\}$  are convex set in the sense of definition.

**Theorem 6.1 (A beautiful Theorem from Georgi Shilov's Elementary Real & Complex Analysis).** Given a closed set  $F$  and an open set  $G$  such that  $F \subset G$  there exists another open set  $F \subset H$  and  $F \subset H \subset \bar{H} \subset G$ .

PROOF. Our proof hinges upon the fact that any open set can be written as union of open balls, refer to Definition 5.1 and Theorem 5.2

$$\text{Let } x \in G, \quad D(x) := \inf_{y \in M \setminus G} \{d(x, y)\}$$

$D(x)$  is a distance from  $x$  to the set  $A = M \setminus G$ : the complement of  $G$  in metric space  $M$  where  $G$  is open in  $M$ .

**We claim that**  $D(x) > 0$  is positive for all  $x \in G$ . You can not argue that infimum of set

$$\inf\{d(x, y) \mid y \in M \setminus G\} > 0$$

of positive numbers  $d(x, y)$  will always be positive. We have a counter example to make our case -

$$S = \left\{ \frac{1}{n} \mid n \in \mathbb{N}, n = 1, 2, 3, \dots \right\}$$

$$\inf S = 0$$

even though each term of the set is positive. To see why this is true, let us assume that  $\inf S = \lambda > 0$ . Now, by the **Archimedean principle**,  $\exists n_0 \in \mathbb{N}$  such that

$$\lambda > \frac{1}{n_0} > 0.$$

Therefore, no positive  $\lambda$  can be infimum of the set  $S = \left\{ \frac{1}{n} \mid n \in \mathbb{N}, n = 1, 2, 3, \dots \right\}$ .

Now, we prove that  $D(x) := \inf\{d(x, y) \mid y \in M \setminus G\}$  is indeed a positive real number.

- Since  $x \in G$  and  $G$  is an open set, so  $\exists \epsilon_x > 0$  s.t.

$$B(x, \epsilon_x) \subseteq G \quad \text{and now let } y \in M \setminus G. \text{ Then}$$

$$y \notin B(x, \epsilon_x) \implies d(y, x) \geq \epsilon_x \text{ for any } y \in M \setminus G$$

$$\text{Therefore, } \inf_{y \in M \setminus G} d(y, x) \geq \epsilon_x > 0.$$

This closes the proof that  $D(x) > 0$  for all  $x \in G$ . **Now, our goal is to construct the open set  $H$  that satisfies the stated properties, and we do this in what follows.**

- Since we did this construction using  $x \in G$  and  $F \subseteq G$ , the same construction follows. Consider basic open set for each  $x \in F$ :

$$U(x) = B(x, \frac{D(x)}{2})$$

Let  $H$  be the union of all such balls, that is,

$$H = \bigcup_{x \in F} B(x, \frac{D(x)}{2}) \quad \text{clearly, } H \supseteq F.$$

Now, our goal is to prove that  $\overline{H} \subseteq G$ . Suppose it is not the case. Then this means  $\exists y \in \overline{H}$  but  $y \notin G \implies y \in M \setminus G = G^c \implies y \in \overline{H} \cap (M \setminus G)$

But by definition of limit/adherent/cluster points, we know that there exists a sequence of elements  $z_n \in H$  s.t.  $z_n \rightarrow y$ , since  $y \in \overline{H}$ .

$$z_n \in H = \bigcup_{x \in F} B(x, \frac{D(x)}{2}).$$

For each  $n$ , corresponding to  $z_n$ ,  $\exists x_n \in F$  s.t.

$$z_n \in B(x_n, \frac{D(x_n)}{2}).$$

This means

$$d(x_n, z_n) \leq \frac{D(x_n)}{2}.$$

But  $D(x_n) = \inf\{d(x_n, y) \mid y \in M \setminus G\}$ , since  $x_n \in F \subseteq G$ .

For each  $n$ ,

$$D(x_n) = \inf_{y \in M \setminus G} \{d(x_n, y)\} \leq d(x_n, y) \leq d(x_n, z_n) + d(z_n, y) \leq \frac{D(x_n)}{2} + d(z_n, y)$$

$$D(x_n) \leq \frac{D(x_n)}{2} + d(z_n, y)$$

$$D(x_n) \leq 2d(z_n, y)$$

$$\lim_{n \rightarrow \infty} D(x_n) \leq \lim_{n \rightarrow \infty} 2d(z_n, y) = 0$$

This contradicts our previous results that  $D(x) > 0$  for all  $x \in G$ .

□

**Example 6.1.** Prove that when we give a set  $X$  the discrete metric  $d_{disc}$ , and  $E$  is any subset of  $X$ , then every element of  $E$  is an interior point of  $E$ , every point not contained in  $E$  is an exterior point of  $E$ , and there are no boundary points.

**Solution 6.1.** Let  $e$  be an arbitrary element of  $E$ . We have to prove that  $e$  is an interior point of  $E$ . We have to find a  $r$  such that  $B(e, r) \subseteq E$ . Now,

$$B(e, r) = \{x \mid d_{disc}(x, e) < r\}$$

Let be  $r > \frac{1}{2}$  and any  $x \in X$ . We will always have  $B(x, r) = \{x\}$  by definition of the discrete metric  $d_{disc}$ . So, each point of  $E$  is an interior point of  $E$ . Now assume that  $u \notin E$ . Consider the ball  $B(u, \frac{1}{2})$ . This is just  $\{u\} \subset X - E$ . Hence, each point  $u \in X - E$  is an exterior point of  $E$ . This makes is clear that there are no boundary point of  $E$ .

**Example 6.2.** Let  $(X, d)$  be a metric space,  $x_0$  be a point in  $X$ , and  $r > 0$ . Let  $B$  be the open ball  $B := B(x_0, r) = \{x \in X : d(x, x_0) < r\}$ , and let  $C$  be the closed ball  $C := \{x \in X : d(x, x_0) \leq r\}$ .

(1) Show that  $\bar{B} \subseteq C$ .

(2) Give an example of a metric space  $(X, d)$ , a point  $x_0$ , and a radius  $r > 0$  such that  $B$  is not equal to  $C$ .

**Solution 6.2.** (1) Let  $\bar{B}$  = collection of adherent points of  $B$ ,

let  $x \in \bar{B}$ . We have to prove that  $x \in C$ , that is,  $x$  is an adherent point of  $B$ . We prove it by contradiction. Let us assume that  $x \notin C$ . Then  $r' = d(x, x_0) > r$ . Consider  $r_0 = \frac{r'-r}{3}$ . Then the ball  $B(x, r_0) \cap B(x_0, r) = \emptyset$ , a contradiction to the fact that  $x$  is an adherent point of  $B(x_0, r)$ .

(2) Consider the discrete metric space  $(X, d_{disc})$ . Now, take  $B(x_0, 1) = \{x \in X \mid d_{disc}(x_0, x) < 1\} = \{x_0\}$ . Closure of  $B(x_0, 1)$  with respect to discrete metric is essentially collection of all limit points of all possible sequences in  $B(x_0, 1)$ . There is only one possible sequence here, namely, the constant sequence  $x_n = x_0$  for all  $n \in \mathbb{N}$ . Then  $\bar{B}(x_0, 1) = \{x_0\}$ . However,  $C = \{x \in X \mid d_{disc}(x, x_0) \leq 1\} = X$ .

## 7. Cauchy sequences and complete metric spaces

Complete spaces were first studied by M. Frechet. The high point of this section is the general scheme of constructing the completion of an arbitrary metric space. Cauchy and Dedekind played a fundamental role in filling gaps in the Newton's - Leibnitz calculus and cementing calculus on the firm of rigour. They were the first ones to give a rigorous definition of real numbers as 'limit of cauchy sequences of real number and Dedekind cut' respectively. Efforts of Cantor, Cauchy, Weierstrass, Dedekind were monumental in the foundation of modern analysis.

**Definition 7.1 (Subsequence).** Suppose that  $(x^{(n)})_{n=m}^{\infty}$  is a sequence of points in a metric space  $(X, d)$ . Suppose that  $n_1, n_2, n_3, \dots$  is an

$$(1) \quad m \leq n_1 < n_2 < n_3 < \dots$$

Then we call the sequence  $(x^{(n_j)})_{j=1}^{\infty}$  a subsequence of the original sequence  $(x^{(n)})_{n=m}^{\infty}$

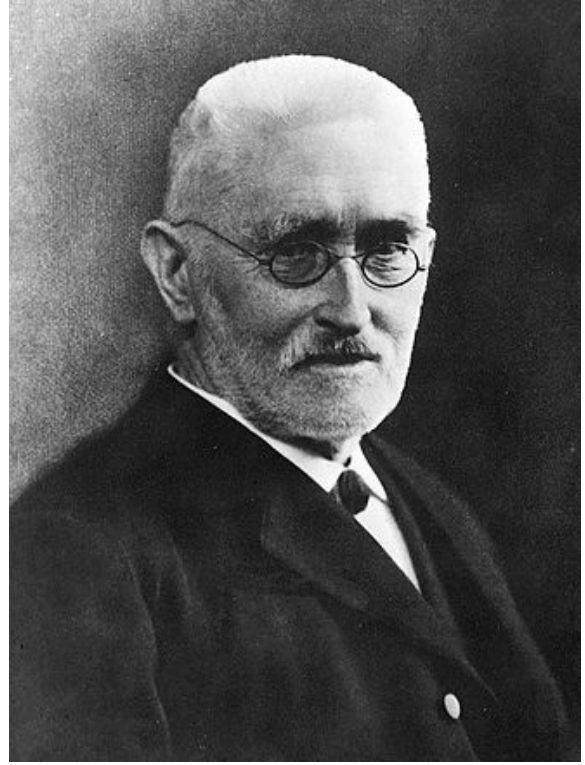
**Definition 7.2 (Limit points).** Suppose that  $(x_n)_{n=m}^{\infty}$  is a sequence of points in a metric space  $(X, d)$ , and let  $L \in X$ . We say that  $L$  is a limit point of  $(x_n)_{n=m}^{\infty}$  if for every  $N \geq m$  and  $\varepsilon > 0$  there exists an  $n \geq N$  such that  $d(x_n, L) \leq \varepsilon$ .

**Proposition 7.1 (Characterization of limit points).** Let  $(x^{(n)})_{n=m}^{\infty}$  be a sequence of points in a metric space  $(X, d)$ , and let  $L \in X$ . Then the following are equivalent:

- $L$  is a limit point of  $(x_n)_{n=m}^{\infty}$ .



FIGURE 6. Augustin-Louis Cauchy

FIGURE 7. Julius Wilhelm  
Richard Dedekind

- There exists a sub-sequence  $(x_{n_j})_{j=1}^{\infty}$  of the original sequence  $(x_n)_{n=m}^{\infty}$  which converges to  $L$ .

PROOF. • At first, we establish the implication in the forward direction. Let  $L$  be a limit point of  $(x^{(n)})_{n=m}^{\infty}$ . Then for each  $k \in \mathbb{N}$ , and for  $\epsilon = \frac{1}{k}$ , in view of Definition 7.2, there is a  $n_k > k$  such that

$$d(x_{(n_k)}, L) \leq \epsilon = \frac{1}{k}$$

Following this procedure way, we can construct a subsequence  $(x_{(n_k)})_{k=1}^{\infty}$  of the original sequence  $(x_{(n)})_{n=m}^{\infty}$  such that

$$d(x_{(n_k)}, L) \leq \frac{1}{k}$$

for all  $k \in \mathbb{N}$ . This implies that this subsequence converges to  $L$ . This clearly means that subsequence  $(x_{(n_j)})_{j=1}^{\infty}$  converges to  $L$

- Let us assume that the subsequence  $(x_{(n_j)})_{j=1}^{\infty}$  converges to  $L$ , and we need to prove that  $L$  is the limit point of this sequence in the sense of Definition 7.2. Let  $\epsilon > 0$  and  $i \geq 1$  be given. We know that for this  $\epsilon > 0$ , there exists a  $N_{\epsilon} \geq m$  such that for all  $n_j \geq N_{\epsilon}$  we get

$$d(x_{n_j}, L) \leq \epsilon$$

and given any  $i \in \mathbb{N}$ , previous convergence relation implies that  $d(x_{n_i}, L) \leq \epsilon$ . This closes the proof. □

**Lemma 7.1 (Limit is a limit point).** Let  $(x_n)_{n=m}^{\infty}$  be a sequence in  $(X, d)$  which converges to some limit  $L$ . Then every subsequence  $(x_{n_j})_{j=1}^{\infty}$  of that sequence also converges to  $L$ .

PROOF. Let  $(x_{n_j})_{j=1}^{\infty}$  be a subsequence of the original sequence. We have to prove that this subsequence also converges to  $L$  - the limit of the original sequence  $(x_n)_{n=m}^{\infty}$ .

Let us choose an arbitrary  $\epsilon_0 > 0$ . Our goal is to find a  $N_{\epsilon_0}$  such that

$$d(x_{n_j}, L) \leq \epsilon_0 \quad \forall j \geq N_{\epsilon_0}$$

Since  $(x_n)_{n=m}^{\infty}$  converges, so it is also a Cauchy sequence and hence, for this chosen  $\epsilon_0 : \exists N$  such that  $d(x_i, x_j) \leq \frac{\epsilon_0}{2} \quad \forall i, j \geq N$ .

This statement will remain true if  $x_j$  gets replaced with the term  $x_{n_j}$  of the subsequence, where  $n_j > j$ . That is,

$$d(x_i, x_{n_j}) \leq \frac{\epsilon_0}{2} \quad \forall i, n_j \geq N$$

Furthermore, since  $L$  is the limit of  $(x_n)_{n=m}^{\infty}$ , therefore, there exists a  $\exists M$  such that

$$d(x^i, L) \leq \frac{\epsilon_0}{2} \quad \forall i \geq M.$$

Now, set  $N_{\epsilon_0} = \max\{N, M\}$

Then,  $\forall i, j \geq N_{\epsilon_0} = \max\{N, M\}$ :

$$d(x_{n_j}, L) \leq d(x_{n_j}, x_i) + d(x_i, L) \leq \frac{\epsilon_0}{2} + \frac{\epsilon_0}{2} = \epsilon_0$$

That is,

$$d(x_{n_j}, L) \leq \epsilon_0 \quad \forall n_j \geq N_{\epsilon_0}$$

and this proves the claim.  $\square$

**Definition 7.3 (Cauchy or fundamental sequences).** Let  $(x_n)_{n=m}^{\infty}$  be a sequence of points in a metric space  $(X, d)$ . We say that this sequence is a Cauchy (or fundamental) sequence if for every  $\epsilon > 0$ , there exists an  $N \geq m$  such that  $d(x_j, x_k) < \epsilon$  for all  $j, k \geq N$ .

**Proposition 7.2 (Convergent sequences are Cauchy sequences).** Let  $(x_n)_{n=m}^{\infty}$  be a sequence in  $(X, d)$  which converges to some limit  $x_0$ . Then  $(x_n)_{n=m}^{\infty}$  is also a Cauchy sequence.

PROOF. Suppose that  $(x_n)_{n=m}^{\infty}$  is a convergent sequence of points in a metric space  $(X, d)$ , and that its limit is  $x_0$ .

We prove that it is a Cauchy sequence  $\iff \forall \epsilon > 0, \exists N$  such that for all  $j, k \geq N, d(x_j, x_k) \leq \epsilon$ . This  $N$  is dependent on the choice of  $\epsilon$ . Let  $\epsilon > 0$  be given. For this  $\epsilon$ , there exists an  $N$  such that for  $k \geq N$ , the term of sequence  $(x_n)_{n=1}^{\infty}$  satisfy  $d(x_k, x_0) \leq \frac{\epsilon}{2}$ . Now due to triangle inequality, we know that for all  $j, k$ , we have:

$$d(x_j, x_k) \leq d(x_j, x_0) + d(x_0, x_k).$$

In particular, previous relation holds true for  $j, k \geq N$  and which results in -

$$d(x_j, x_k) \leq d(x_j, x_0) + d(x_0, x_k) \leq \epsilon/2 + \epsilon/2 = \epsilon$$

which precisely means that  $(x^{(n)})_{n=m}^{\infty}$  is a Cauchy sequence.  $\square$

**Proposition 7.3 (For Cauchy sequences, limit and limit points are same).** Let  $(x_n)_{n=m}^{\infty}$  be a Cauchy sequence in  $(X, d)$ . Suppose that there is some sub-sequence  $(x_{n_j})_{j=1}^{\infty}$  of this sequence which converges to a limit  $x_0$  in  $X$ . Then the original sequence  $(x_n)_{n=m}^{\infty}$  also converges to  $x_0$ .

PROOF. Let  $\epsilon > 0$ . We have to find a natural number  $N \geq m$  such that for all  $k \geq N$  we have  $d(x_k, x_0) \leq \epsilon$ . Now, since  $\frac{\epsilon}{2} > 0$ , therefore, there exists a natural number  $M \geq m$  such that  $d(x_{n_j}, x_0) \leq \frac{\epsilon}{2}$  whenever  $n_j \geq M$ .

However, note that the given sequence  $(x_n)_{n=m}^{\infty}$  is a Cauchy sequence. Therefore, there exists an  $M' \geq m$  such that  $d(x_j, x_k) \leq \frac{\epsilon}{2}$  whenever  $j, k \geq M'$ .

Now, let  $N = \max\{M, M'\}$ . If  $n, n_j \geq N$ , we have

$$d(x_n, x_0) \leq d(x_n, x_{n_j}) + d(x_{n_j}, x_0) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2}$$

Since  $M, M'$  both are greater than or equal to  $m$ , therefore,  $N \geq m$ . This closes the proof.  $\square$

**Definition 7.4 (Complete metric spaces).** A metric space  $(X, d)$  is said to be complete iff every Cauchy sequence in  $(X, d)$  is convergent in  $(X, d)$ .

**Proposition 7.4 (Characterization of complete subset of a complete metric space).** (1) Let  $(X, d)$  be a metric space, and let  $(Y, d|_{Y \times Y})$  be a subspace of  $(X, d)$ . If  $(Y, d|_{Y \times Y})$  is complete, then  $Y$  must be closed in  $X$ .

(2) Conversely, suppose that  $(X, d)$  is a complete metric space, and  $Y$  is a closed subset of  $X$ . Then the subspace  $(Y, d|_{Y \times Y})$  is also complete.

PROOF. (1) Let  $(Y, d|_{Y \times Y})$  is complete. Let  $(y_n)_{n=m}^\infty$  be a convergent sequence converging to limit  $L$ . We need to show that  $L \in Y$ . Since  $Y$  is complete metric space, therefore, every Cauchy sequence converges inside  $Y$ . Since convergent sequences are Cauchy and  $(y_n)_{n=m}^\infty$  is convergent and hence, it converges inside  $Y$ . That is  $L \in Y$ . This precisely means that limit of every convergent sequence of point in  $Y$  lies inside  $Y$ , that is,  $Y$  is closed in  $X$ .

(2) Let  $(X, d)$  is a complete metric space, and  $Y$  is a closed subset of  $X$ . Let  $(y_n)_{n=m}^\infty$  be a Cauchy sequence in  $(Y, d|_{Y \times Y})$ . Since  $(Y, d|_{Y \times Y})$  is a subspace of  $(X, d)$  therefore, due to the completeness of  $X$ , the sequence  $(y_n)_{n=m}^\infty$  converges in  $X$  and let us call the limit to be  $L$ . However, since  $Y$  is closed subset of  $X$ , so  $L \in Y$ . This proves that every Cauchy sequence in  $Y$  is convergent in  $Y$  itself and hence  $Y$  is a complete metric subspace.  $\square$

Next, we give a **general scheme of completion of a metric spaces**<sup>2</sup>

### 7.1. Isometric embedding of a metric space into a complete metric space.

**Definition 7.5 (Equivalent Cauchy sequences).** Let  $(a_n)_{n=m}^\infty$  and  $(b_n)_{n=m}^\infty$  be two Cauchy sequences. We say they are equivalent if for each  $\epsilon > 0$ , there exists some natural number  $N \geq m$  such that for all  $k \geq N$  we have  $d(a_k, b_k) \leq \epsilon$ .

**Proposition 7.5 (Class of equivalent Cauchy sequences).** Let  $(X, d)$  be a metric space. Define  $\sim$  to be a relation on collection  $M$  of all Cauchy sequences in  $X$  as  $(a_n)_{n=m}^\infty \sim (b_n)_{n=m}^\infty$  iff they are equivalent Cauchy sequences in the sense of Definition 7.5. It is easy to see that this relation is indeed an equivalence relation on the set  $M$  of all Cauchy sequences in  $X$ . **Equivalence classes of this equivalence relation will partition the set  $M$ .**

PROOF. Proof of this result follows from the application of definition of equivalence relation on the relation  $\sim$  defined on  $M$ .  $\square$

**Theorem 7.1 (Hausdorff).** Let  $M$  be a metric space, we assume it is incomplete. Then there exists a complete metric space  $\overline{M}$ , called as completion of  $M$ , with the following properties:

- (1)  $M$  is **isometric** to a subset  $M_1 \subset \overline{M}$
- (2)  $M_1$  is **dense** in  $\overline{M}$

Moreover, every pair of complete space  $M, N$  satisfying properties (1) and (2) are isometric (**bijjective distance preserving map** between  $M$  and  $\overline{M}$ ).

PROOF. To begin with, we want to construct a metric on the collection  $\overline{M}$  of equivalence classes of equivalent Cauchy sequence.

Let  $X$  and  $Y$  be two equivalence classes of equivalent Cauchy sequences. Since equivalence classes form partitions of collection of all Cauchy sequences in metric space  $(X, d)$ , therefore, either  $X, Y$  are mutually disjoint or completely overlap. Now, on the collection  $\overline{M}$  of all equivalence classes of Cauchy (also referred as fundamental) sequences we define a metric:

$$(2) \quad \rho'(X, Y) := \lim_{n \rightarrow \infty} \rho(x_n, y_n) \quad \text{where} \quad x_n \in X, y_n \in Y$$

<sup>2</sup>This is one of the high point in the "rigorous analysis course" run at AMMOC - an international math circle.

where  $(x_n)_{n=m}^{\infty}$  and  $(y_n)_{n=m}^{\infty}$  are Cauchy (fundamental) sequences in equivalence classes  $X$  and  $Y$  respectively, see Theorem 7.5.

**Remark 7.1.** Whenever we construct a new function we have to verify its well defined (it yields an unique output for a given input) and that it respects the axiom of substitution.

We begin to verify that that metric defined in (2) is indeed a metric on  $\bar{M}$ . Set

$$(3) \quad d_n := \rho(x_n, y_n) \geq 0, \quad d_n \in [0, \infty)$$

In order to prove that this limit exists, we need to show that  $\{d_n\}$  is a fundamental sequence in  $\mathbb{R}$ , and then by completeness of  $\mathbb{R}$ , we would conclude that limit  $d_n$  exists and is unique. Let  $\epsilon > 0$  be given. We need to find a  $N \in \mathbb{N}$  such that for all  $n, m \geq N$ , we have  $|d_n - d_m| < \epsilon$  which is precisely the definition of a Cauchy sequence. To achieve this we need a lemma first

**Lemma 7.2 (Quadrilateral inequality).** Let  $x, y, z, u$  be elements of a metric space  $(X, \rho)$ . Then  $|\rho(x, y) - \rho(z, u)| \leq \rho(x, z) + \rho(y, u)$

PROOF.

$$(4) \quad \rho(x, y) \leq \rho(x, z) + \rho(z, u) + \rho(u, y)$$

$$(5) \quad \rho(z, u) \leq \rho(z, x) + \rho(x, y) + \rho(y, u)$$

These inequations can be re-arranged as  $\rho(x, y) - \rho(z, u) \leq \rho(x, z) + \rho(u, y)$  and  $\rho(z, u) - \rho(x, y) \leq \rho(z, x) + \rho(y, u)$  respectively. These two inequation give us  $|\rho(x, y) - \rho(z, u)| \leq \rho(x, z) + \rho(y, u)$   $\square$

Now, we proceed with the proof of Theorem 7.1.

**Step 1 :** To begin with we prove that the sequence  $(d_n)_{n=1}^{\infty}$ , defined in (3) is indeed a Cauchy sequence in the sense of Definition 7.3. Let  $\epsilon > 0$  be given and note that due to quadrilateral inequality 7.2, we have

$$(6) \quad |d_n - d_m| = |\rho(x_n, y_n) - \rho(x_m, y_m)| \leq \rho(x_n, x_m) + \rho(y_n, y_m)$$

But  $(x_n)_{n=m}^{\infty}$  and  $(y_n)_{n=m}^{\infty}$  are fundamental (Cauchy sequences), therefore, for this arbitrarily chosen  $\epsilon > 0$

$$\exists N_1 \text{ such that } \rho(x_n, x_m) \leq \frac{\epsilon}{2} \quad \forall n, m \geq N_1$$

$$\exists N_2 \text{ such that } \rho(y_n, y_m) \leq \frac{\epsilon}{2} \quad \forall n, m \geq N_2$$

Set  $N_0 = \max\{N_1, N_2\}$  and for this  $N_0$ , the inequality (6) becomes

$$|d_n - d_m| \leq \epsilon \quad \forall n, m \geq N_0$$

That is, the sequence  $(d_n)_{n=1}^{\infty}$  defined in (3) becomes a Cauchy sequence in  $\mathbb{R}$  and due to completeness of  $\mathbb{R}$ , this sequence  $(d_n)_{n=1}^{\infty}$  will have an unique limit. Hence,

$$\rho(X, Y) = \lim_{n \rightarrow \infty} d(x_n, y_n)$$

is a unique number. Thus  $\rho(X, Y) = \lim_{n \rightarrow \infty} d(x_n, y_n)$  is indeed a function on  $\bar{M}$ . **To prove that this definition of distance on  $\bar{M}$  respects the axiom of substitution**, let  $(x_n)_{n=1}^{\infty}$  and  $(x'_n)_{n=1}^{\infty}$  be equivalent sequences in the equivalence class  $X$  and  $(y_n)_{n=1}^{\infty}$  and  $(y'_n)_{n=1}^{\infty}$  be equivalent sequences in  $Y$ . We need to show that  $d(x_n, y_n) = d(x'_n, y'_n)$  which precisely is equivalent to show that

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq \epsilon \quad \forall \epsilon > 0$$

Let  $\epsilon > 0$  be given. By the quadrilateral inequality in Lemma 7.2 we get

$$0 \leq |d(x_n, y_n) - d(x'_n, y'_n)| \leq d(x_n, x'_n) + d(y_n, y'_n)$$

Now, for  $\frac{\epsilon}{2}$ , the pairs of equivalent Cauchy sequences eventually (for sufficiently large values of  $n$ ) satisfy - Since  $d(x_n, x'_n) \leq \frac{\epsilon}{2}$  and  $d(y_n, y'_n) \leq \frac{\epsilon}{2}$ , and hence, eventually we have

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

This proves that  $\rho(X, Y)$ , as stated in (2) and (3) is indeed a non-negative function as  $d_n \geq 0$  and **non-negative reals are closed**.

**Step 2:** To prove that  $\rho(X, Y) = \lim_{n \rightarrow \infty} d(x_n, y_n)$  follows the three axioms of a metric. This is a straight forward verification. So, we skip this.

**Step 3:** To construct a subset  $M_1$  (isometric to  $M$ ) and contained in  $\overline{M}$ .  $M_1$  is a collection of all equivalence classes of fundamental sequences of elements of  $M$ . So the elements of  $M_1$  must be equivalence classes (themselves) of fundamental sequences.

For  $X \in \overline{M}$ , define class  $X$  to be of those Cauchy sequences which are equivalent to the sequence  $(x, x, x, \dots)$ , where  $x \in M$ . The limit of this sequence is  $x$  itself.

Therefore, in simple terms,  $X$  is the equivalence class of those sequences which converge to  $x$ . Note that  $X \neq \emptyset$ . **We set  $M_1$  to be the collection of all such equivalence classes.** Now, by its very definition,  $M_1$  is a subset of  $\overline{M}$ . We claim  $M_1$  is isometric to  $M$ , and this is easy to see because

$$\rho(X, Y) = \lim_{n \rightarrow \infty} d(x_n = x, y_n = y) = d(x, y).$$

Note that  $\rho$  is an induced metric on  $M_1$  and  $(M_1, \rho)$  becomes the subspace of  $(M, \rho)$ .

**Step 4:** Now we prove that  $M_1$  is dense in  $\overline{M}$ . This means we need to prove that each point of  $\overline{M}$  is the limit point of  $M_1$ . This by definition 7.2 requires us to prove that given a  $X \in \overline{M}$ , we can find at least one sequence  $X_1, X_2, X_3, \dots$  of elements in  $M_1$  that converges to  $X$ .

**New mathematical objects are constructed using old objects.** Now,  $X$  is an equivalence class in  $\overline{M}$ , so  $X$  contains some Cauchy sequence  $x_1, x_2, x_3, \dots$ . Let  $X_n$  be equivalence classes of those sequences in  $M$  whose limit is  $x_n$ . That is,  $X$  is the equivalence class of those sequences which are equivalent to constant sequence  $x_n, x_n, x_n, \dots$ .

We claim  $X_1, X_2, \dots, X_n$  will converge to  $X \in \overline{M}$ .

Now note that  $x_1, x_2, x_3, \dots$  is a Cauchy sequence, therefore,

$$\forall \epsilon > 0, \exists N_\epsilon \text{ such that } d(x_n, x_m) < \epsilon \quad \forall n, m > N_\epsilon$$

Now, for this  $N_\epsilon$  we get

$$\rho(X, X_n) = \lim_{k \rightarrow \infty} d(x_k, x_n) \leq \epsilon$$

because  $x_k$  is the Cauchy sequence  $x_1, x_2, \dots$  in the equivalence class  $X$ .

So, we proved that  $X_1, X_2, X_3, \dots, X_n$  converges to  $X$  in  $M$  and therefore, each  $X_n$  is limit of some Cauchy sequence in  $M$ .

Hence,  $M_1$  is dense in  $\overline{M}$ .

**Step 5:** We set to prove that  $\overline{M}$  is complete and for this, we need to show that every Cauchy sequence in  $\overline{M}$  converges in  $\overline{M}$ . Recall that Cauchy sequence in metric space  $\overline{M}$  are sequences of equivalence classes. Let  $X_1, X_2 \dots X_n$  be a Cauchy sequence in  $\overline{M}$ . Since  $M_1$  is dense in  $\overline{M}$ . So, these  $X_i$  are the limit point of  $M_1$ . This means given for each  $i \in \mathbb{N}$ , and for  $\epsilon = \frac{1}{i} > 0$ ,  $\exists Y_i \in M_1$  (and  $X_i \neq Y_i$ ) such that

$$d(X_i, Y_i) \leq \frac{1}{i}$$

$Y_i \in M_1$  and is equivalence class of Cauchy sequences in  $M_1$  which converge to  $y_i \in M$ . That is,  $y_i$  is the common limit of all Cauchy sequences in  $Y_i$ .

Now, corresponding to  $Y_i$ , we choose  $y_i \in M$ .

Corresponding to  $Y_1$ , we choose  $y_1 \in M$

Corresponding to  $Y_2$ , we choose  $y_2 \in M$

Corresponding to  $Y_n$ , we choose  $y_n \in M$

So,  $(y_n)_{n=1}^\infty$  is a sequence of elements in  $M$ . We claim  $(y_n)_{n=1}^\infty$  is a Cauchy sequence. To do so, let  $\epsilon > 0$  be given. We need to show that  $(y_n)_{n=1}^\infty$  is eventually  $\epsilon$ -close.

$$\begin{aligned} d(y_n, y_m) &= d(Y_n, Y_m) \leq d(Y_n, X_n) + d(X_n, X_m) + d(X_m, Y_m) \\ &\leq \frac{1}{n} + d(X_n, X_m) + \frac{1}{m} \end{aligned}$$

Since  $Y_1, Y_2, Y_3, \dots$  is a Cauchy sequence of elements in  $M_1$  and as  $n, m \rightarrow \infty$ , RHS:  $\frac{1}{m} + d(X_n, X_m) + \frac{1}{n} \rightarrow 0$ . Corresponding Cauchy sequence  $(y_n)_{n=1}^\infty$ , there exists an equivalence class in  $\overline{M}$  corresponding to this Cauchy sequence. Label this equivalence class as  $Y$ . We claim that

$$\lim_{n \rightarrow \infty} X_n = Y$$

and we prove it in what follows.

$$d(Y, X_n) \leq d(Y, Y_n) + d(Y_n, X_n) \leq d(Y, Y_n) + \frac{1}{n}$$

. To define  $d(Y, Y_n)$ , we are choosing sequence  $y_1, y_2, \dots, y_m \in Y$  and the constant sequence  $y_n, y_n, \dots$  in  $Y_n$ . With this in place,  $d(Y, Y_n) = d(y_m, y_n)$  and last inequation reduces to -

$$d(Y, X_n) \leq d(y_m, y_n) + \frac{1}{n}$$

As the sequence  $y_1, y_2, \dots, y_m$  is Cauchy,  $d(y_m, y_n) + \frac{1}{n} \rightarrow 0$  for sufficiently large of  $n$ . So, Cauchy sequence  $X_1, X_2 \dots X_n$  converges to  $Y \in \overline{M}$

Hence,  $\overline{M}$  is a complete metric space.

**Final Step:** At the end we show that metric space  $\overline{\overline{M}}$  with the Properties 1 and 2, in the Theorem 7.1, is isometric to the metric space  $\overline{M}$ . Suppose  $M_1$  and  $M_2$  are subsets of the spaces  $\overline{M}$  and  $\overline{\overline{M}}$ , respectively, and isometric to the space  $M$ . By the transitive property of "isometry",  $M_1$  and  $M_2$  are isometric to each other. Let  $X \in \overline{M}$ . Since  $M_1$  is dense in  $\overline{M}$ , there exists a sequence  $X_1, X_2, \dots$  of elements in  $M_1$  that converges to  $X$ . The isometry between  $M_1$  and  $M_2$  gives us the corresponding sequence  $Y_1, Y_2, \dots$  in  $M_2$  and this isometry  $\varphi : M_1 \rightarrow M_2$  guarantees that sequence  $(\varphi(X_i) = Y_i)_{i \in \mathbb{N}}$  is indeed a Cauchy sequence. Now,  $\overline{\overline{M}}$  is complete and  $M_2 \subset \overline{\overline{M}}$ , therefore,  $(\varphi(X_i) = Y_i)_{i \in \mathbb{N}}$  will converge to some point  $Y \in \overline{\overline{M}}$ .  $\square$

**Example 7.1.** Let  $(x_n)_{n=m}^\infty$  be a sequence of points in a metric space  $(X, d)$ , and let  $L \in X$ . Show that if  $L$  is a limit point of the sequence  $(x_n)_{n=m}^\infty$ , then  $L$  is an adherent point of the set  $\{x^{(n)} : n \geq m\}$ . Is the converse true?

**Solution 7.1.** Suppose that  $L$  is a limit point of  $(x_n)_{n=m}^\infty$ . This by definition 7.2 means that

$$(7) \quad \forall \epsilon > 0, \forall N \geq m, \exists n \geq N : d(x_n, L) \leq \epsilon$$

Now, let us consider a ball  $B(L, \epsilon_0)$ . For this  $\epsilon_0 > 0$  and any  $N \in \mathbb{N}$  there exists an  $n \geq N$  such that  $d(x_n, L) \leq \epsilon_0$  and this implies that  $x_n \in B(L, \epsilon_0)$  and also lies in the set  $\{x_n : n \geq m\}$ . This essentially means that  $B(L, \epsilon) \cap \{x_n : n \geq m\}$  is always non-empty. Hence,  $L$  is an adherent/limit point of  $\{x_n : n \geq m\}$ .

**Example 7.2.** Show that every Cauchy sequence can have at most one limit point.

**Remark 7.2.** This is essentially the restatement of Proposition 7.3.

**Solution 7.2.** Suppose  $L, L'$  are the two limit points of the sequence  $(x_n)_{n=m}^\infty$ .

**Now, for any given  $\epsilon > 0$ , we can always divide it into smaller positive parts and this is very useful technique in mathematical analysis.** Now for  $\frac{\epsilon}{3}$ , the sequence  $(x_n)_{n=m}^\infty$  is Cauchy, and therefore, there exists  $N \geq m$  such that  $d(x_k, x_j) \leq \epsilon/3$  for all  $j, k \geq N$ . Moreover,  $L$  is a limit point, so for this  $N \geq m$ , and  $\frac{\epsilon}{3}$ , there exists  $i \geq N$  such that  $d(x_i, L) \leq \epsilon/3$ .

Also,  $L'$  is a limit point, so again for this  $N \geq m$  and  $\frac{\epsilon}{3}$ , there exists  $l \geq N$  such that  $d(x_l, L') \leq \epsilon/3$ .

Since  $i, j, k, l \geq N$ , therefore we can apply triangle inequality to obtain the following -

$$\begin{aligned} d(L, L') &\leq d(L, x_k) + d(x_k, x_l) + d(x_l, L') \\ &\leq \epsilon/3 + \epsilon/3 + \epsilon/3 \\ &\leq \epsilon \end{aligned}$$

$d(L, L') \leq \epsilon$  for all  $\epsilon > 0$ , which precisely is how we establish the equality of between two points of a metric space. Thus  $L = L'$ .

**Theorem 7.2 (Banach Fixed point theorem).** If  $f : X \rightarrow X$  is a contraction on a complete metric space  $X$ , then  $f$  has a unique **fixed point**, namely a unique point  $x$  with  $f(x) = x$ .

PROOF. Proof of the uniqueness of a mathematical object that satisfies some properties, naturally goes well with the method of proof by contradiction. Suppose that  $x \neq y$  are both fixed points. Given that  $f$  is a contraction, therefore, we have

$$d(x, y) = d(f(x), f(y)) \leq \alpha \cdot d(x, y)$$

for some  $0 \leq \alpha < 1$ . This leads to the contradiction  $\alpha \geq 1$ .

Now, we show the existence. Let  $x \in X$  be arbitrary and define a sequence by setting  $x_1 = x$  and  $x_{n+1} = f(x_n)$  for each  $n \geq 1$ . If this sequence is actually Cauchy, then by the completeness of  $X$ , it converges to some limit say,  $y$ . In such situation we claim,  $y$  is a fixed point of  $f$  because

$$y = \lim_{n \rightarrow \infty} x_n \implies f(y) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = y.$$

It remains to show that our sequence is Cauchy. Since  $f$  is a contraction and  $x_{n+1} = f(x_n)$  for each  $n$ , we have

$$\begin{aligned} d(x_n, x_{n+k}) &\leq d(x_n, x_{n+1}) + \cdots + d(x_{n+k-1}, x_{n+k}) \\ &= \sum_{i=n}^{n+k-1} d(x_i, x_{i+1}) \leq \sum_{i=n}^{n+k-1} \alpha^{i-n} \cdot d(x_1, x_2) \\ &\leq \sum_{i=n}^{\infty} \alpha^{i-n} \cdot d(x_1, x_2) = \frac{\alpha^{n-1}}{1-\alpha} \cdot d(x_1, x_2). \end{aligned}$$

The right-hand side goes to zero as  $n \rightarrow \infty$ , so the same is true for the left-hand side. This shows that our sequence is Cauchy.  $\square$

**Definition 7.6 (Banach and Hilbert Space).** A normed linear space  $(E, \|\cdot\|)$  is a **Banach space** if every Cauchy sequence in  $E$  is convergent in  $E$  with respect to metric induced by the norm  $\|\cdot\|$ . An inner product space  $(E, \langle \cdot, \cdot \rangle)$  is known as **Hilbert space** if every Cauchy sequence in  $E$  is convergent in  $E$  with respect to norm  $\|x\| = \sqrt{\langle x, x \rangle}$ .

## 8. Compactness

The concept of Compact metric spaces is due to M. Frechet (1906). However, the concept of compactness for more general spaces was introduced by independently by L. Vietoris in 1921, P. Alexandorov (1921), and P. Uryshon (1923)

**Definition 8.1 (Sequential compactness).** A metric space  $(X, d)$  is said to be compact iff every sequence in  $(X, d)$  has at least one convergent sub-sequence. A subset  $Y$  of a metric space  $X$  is said to be compact if the subspace  $(Y, d|_{Y \times Y})$  is compact.

**Proposition 8.1 (Sequential characterization of compactness).** Show that a metric space  $(X, d)$  is compact if and only if every sequence in  $X$  has at least one limit point.

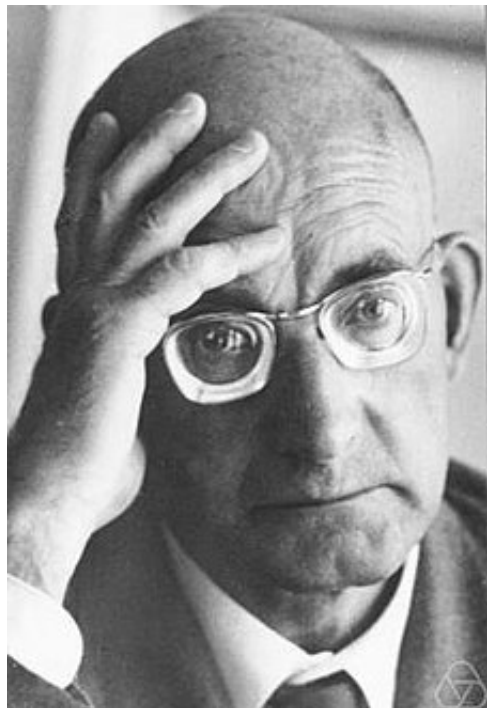


FIGURE 8. Pavel Sergeyevich Alexandrov



FIGURE 9. Pavel Samuilovich Urysohn

**PROOF. Forward direction:** Assume  $(X, d)$  is compact. Consider a sequence  $x_n \in X$ . By definition this sequence has a convergent sub-sequence and by proposition 2.5, a limit point.

**Backwards direction:** Now assume that every sequence in  $X$  has a limit point. Then by definition, of limit point, there exists a subsequence of the given sequence which converges to this limit point. This proves the claim.  $\square$

**Proposition 8.2 (Compactness implies completeness and boundedness).** Let  $(X, d)$  be a compact metric space. Then  $(X, d)$  is both complete and bounded.

**PROOF.** • **To prove compactness  $\implies$  completeness** Let  $(X, d)$  be a compact metric space. We need to prove that  $X$  is complete. So, let  $(x_n)_{n=m}^{\infty}$  be a Cauchy sequence in  $X$  and we will prove that this sequence converges in  $X$ . Since  $X$  is compact, there exists at least one subsequence of  $(x_n)_{n=m}^{\infty}$  which has a limit, say  $L$ , in  $X$ . Now, we have already proven in Proposition 7.3, that for Cauchy sequences, if its one of subsequence converges to a point  $L$ , then the original sequence also converges to the same point  $L$ . Hence, by the application of this proposition, the chosen Cauchy sequence  $(x_n)_{n=m}^{\infty}$  converges to  $L$ , and therefore,  $X$  is complete metric space.  $\square$

**Corollary 8.1 (Compact sets are closed and bounded).** Let  $(X, d)$  be a metric space, and let  $Y$  be a compact subset of  $X$ . Then  $Y$  is closed and bounded.

**PROOF.** Let  $K$  be a compact set of  $(X, d)$ . Let  $(a_n)_{n=1}^{\infty}$  be a convergent sequence of points in  $K$ . Let  $L$  be the limit of this sequence and  $L \in X$ . We need to show that  $L \in K$ . Due to the compactness of  $K$ , we know that there exists a subsequence  $(a_{n_j})$  of original sequence such that limit of this subsequence lies in  $K$ . However, the original sequence is converging to  $L$ . If a sequence is convergent then all its subsequence converge to the same limit. Hence,  $L$  must be the limit of  $(a_{n_j})$  and due to the compactness of  $K$ , the point  $L \in K$ . Thus, compact sets are closed.  $\square$

**Theorem 8.1 (Heine-Borel theorem).** Let  $(\mathbb{R}^n, d)$  be an Euclidean space with either the Euclidean metric, the taxicab metric, or the sup norm metric. Let  $E$  be a subset of  $\mathbb{R}^n$ . Then  $E$  is compact if and only if it is closed and bounded.

PROOF. Due to Proposition 8.2 and Corollary 8.1 we know that compact sets of  $\mathbb{R}^n$  are closed and bounded. Now, we prove that if  $K \subset \mathbb{R}^n$  is closed and bounded then  $K$  is compact in  $\mathbb{R}^n$ . Let  $(a_n)_{n=1}^\infty$  be a sequence in  $K$ . Since  $K$  is bounded, we know by Theorem 5.3 that  $K$  contains a convergent subsequence  $(a_{n_j})$  in  $\mathbb{R}^n$ . Since  $K$  is closed, therefore, the limit of the sequence must be inside  $K$ . Hence, every sequence in  $K$  has at least one limit point inside  $K$ . That is,  $K$  is compact.  $\square$

**Theorem 8.2 (Topological characterization of compactness).** Let  $(X, d)$  be a metric space, and let  $Y$  be a compact subset of  $X$ . Let  $(V_\alpha)_{\alpha \in I}$  be a collection of open sets in  $X$ , and suppose that

$$Y \subseteq \bigcup_{\alpha \in I} V_\alpha.$$

(i.e., the collection  $(V_\alpha)_{\alpha \in I}$  covers  $Y$ ). Then, there exists a finite subset  $F$  of  $I$  such that

$$Y \subseteq \bigcup_{\alpha \in F} V_\alpha.$$

**Theorem 8.3 (Converse).** Let  $(X, d)$  have the property that every open cover of  $X$  has a finite sub-cover. Show that  $X$  is compact.

PROOF. Assume for the sake of contradiction,  $X$  is not compact. So there is a sequence  $(x^{(n)})_{n=1}^\infty$  with no limit points. Thus for each  $x$  the ball  $B(x, \varepsilon)$  contains finitely many points of the sequence  $(x^{(n)})_{n=1}^\infty$ . Also  $x \in B(x, \varepsilon)$  and the ball  $B(x, \varepsilon)$  is open. Now for a fixed  $\varepsilon$ , we have

$$X \subseteq \bigcup_{x \in X} B(x, \varepsilon)$$

Note that the ball  $B(x, \varepsilon)$  only contains finitely many points of the sequence. So the finite sub-cover will only contain finitely many points of the sequence  $(x^{(n)})_{n=1}^\infty$ . Thus, it cannot contain  $X$  and therefore there is a no finite sub-cover.  $\square$

**Corollary 8.2.** Let  $(X, d)$  be a metric space, and let  $K_1, K_2, K_3, \dots$  be a sequence of non-empty compact subsets of  $X$  such that  $K_1 \supset K_2 \supset K_3 \supset \dots$ . Then the intersection  $\bigcap_{n=1}^\infty K_n$  is non-empty.

PROOF. Assume for the sake of contradiction  $\bigcap_{n=1}^\infty K_n$  is empty. Let  $V_n = K_1 - K_n$ . Now,

$$K_1 = K_1 - \bigcap_{n=1}^\infty K_n = \bigcup_{n=1}^\infty (K_1 - K_n) = \bigcup_{n=1}^\infty V_n$$

The collection  $(V_n)_{n=1}^\infty$  covers  $F_1$  and as  $F_1$  is a compact. There is a finite subcover  $(V_j)_{j=1}^{j_k}$

$$\begin{aligned} K_1 &= \bigcup_{j=1}^{j_k} V_j = K_1 - \bigcap_{j=1}^{j_k} K_j \\ &\quad \bigcap_{j=1}^{j_k} K_j = \emptyset. \end{aligned}$$

Note that  $K_1, K_2, K_3, \dots$  are non empty and  $K_1 \supset K_2 \supset K_3 \supset \dots$  and we can arrange their indices  $j_1, j_2, j_3 \dots j_k$  in increasing order such that their finite intersection must be non-empty, which is clear contradiction to the last statement.  $\square$

**Theorem 8.4.** Let  $(X, d)$  be a metric space.

a If  $Y$  is a compact subset of  $X$ , and  $Z \subseteq Y$ , then  $Z$  is compact if and only if  $Z$  is closed.

- b If  $Y_1, \dots, Y_n$  are a finite collection of compact subsets of  $X$ , then their union  $Y_1 \cup \dots \cup Y_n$  is also compact.
- c Every finite subset of  $X$  (including the empty set) is compact.

PROOF. (1) Forward direction proof is true by Corollary 8.1.

**To prove  $Z$  is closed  $\implies Z$  is compact.** Let  $z_n \in Z$  be a sequence of elements in  $Z$ . As  $Z \subseteq Y$ , we have  $z_n \in Y$ . As  $Y$  is compact,  $z_n$  has a convergent sub-sequence  $(z_{n_j})_{j=1}^{\infty}$  which converges to point, say  $l \in Y$ . As  $Z$  is closed, convergent sequence  $(z_{n_j})$  converges to point in  $Z$ , and let that point  $l_1 \in Z$ . However, the limit is always unique, hence,  $l = l_1 \in Z$ . So  $l \in Z$  and therefore  $z_n$  has a convergent sub-sequence.

- (2) Let  $Y = \bigcup_{i=1}^n Y_i$ . Let  $y_n \in Y$ . There exists at least one  $1 \leq k \leq n$  such that infinitely many terms of  $(y_n)_{n=1}^{\infty}$  are in  $Y_k$ . Since  $Y_k$  is compact, therefore,  $Y_k$  contains a convergent subsequence of  $(y_n)_{n=1}^{\infty}$  with its limit in  $Y_k \subset Y$ . Hence,  $Y$  is compact.
- (3) Now, we choose an infinite number of elements in a sequence from a finite set. So, there must be a point which repeats infinitely many times in the sequence. Let that point be  $a$ . Clearly, the subsequence  $a, a, a \dots a$  is a convergent subsequence of original sequence with its limit  $a$  in the given finite set.

□

**Theorem 8.5 (Finite intersection property).** Let  $(X, d)$  be a compact metric space. Suppose that  $(K_\alpha)_{\alpha \in \mathcal{I}}$  is a collection of closed sets in  $X$  with the property that any finite subcollection of these sets necessarily has non-empty intersection, thus  $\bigcap_{\alpha \in F} K_\alpha \neq \emptyset$  for all finite  $F \subseteq \mathcal{I}$ . (This property is known as the finite intersection property.) Show that the entire collection has non-empty intersection, thus  $\bigcap_{\alpha \in \mathcal{I}} K_\alpha \neq \emptyset$ . Show by counterexample that this statement fails if  $X$  is not compact.

PROOF. Assume for the sake of the contradiction that  $\bigcap_{\alpha \in \mathcal{I}} K_\alpha = \emptyset$ . Note that  $K_\alpha^C = X - K_\alpha$  is open. Also,

$$X = X - \bigcap_{\alpha \in \mathcal{I}} K_\alpha = \bigcup_{\alpha \in \mathcal{I}} K_\alpha^C.$$

Thus, we found an open cover of  $X$ . Now notice that as  $X$  is compact, therefore, there is a finite subcover of  $X$ . That is, there exists a finite subset  $F$  of  $\mathcal{I}$  such that

$$X = \bigcup_{\alpha \in F} K_\alpha^C = X - \bigcap_{\alpha \in F} K_\alpha.$$

So ,

$$\bigcap_{\alpha \in \mathcal{I}} K_\alpha = \emptyset.$$

which is a contradiction as intersection of finite family of sets  $(K_\alpha)_{\alpha \in \mathcal{I}}$  is non-empty. □

**Proposition 8.3 (A special case of compactness in product spaces).** Let  $E$  and  $F$  be two compact subsets of  $\mathbb{R}$  (with the standard metric  $d(x, y) = |x - y|$ ). Show that the Cartesian product  $E \times F := \{(x, y) : x \in E, y \in F\}$  is a compact subset of  $\mathbb{R}^2$  (with the Euclidean metric  $d_{l_2}$ ).

PROOF. Let  $(e_n, f_n)$  be a sequence in  $E \times F$ . We will now prove that this sequence has a convergent sub-sequence. Now, as  $E$  is compact, so  $(e_n)_{n \in \mathbb{N}} \in E$  has a convergent sub-sequence  $e_{n_j}$  which converges to a limit, say  $l_1 \in E$ . Moreover, as  $F$  is compact  $(f_n)_{n \in \mathbb{N}} \in F$  has a convergent sub-sequence, say  $f_{n_j}$  which converges to its limit, say  $l_2 \in F$ .

We claim that  $(e_{n_j}, f_{n_j})$  will converge to  $(l_1, l_2)$ . To do so, let  $\epsilon > 0$  be given. Then  $\frac{\epsilon}{\sqrt{2}} > 0$ , and for this setup, there exists a  $N_1$  such that

$$d(e_{n_j}, l_1) = |e_{n_j} - l_1| < \frac{\epsilon}{\sqrt{2}} \text{ for } n_j > N_1$$

Also,  $f_{n_j}$  converges to  $l_2$ , so for  $\frac{\epsilon}{\sqrt{2}}$ , there is a  $N_2$  such that

$$d(f_{n_j}, l_2) = |f_{n_j} - l_2| < \frac{\epsilon}{\sqrt{2}} \text{ for } n_j > N_2$$

Consider  $N = \max\{N_1, N_2\}$ . Then for  $n_j > N$  we have

$$d((e_{n_j}, f_{n_j}), (l_1, l_2)) = \sqrt{(e_{n_j} - f_{n_j})^2 + (l_1 - l_2)^2} < \sqrt{\frac{\epsilon^2}{2} + \frac{\epsilon^2}{2}} = \epsilon$$

This proves that every sequence in the product space  $E \times F$  has a convergent subsequence that converges in  $E \times F$ . hence, the product of the compact spaces is also a compact spaces. Through induction we can prove that this holds true at least in case of finite product. That is, by induction, it is easy to see that product of finite number of compact spaces is a compact space.  $\square$

**Example 8.1.** Let  $(X, d)$  be a metric space, let  $E$  be a non-empty compact subset of  $X$ , and let  $x_0$  be a point in  $X$ . Show that there exists a point  $x \in E$  such that

$$d(x_0, x) = \inf\{d(x_0, y) : y \in E\},$$

i.e.,  $x$  is the closest point in  $E$  to  $x_0$ .

**Solution 8.1.** Let  $R$  be the quantity  $R := \inf\{d(x_0, y) : y \in E\}$ . We know that if  $R$  is the infimum, then  $R + \epsilon$  is not the lower bound of the set, that is, there exists a  $x$  such that  $x \leq R + \epsilon$ . Choose  $\epsilon = \frac{1}{n}$  and then as we vary  $n \in \mathbb{N}$ , we can construct a sequence  $(x_n)_{n=1}^{\infty}$  in  $E$  such that

$$d(x_0, x_n) \leq R + \frac{1}{n}.$$

Also, as  $R$  is the infimum

$$R \leq d(x_0, x_n) \leq R + \frac{1}{n}.$$

So now  $\lim_{n \rightarrow \infty} d(x_0, x_n) = R$ . Since  $(x_n)_{n=1}^{\infty}$  in  $E$  and  $E$  is compact, there exists a subsequence  $(x_{n_j})$  that converges to some limit  $l \in E$  and we know that  $\lim_{n \rightarrow \infty} d(x_0, x_n) = \lim_{j \rightarrow \infty} d(x_0, x_{n_j}) = d(x_0, l)$ . This gives  $d(x_0, l) = R$  and  $l \in E$ . This proves the claim.

**Example 8.2.** Let  $(\mathbb{R}, d)$  be the real line with the standard metric. Give an example of a continuous function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , and an open set  $V \subseteq \mathbb{R}$ , such that the image  $f(V) := \{f(x) : x \in V\}$  of  $V$  is not open.

**Solution 8.2.** Consider the function  $f : V = (a, b)$  given by  $f(x) = c$ , for some constant real number  $c$ . Note that  $f(V) = \{c\}$ . As singleton sets are always closed, so  $f(V)$  is closed.

**Example 8.3.** Let  $(\mathbb{R}, d)$  be the real line with the standard metric. Give an example of a continuous function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , and an closed set  $F \subseteq \mathbb{R}$ , such that  $f(F)$  is not closed.

**Solution 8.3.** Consider the function  $f(x) = e^x$  and  $F = \mathbb{R}$ .  $f(F) = (0, \infty)$ . This set is not closed. However,  $\mathbb{R}$  is closed. **The essence of this construction lies in the fact that such closed set can not be bounded because otherwise, continuous image of closed and bounded set is closed and bounded. In topological language, we say image of a compact set under a continuous map is a compact set.**

## 9. History of Compactness

The pursuit of firm real analysis foundations in the 19th century marked the beginning of the history of what is now known as the Heine–Borel theorem. The idea of uniform continuity and the theorem that every continuous function on a closed and bounded interval is uniformly continuous were fundamental to the theory. Peter Gustav Lejeune Dirichlet was the first to prove this, and in his proof, he implicitly used the existence of a finite sub-cover of a given open cover of a closed interval. He used this proof in his 1852 lectures, which were published only in 1904. Later Eduard Heine, Karl Weierstrass and Salvatore Pincherle used similar techniques. Émile Borel in 1895 was the first to state and prove a form of what is now called the



FIGURE 10. Peter Gustav Lejeune Dirichlet



FIGURE 11. Heinrich Eduard Heine



FIGURE 12. Salvatore Pincherle

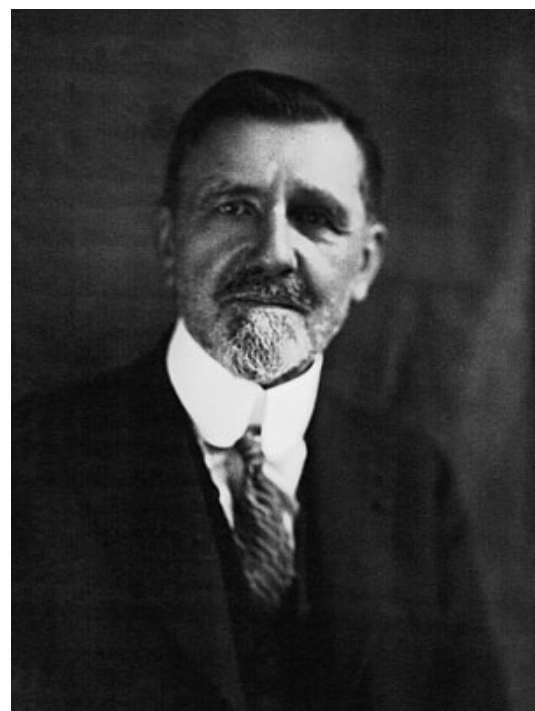


FIGURE 13. Émile Borel

Heine–Borel theorem. His formulation was restricted to countable covers. Pierre Cousin (1895), Lebesgue (1898) and Schoenflies (1900) generalized it to arbitrary covers



FIGURE 14. Henri Lebesgue



FIGURE 15. Arthur Moritz Schoenflies

### 10. Future Works

Due to constraints on time, I am not able to type more work that I have read and understood during my apprenticeship under my mentor. However, I plan to expand it in the coming months before starting my UG. More specifically, I am keen on the typing study of Analysis on Manifold, for which I am deeply fascinated by the contents of . Multidimensional Real Analysis by J. J. Dieustermaat



## CHAPTER 3

### Introduction to groups

The objective of this chapter is to provide a rigorous overview of the theory of groups, classical examples and rudimentary results about important topics which are studied in any standard course on group theory. This sets the direction for the detailed study in subsequent chapters. This and other expository works are outgrowth of my ongoing rigorous apprenticeship at [AMMOC - an international math circle](#)<sup>1</sup>. A point worth mentioning is that the [strong appreciation](#) that I received from Prof. Krishnasamy Arasu (Wright State University) for my own proof of the group theory problem 1.1 that he had asked on LinkedIn on that very day itself.

#### 1. Groups - Definitions and Elementary Properties

**History.** The theory of group is deeply rooted in the symmetries of the following objects.

- Sets whose symmetries are permutations and groups are denoted by  $S_n$ .
- Polynomials  $\prod_{1 \leq i < j \leq n} (x_i - x_j)$  whose symmetries are permutations of variables and the group is known as *alternating group*,  $A_n$ .
- Regular  $n$ -gon whose symmetries are rotation and reflection and the group is known as *Dihedral group*,  $D_n$ .
- Vector space  $V$  whose symmetries are bijective. Linear transformation groups will be known as *General linear group*  $GL_n(V)$ , where  $K$  is a field in the sense of definition 2.1.
- Inner product spaces  $V$  whose symmetries will be *orthogonal linear transformation*,  $O(V)$ .

The first appearance of the definition of an *abstract group* can be traced to the work of [Walther Von Dyck](#) in 1882. Necessary prerequisites on sets, relations, partitions, and results on bijective maps are covered in Appendix A.

**Definition 1.1 (Transformations of a set).** Let  $A$  be an abstract set. By a transformation of  $A$ , we mean a bijective map of  $A$  onto itself.

From our knowledge of set theory, see Appendix A, we know that on the collection  $S_A$  of all transformations of the non-empty set  $A$ , the following facts hold.

- If  $\sigma$  and  $\tau$  are two transformations in  $S_A$ , then their set theoretical composition  $\sigma \circ \tau \in S_A$ .
- $\sigma \in S_A$  then its set-theoretical inverse  $\sigma^{-1} \in S_A$ .
- Identity function/map  $i_A : A \rightarrow A$  is also an element of  $S_A$ .
- The set-theoretical composition of transformations of  $A$  satisfies the associative law.

We call this collection  $S_A$ , the **group of transformations** of  $A$ .

In the systematic study of sets, refer to Appendix A, we have given a detailed discussion of the following.

- (1) Characterization of injective (monomorphism), and surjective (epimorphism) maps on sets, see 2.2 and 2.1
- (2) Notion of 'equivalence relation', 'quotient set' and 'canonical decomposition', see 1.1, 1.1, 2.1, and 2.5 in the Appendix A.

**Goal.** We want to discuss these aspects (1) and (2) in the context of a new object known as **Group 1.1**.

---

<sup>1</sup>AMMOC - a premier international math circle, where I have been receiving my mathematical apprenticeship since **October 2020**, is named in honour of Vladimir Arnold and Jerrold Eldon Marsden - an eminent mathematician and revered faculty of CALTECH



FIGURE 1. Walther Von Dyck

FIGURE 2. Felix Klein - Mentor of  
Walter Von Dyck

Before we introduce the definition of an abstract group, we first define two algebraic structures, namely semi-group and monoid, which are more general than a group.

**Definition 1.2 (Semi-group).** A semi-group is a set  $A$  endowed with a binary operation  $\star$  that satisfies  $a \star (b \star c) = (a \star b) \star c$  for all  $a, b, c \in A$ .

This associativity can be visualized by the following commutative diagram.

$$\begin{array}{ccc}
 A \times A \times A & \xrightarrow{Id_A \times \star} & A \times A \\
 \downarrow \star \times Id_A & & \downarrow \star \\
 A \times A & \xrightarrow{\star} & A
 \end{array}$$

Here  $Id_A : A \rightarrow A$  is the identity map on  $A$ . The above diagram commutes means that the maps  $\star \circ (Id_A \times \star)$  coincides with the map  $\star \circ (\star \times Id_A)$ . That is, these are the two identical ways to go from the upper left corner to the lower right corner.

**Definition 1.3 (Monoid).** A semi-group  $A$  is called a monoid if there exists an element  $e \in A$ , called an *identity*, such that  $e \star a = a \star e = a$  for all  $a \in A$ .

**Lemma 1.1 (Uniqueness).** The identity element  $e$  of the monoid  $A$  is unique.

**PROOF.** Suppose that there are two identity elements  $e, e' \in A$ . Then, for  $e$  being identity, we get  $e \star e' = e'$ . Similarly, for  $e'$  being identity, we get  $e \star e' = e$ . Thus,  $e = e'$ .  $\square$

**Definition 1.4 (Inverse).** Let  $A$  be a monoid and  $a \in A$ . An element  $b \in A$  is called an inverse of  $a$  if  $a \star b = b \star a = e$ . Conventionally we set  $a^{-1} := b$ .

By method of contradiction, it is easy to see that if the inverse of an element in  $A$  exists, then it is unique.

**Definition 1.5 (Abstract group).** A non-empty set  $G$  together with a binary operation  $\star$  is called a group if the following properties hold.

- $(a \star b) \star c = a \star (b \star c)$  for all  $a, b, c \in G$ .
- There exist an element  $e_G \in G$  such that  $e_G \star a = a \star e_G = a$  for all  $a \in G$ . The element  $e_G$  is called the identity of  $G$  with respect to  $\star$ .
- For each  $a \in G$  there exists an element  $b \in G$  such that  $a \star b = b \star a = e_G$ . We denote this element as  $a^{-1}$  and call it as inverse of  $a$  with respect to  $\star$ .

**Lemma 1.2.** A monoid  $G$  is a group if it contains the inverse of every element of  $G$ .

PROOF. This equivalent characterization follows from the definitions of monoid and group.  $\square$

In general, it is not true that  $a \star b = b \star a$  for all  $a, b \in G$ . However, if this holds, then we have a special name (in honour of mathematical genius Neils Henrick Abel) for such groups. In June 2024, a mathematician, Prof. Krishnasamy Arasu (Wright State University), asked a very good problem at a LinkedIn forum. I was the first one to solve that problem. Prof. Arasu wrote a LinkedIn post appreciating my knowledge of the group theory.

**Definition 1.6 (Commutative structure).**  $(G, \star)$  is known as an abelian or commutative semi-group/monoid/group if  $a \star b = b \star a$  for all  $a, b \in G$ .



FIGURE 3. Camille Jordan



FIGURE 4. Niels Henrick Abel

**History.** Camille Jordan named abelian groups after Norwegian mathematician Niels Henrik Abel, as Abel had found that the commutativity of the group of a polynomial implies that the roots of the polynomial can be calculated by using radicals. A very profoundly relevant textbook is [Arnold & Alekseev].

**On the notations.** Generally, we will omit the operation  $\star$  in  $a \star b$  and write this as  $ab$ . The identity element would be denoted by 1 or  $e_G$ . As long as group operations are clear, we would take the liberty to skip their explicit mention in the computations involved.

**Proposition 1.1** (Fundamental arithmetic moves). We have already proved the uniqueness of the identity element and the inverse of an element. The following are direct consequences of associative axiom, left/right multiplication by a suitable element, using associativity, definition of identity element and induction.

- Cancellation laws:

$$gx = gy \implies x = y \text{ and } xg = yg \implies x = y$$

for all  $x, y, g \in G$ .

- Mirror rule:  $(xy)^{-1} = y^{-1}x^{-1}$  for all  $x, y \in G$ .
- Telescopic product:  $(gxg^{-1})^n = gx^n g^{-1}$  for all  $x, y \in G$  and  $n \in \mathbb{Z}$ .
- $g^m g^n = g^{m+n}$  and  $(g^m)^n = g^{mn}$  for all  $g \in G$  and  $m, n \in \mathbb{Z}$ .

**Problem 1.1.** Let  $g \in G$ . If  $g^5 = g^7 = 1$ , then  $g = 1$ .

**Solution 1.1.** Note that,  $g^7 = g^5 = 1 \implies g^2 = 1$ . Then  $g^5 = g^3 \implies g^3 = 1$ . Using this  $g^7 = g^4 = 1 \implies g^4 = 1$ . From this, we get  $g^5 = g \implies g = 1$ . This proves the claim.

**Alternate solution.** Note that,  $(7, 5) = 1$ . Therefore, by *Bezout's lemma*, we know that there exist integers,  $x, y$  such that  $7x + 5y = 1$ . In our case, a direct observation gives  $x = 2, y = -3$ . Using this,  $g^1 = g^{7 \cdot 2 - 5 \cdot 3} = (g^7)^2 (g^5)^{-3} = 1^{2-3} = 1^{-1} = 1 \implies g = 1$ .

The above alternate solution is the motivation for the following lemma.

**Lemma 1.3.** Let  $(m, n) = 1$  and  $g^m = g^n = 1$ . Prove that  $g = 1$ .

PROOF. By *Bezout's lemma*, we know that there exist integers  $x, y$  such that  $mx + ny = 1$ . Using this,  $g = g^{mx+ny} = g^{mx} g^{ny} = 1^{x+y} = 1$ .  $\square$

The concept of an abelian group underlies algebraic structures such as vector spaces, rings, fields and algebra.

**Problem 1.2.** Let  $G$  be a group such that each non-identity element is of the order 2. Then  $G$  is an abelian group.

**Solution 1.2.** The given information can be interpreted as  $g = g^{-1}$  for all  $g \in G$ . Now, let  $a, b \in G$ . Then, we have  $ab = (ab)^{-1} = b^{-1}a^{-1} = ba$ , where the last step follows from the given fact that  $b^{-1} = b$  and  $a^{-1} = a$ .

**Theorem 1.1** (The question asked by Prof. Arasu). Let  $(G, \star)$  be any group. For any fixed positive integer  $n$ , let  $P(n)$  denote the premise:  $(ab)^n = a^n b^n$  for all elements  $a$  and  $b$  in  $G$ .

- (1) Prove that  $G$  is abelian if and only if  $P(2)$  is true.
- (2) If  $P(n)$  is true for any three consecutive integers (say  $k, k+1$ , and  $k+2$ ), prove that  $G$  is abelian.
- (3) If  $P(8)$  and  $P(11)$  are true, prove that  $G$  is abelian.
- (4) Generalization: If  $P(k)$  and  $P(k+3)$  are true, prove that  $G$  is abelian

PROOF. (1) Assume that  $G$  is an abelian. Then  $(ab)^2 = abab = aabb = a^2b^2$ . Conversely, we assume that  $(ab)^2 = a^2b^2$ . Then, we have  $abab = aabb$ . By cancellation law, we have  $ab = ba$ .

- (2) Since  $(ab)^k = a^k b^k$ ,  $(ab)^{k+1} = a^{k+1} b^{k+1}$ ,  $(ab)^{k+2} = a^{k+2} b^{k+2}$ , we have  $(ab)^{k+1} = (ab)(ab)^k = aba^k b^k = a^{k+1} b^{k+1}$ . By cancellation law, we have

$$(8) \quad a^k b = ba^k.$$

Next,  $(ab)^{k+2} = (ab)^2 (ab)^k = ababa^k b^k$ . By cancellation law,  $a^{k+1} b^2 = baba^k = baa^k b$ . By the equation (8) and By cancellation law,  $a^{k+1} b = ba^{k+1} = ba^k a = a^k ba$ . So,  $a^{k+1} b = a^k ba$  and applying cancellation law, we get  $ab = ba$ .

- (3)  $P(8)$  and  $P(11)$  are both true. So,

$$(9) \quad (ab)^8 = a^8 b^8$$

$$(10) \quad (ab)^{11} = a^{11} b^{11}.$$

Now, consider  $(ab)^{11} = (ab)^8 (ab)^3 = a^8 b^8 (ab)^3 = a^{11} b^{11}$ . So  $a^3 b^{11} = b^8 (ab)^2 ab$ .  $(ab)^2 = b^{-8} a^3 b^{10} a^{-1}$ .  $(ab)^{10} = (ab)^8 (ab)^2 = a^8 b^8 b^{-8} a^3 b^{10} a^{-1} = a^{11} b^{10} a^{-1}$ .

$$(ab)^9 = (ab)^{11}(ab)^{-2} = a^{11}b^{11}b^{-2}a^{-2} = a^{11}b^9a^{-2}.$$

Therefore,  $(ab)^{10} = (ab)^9(ab)$ . Thus,  $a^{11}b^{10}a^{-1} = a^{11}b^9a^{-2}ab$  and  $ba^{-1} = a^{-1}b \implies ab = ba$ .

(4) As  $P(k), P(k+3)$  are true we have  $(ab)^k = a^k b^k$  and  $(ab)^{k+3} = a^{k+3} b^{k+3}$ . This implies that

$$a^{k+3}b^{k+3} = (ab)^{k+3} = (ab)^k(ab)^3 = a^k b^k (ab)^2 ab$$

and

$$a^3 b^{k+2} = b^k (ab)^2 a.$$

Multiplying the left side by  $b^{-k}$  and the right side by  $a^{-1}$ , we get

$$(ab)^2 = b^{-k} a^3 b^{k+2} a^{-1},$$

$$(ab)^{k+2} = (ab)^k (ab)^2 = a^k b^k b^{-k} a^3 b^{k+2} a^{-1}$$

$$(ab)^{k+2} = a^{k+3} b^{k+2} a^{-1},$$

$$(ab)^{k+1} = (ab)^{k+3} (ab)^{-2} = a^{k+3} b^{k+3} b^{-2} a^{-2}$$

$$(ab)^{k+1} = a^{k+3} b^{k+1} a^{-2}$$

$$a^{k+3} b^{k+2} a^{-1} = (ab)^{k+2} = (ab)^{k+1} (ab) = a^{k+3} b^{k+1} a^{-2} \cdot ab.$$

By cancellation law, we get

$$ba^{-1} = a^{-1}b \implies ab = ba.$$

□

**Lemma 1.4 (Translations and shuffle).** We have two very innocuous bijections  $G \hookrightarrow G$  is given by

- Left translation by  $g : x \mapsto gx$
- Right translation by  $g : x \mapsto xg$

PROOF. The inverse of these two maps is given by left (respectively, right) translation by  $g^{-1}$ . □

This is the right time to introduce the notion of conjugation.

**Definition 1.7 (Conjugation).** Let  $G$  be a group and  $g \in G$  be an arbitrarily fixed element. The map  $G \rightarrow G$  given by  $x \mapsto xg^{-1}$  is known as the conjugation of  $x$  by  $g$ .

Note that, the conjugate of the product  $xy$  is the product of the conjugate. That is  $g(xy)g^{-1} = gxg^{-1}gyg^{-1}$ . Also, a conjugate of  $x$  by the product  $gh$  is the same as taking the first conjugate by  $h$  and then conjugating the outcome by  $g$ . That is  $ghx(gh)^{-1} = ghxh^{-1}g^{-1} = g(hxh^{-1})g^{-1}$ .

It is natural to ask the following question.

Given a structure  $S$ , does there exist a sub-structure  $H$ , whereby sub-structure we mean to say a subset of the original structure that obeys all axioms of  $S$ . For example, given a group  $G$ , does there exist a subset  $H$  of  $G$  such that all defining axioms of the group are also true when we restrict them on  $H$ ? This question motivates the following definition.

**Definition 1.8 (Subgroup).** Let  $(G, \star)$  be a group. Let  $H$  be a subset of  $G$ . We say that  $H$  is a subgroup of  $G$  and write this as  $H \leq G$  if  $(H, \star)$  is a subgroup in its own right. That is,  $H$  is closed under the binary operation  $\star$  and inverses.

**Definition 1.9 (Normal subgroup).** Let  $H$  be a subgroup of group  $G$ . Define  $gHg^{-1} = \{ghg^{-1} \mid h \in H\}$ . This is known as the conjugation of  $H$  by  $g$ . We say that  $H$  is a *normal subgroup* of group  $G$  if  $gHg^{-1} = H$  for all  $g \in G$  and denoted by  $H \trianglelefteq G$ .

**Definition 1.10 (Order of an element).** Let  $G$  be a group and  $g \in G$ . The smallest positive integer  $n$  for which  $g^n = e_G$  is known as the order of  $g$ , and it is denoted by  $|g|$ . If there is no such positive integer  $n$ , then we say that  $g$  has an 'infinite' order.

**Lemma 1.5.** Let  $n$  be the order of an element  $a \in G$ . Prove that  $a^m = e$  if and only if  $m = nd$ , where  $d$  is an integer.

PROOF. Assume that  $a^m = e$  and the order of  $a$  is  $n$ . By Euclid division algorithm, we have  $m = nq + r$  where  $0 \leq r < n$ .  $a^m = a^{nq+r} = a^{nq}a^r = ea^r = a^r$ . So,  $a^r = e$ , but  $r < n$  and  $n$  is the order of  $a$ . So we must have  $r = 0$  and  $n = mq$  for some integer  $q$ .

Conversely, we assume that  $m = nd$  for some integer  $d$ . Then  $a^m = a^{nd} = (a^n)^d = e^d = e$ .  $\square$

**Corollary 1.1.** Suppose that the order of an element  $a$  is equal to a prime number  $p$  and  $m$  is an arbitrary integer. Prove that either  $a^m = e$  or  $a^m$  has order  $p$ .

PROOF. Case 1) If  $p$  divides  $m$ , then  $a^m = e$  (As  $a^p = e$ ).

Case 2) If  $p$  does not divide  $m$ , then  $\gcd(m, p) = 1$ . Now, prove that the order of  $a^m$  is  $p$ . Clearly,

$$(a^m)^p = a^{mp} = (a^p)^m = e^m = e.$$

Assume  $(a^m)^x = e$  for some positive integer  $x$ . It is enough to prove that  $p$  divides  $x$ . By the previous problem  $p \mid mx$ , but as  $p$  does not divide  $m$ , we get that  $p$  divides  $x$ .  $\square$

**Corollary 1.2.** Suppose that  $d$  is the maximal common divisor of the integers  $m$  and  $n$  and that  $a$  has order  $n$ . Prove that the element  $a^m$  has order  $\frac{n}{d}$ .

PROOF. First, note that

$$(a^m)^{\frac{n}{d}} = a^{\frac{mn}{d}} = a^{\frac{mn}{\gcd(m,n)}} = a^{\text{lcm}(m,n)} = e.$$

Let  $m = dp$ ,  $n = dq$  where  $\gcd(p, q) = 1$ . Let  $(a^m)^x = e$ . It is sufficient to prove that  $\frac{n}{d}$  divides  $x$ . i.e to prove that  $q$  divides  $x$ . Now,  $a^{mx} = e$ , and order of  $a$  is  $n$ . So we have  $n$  divides  $mx$ . So  $dq$  divides  $dpqx$ . So  $q$  divides  $px$ . But as  $\gcd(p, q) = 1$ , we get that  $q$  divides  $x$ .  $\square$

**Problem 1.3.** Let  $G$  be a finite abelian group with exactly one element  $f$  of order 2. Prove that  $\prod_{g \in G} g = f$ .

**Solution 1.3.** An element of order 2 is its own inverse. Using the commutativity of group  $G$ , the given product  $\prod_{g \in G} g = f$  can be rearranged to keep an element and its inverses together. Using associative law, the whole product becomes  $f \star e_G \star \prod_{g \in G, g \neq f} g = f \star e_G = f$ .

**Problem 1.4.** Let  $G$  be a finite group of order  $n$ . Let  $m$  be the number of elements of order 2 in  $G$ . Prove that  $n - m$  is odd. Deduce that if  $n$  is even, then  $G$  necessarily contains an element of order 2. Now, if  $n$  is even and we know that  $n - m$  is odd, this implies that  $m$  is odd. That means  $m \geq 1$ . Hence,  $G$  necessarily contains an element of order 2.

**Solution 1.4.** We know that all the elements of order 2 are their own inverse. If  $x \in A := G - \{H, 1\}$ , where  $H$  is the collection of all elements of order two, then for  $y \in A$ , no elements will be its own inverse. Given that the inverse for each element is unique,  $\{y, y^{-1}\}$  is always a pair. Therefore, the cardinality  $|A|$  of  $A$  is always even. Since we are only counting those elements which are not of order 2, it is basically  $|A| + 1$  because we have to exclude identity from  $A$ . Hence,  $|A| + 1$  is always odd.

We first discuss some examples, which we call as Toy group (basically some select groups which are of interests to a beginner algebraist), of classical groups which will appear through out our discourse to discuss standard results in theory of groups.

## 2. Quaternionic Group

**Definition 2.1 (Field).** A field  $K$  is a non-empty set together with two binary operations  $+$  and  $\times$  such that the following hold.

- (1)  $(K, +)$  is an additive abelian group,
- (2)  $(K^\times, \times)$  is a multiplicative abelian group, where  $K^\times$  is the set  $K - \{0\}$ , and
- (3) Distributive law holds true:  $a \times (b + c) = a \times b + a \times c$  and  $(a + b) \times c = a \times c + b \times c$  for all  $a, b, c \in K$ .

**Definition 2.2 (Skew-field).** In definition 2.1 of Field, if the multiplicative group is non-abelian, then we say  $K$  is skew-field.

**Motivation behind Quaternion.**  $\mathbb{Q}$  and  $\mathbb{R}$  are fields. The Cartesian plane  $\mathbb{R}^2 := \{(x, y) | x, y \in \mathbb{R}\}$  is not a field under the following definitions of addition and multiplication

- Point-wise addition:  $(a, b) + (c, d) := (a + c, b + d)$ ,
- Point-wise multiplication:  $(a, b) \cdot (c, d) := (ac, bd)$ .

First, note that  $(1, 1)$  is the only possible candidate for multiplicative identity with respect to point-wise multiplication. Now, for  $\mathbb{R}^2$  to be a field under 'point-wise addition and point-wise multiplication', we would need multiplicative inverses, and there is no multiplicative inverse for any non-zero element of the form  $(a, 0)$  where  $a \neq 0$ . **To achieve our objective of endowing a field structure on  $\mathbb{R}^2$** , we re-define multiplication



FIGURE 5. Sir William Rowan Hamilton

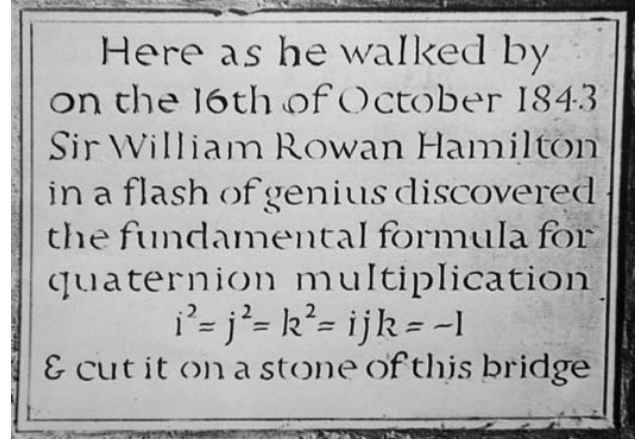


FIGURE 6. Quaternion Plaque on Brougham Bridge in Dublin

on  $\mathbb{R}^2$  as  $(a, b)(c, d) = (ac - bd, ad + bc)$ . If we represent each pair  $(a, b)$  in  $\mathbb{R}^2$  by  $a + \mathbf{i}b$ , then multiplication becomes

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc) \equiv (ac - bd) + \mathbf{i}(ad + bc)$$

with a 'constraint' that  $\mathbf{i}^2 = -1$ . The field obtained by endowing such an axiomatic structure on  $\mathbb{R}^2$  is known as the field of complex numbers.

**Discovery of Quaternions.** The difficult question that Sir Hamilton pursued was "Is it possible to create a multiplication operation which, along with component-wise addition, makes  $\mathbb{R}^n$  into a skew-field for  $n > 2$ ". In 1843, while taking a long walk along with his wife, Hamilton discovered that the answer is yes for  $n = 4$ , and he carved his discovery into the stone as he paused on it, see Fig. 6. Now, we define the rule of multiplication in this skew field  $\mathbb{R}^4$ . An element of which will be denoted as  $a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d$  and multiplication rule for  $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$  is what occurred as a flash of genius to Sir Hamilton:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$$

$$\mathbf{i}\mathbf{j} = \mathbf{k}, \mathbf{j}\mathbf{k} = \mathbf{i}, \mathbf{k}\mathbf{i} = \mathbf{j}$$

$$\mathbf{j}\mathbf{i} = -\mathbf{k}, \mathbf{k}\mathbf{j} = -\mathbf{i}, \mathbf{i}\mathbf{k} = -\mathbf{j}$$

For the sake of convenience, let us call these relations as **Hamilton's algebraic relations**. Now, we define the **multiplication** as

$$(a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d) + (u + \mathbf{i}v + \mathbf{j}w + \mathbf{k}x) = (au - bv - cw - dx) + (av + bu + cx - dw)\mathbf{i} \\ + (aw + cu + dv - bx)\mathbf{j} + (ax + du + bw - cv)\mathbf{k}.$$

The set  $\mathbb{R}^4$  with this definition of component-wise addition and multiplication is known as a set of **quaternion**, and we denote this set as  $\mathbb{H}$ . Now, we define **multiplicative inverse** of a non-zero quaternion, and this will be the final step in endowing a **skew field** structure on  $\mathbb{H}$ . Denote  $\bar{q}$  as conjugate of  $q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d$  and define  $\bar{q} := a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$ . Let  $|q| := \sqrt{a^2 + b^2 + c^2 + d^2}$  be known as norm of  $q$ . Note that,  $|\bar{q}| = |q|$ . Also,  $q \neq 0$  if and only if at least one of the real numbers  $a, b, c, d$  is non-zero. Hence, the norm of a non-zero quaternionic number is always positive. Define inverse of non-zero  $q$  to be  $\frac{\bar{q}}{|q|^2}$ .

If we only consider quaternions without  $j, k$  components, then we see that

$$\mathbb{R} \subset \mathbb{C} \subset \mathbb{H}$$

as skew fields.

**Remark 2.1.** Multiplication between a real number and a quaternion is commutative. However, multiplication between every purely imaginary complex number and any quaternion is always non-commutative. These claims are obvious from the definition of multiplication and Hamilton's algebraic relations.

**Definition 2.3 (The quaternion group).** The set  $\mathbf{Q}_8 : \{1, \mathbf{i}, \mathbf{j}, \mathbf{k}, -1, -\mathbf{i}, -\mathbf{j}, -\mathbf{k}\}$  with the product defined via **Hamilton's algebraic relations** is known as quaternion group  $\mathbf{Q}_8$ . This is our first example of *non-abelian group*.

### 3. General Linear Groups

**Definition 3.1 (General and special Linear group).** Let  $F$  be a field in the sense of definition 2.1. The *General Linear Group*  $GL_n(\mathbb{F})$  over a field  $\mathbb{F}$ , is a set of all invertible square matrices with entries from  $\mathbb{F}$ . *Special Linear Group*  $SL_n(\mathbb{F})$  is a subgroup of  $GL_n(\mathbb{F})$  such that each matrix in  $SL_n(\mathbb{F})$  has a determinant of 1.

If the Field  $\mathbb{F}$  has infinite entries, then  $GL_n(\mathbb{F})$  has infinite entries. For example. for all  $a \in \mathbb{F} - \{0\}$ , consider the matrix  $aI_n$ , then this matrix is invertible with inverse  $a^{-1}I_n$  and its determinant is  $\frac{1}{a}$ . Hence,  $GL_n(\mathbb{F})$  is an infinite group if  $\mathbb{F}$  is an infinite group. In fact, such matrices have special name **scalar matrices** and they form a subgroup of  $GL_n(\mathbb{F})$  and moreover this subgroup is isomorphic (see definition 7.2) to  $\mathbb{F} - \{0\}$ , the set of non-zero entries in the field  $\mathbb{F}$ .

**Lemma 3.1.** Let  $\mathbb{F}_q$  be a finite field of order  $q$ . Then the number of elements of  $GL_n(\mathbb{F}_q)$  is  $\prod_{k=0}^{n-1} (q^n - q^k)$ .

Here we use a bit clever argument involving linear independence of vectors in context of vector spaces and required material is covered systematically in my own work - [Linear algebra notes](#). However, we just quickly recall the characterization of an invertible matrix.

**Definition 3.2 (Characterization of an invertible matrix).** A matrix is invertible if and only if all of its rows are linearly independent.

Now, we prove the lemma.

**PROOF.** So, a matrix with entries from  $\mathbb{F}_q$  is an element of  $|GL_n(\mathbb{F}_q)|$  if and only if rows of this matrix form a linearly independent set of vectors. In particular, we will use the fact that if a matrix is invertible, then none of its rows (or columns) can be a zero row (or column). We will work with rows. So what are the possibilities for the first row of a  $n \times n$  matrix to be non-zero? Well, this essentially means that at least one element of the first row must be non-zero, and the rest can be any element of  $\mathbb{F}_q$ . Using the multiplication principle of Combinatorics, we see that this can be done in  $q^n - 1$  ways. The second row must not be any scalar (taken from  $\mathbb{F}_q$ ) multiple of the first row. There are  $q$   $\mathbb{F}_q$ -multiple of the first row, and we can't take them. Therefore, for the second row, we have  $q^n - q$  possibility. Now, the third row can not be  $\mathbb{F}$ -linear combination of the first two rows, say  $R_1, R_2$ . That is,  $R_3 \neq x_1 R_1 + x_2 R_2$  for any  $x_1, x_2 \in \mathbb{F}_q$ . Moreover,  $x_1, x_2$  total  $q^2$  choices. Hence,  $R_3$  has  $q^n - q^2$  choices. In particular,  $k^{th}$  row has to be linearly independent of the first  $k - 1$  rows and therefore, generalizing our previous argument gives us  $q^n - q^{k-1}$  possibilities. Thus, total number of possible invertible square matrices are  $(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{n-1}) = \prod_{k=0}^{n-1} (q^n - q^k)$ . □

**Corollary 3.1 (Estimate of  $|GL_n(\mathbb{F}_q)|$ ).** Let  $\mathbb{F}_q$  be a finite field of order  $q$ . Then, we have  $|GL_n(\mathbb{F}_q)| < q^{n^2}$ .

**PROOF.** Note that,  $q^n - q^k < q^n$ . For any  $n > 2$ ,  $2n < n^2$ . Therefore,  $\prod_{k=0}^{n-1} (q^n - q^k) < q^{2n} < q^{n^2}$ . □

**Proposition 3.1** (Heisenberg group in quantum mechanics). Let

$$H(F) = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \mid a, b, c \in F \right\}.$$

We call  $H(F)$  is the Heisenberg group over  $F$ . Let  $X = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$  and  $Y = \begin{pmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix}$  be elements of  $H(F)$ .

- (1) Compute the matrix  $XY$  and deduce that  $H(F)$  is closed under matrix multiplication. Exhibit explicit matrices such that  $XY \neq YX$  (so that  $H(F)$  is always non-abelian).
- (2) Find an explicit formula for the matrix inverse  $X^{-1}$  and deduce that  $H(F)$  is closed under inverses.
- (3) Prove the associative law for  $H(F)$  and deduce that  $H(F)$  is of the order  $|F|^3$  (Do not assume that matrix multiplication is associative).
- (4) Find the order of each element of the finite group  $H(\mathbb{Z}/2\mathbb{Z})$ .
- (5) Prove that every non-identity element of the group  $H(\mathbb{R})$  has infinite order.

**Solution 3.1.** (1) Since

$$XY = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+d & af+b+e \\ 0 & 1 & c+f \\ 0 & 0 & 1 \end{pmatrix} \in H(F),$$

we have  $H(F)$  is closed under multiplication.

- (2) Let

$$V = \begin{pmatrix} 1 & -a & ca-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}.$$

It is easy to verify that  $VZ = ZV = I$ . So,  $V$  is the inverse of  $I$ .  $H(F)$  is closed under inverses.

- (3) Let  $X = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$ ,  $Y = \begin{pmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix}$  and  $Z = \begin{pmatrix} 1 & g & h \\ 0 & 1 & i \\ 0 & 0 & 1 \end{pmatrix}$ . Now, we compute  $(XY)Z$ .

First, we compute  $XY$ . We have

$$XY = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+d & af+b+e \\ 0 & 1 & c+f \\ 0 & 0 & 1 \end{pmatrix}.$$

Next, we consider

$$(XY)Z = \begin{pmatrix} 1 & a+d & af+b+e \\ 0 & 1 & c+f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & g & h \\ 0 & 1 & i \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+d+g & ai+di+af+b+e+h \\ 0 & 1 & c+f+i \\ 0 & 0 & 1 \end{pmatrix}.$$

Now, we find  $X(YZ)$ . Note that,

$$YZ = \begin{pmatrix} 1 & d & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & g & h \\ 0 & 1 & i \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & d+g & di+h+e \\ 0 & 1 & f+i \\ 0 & 0 & 1 \end{pmatrix}.$$

Now,  $X(YZ)$  is

$$X(YZ) = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & d+g & di+h+e \\ 0 & 1 & f+i \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+d+g & af+ai+b+di+h+e \\ 0 & 1 & c+f+i \\ 0 & 0 & 1 \end{pmatrix}.$$

So  $X(YZ) = (XY)Z$  and  $H(F)$  is associative. Since  $a, b, c$  can take  $|F|$  values, we have the multiplication law order of  $H(F)$  is  $|F|^3$ .

(4) Now,  $H(\mathbb{Z}/2\mathbb{Z})$  has by the previous part  $2^3 = 8$  elements.

(5) Let  $X$  be a non-identity element of  $H(\mathbb{R})$ . Let  $X = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$  Now,

$$X^2 = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2a & ac + 2b \\ 0 & 1 & 2c \\ 0 & 0 & 1 \end{pmatrix}.$$

In general, it is easy to see that in  $X^n$ , the first-row second column has value  $na$ . So for any  $n \neq 1$ , we cannot have  $na = a$ , i.e.  $X^n = X$ . So  $X$  has an infinite order.



FIGURE 7. Werner Heisenberg - Father of Quantum Matrix Mechanics and winner of the Nobel Prize in physics.



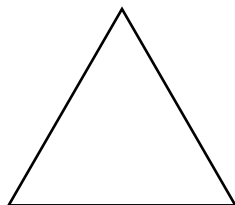
FIGURE 8. Hermann Weyl (left) - last great universal mathematician of the nineteenth century.

#### 4. Dihedral group

For now, we only introduce the idea of the dihedral group with its 'elementarily' most important algebraic properties proved in Theorem 4.1.

**Definition 4.1 (Isometries of the Euclidean plane and space).** An isometry of the Euclidean plane  $\mathbb{R}^2$  or space  $\mathbb{R}^3$  is a distance preserving map  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  (respectively,  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ ). That is, the distance  $d$  between points  $a, b \in \mathbb{R}^2$  and distance  $d'$  between  $f(a), f(b)$  are the same.

**Example 4.1** (Group of symmetries of the equilateral triangle). We consider all the isometric transformations (distance preserving mappings) of the triangle  $ABC$  onto itself.



We label the vertices of the triangle as  $A, B, C$  in the anti-clockwise direction. Let  $r_{120}$  and  $r_{240}$  be rotations of this triangle in a counter-clockwise direction by an angle of 120 and 240 degrees, respectively. Let  $t_a, t_b, t_c$  be the reflections about an axis passing through  $A, B$  and  $C$ . Elements  $r_{120}$  and  $r_{240}$  are inverses of each other and  $t_a, t_b, t_c$  are their own inverses. Thus, the collection  $D_6 = \{e, r_{120}, r_{240}, t_a, t_b, t_c\}$  has a group structure is known as the group of symmetries of an equilateral triangle. Basically, the rotation of an equilateral triangle by 120 degrees sends vertices  $A \mapsto B, B \mapsto C, C \mapsto A$ . Any isometry of the  $\triangle$  will send vertices to vertices, and any one-to-one correspondence  $f : \{A, B, C\} \rightarrow \{A, B, C\}$  determines an isometry. Therefore,  $D_6$  is an exhaustive collection of all isometric transformations of an equilateral triangle  $ABC$ . It is easy to verify that

- rotations do commute among themselves
- rotations and reflections do not commute
- reflections are elements of order 2.

So,  $D_6$  is a nice model of a non-abelian group.

Now, we extend this concept to any regular  $n$ -sided polygon.

**Definition 4.2** (Dihedral group). Consider a regular  $n$ -gon. We want to discuss the group of symmetries - rotations and reflections.

- **Rotations.** There are total  $n$  rotations by  $\frac{2\pi}{n}$  about the center (origin) of the regular  $n$ -gon. Since rotations are isometric automorphism, therefore, we conclude that  $\{r^0, r^1, \dots, r^{n-1}\}$  is indeed a cyclic group under the composition of isometric transformation as functions.
- **Reflections.** Reflection about a line that passes through the centre and a vertex or through the midpoint of a side. There are  $n$  such possible symmetries.
- **To define the composition of reflection and rotation.** Let  $s$  be a reflection and  $t$  be a rotation. The composition  $s \circ t$ , written as  $st$ , means that first, we apply the rotation  $t$  on the polygon, and then we apply the reflection  $s$  on the outcome of the previous step. Note that, the composition of any number of isometric automorphisms is an isometric automorphism, and hence,  $st$  is indeed a symmetry of regular  $n$ -gon. Thus,  $D_{2n} = \{1, r, r^2, \dots, r^{n-1}, s, sr, sr^2, \dots, sr^{n-1}\}$  is the exhaustive set of symmetries of a regular  $n$ -gon which forms the group structure under the operation of 'composition of reflection and rotations as an isometric automorphism of the regular  $n$ -gon'.

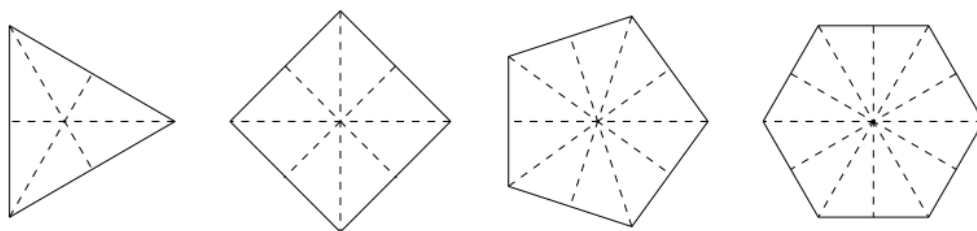


FIGURE 9. Regular  $n$ -gons for  $n = 3, 4, 5, 6$

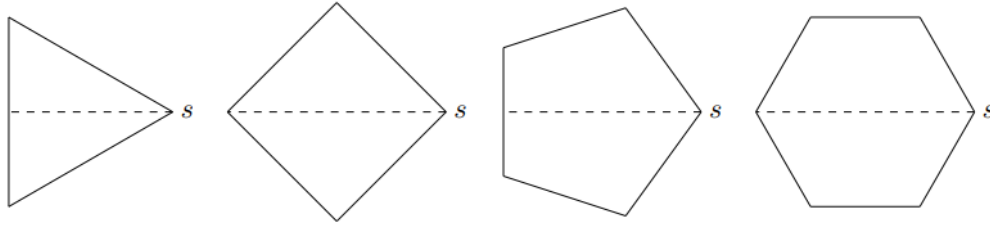


FIGURE 10. Lines of reflections for some n-gon

**Theorem 4.1 (Properties of Dihedral group).**

- (1) In dihedral group  $D_{2n}$ ,  $sr s^{-1} = r^{-1}$   
 (2)  $r^i s(r^j) = sr^{-1}r^j = sr^{j-i}$ .

**PROOF.** (1) This follows from the fact that  $rs$  is a reflection element and therefore,  $(rs)^2 = rsrs = 1 \implies sr s = r^{-1}$ . But a reflection is its own inverse. Therefore,  $s = s^{-1}$ . This gives  $sr s^{-1} = r^{-1} \iff sr = r^{-1}s \iff rs = sr^{-1}$ . In essence, it says that you can move  $r$  to the other side of  $s$  by merely inverting it. Though it is a very simple result, it is extremely effective in carrying out computations in very efficient algebraic ways without appealing to underlying geometric transformations. Now, let us generalise it. Let  $k$  be a positive integer, then by induction, it is obvious that  $r^k s = sr^{-k}$ . Then, we can extend this to all integers. From this, we conclude that *we can move any power of  $r$  to the other side of  $s$  by merely inverting it.*

- (2) This part follows immediately from the discussion in part (1). □

## 5. Cyclic Groups

**Motivation.** Recall that a ‘symmetry’ is a geometrical operation that maps the object onto itself. This basically means that the object is invariant under such transformations. Some examples of such symmetries are translation, scaling, rotation and reflection. For example, a regular  $n$ -gon rotated by an angle of  $\frac{2\pi}{n}$  about its centre maps the polygon into itself. More precisely, it is an isometric automorphism of the regular  $n$ -gon. A circle is thus said to be symmetric under rotation or to have rotational symmetry. If we consider the collection of rotational symmetry of geometrical objects, say of a regular  $n$ -gon, they form a group under the usual composition of functions, and each rotational symmetry is essentially obtained by one carrying out a rotation of  $\frac{2\pi}{n}$ .

Such groups have special names in group theory.

**Definition 5.1 (Cyclic group).** If  $a$  is an element of order  $n$  in  $G$ , in the sense of Definition 1.10, and there are no elements in  $G$  other than  $e_G, a, a^2, \dots, a^{n-1}$  then we say that  $G$  is a cyclic group of ‘order  $n$ ’ and generated by ‘the generator  $a$ ’.

### 5.0.1. Cyclic groups - some solved problems.

**Example 5.1.** (1) Prove that rotations of a regular  $n$ -gon form a cyclic group of order  $n$ .

- (2) Find all generators in the group of rotations of the equilateral triangle and group of rotations of the square.  
 (3) Prove that, if  $a$  is an element of infinite order in the group  $G$ , then

$$, \dots, a^{-3}, a^{-2}, a^{-1}, a^0, a^1, a^2, a^3, \dots,$$

are all unique.

**Solution 5.1.** (1) Denote  $r$  as rotation by  $\frac{2\pi}{n}$  degree in anticlockwise direction. We prove that all the rotations are generated by  $r$ . Denote  $G$  as the rotations of the plane that transform  $n$ -gon into itself. i.e.,  $G = \langle r \rangle$ . Clearly,  $\langle r \rangle \subseteq G$ . To prove that  $G \subseteq \langle r \rangle$ . Let  $g \in G$ . So  $g$  transforms  $n$ -gon

into itself. It is easy to see that  $g$  must be a multiple of  $\frac{2\pi}{n}$  for it to transform  $n$ -gon into itself. Thus,  $g \in \langle r \rangle$ .

- (2) By the previous problem, the group of rotations of an equilateral triangle form a cyclic group of order 3, which is generated by  $r$ , where  $r$  denotes rotation by 120 degrees acw. Let  $G = \{e, r, r^2\}$ . Then  $G = \langle r \rangle$ . and  $\langle r^2 \rangle = \{e, r^2, r^4\} = \{e, r, r^2\} = G$ .  $G$  is generated by  $r, r^2$ .  $r$  denotes rotation by 120 degrees acw and  $r^2$  denotes rotation by 240 degrees acw. By the previous problem, the group of rotations of a square form a cyclic group of order 4, which is generated by  $r$ , is the rotation by 90 acw. Also, it is isomorphic to  $Z_4$ . We know that 1, 3 generates  $Z_4$ . So  $r, r^3$  generates the group of rotation of squares, where  $r$  denotes rotation by 90 degrees acw and  $r^3$  denotes rotation by 270 degrees acw.
- (3) Assume for the sake of contradiction, there exists integers  $i, j$  such that  $a^i = a^j$ . So  $a^{i-j} = e$ . This means,  $|a|$  divides  $i - j$ . So, the order of  $a$  is finite, and this is a clear contradiction.

## 6. Permutation Groups

The collection of permutation of set  $A = \{1, 2, 3, 4, \dots, n\}$ . Usual properties of permutations as bijections of underlying set  $A$  guarantee that this collection  $S_n = \{\sigma \mid \sigma \text{ is a permutation of the set } A\}$  is indeed a group under the composition of functions.

**Cycle notation.** A permutation  $\sigma$  that maps  $i \mapsto a_1$  will be denoted using the following notation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ a_1 & a_2 & a_3 & a_4 & \dots & a_n \end{pmatrix}.$$

A  $n$ -cycle  $\sigma := (a_1 a_2 a_3 \dots a_n)$  is the permutation of the set  $A = \{1, 2, 3, \dots, n\}$  defined by

$$(11) \quad \sigma(x) = \begin{cases} x & x \notin \{a_1, a_2, \dots, a_n\} \\ a_1 & x = a_n \\ a_{i+1} & x = a_i \text{ and } 1 \leq i < n. \end{cases}$$

A 2-cycle is generally known as *transposition*. It is obvious that a  $n$ -cycle has the order  $n$  in the group  $S_n$ . Also, from the definition of a  $n$ -cycle, it is clear that  $(a_1 a_2 a_3 a_4 \dots a_n) = (a_2 a_3 \dots a_n a_1) = (a_n a_1 \dots a_{n-1})$ .

**Proposition 6.1 (Main results on the cycle decomposition).** (1) Every permutation can be uniquely represented (up to different orderings of factors) by a product of independent cycles.

- (2) Every cycle can be represented as a product of transpositions (not necessarily independent).
- (3) The transposition  $(12), (13), \dots, (1n)$  generate  $S_n$ .
- (4) The transpositions  $(1, 2), (2, 3), \dots, (n-1, n)$  are called elementary transpositions. Every transposition can be represented as the product of elementary transpositions.

**PROOF.** (1) Consider  $\sigma^0(1) = 1, \sigma^1(1), \sigma^2(1), \sigma^3(1), \sigma^4(1), \dots$ . Since  $S$  has only  $n$  therefore, there exists a smallest positive integer, say  $r$  such that  $\sigma^r(1) = 1$ . By definition of minimality of  $r$ , the elements  $1, \sigma(1), \sigma^2(1), \sigma^3(1), \dots, \sigma^{r-1}(1)$  are all distinct. This gives us a  $r$ -cycle  $(1, \sigma(1), \sigma^2(1), \sigma^3(1), \dots, \sigma^{r-1}(1))$ . If  $r < n$ , then choose an integer  $a \in S$ , which does not appear in the  $r$ -cycle found above. Now,  $S$  is finite, therefore, there exists a least positive integer  $q$  such that  $\sigma^q(s) = s$ . Now, consider the cycle  $(a, \sigma(a), \sigma^2(a), \dots, \sigma^{q-1}(a)(s))$ . If  $q + r < n$ , then we continue in this way until we have exhausted  $S$ .

- (2) We prove this using induction on  $m$ . Let us take a  $m$ -cycle  $(a_1 a_2 a_3 \dots a_m)$ . Our claim  $(a_1 a_2 a_3 \dots a_m) = (a_1 a_2)(a_2 a_3) \dots (a_{m-1} a_m)$ .

We do induction on  $m$ . For  $m = 3$ , clearly  $(a_1 a_2 a_3) = (a_1 a_2)(a_2 a_3)$  - the product of transpositions.

Now, assume that  $(a_1 a_2 a_3 \dots a_m) = (a_1 a_2)(a_2 a_3) \dots (a_{m-1} a_m)$  and using this we have to prove that  $(a_1 a_2 a_3 \dots a_m a_{m+1}) = (a_1 a_2 a_3 \dots a_{m-1} a_m a_{m+1})$ .

Now,  $(a_1 a_2 a_3 \dots a_m a_{m+1}) = (a_1 a_2 a_3 \dots a_m)(a_m a_{m+1}) = (a_1 a_2)(a_2 a_3) \dots (a_{m-1} a_m)(a_m a_{m+1})$  (by induction hypothesis). This proves our claim.

- (3) In part (2), we proved that  $S_n$  is generated by transposition. Therefore, it is enough to prove that  $(ij)$  can be written as the product of given transpositions. Without loss of generality, assume that  $i < j$ . Then the transposition  $(ij) = (1i)(1j)(1i)$ .
- (4) First, note that  $(1k) = (k-1k) \dots (34)(23)(12)(34)(45) \dots (k-1k)$ . Now, each of these can be simplified using part (3). Hence, the claim follows.  $\square$

**Sign of a permutation.** Due to parts (2) - (4) of Proposition 6.1, we know that each permutation can be written as the product of transpositions in many ways. However, the number of transpositions that appear in such decomposition is always the same. In particular, *a permutation is always a product of an odd or even number of transpositions but can't be both at the same time.*

**Definition 6.1 (Even and odd permutation).** A permutation is even if it can be written as a product of an even number of transpositions and, odd if it can be written as an odd number of transpositions.

**Theorem 6.1.** Every permutation is either odd or even, but not both.

PROOF. Thanks to Proposition 6.1, we know every permutation can be written as a product of transpositions. For the rest of the claim, we introduce the polynomial

$$P = P(x_1, x_2, \dots, x_n) \\ = (x_1 - x_2)(x_1 - x_3) \dots (x_1 - x_n)(x_2 - x_3)(x_3 - x_4) \dots (x_{n-1} - x_n).$$

That is,  $P = \prod_{1 \leq i < j \leq n} (x_i - x_j)$ . Now, let  $\sigma \in S_n$ . Definition  $\sigma P = \prod_{1 \leq i < j \leq n} (x_{\sigma(i)} - x_{\sigma(j)})$ . So, in essence,  $\sigma$  permutes the term of the product  $\prod_{1 \leq i < j \leq n} (x_i - x_j)$  and also possibly changes the sign of some of these terms. Therefore,  $\sigma P$  is either  $+P$  or  $-P$ . We say **sign** of  $\sigma$  is  $+1$  if  $\sigma P = +P$ , and  $-1$  if  $\sigma P = -P$ . Furthermore, if  $\sigma = (i, i+1)$  then  $\sigma P = -P$  because only the pair  $i$  and  $i+1$  get swapped sending  $(x_1 - x_{i+1})$  to  $(x_{i+1} - x_i)$ . From this, we conclude that every permutation is either the product of an even number or an odd number of transpositions but not both at the same time.  $\square$

**Definition 6.2 (Sign and parity of permutation).** Two permutations are said to be of the same parity if they have the same sign, and the sign is defined by

$$(12) \quad \text{sign}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is product of even number of transpositions} \\ -1 & \text{if } \sigma \text{ is the product of odd number of transpositions.} \end{cases}$$

Thus, the  $\text{sign}(a_1 a_2 \dots a_m) = (-1)^m - 1$ . It is easy to see that  $\text{sign}(\sigma\tau) = \text{sign}(\sigma) \text{sign}(\tau)$ .

**Sign of product and inverses.** To determine whether a permutation is even or odd, just look at its cycle structure. To do this, note that we can write a  $k$ -cycle as a product of  $k-1$  transpositions. So, if  $k$  is even, then  $k-1$  would be odd. *That is, a cycle of even length is an odd permutation, and a cycle of odd length is an even permutation.* Suppose that  $\sigma$  and  $\tau$  are the product of  $l$  and  $m$  transposition, then  $\sigma\tau$  would be the product of  $l+m$  transposition. Therefore, the following theorem is now an obvious consequence of this discussion.

**Proposition 6.2 (Sign of product, inverse and identity).** (1) Prove that the result of the multiplication of two permutations of the same parity is an even permutation, whereas the result of the multiplication of two permutations of opposite parities is an odd permutation.

(2) The identity of  $S_n$  is an even permutation.

(3) Let  $a$  be an arbitrary permutation. Prove that  $a$  and  $a^{-1}$  have the same parity.

PROOF. (1) Assume that  $\sigma_1$  and  $\sigma_2$  have the same parity. Then  $\sigma_1 = a_1 a_2 \dots a_m$ . and  $\sigma_2 = b_1 b_2 \dots b_n$  where  $m$  and  $n$  have the same parity. Note that,  $\sigma_1 \circ \sigma_2 = a_1 a_2 \dots a_m b_1 b_2 \dots b_n$ . This is the product of  $m+n$  number of transpositions. As  $m$  and  $n$  have the same parity, we get that  $m+n$  is even. If  $m$  and  $n$  had different parity, then  $m+n$  is odd.

- (2) Let  $n \geq 2$ , because for  $n = 1$  there is nothing to prove. Now, in this set-up,  $(1) = (12)(12)$ . As we discussed earlier, the number of transpositions in any permutation is invariant; hence, whatever

transpositions we choose to write the *identity permutation*, by previous fact, it would always be a product of two transpositions. Hence, the identity is an even permutation.

**Alternate proof.** Assume that the identity,  $e$  of  $S_n$  is an odd permutation. Let  $\sigma \in S_n$  be an odd permutation. Then the product  $\sigma = \sigma e = e\sigma$  would become an even permutation, a contradiction.

- (3) Assume for the sake of contradiction,  $a$  and  $a^{-1}$  have different parity. Then  $aa^{-1} = e$  will be an odd permutation, which is a contradiction of what we just proved in part (2).  $\square$

**Definition 6.3 (Alternating group).** The collection  $A_n = \{\sigma \in S_n \mid \text{sign}(\sigma) = 1\}$  has a group structure due to the fact that  $\text{sign}(\sigma\tau) = \text{sign}(\sigma) \text{sign}(\tau)$ . This group is called an **alternating group** and is a subgroup of  $S_n$ .

Now, we state a very fundamental result. However, this becomes an absolutely obvious result at the end of the next section on group homomorphism.

As we have defined the concept of *normal subgroup*, see definition 1.9, we have the following important result.

**Proposition 6.3.** Prove that the alternating group  $A_n$  is a normal subgroup of the symmetric group  $S_n$  and find the partition of  $S_n$  by  $A_n$ .

PROOF. Let  $g \in G$  and  $n \in A_n$ . To prove  $gng^{-1} \in A_n$ . Note that,  $g$  and  $g^{-1}$  have the same parity. Assume  $g$  is a product of  $m$  transpositions.  $g^{-1}$  is a product of  $n$  transpositions and  $n$  is a product of  $k$  transpositions.  $gng^{-1}$  is a product of  $m+n+k$ . As  $m$  and  $n$  have the same parity and  $k$  is even ( $n \in A_n$ ),  $m+n+k$  is even. Thus,  $gng^{-1} \in A_n$ .

$S_n/A_n = \{gA_n \mid g \in S_n\}$ .  $g \in S_n$ . So  $g \in A_n$  or  $g \in S_n - A_n$ .

- **Case 1:**  $g \in A_n$ , then  $gA_n = A_n$ .
- **Case 2:**  $g \in S_n - A_n$ . **Claim:**  $gA_n$  contains all odd permutations. Let  $b$  be an odd permutation. Note that,  $g(g^{-1}b) = gg^{-1}b = b$ . As  $g^{-1}$  and  $b$  are both odd permutations,  $g^{-1}b$  is an even permutation.  $g^{-1}b \in A_n$ . So  $b \in gA_n$ .  $\square$

**Theorem 6.2.** The map  $\varphi : S_n \rightarrow \{-1, +1\}$  given by  $\sigma \mapsto \text{sign}(\sigma)$  is a group homomorphism in the sense of definition 7.1. The alternating group is the kernel (7.3) of this group homomorphism.

When I was in grade 11, I appeared for the Caltech Math Meet and led my team (Darsan - AMMOC) to a remarkable victory while achieving third place internationally. Moreover, in the overall final ranking, we were placed twelfth even though we missed the final round that was held in person at Caltech. For the sake of completeness, I state a result that was used in CMM, and I believe it is important to record it here.

**Definition 6.4 (Simple group).** A non-trivial group  $G$  is said to be simple if the only normal subgroups of  $G$  are  $\{e\}$  and  $G$  itself.

**Theorem 6.3 (Simplicity of  $A_n$ ).** Prove that for  $n \geq 5$ ,  $A_n$  is a simple group.

PROOF. Proof involves more advanced mathematical machinery and therefore, it is relegated to the last chapter of this work.  $\square$

**Proposition 6.4 (The size of  $A_n$ ).** Prove that the size of  $S_n$  is  $n!$  and the size of  $A_n$  is  $\frac{n!}{2}$ .

PROOF. **The size of  $S_n$ :** Every permutation can be written as

$$P = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & n \\ i_1 & i_2 & i_3 & i_4 & \dots & i_n \end{pmatrix},$$

where  $i_n$  is the image of  $n$  under the permutation. So  $i_1$  has  $n$  choices. As  $i_2 \neq i_1$ ,  $i_2$  has  $n-1$  choices. Similarly  $i_3$  has  $n-2$  choices,  $\dots$ ,  $i_n$  has 1 choice. By multiplication rule no of permutations is  $n \times (n-1) \times \dots \times 1 = n!$ .

**The size of  $A_n$ :** Let  $B_n = S_n - A_n$ . Claim:  $f : A_n \rightarrow B_n$ , where  $f(a) = (12)a$  is a bijection. For injection, let  $f(a) = f(b)$ .  $(12)a = (12)b$ . As the set of all permutations, form a group under the binary operation of composition. Thus, the cancellation law is true. Hence,  $a = b$ . For surjection, for any  $u \in S_n - A_n$  its preimage is  $(12)u$ . As  $f((12)u) = (12)(12)u = u$ . Therefore, it is a bijection and so  $|A_n| = |S(n) - A_n|$ . Thus,  $|A_n| = n! / 2$ .  $\square$

## 7. Homomorphism, Isomorphism, Automorphism

**Definition 7.1 (Homomorphism).** Let  $G, H$  be two groups. Let  $\varphi : G \rightarrow H$  be a set-theoretical function between the underlying sets of two groups. We say  $\varphi$  is a group homomorphism if  $\varphi(a \star b) = \varphi(a) \bullet \varphi(b)$  for each  $a, b \in G$ .

Here, there are three functions in the play. Namely, group operations  $\star, \bullet$  and map  $\varphi$ . The definition of group homomorphism essentially says that the following diagrams commute.

$$\begin{array}{ccc} (a, b) & \xrightarrow{\quad \quad} & (a, b) \\ \downarrow \star_G & & \downarrow \bullet_H \\ a \star b & \xrightarrow{\varphi} & \varphi(a) \bullet \varphi(b) \end{array}$$

7.0.1. *Category Grp.* We invoke the knowledge of Category theory that we have documented in our separate note [here](#).

Let  $G, H$  be groups. Define

$$\text{Hom}_{Grp}(G, H) := \{f : G \rightarrow H \mid f \text{ is a group homomorphism}\}.$$

If  $(G, \star), (H, \bullet), (K, \Delta)$  are groups and  $f : G \rightarrow H, g : H \rightarrow K$  are group homomorphisms between respective groups. We note that the following

- $g \circ f$  is a set-theoretical map.
- $g \circ f$  respects the underlying group structures. That is,  $g \circ f$  is a group homomorphism. This can be conveyed via the following commutative diagram.

$$\begin{array}{ccccc} & & (g \circ f) \times (g \circ f) & & \\ & \searrow & \downarrow & \searrow & \\ G \times G & \xrightarrow{f \times f} & H \times H & \xrightarrow{g \times g} & K \times K \\ \downarrow \star & & \downarrow \bullet & & \downarrow \Delta \\ G & \xrightarrow{f} & H & \xrightarrow{g} & K \\ & \searrow & \downarrow & \searrow & \\ & & g \circ f & & \end{array}$$

- It is easy to verify that the composition of group homomorphisms respects associative law.

**Proposition 7.1 (Examples of homomorphism).** Let  $G, H$  be two groups.

- The logarithm and exponential functions

$$\log : \mathbb{R}^+ \rightarrow \mathbb{R} \quad \text{and} \quad \exp : \mathbb{R} \rightarrow \mathbb{R}^+$$

are homomorphism between the underlying additive group  $(\mathbb{R}, +)$ .

- The absolute value

$$\begin{aligned} \mathbb{C}^\times &\rightarrow \mathbb{R}^+ \\ z &\mapsto |z| \end{aligned}$$

is a morphism between the underlying multiplicative groups.

- The determinant map

$$GL_n(\mathbb{R}) \rightarrow \mathbb{R}^+ \\ M \mapsto \det(M)$$

is a group homomorphism of the underlying multiplicative group.

- Let  $C$  be an additive group of real-valued continuous function on the compact set  $[0, 1]$ . Define the map  $C \rightarrow \mathbb{R}$  given by  $f \mapsto \int_0^1 f(t)dt$  is a group homomorphism between the underlying additive group.
- The map  $\varphi : G \rightarrow G$  given by  $g \mapsto g^2$  is a group homomorphism if and only if  $G$  is abelian.
- The map  $\varphi : G \rightarrow G$  defined by  $g \mapsto g^{-1}$  is a group homomorphism if and only if  $G$  is abelian.
- Any group homomorphism  $\varphi : \mathbb{Z} \rightarrow G$  is uniquely determined by the value of  $\varphi(1)$  in  $G$ . This follows from the fact that in the additive abelian group  $\mathbb{Z}$ ,  $\varphi(n) = \varphi(\sum_{i=1}^n 1) = \varphi(1)$ . Hence there is a one-to-one correspondence  $\Psi : \text{Hom}(\mathbb{Z}, G) \rightarrow G$  given by  $\varphi \mapsto \varphi(1)$ .

**Definition 7.2 (Isomorphism).** Let  $\varphi : G \rightarrow H$  be a group homomorphism. We say that  $\varphi$  is an isomorphism if  $\varphi$  is a bijective homomorphism.

Intuitively, we can take isomorphism to be ‘relabelling’ of elements of  $G$ . That is, if  $\varphi : G \rightarrow H$  is a group isomorphism, then  $g \in G$  gets relabelled as  $\varphi(g) \in H$ . So, any property that depends on the ‘underlying group axioms’ alone would remain invariant (up to relabelling) under any isomorphism.

In Theorem 7.1 and Theorem 7.2, we see that an isomorphism between two groups is a dictionary that allows us to translate elements and operations from one group to the other without losing any essential information. For example, we will see in Proposition 7.4 that all cyclic groups of the same size are isomorphic. This essentially means that if we understand one cyclic group, then we can transfer that understanding to all the other cyclic groups of the same size.

**Proposition 7.2 (Inverse of an isomorphism).** Let  $\varphi : G \rightarrow H$  be a group isomorphism. Then  $\varphi^{-1}$  is also a group isomorphism.

PROOF. We first prove that  $\varphi^{-1}$  is a homomorphism. Let  $x, y \in H$  and  $a, b \in G$  such that  $\varphi^{-1}(x) = a \implies \varphi(a) = x$  and  $\varphi^{-1}(y) = b \implies \varphi(b) = y$ . Now,  $\varphi^{-1}(xy) = \varphi^{-1}(\varphi(a)\varphi(b)) = \varphi^{-1}(\varphi(ab)) = (\varphi^{-1} \circ \varphi)(ab) = ab\varphi^{-1}(x)\varphi^{-1}(y)$ . Therefore,  $\varphi^{-1}$  is indeed a group homomorphism. Now, to prove that  $\varphi^{-1}$  is *injective*. Let  $\varphi^{-1}(x) = a = \varphi^{-1}(y) = b$ . But  $\varphi$  is a function, therefore,  $\varphi(a) = \varphi(b) \implies x = y$ . To prove that  $\varphi^{-1}$  is *surjective*, let  $a \in G$ . Then  $\varphi(a) = x \in H \implies \varphi^{-1}(x) = a$ . Thus,  $\varphi^{-1}$  is a surjection. Hence, the inverse of an isomorphism is also an isomorphism.  $\square$

**Proposition 7.3 (Composition of isomorphisms).** Let  $\phi_1 : G_1 \rightarrow G_2$  and  $\phi_2 : G_2 \rightarrow G_3$  be two isomorphisms. Prove that the composite mapping  $\phi_2 \circ \phi_1 : G_1 \rightarrow G_3$  is also an isomorphism.

PROOF. We prove that  $\phi_3$  is injective, where  $\phi_3 = \phi_2 \circ \phi_1$ . Consider  $\phi_3(a) = \phi_3(b)$  for some  $a, b \in G_1$ . This gives  $\phi_2 \circ \phi_1(a) = \phi_2 \circ \phi_1(b)$ . Since  $\phi_2$  is injective, therefore, previous equality gives  $\phi_1(a) = \phi_1(b)$ . Further, as even  $\phi_1$  is an injective map, therefore, the last equation yields  $a = b$ . This proves that  $\phi_2 \circ \phi_1$  is injective.

Now, we prove that  $\phi_3$  is surjective. Let  $y \in G_3$ . Since  $\phi_2$  is a surjection, therefore, there exists an  $x' \in G_2$  such that  $\phi_2(x') = y$ . Since  $\phi_1$  is surjection, there exists an  $x \in G_1$  such that  $x' = \phi_1(x)$ . Hence  $\phi_2(x') = \phi_2(\phi_1(x)) = y$ . This proves the surjectivity of  $\phi_3 = \phi_2 \circ \phi_1$ . Now, we prove  $\phi_3$  is a homomorphism. Consider  $*$ ,  $\bullet$ ,  $\star$  to be binary operations on  $G_1, G_2, G_3$ , respectively.

Next, we prove that

$$\phi_3(a * b) = \phi_3(a) \star \phi_3(b).$$

As  $\phi_1$  is a isomorphism  $\phi_1(a * b) = \phi_1(a) \bullet \phi_1(b)$ , we have  $\phi_2(\phi_1(a * b)) = \phi_2(\phi_1(a) \bullet \phi_1(b)) = \phi_2(\phi_1(a)) * \phi_2(\phi_1(b)) = \phi_2 \circ \phi_1(a) = \phi_2 \circ \phi_1(a) \star \phi_2 \circ \phi_1(b)$ . Thus,  $\phi_3$  is an isomorphism.  $\square$

**Corollary 7.1 (Aut(G)).** Let  $G$  be a group. An isomorphism of  $G$  onto itself is known as an **automorphism** of  $G$ . Prove that collection  $\text{Aut}(G)$  of all automorphisms of  $G$  is a group under the usual composition of functions.

PROOF. This follows easily from previous two propositions.  $\square$

**Theorem 7.1 (Group isomorphism and order).** let  $\phi : (G, \star) \rightarrow (H, \bullet)$  be a group isomorphism. Then the following hold.

- (1)  $\phi(e_G) = e_H$ .
- (2)  $\phi(g^m) = \phi(g)^m$  for all  $m \in \mathbb{N}$ .
- (3)  $\phi(g^{-1}) = \phi(g)^{-1}$  for all  $g \in G$  and  $m \in \mathbb{Z}$ .
- (4) Using part (2), (3), prove that  $\phi(g^k) = \phi(g)^k$  for all  $k \in \mathbb{Z}$ .
- (5)  $|g| = |\phi(g)|$ .
- (6) Two isomorphic groups have the same number of elements of the same order.

PROOF. (1) Note that,  $\phi(e_G) = \phi(e_G \star e_G) = \phi(e_G) \bullet \phi(e_G)$ . Now, multiply both sides by  $\phi(e_G)^{-1}$ .  
 $\phi(e_G) \bullet \phi(e_G)^{-1} = \phi(e_G) \bullet \phi(e_G) \bullet \phi(e_G)^{-1}$ . Now, apply associativity to get  $e_H = \phi(e_G) \bullet e_H$ .  
 Thus,  $\phi(e_G) = e_H$ .

- (2) This is an easy consequence of the definition of group homomorphism and the Principle of mathematical induction.
- (3) We have to prove that  $\phi(g^{-1}) = \phi(g)^{-1}$ . We know that  $\phi(g \star g^{-1}) = \phi(e_G) = e_H$ . As  $\phi$  is an isomorphism we have  $\phi(g) \bullet \phi(g^{-1}) = \phi(g \star g^{-1}) = e_H$ . We do left multiplication on both sides by  $\phi(g)^{-1}$ , use associativity and arrive at  $(\phi(g)^{-1} \bullet \phi(g)) \bullet \phi(g^{-1}) = \phi(g)^{-1} \bullet e_H \implies e_H \bullet \phi(g^{-1}) = \phi(g)^{-1} \bullet e_H \implies \phi(g^{-1}) = \phi(g)^{-1}$ .
- (4) For positive integers, it is true due to part (2). For negative integers, it follows from part (3) and the definition of a group homomorphism.
- (5) Let  $m$  and  $n$  be the orders of  $g$  and  $\phi(g)$ , respectively. let  $\phi(g) = h$ .  
 Thus, we have  $g^m = e_G$  and  $h^n = \phi(g)^n = e_H$  where  $\phi(g) = h$ . As  $\phi$  is an isomorphism and hence a homomorphism, therefore,

$$e_H = \phi(e_G) = \phi(g^m) = \underbrace{\phi(g) \bullet \dots \bullet \phi(g)}_{m \text{ times}} = \phi(g)^m = h^m.$$

Since order is the smallest positive integer, therefore,  $m \geq n = |h = \phi(g)|$ . Moreover, since  $n$  is the order of  $h$ , therefore,

$$e_H = h^n = \phi(g)^n = \phi(g^n).$$

However,  $\phi$  is an injection (by virtue of being an isomorphism); hence, only one element  $e_G$  can be mapped to  $e_H$ . This simply means that  $g^n$  must be an identity  $e_G$ . This means that  $n \geq m$ . Combining with the previous argument, we conclude that  $m = n$ .

- (6) Suppose that  $G$  and  $H$  are isomorphic groups with isomorphism  $\varphi$ . Since under the isomorphism, we have  $|g| = |\varphi(g)|$ . If  $G$  group has exactly  $m$  elements,  $g_1, g_2, \dots, g_m$  of order  $k$ , then images of these elements would also have order  $k$ , and if  $H$  has an additional element  $x \notin \{\varphi(g_i) \mid i = 1, 2, \dots, m\}$  of order  $k$ , then  $\varphi^{-1}(x)$  will also be an element of order  $k$ , contradicting that  $G$  has exactly  $m$  elements of order  $k$ .  $\square$

**Theorem 7.2 (Isomorphism preserves commutativity and cyclicity).** If  $\varphi : G \rightarrow H$  is an isomorphism, prove that  $G$  is abelian if and only if  $H$  is abelian. Also, prove that  $G$  is cyclic if and only if  $H$  is cyclic.

PROOF. First, assume that  $G$  is abelian. Let  $h_1, h_2$  be two arbitrary elements of  $H$ . As  $\varphi$  is bijective and hence surjective, therefore, there exists  $g_1, g_2$  such that  $\varphi(g_1) = h_1$  and  $\varphi(g_2) = h_2$ . Now, consider

$$h_1 h_2 = \varphi(g_1) \varphi(g_2) = \varphi(g_1 g_2) = \varphi(g_2 g_1) = \varphi(g_2) \varphi(g_1) = h_2 h_1.$$

Thus,  $H$  is an abelian group. Now, we assume that  $H$  is an abelian group. Let  $g_1, g_2$  be two arbitrary elements of  $G$ . Thus,  $\varphi(g_1) \in H, \varphi(g_2) \in H$ . Since  $H$  is abelian, we have  $\varphi(g_1) \varphi(g_2) = \varphi(g_2) \varphi(g_1)$ . So  $\varphi(g_1 g_2) = \varphi(g_2 g_1)$ . However, as  $\varphi$  is injective, therefore,  $\varphi(g_1 g_2) = \varphi(g_2 g_1) \implies g_1 g_2 = g_2 g_1$ . To prove that isomorphism preserves the cyclicity, let  $\varphi : G \rightarrow H$  be an isomorphism. Let us assume that  $G = \langle g \rangle$ . Let  $x \in H$ . Since  $\varphi : G \rightarrow H$  is a surjection, therefore, there exists a positive integer  $k$  such

that  $\varphi(g^k) = (\varphi(g))^k = x$ . So each  $x \in H$  is some power of  $\varphi(g)$ . That is,  $H = \langle \varphi(g) \rangle$ . From this, we conclude that an isomorphism sends a generator to a generator. Now, assume that  $H$  is a cyclic group generated by  $x$ . Let  $a \in G$ . Then as  $\varphi^{-1}$  is also an isomorphism, so for a  $g \in G$ , there exist some  $i \in \mathbb{Z}$  such that  $\varphi^{-1}(x^i) = (\varphi^{-1}(x))^i = g$ . So, each element in  $G$  is some power of  $\varphi^{-1}(x)$ . Thus,  $G$  is cyclic with the generator  $\varphi^{-1}(x)$ . **Isomorphism is a property between two groups and not a property of an individual group, like being cyclic or commutative.** *The Concept of isomorphism is fundamental to the classification problem of the group, which deals with questions like 'find all groups satisfying so and so properties.'* As we say, up to isomorphism, every cyclic group of finite order is isomorphic to  $\mathbb{Z}_n$ . Up to isomorphism, every non-abelian group of order 6 is isomorphic to  $S_3$ , the group of permutations of 3 letters.  $\square$

**Proposition 7.4 (Characterization of finite cyclic group).** Prove that every cyclic group of order  $n$  is isomorphic to additive group  $\mathbb{Z}_n = \{[0]_n, [1]_n, \dots, [n-1]_n\}$  of integer modulo  $n$ .

PROOF. Let  $G_n = \{e, a, a^2, a^3, \dots, a^{n-1}\}$  be a finite cyclic group of order  $n$  with  $a$  being its generator. Let  $\mathcal{G} = \{[0]_n, [1]_n, \dots, [n-1]_n\}$  be the additive group of remainders modulo  $n$ , wherein addition is defined as  $[a]_n + [b]_n = [a+b]_n$ . Let  $\varphi: G_n \rightarrow \mathbb{Z}_n$  defined as  $\varphi(a^m) = [m]_n$  - the congruence class of integer  $m$ . This is a well-defined map. we prove  $\varphi$  is injective. Let  $\varphi(a^i) = \varphi(a^j)$  for some  $1 \leq i, j \leq n$ . By definition,  $[i]_n = [j]_n$ . But note that  $1 \leq i, j \leq n$ , and therefore,  $[i]_n = [j]_n \implies i = j$ . Now, for surjection, for a given congruence class  $[k]_n$ , consider the element  $a^k$  then  $\varphi(a^k) = [k]_n$ . So, this map is a surjection. Now, we prove  $\varphi$  is a homomorphism. To do this,  $\varphi(a^i \times a^j) = \varphi(a^{i+j}) = [i+j]_n = [i]_n + [j]_n = \varphi[a^i] + \varphi[j^i]$ . This proves that  $\mathbb{Z}_n \simeq G_n$ .  $\square$

**Corollary 7.2.** For a prime  $p$ , every group of order  $p$  is isomorphic to  $\mathbb{Z}_p$ .

PROOF. Let  $G$  be a group of the prime order  $p$ . By Lagrange theorem, we know that the order of an element divides the order of group  $G$ . So, if  $g$  is non-identity element of  $G$ , then  $|g| = p$ . Hence,  $G = \langle g \rangle$ . By the previous proposition,  $G \simeq \mathbb{Z}_p$ .  $\square$

**Proposition 7.5 (Characterization of infinite cyclic group).** Every infinite cyclic group is isomorphic to  $\mathbb{Z}$ , the group of integers under addition.

PROOF. Let  $G = \langle g \rangle$ . Define a map  $f: \mathbb{Z} \rightarrow G$  by  $n \mapsto g^n$ . We claim  $f$  is a group isomorphism. Let  $n \neq m$ . Without loss of generality, let  $n < m$  and  $g^n = g^m$ . Then  $g^{m-n} = 1$  and  $m-n$  is a positive integer. Therefore,  $g$  is an element of finite order. This contradicts that  $G$  is an infinite group. Therefore,  $f$  is an injection. It is obvious that this map is a surjection because every element of  $G$  is of the form  $g^k$  for some  $k \in \mathbb{Z}$ . It is easy to see that  $f$  is a homomorphism. Thus,  $\mathbb{Z} \simeq G$ .  $\square$

**Definition 7.3 (Kernel).** Let  $\varphi: G \rightarrow H$  be a group homomorphism. The set  $\{g \in G \mid \varphi(g) = e_H\}$  is known as kernel of the given group homomorphism,  $\varphi$ . We denote this set as  $\text{Ker}\varphi$ .

**Lemma 7.1.** Prove the following

- (1) Prove that  $\text{ker}\varphi$  is a subgroup of group  $H$ .
- (2) Prove that  $\text{Im}\varphi$  is a subgroup of  $H$ .
- (3)  $\varphi$  is an injection iff  $\text{ker}\varphi = \{e_G\}$ .

PROOF. (1) First, note that  $\varphi(e_G) = e_H$ . So, the kernel of any group homomorphism is always non-empty. From Theorem 7.1, we know that  $\varphi(g^{-1}) = \varphi(g)^{-1}$ . Therefore, if  $\varphi(g) = e_H$ , then  $\varphi(g^{-1}) = \varphi(g)^{-1} = e_H^{-1} = e_H$ . Also, if  $a, b \in \text{ker}\varphi$ , then  $\varphi(ab) = \varphi(a)\varphi(b) = e_H$ . Hence,  $\text{ker}\varphi$  is indeed a subgroup of  $G$ .

- (2) First, note that  $e_H = \varphi(e_G) \in \text{Im}\varphi$ . Let  $h \in \text{Im}\varphi$ . Therefore, there exists a  $g \in G$  such that  $\varphi(g) = h$ . Since  $g \in G$ , therefore,  $g^{-1} \in G$ . Hence,  $\varphi(g^{-1}) = \varphi(g)^{-1} \in \text{Im}\varphi$ . Let  $h_1, h_2 \in \text{Im}\varphi$ . Therefore, there exist  $g_1, g_2 \in G$  such that  $\varphi(g_1) = h_1$  and  $\varphi(g_2) = h_2$ . Since,  $h_1 h_2 = \varphi(g_1)\varphi(g_2) = \varphi(g_1 g_2) \in \text{Im}\varphi$ . Hence,  $h_1 h_2 \in \varphi$ . This completes the proof.

- (3) Assume that  $\varphi$  is an injection. Let  $g \in \ker \varphi$ . Then  $g = e_G$ , otherwise it will contradict the injectivity of  $\varphi$ . To prove the converse, assume that  $\ker \varphi$  is the singleton set  $\{e_G\}$ . Let  $\varphi(g_1) = \varphi(g_2)$ . This gives  $\varphi(g_1)\varphi(g_2)^{-1} = \varphi(g_1g_2^{-1}) = e_G \implies g_1g_2^{-1} = e_G \iff g_1 = g_2$ . Thus,  $\varphi$  is indeed an injection.  $\square$

**Problem 7.1.** Prove that the multiplicative groups  $\mathbb{R} - \{0\}$  and  $\mathbb{C} - \{0\}$  are not isomorphic.

**Solution 7.1.** By Theorem 7.1, we know that two isomorphic groups share an exact set of algebraic properties, with just one element being relabelled. In particular, by part (6) of Theorem 7.1, two isomorphic groups have the same number of elements of a given order. Now, note that the order of  $i$  is 4. However, every element in  $\mathbb{R} - \{0\}$  has infinite order apart from  $-1, 1$  which have order 1 and 2 respectively. So, no element of  $\mathbb{R} - \{0\}$  has order 4. Therefore, these two multiplicative groups cannot be isomorphic.

**Problem 7.2.** Prove that additive groups  $\mathbb{R}$  and  $\mathbb{Q}$  are not isomorphic.

**Solution 7.2.** An isomorphism is a bijection that preserves the group operations of underlying groups. So, in particular, any two isomorphic groups must have the same cardinality. Using Cantor's countability argument, we know that there can not exist any bijection between  $\mathbb{Q}$  and  $\mathbb{R}$ . The cardinality of  $\mathbb{Q}$  is a countable infinite, and the cardinality of  $\mathbb{R}$  is an uncountable infinite. There are different sizes of infinities. 'No one can expect us from the paradise that Cantor's created' remarked the last universal mathematician David Hilbert.



FIGURE 11. David Hilbert

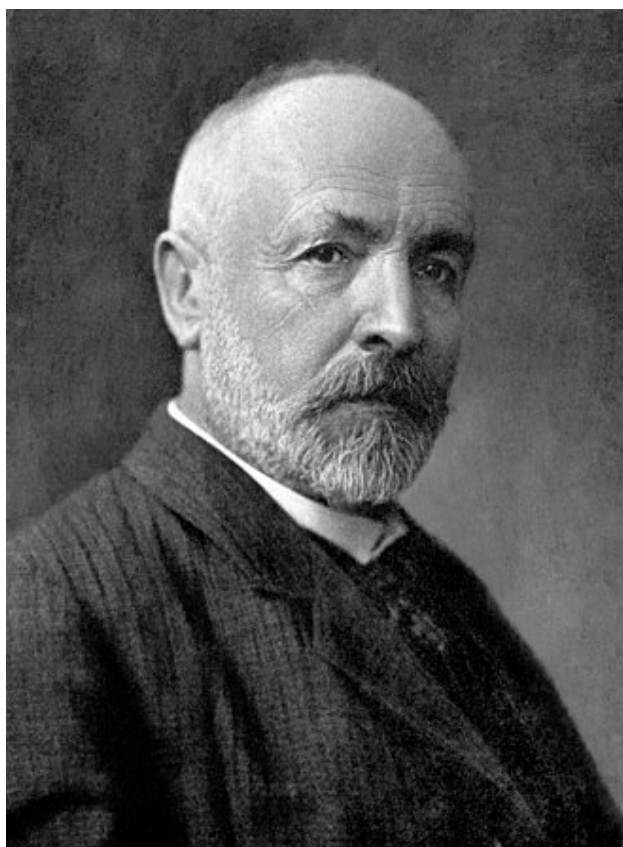


FIGURE 12. George Cantor

**Problem 7.3.** Prove that  $Q_8$  and  $D_8$  are not isomorphic.

**Solution 7.3.** Again, we use part (6) of Theorem 7.1. First, note that 1 has order 1 and  $-1$  has order 2.  $i, j, k$  each have order 4 also  $-i, -j, -k$  have order 4. Now, in  $D_8$ , there are two types of elements. There

are 4 rotational elements, and each one of them has an order of 4. So,  $D_8$  has 4 elements with order 4. However,  $Q_8$  has 6 elements with order 4. Thus, they are not isomorphic.

**Problem 7.4.** Let  $G$  be any group. Prove that the map from  $G$  to itself defined by  $g \mapsto g^{-1}$  is a homomorphism if and only if  $G$  is abelian.

**Solution 7.4.** Suppose  $G$  is abelian. To prove the map  $\varphi : g \mapsto g^{-1}$  is a homomorphism. Now,  $\varphi(g_1g_2) = (g_1g_2)^{-1} = g_2^{-1}g_1^{-1}$ . As  $G$  is abelian,  $g_2^{-1}g_1^{-1} = g_1^{-1}g_2^{-1} = \varphi(g_1)\varphi(g_2)$ . So we get  $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$ . Suppose  $\varphi : g \mapsto g^{-1}$  is a homomorphism. Let  $g_1, g_2$  be arbitrary elements of  $G$ . So we have  $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$ . Thus, we get  $(g_1g_2)^{-1} = g_1^{-1}g_2^{-1}$ . So by the definition of an inverse,  $g_1^{-1}g_2^{-1}g_1g_2 = e$ . Multiply by  $g_1$  on left side to get  $g_2^{-1}g_1g_2 = g_1$ . Multiply by  $g_2$  on the left side to get  $g_1g_2 = g_2g_1$ .

**Problem 7.5.** Let  $G$  be any group. Prove that the map from  $G$  to itself defined by  $g \mapsto g^2$  is a homomorphism if and only if  $G$  is abelian.

**Solution 7.5.** Suppose  $G$  is abelian. To prove the map  $\varphi : g \mapsto g^2$  is a homomorphism. Let  $g_1, g_2$  be arbitrary elements of  $G$ .  $\varphi(g_1g_2) = (g_1g_2)^2 = (g_1g_2)(g_1g_2)$ . As  $G$  is abelian, we have

$$(g_1g_2)(g_1g_2) = (g_1g_2g_1)(g_2) = (g_1g_1)(g_2g_2)g_1^2g_2^2 = \varphi(g_1)\varphi(g_2).$$

Suppose the map  $\varphi : g \mapsto g^2$  is a homomorphism. Let  $g_1, g_2$  be arbitrary elements of  $G$ . Then  $\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$ . So

$$\begin{aligned} (g_1g_2)^2 &= g_1^2g_2^2 = g_1g_1g_2g_2 \\ g_1g_2g_1g_2 &= g_1g_1g_2g_2. \end{aligned}$$

By Cancellation law, we can cancel  $g_1$  on the left side and  $g_2$  on the right-hand side to get  $g_2g_1 = g_1g_2$ . i.e  $G$  is abelian.

**Problem 7.6.** Let  $G = \{z \in \mathbb{C} \mid z^n = 1 \text{ for some } n \in \mathbb{Z}^+\}$ . Prove that for any fixed integer  $k > 1$ , the map from  $G$  to itself defined by  $z \mapsto z^k$  is a surjective group homomorphism but is not an isomorphism.

**Solution 7.6.** Note that,  $G \subseteq \mathbb{C}$ . As complex Multiplication is commutative,  $G$  is abelian. First, we prove that the given map is indeed surjective. Let  $z \in G$  be an arbitrary element of  $G$ . So we have  $z^n = 1$ , and Therefore,  $(z^{-1})^n = 1$ . First, we claim that  $\sqrt[k]{z}$  exists in  $G$ .

As  $(\sqrt[k]{z})^n = (\sqrt[n]{z})^k = ((z^{-1})^n)^k = 1^k = 1$ , and

$$\varphi(\sqrt[k]{z}) = z.$$

So every element of  $G$  has a preimage  $(\sqrt[k]{z})$  under  $\varphi$ . For homomorphism, we have

$$\varphi(z_1z_2) = (z_1z_2)^k = z_1^kz_2^k = \varphi(z_1)\varphi(z_2).$$

**Problem 7.7.** Prove that for each fixed nonzero  $k \in \mathbb{Q}$  the map from  $\mathbb{Q}$  to itself defined by  $q \mapsto kq$  is an automorphism of  $\mathbb{Q}$ .

**Solution 7.7.** Denote the map as  $\varphi : \mathbb{Q} \rightarrow \mathbb{Q}$  such that  $\varphi(q) = kq$ . It is clearly injective. For surjection, take an arbitrary element  $q_1 \in \mathbb{Q}$ , its preimage is clearly  $\frac{q_1}{k}$  as  $\varphi(\frac{q_1}{k}) = k\frac{q_1}{k} = q_1$ . Note that,

$$\varphi(q_1 + q_2) = k(q_1 + q_2) = kq_1 + kq_2 = \varphi(q_1) + \varphi(q_2).$$

Hence,  $\varphi$  is a homomorphism. Thus,  $\varphi$  is an automorphism.

**Problem 7.8.** Let  $G$  be a finite group which possesses an automorphism  $\sigma$  such that  $\sigma(g) = g$  if and only if  $g = 1$ . If  $\sigma^2$  is the identity map from  $G \rightarrow G$ , prove that  $G$  is abelian (such an automorphism  $\sigma$  is called a fixed point free of order 2).

**Solution 7.8.** First, we show that every element  $g \in G$  can be written as  $x^{-1}\sigma(x)$ .

Given that we now know the concept of the kernel of a group homomorphism, we can introduce some subgroups of the general linear group.

## 8. The Special orthogonal and unitary group

**Definition 8.1 (Orthogonal square matrices).** A square matrix  $A$  is an *orthogonal matrix* if its inverse is given by its transpose,  $A^T$ . Now, suppose that  $A, B$  are orthogonal matrices. The following properties are straightforward deductions.

- $I_n^T = I_n = I_n^{-1}$ .
- $(AB)^{-1} = B^{-1}A^{-1} = B^T A^T = (AB)^T$ .
- If  $A$  is orthogonal then so is  $A^{-1}$ .

**Definition 8.2 (Real orthogonal group).** The collection

$$O(n, \mathbb{R}) = \{A \in GL_n(\mathbb{R}) \mid AA^T = A^T A = I\}$$

is a subgroup of  $GL_n(\mathbb{R})$  and is called *real orthogonal group*. If  $\mathbb{R}$  is replaced by any arbitrary field  $\mathbb{F}$ , then we say *special orthogonal group* over a field  $\mathbb{F}$ .

Since for orthogonal matrices we have  $AA^T = A^A = I \implies \det(A^T A) = \det(AA^T) = \det(A)^2 = \det(I_n) = 1 \implies \det(A) = \pm 1$ . If  $\det(A) = 1$ , then we say that  $A$  is a *orientation preservation isometry*. If  $\det(A) = -1$ , then we say  $A$  is a *orientation reversing isometry*.

**Definition 8.3 (Special orthogonal group).** The kernel of group homomorphism  $\det : O(n, \mathbb{R}) \rightarrow -1, +1$  is known as *special orthogonal group*, it is a subgroup of  $O(n, \mathbb{R})$ , and is denoted by  $SO(n, \mathbb{R})$ .

**Definition 8.4 (SO(2) and SU(2)).**

$$SO_{2 \times 2} = \left\{ \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \mid \theta \in \mathbb{R} \right\}$$

is subgroup of general linear group  $GL_2(\mathbb{R})$ , and is known as group of rotation in the dimension 2.

$$SO_{2 \times 2} = \left\{ \begin{bmatrix} u & -\bar{v} \\ v & \bar{u} \end{bmatrix} \mid u, v \in \mathbb{C}, |u|^2 + |v|^2 = 1 \right\}$$

is a subgroup of general linear group  $GL_2(\mathbb{C})$  of  $2 \times 2$  invertible square matrices over the field of complex numbers. This subgroup is known as *special unitary group*.

In the next section, we would study the *direct product* of groups and then use them further to study the group of *platonic solids*.

## 9. Direct Product, Polyhedra, and Platonic solids

**9.1. Direct product of groups.** Let  $(G, \star)$  and  $(H, \bullet)$  be two groups. We construct a new group  $G \times H$  whose elements are those of the cartesian product of the sets of elements of the two underlying original groups. The group operation  $\Delta$ , is defined by:

$$(g, h)\Delta(g', h') := (g \star g', h \bullet h')$$

**Theorem 9.1 (Properties of direct product).** (1) Prove that  $G \times H$  is indeed a group.

- (2) If  $|G| = m$  and  $|H| = n$  then  $|G \times H| = mn$ .
- (3)  $G \times H \simeq H \times G$ .
- (4)  $G$  is embedded in  $G \times H$  by the map  $g \mapsto (g, e_H)$  and same holds true for  $H$ .
- (5) If  $G$  and  $H$  are abelian group then so is their direct product  $G \times H$ .
- (6) If  $G_1 \leq G$  and  $H_1 \leq H$  then  $G_1 \times H_1 \leq G \times H$ .
- (7) Let  $H, K$  be subgroup of  $G$ . Define  $(HK) = \{hk \mid h \in H, k \in K\}$ . If  $HK = G$ ,  $H \cap K = \{e\}$  and if  $hk = kh$  for all  $h \in H, k \in K$ , then  $G \simeq H \times K$ .

PROOF. (1) Note that, by the definition of the binary operation

$$(g, h)\Delta(g', h') := (g \star g', h \bullet h')$$

on  $G \times H$ , it is clear that  $(e_G, e_H)$  is its identity. Further,  $(g^{-1}, h^{-1})$  is the inverse of  $(g, h)$ . Associativity of  $\Delta$  is inherited from respective operations on  $G$  and  $H$ .

- (2) Basically, for each  $g \in G$ , there are total  $n$  choices of elements in  $H$  to make an order tuple with  $g$ . Hence, the claim follows.
- (3) Let  $\varphi : G \times H \rightarrow H \times G$  be a map given by  $(g, h) \mapsto (h, g)$ . By its very definition, it is clear that it is a bijection. Note that,

$$\begin{aligned}\varphi((g_1, h_1)\Delta(g_2, h_2)) &= \varphi((g_1 \cdot g_2, h_1 \cdot h_2)) \\ &= (h_1 \cdot h_2, g_1 \cdot g_2)\end{aligned}$$

and

$$\begin{aligned}\varphi((g_1, h_1)) \cdot \varphi((g_2, h_2)) &= (h_1, g_1) \cdot (h_2, g_2) \\ &= (h_1 \cdot h_2, g_1 \cdot g_2)\end{aligned}$$

Hence,  $\varphi((g_1, h_1) \cdot (g_2, h_2)) = \varphi((g_1, h_1)) \cdot \varphi((g_2, h_2))$ .

- (4) Define a map  $\varphi : G \rightarrow G \times H$  given by  $g \mapsto (g, e_H)$ . This map is a bijection. We just prove that it is a homomorphism:  $\varphi(a \star b) = (a \star b, e_H) = (a \star b, e_H \bullet e_H) = (a, e_H)\Delta(b, e_H) = \varphi(a)\varphi(b)$ .
- (5) Let  $(g_1, h_1)$  and  $(g_2, h_2)$  be the two elements of  $G \times H$ . Then  $(g_1, h_1)\Delta(g_2, h_2) = (g_1 \star g_2, h_1 \bullet h_2) = (g_2 \star g_1, h_2 \bullet h_1) = (g_2, h_2)\Delta(g_1, h_1)$  wherein the second last equality holds true because  $G$  and  $H$  are abelian. Hence,  $G \times H$  is abelian.
- (6) Let  $(g_1, h_1)$  and  $(g_2, h_2)$  be the two elements of  $G_1 \times H_1$ . Then  $(g_1, h_1)\Delta(g_2, h_2) = (g_1 \star g_2, h_1 \bullet h_2)$ . As  $g_1 \star g_2 \in G_1$  and  $h_1 \bullet h_2 \in H_1$ . Thus,  $(g_1, h_1)\Delta(g_2, h_2) \in G_1 \times H_1$ . Notice that  $(e_G, e_H)$  is in  $G_1 \times H_1$  because  $e_G \in G_1$  and  $e_H \in H_1$ . Let  $(g_1, h_1) \in G_1 \times H_1$ . Then it is easy to see that  $(g_1^{-1}, h_1^{-1})$  is the inverse of  $(g_1, h_1)$ , in  $G \times H$ , with respect to the operation  $\Delta$  on  $G \times H$ . However, as  $g_1^{-1} \in G_1$  and  $h_1^{-1} \in H_1$ , therefore,  $(g_1^{-1}, h_1^{-1}) \in G_1 \times H_1$ . Thus,  $G_1 \times H_1$  is the subgroup of  $G \times H$ .
- (7) Define a map  $\varphi : H \times K \rightarrow G$  by  $(h, k) \mapsto hk$ . By definition of the direct product and given hypothesis on the commutativity of product  $hk$ , we get  $\varphi((g, h), (g', h')) = \varphi(gg', hh') = gg'hh' = ghg'h' = \varphi(g, h)\varphi(g', h')$ . So,  $\varphi$  is a group homomorphism. Since  $HK = G$ , therefore, each element of  $G$  is written as the product of elements in  $H$  and  $K$ . Hence,  $\varphi$  is a surjective map. Now, suppose that  $\varphi(g, h) = \varphi(g', h')$ . Then,  $gh = g'h' \implies (g')^{-1}gg = h'(h)^{-1}$ . However,  $H \cap K = \{e\}$ . Therefore,  $(g')^{-1}gg = h'(h)^{-1} = e \iff x = x', y = y'$ . This proves that  $\varphi$  is an injective map.

□

**9.2. Group of Platonic solids.** A *Platonic Solid* is a regular, convex polyhedron with congruent faces, edges, and angles. In this section, we study four of five platonic solids - the tetrahedron (four triangular faces), the octahedron (eight triangular faces), the dodecahedron (twelve pentagonal faces), and the icosahedron (twenty triangular faces). We will denote the group of rotational symmetry of the solid as  $R$  and the group of reflectional symmetry of the solid as  $S$ .

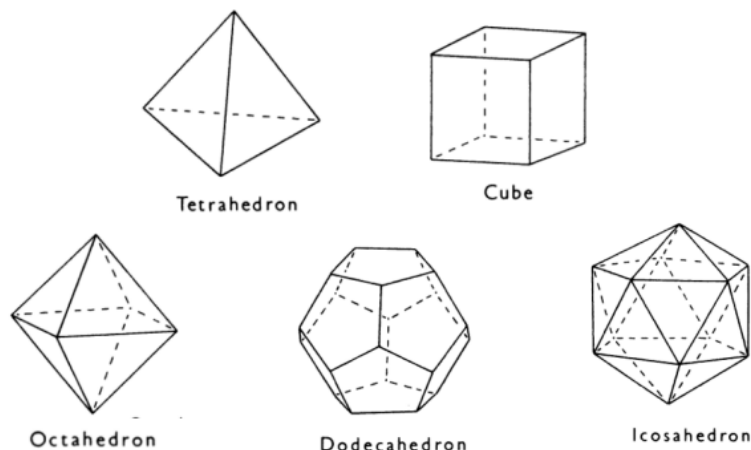


FIGURE 13. The five Platonic Solids

**Definition 9.1 (Dual polyhedra).** The dual polyhedron of a regular solid is another polyhedron such that the faces and vertices of the two occupy complementary locations. This can be done by connecting the center of each face of the solid to that of adjacent faces. Dual of the dual is the original platonic solid. Tetrahedron is its own dual. Cube and octahedron are dual. By very construction of dual, two dual solid have the same set symmetry groups.

**Theorem 9.2.** (1) Show that the group of rigid motions of a tetrahedron is isomorphic to a subgroup of  $S_4$ .  
 (2) The Group of rotational symmetries of a cube and of an octahedron is isomorphic to  $S_4$ , and their respective total group of symmetries is a group of symmetries of a cube and of an octahedron are isomorphic to  $S_4 \times \mathbb{Z}_2$ .

**PROOF.** (1) Refer to Figure 1. We prove that the group of rigid motions of a tetrahedron is isomorphic to  $A_4$ . Now, the group  $G$  of the rigid motion of a tetrahedron acts on the vertices of the tetrahedron  $\{1, 2, 3, 4\}$ . Clearly,  $\varphi : G \rightarrow A_4$  is a homomorphism. We first find  $|G|$  by orbit stabilizer theorem. Consider an arbitrary vertex 1. Clearly,  $|orb(1)| = 4$  (Rotational elements can send 1 to any of the four vertices). Also, there are 3 rotational elements that fix 1. (2 rotations that fix 1 and rotate around the axis connecting vertex 1 to the centre of the opposite face). So,  $stab(1) = 3$ . Thus,  $|G| = 12 = |A_4|$ . Now, it is enough to prove that  $\varphi$  is surjective. There are 8 rotational elements that fix vertices and rotate around the axis connecting the vertex to the centre of the face by  $120^\circ 240^\circ$ . Permutation representation of these (images of these elements under  $\varphi$ ) are  $(123), (124), (132), (142), (143), (134), (234), (243)$ . There are 3 rotations around the axis between the midpoints of the side. These rotations are  $(12)(34), (24)(13), (14)(23)$  and the last element is  $e$ . So we found the preimage image of each element of  $A_4$ . So,  $\varphi$  is surjective and  $|G| = |A_4|$  this means  $\varphi$  is injective and thus isomorphism.

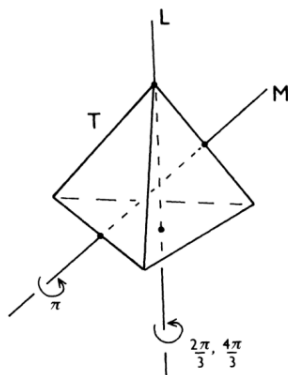


FIGURE 14. Symmetries of a Tetrahedron

(2) Refer to Figure 2

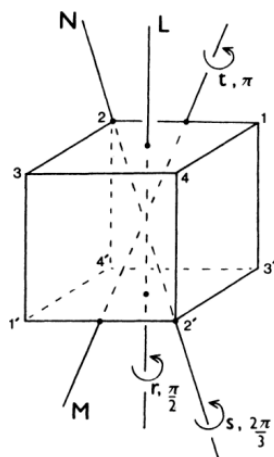


FIGURE 15. Symmetries of a Cube

To count the rotational symmetry of the cube, we first find out all axes of symmetry and then count the number of distinct rotations. We label these axes as  $L$ ,  $M$ , and  $N$ . There are four axes, like diagonal axes  $N$  and about them, there are two distinct rotation of  $\frac{2\pi}{3}$  and  $\frac{4\pi}{3}$ . This gives us 8 distinct rotations about four diagonal axes. There are three axes of the type  $L$  and about each of these axis three are three distinct rotations of  $\frac{\pi}{2}$ ,  $\pi$ , and  $\frac{3\pi}{2}$ . Hence, there are a total of 9 distinct rotations about the axis of  $L$  type. There are six rotations of the category  $M$ , and we have one distinct rotation about each one of them. At the last, we include identity and get  $8 + 9 + 6 + 1 = 24$ . With the labelling of the vertices as shown in Figure 2, rotation about  $L$  is a 4-cycle, rotation about axes  $M$  gives a transposition and rotation about gives a three-cycle. In essence, any permutation of the vertices of the cube is obtained by permuting four diagonal axes of  $N$  type. So, if we label this set as  $A = \{N_1, N_2, N_3, N_4\}$ , then each rotational symmetry affects the permutation of these four diagonals and Therefore, of the vertices 1, 2, 3, 4.

□

## 10. Group Actions

In mathematics, when one algebraic structure acts on another, it reveals itself. So, the action of an object on another is a recurring theme in the study of algebraic structures.

**Definition 10.1 (Group action).** Let  $(G, \star)$  be a group and  $A$  be an abstract set. An action of group  $G$  on set  $A$  is a function  $G \times A \rightarrow A$  given by  $(g, a) \mapsto g \cdot a$  such that the following properties hold.

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

**A remark on the notation.** For each  $g \in G$ ,  $g \cdot a$  is an element of  $A$ . So,  $g_1 \cdot a \in A$  and it makes sense to act on it by  $g_2$ . Also,  $g_1 \star g_2 \in G$ , so again it makes sense to act on  $a$  by  $g_1 \star g_2$ .

**Definition 10.2 (Kernel of group action).** Let  $G$  acts on  $A$ . The kernel of the group action is the set  $\{g \in G : g \cdot a = a \text{ for all } a \in A\}$ .

**Theorem 10.1 (Permutation representation associated with the group action).** Let  $G$  act on set  $A$  in the sense of Definition 10.1. For each fixed  $g \in G$ , define a map  $\sigma_g : A \rightarrow A$  given by  $\sigma_g(a) := g \cdot a$ . Then

- (1)  $\sigma_g$  is a permutation of  $A$  for each  $g \in G$ .
- (2) The map  $\varphi : G \rightarrow S_A$  given by  $g \mapsto \sigma_g$  is a group homomorphism, and  $\varphi$  is known as permutation representation associated with the given group action.

This theorem proves that given a group action of  $G$  on  $A$ , there is an associated permutation representation, technically a group homomorphism of  $G$  into  $S_A$ . Converse is also true. That is,  $f : G \rightarrow S_A$  is a group homomorphism, then we can induce a group action of  $G$  on  $A$  given by  $g \cdot a = f(g)(a)$  for all  $a \in A$  and  $g \in G$ . This makes sense because  $f(g) \in S_A$ . Therefore,  $f(g)(a) \in A$ .

PROOF. (1) Recall that a map  $f : A \rightarrow A$  is a permutation if and only if it has two side inverse. For more details, refer to Appendix A. We claim that for each  $g \in G$ ,  $\sigma_{g^{-1}}$  is the inverse of  $\sigma_g$ . To see it,  $\sigma_g \circ \sigma_{g^{-1}}(a) = (\sigma_g(\sigma_{g^{-1}}(a))) = \sigma_g(g^{-1} \cdot a) = g(g^{-1}) \cdot a = (g \star g^{-1}) \cdot a = e_G \cdot a = a$ . So,  $\sigma_{g^{-1}}$  is the right inverse of  $\sigma_g$ . In the same way, we can also prove that  $\sigma_{g^{-1}}$  is the left inverse of  $\sigma_g$ . This claim is true for all  $g \in G$ . Hence,  $\sigma_g$  is the permutation of  $A$ .

(2) From the previous part, we know that for each  $g \in G$ ,  $\sigma_g \in S_A$ . Now, we prove that  $\varphi : G \rightarrow S_A$  given by  $g \mapsto \sigma_g$  is indeed a group homomorphism since  $\varphi(g_1 g_2)(a) = \sigma_{g_1 g_2}(a) = (g_1 \star g_2) \cdot a = g_1(g_2 \cdot a) = \sigma_{g_1}(\sigma_{g_2}(a)) = (\varphi(g_1) \circ \varphi(g_2))(a)$ .

□

The idea here is that the action of a group on a set is basically a permutation of the element of the set, which also respects the group operation in  $G$ .

**Definition 10.3 (Trivial and faithful action).** The action of a group  $G$  on a set  $A$  is called trivial if  $g \cdot a = a$  for all  $g \in G$  and  $a \in A$ . In this case, each element  $g$  induces the same permutation, namely, the identity map of  $A$ . The action of the group is said to be faithful if distinct elements of  $G$  induce distinct permutations of  $A$ .

**Lemma 10.1.** The kernel of the group action is essentially the kernel of the associated permutation representation.

PROOF. This follows immediately from definitions of associated permutation representation as group homomorphism of  $G$  into  $S_A$ , the kernel of a homomorphism and that of group action, see definition 10.2. □

**On the kernel of trivial and faithful action.** In the case of trivial group action, the permutation representation  $\varphi : G \rightarrow S_A$  is just the trivial group homomorphism mapping each element of  $G$  to  $i_A$  - the identity permutation of  $A$ . Basically, in trivial group action, the entire group  $G$  is the kernel of  $\varphi$ . By definition of faithful action, the associated permutation representation is an injective map because distinct elements of  $G$  induce distinct permutations of  $A$ .

**Definition 10.4 (Left regular action).** Let  $(G, \star)$  be a group. Let  $G$  act on itself by left multiplication. That is, the action of an element  $g \in G$  on every element  $x \in G$  is defined by  $g \cdot x = g \star x$ . This action permutes the element of  $G$  through left multiplication and is known as the *left regular action of  $G$  on itself*.

**Lemma 10.2.** The left regular action of  $G$  on itself is a faithful action.

PROOF. The kernel of this group action is the set  $\{g \in G \mid g \cdot x = g \star x = x \text{ for all } x \in G\}$  and this set is precisely the singleton  $\{e_G\}$ . Thus, the left regular action is faithful.  $\square$

**Lemma 10.3 (Action by conjugation).** Let  $G$  be a group and  $A = G$ . Show that the maps defined by  $g \cdot a = gag^{-1}$  for all  $a, g \in G$  do satisfy axioms of group action of  $G$  on itself.

PROOF. Using the axiom of group action and the definition of given maps, we get

$$\begin{aligned} g_1 \cdot (g_2 \cdot a) &= g_1 \cdot g_2 a g_2^{-1} \\ &= g_1 (g_2 a g_2^{-1}) g_1^{-1} \\ &= (g_1 g_2) a (g_2^{-1} g_1^{-1}) \\ &= (g_1 g_2) a (g_1 g_2)^{-1} \\ &= (g_1 g_2) \cdot a. \end{aligned}$$

So, the First, axiom is satisfied. Also

$$e_G \cdot a = e_G a e_G^{-1} = e_G a e_G = a.$$

Therefore, the given map satisfies the axioms of group action.  $\square$

**Corollary 10.1.** Let  $G$  be a group and  $G$  act on itself by left conjugation, so each  $g \in G$  maps  $G$  to  $G$  by

$$x \mapsto gxg^{-1}.$$

For a fixed  $g \in G$ , prove that conjugation by  $g$  is an isomorphism from  $G$  onto itself (i.e., is an automorphism of  $G$ ). Deduce that  $x$  and  $gxg^{-1}$  have same order for all  $x \in G$  and that for any subset  $A$  of  $G$ ,  $|A| = |gAg^{-1}|$ .

PROOF. We first prove that the map  $\varphi : G \rightarrow G$  given by  $\varphi(x) = gxg^{-1}$  is an injective map. To do so let us assume that  $\varphi(x_1) = \varphi(x_2)$  for some  $x_1, x_2 \in G$ . So,  $gx_1g^{-1} = gx_2g^{-1}$ . Multiply left side by  $g^{-1}$  and right hand side by  $g$  to get  $g^{-1}gx_1g^{-1}g = g^{-1}gx_2g^{-1}g$ . That is,  $x_1 = x_2$ . Now, we prove that the given map is also a surjection. For this, take an arbitrary element  $x \in G$ , and we claim that its preimage is  $g^{-1}xg \in G$ . To see it, note that

$$\varphi(g^{-1}xg) = gg^{-1}xgg^{-1} = e_G x e_G = x.$$

Next, we prove that the given map is a group homomorphism.

$$\varphi(x_1x_2) = g(x_1x_2)g^{-1} = gx_1(g^{-1}g)x_2g^{-1} = (gx_1g^{-1})(gx_2g^{-1}) = \varphi(x_1)\varphi(x_2).$$

Hence we conclude that  $\varphi$  is a isomorphism. Thus, by Theorem 7.1, we conclude that  $|x| = |\varphi(x)| = |gxg^{-1}|$ . Now, we prove that there is a bijection from  $A$  to  $gAg^{-1}$ . But this is obvious because we just proved that the map is  $a \mapsto gag^{-1}$  is indeed a bijection. Hence,  $|A| = |gAg^{-1}|$ .  $\square$

**Theorem 10.2 (Precursor to Lagrange Theorem).** (1) Let  $H$  be a group acting on a set  $A$ . Prove that the relation  $\sim$  on  $A$  defined by

$$a \sim b \quad \text{if and only if} \quad a = hb \quad \text{for some } h \in H$$

is an equivalence relation. (For each  $x \in A$  the equivalence classes of  $x$  under  $\sim$  is called the orbit of  $x$  under the action of  $H$ . The orbits under the action of  $H$  partition the set  $A$ .)

- (2) Let  $H$  be a subgroup of the finite group  $G$  and  $H$  act on  $G$  (here  $A = G$ ) by left multiplication. Let  $x \in G$  and  $\mathcal{O}$  be the orbit of  $x$  under the action of  $H$ . Prove that the map

$$H \rightarrow \mathcal{O} \quad \text{defined by} \quad h \mapsto hx$$

is a bijection (hence, all orbits have cardinality  $|H|$ ). From this and the preceding exercises deduce *Lagrange's Theorem*.

**Corollary 10.2 (Lagrange's theorem).** If  $G$  is a finite group and  $H$  is a subgroup of  $G$ , then  $|H|$  divides  $|G|$ .

PROOF. (1) Obviously  $a \sim a$  as  $a = 1a$  for  $1 \in H$ . So, it is reflexive. Now, let  $a \sim b$ . We have to prove that  $b \sim a$ . As  $a \sim b$ ,  $a = hb$  for some  $h \in H$ . Act by  $h^{-1}$  on both sides to get

$$h^{-1}a = h^{-1}(hb) = (h^{-1}h)b = 1b = b.$$

So  $b = h^{-1}a$  and  $h^{-1} \in H$  and  $b \sim a$ . Next, we prove the transitive. Let  $a \sim b$  and  $b \sim c$ . To prove that  $a \sim c$ . As  $a \sim b$  and  $b \sim c$ , we get  $a = h_1b$  and  $b = h_2c$ . So

$$a = h_1b = h_1(h_2c) = (h_1h_2)c$$

$a = (h_1h_2)c$  and  $h_1h_2 \in H$ . So we have  $a \sim c$ . Now, we know that equivalence classes (which are being called as orbit here) do partition the set. Therefore, the set  $A$  can be written as a union of mutually disjoint non-empty equivalence classes. Thus, orbit under the action of  $H$  partition the set  $A$ .

- (2) Let  $\varphi$  be a map between  $H$  and  $\mathcal{O}$  defined by  $\varphi(x) = hx$ . Let  $\varphi(x_1) = \varphi(x_2)$  for some  $x_1, x_2 \in H$ . So  $hx_1 = hx_2$ . By cancellation law,  $x_1 = x_2$ . Consider an element  $x \in \mathcal{O}$ , its preimage is  $h^{-1}x$ . As

$$\varphi(h^{-1}x) = h(h^{-1}x) = (hh^{-1})x = 1x = x.$$

So it is also surjective. Thus, the map is bijective, and each orbit has the cardinality  $|H|$ . Assume that there are  $n$  orbits. Each of them has cardinality as  $|H|$ . By the previous problem, orbits under the action of  $H$  partition  $G$ . So,  $n|H| = |G|$  and  $|H|$  divides  $|G|$ . This proves the Lagrange's theorem.

**History.** Lagrange's theorem is the first combinatorial result in the group theory. □



FIGURE 16. Joseph Luigi Lagrange

**Example 10.1.** (1) Let us consider the action of dihedral group  $D_{2n}$  on the set  $A = \{1, 2, 3, \dots, n\}$  wherein we take the set  $\{1, 2, 3, \dots, n\}$  to be the vertices of the regular  $n$ -gon. First, note that  $D_{2n}$  is the group of symmetries of a regular  $n$ -gon. Secondly, symmetries of geometrical figures are

'isometric isomorphism'. Combining these two facts, we see that the action of a distinct symmetry element in  $D_{2n}$  induces distinct permutations of  $\{1, 2, 3, \dots, n\}$ . Thus, the action is faithful. As a special case, consider the action of  $D_6$  on  $\{1, 2, 3\}$ . This action is faithful. So, the permutation representation  $\varphi : D_6 \rightarrow S_3$  is an injective homomorphism. Since  $|D_6| = |S_3|$ , therefore,  $\varphi$  is an isomorphism. This is one of the classical results in the finite group theory. Moreover, the isomorphic groups have the same algebraic properties up to relabelling by the isomorphism, hence  $D_6$ , in particular, and  $D_{2n}$ , in general, are non-abelian groups because so are  $S_3$  (respectively,  $S_n$ ).

**Problem 10.1.** Let  $G$  be a group and  $A = G$ . Show that if  $G$  is non-abelian, then the maps defined by  $g \cdot a = ag$  for all  $a, g \in G$  do not satisfy axioms of group action of  $G$  on itself.

**Solution 10.1.** Using the definition of group action and the given map, we get

$$\begin{aligned} g_1 \cdot (g_2 \cdot a) &= g_1 \cdot ag_2 \\ &= ag_2g_1 \end{aligned}$$

and  $(g_1g_2) \cdot a = a(g_1g_2) = ag_1g_2$ . Since  $G$  is non-abelian, we have  $g_1 \cdot (g_2 \cdot a) \neq (g_1g_2) \cdot a$  will be true for at least one pair of  $g_1, g_2 \in G$ . Hence, the given map does not define a group action of  $G$  on itself.

**Problem 10.2.** Let  $G$  be a group and  $A = G$ . Show that the maps defined by  $g \cdot a = ag^{-1}$  for all  $a, g \in G$  do satisfy axioms of group action of  $G$  on itself.

**Solution 10.2.** Using the definition of group action and the given map, we get

$$\begin{aligned} g_1 \cdot (g_2 \cdot a) &= g_1 \cdot ag_2^{-1} \\ &= ag_2^{-1}g_1^{-1}. \end{aligned}$$

Also,  $(g_1g_2) \cdot a = a(g_1g_2)^{-1} = ag_2^{-1}g_1^{-1}$ . So,  $g_1 \cdot (g_2 \cdot a) = (g_1g_2) \cdot a$ . Moreover,

$$e_G \cdot a = ae_G^{-1} = ae_G = a.$$

Hence, the given map defines a group action.

**Problem 10.3.** Explain why the action of the group of rigid motions of a cube on a set of three pairs of opposite faces is not faithful. Find the kernel of this action.

**Solution 10.3.** The group of rigid motions of a cube has order 24. Denote the pairs of opposite pairs by integers  $A = \{1, 2, 3\}$ . Then  $\varphi : G \rightarrow S_A$  cannot be injective as  $|G| = 24 > |S_A| = 6$ . We want to find the rotations that fix pairs of opposite faces. The rotation by  $180^\circ$  through the centre of each pair of opposite faces sends each pair of opposite faces to itself. So the kernel consists of three  $180^\circ$  rotations along with the identity transformation.



## CHAPTER 4

### Study of sub-groups

'Contemporary mathematicians, steeped in the ideas of Bourbaki, have had the natural tendency to introduce into secondary and university courses the algebraic theories and structures that have been so useful in their own work and that are uppermost in the mathematical thought of today', said french mathematician, [Rene Thom](#) – the founder of Catastrophe theory.

#### 1. Structure, sub-structure and smallest sub-structure - a philosophical interlude

Two beautiful themes across the spectrum of theoretical mathematics are the following.

**Theme I.** Given a mathematical object  $\mathcal{O}$ , does there exist a sub-object of  $\mathcal{O}$ ? That is, does there exist a subset of  $\mathcal{O}$  which follows the same set of axioms as the original object  $\mathcal{O}$ ?

**Theme II.** Let  $\mathcal{O}$  be a mathematical structure. Let  $S \subseteq \mathcal{O}$ . We want to know if there exists the smallest sub-structure of  $\mathcal{O}$ , which contains  $S$ . Let us give some examples to highlight this theme.

- (1) Given a vector space  $V$  and a subset  $U$  of  $V$ , does there exist a smallest *sub-space*  $W$  of  $V$  such that  $U \subseteq W$ ?
- (2) Given a group  $G$  and subset  $F$  of  $G$ , does there exist a smallest *sub-group*  $H$  of  $G$  such that  $F \subseteq H$ ?
- (3) Given a group  $G$  and a subset  $F$  of  $G$ , does there exist a smallest *normal* subgroup of  $G$  that contains  $F$ ?
- (4) Given a ring  $R$  and a subset  $S$  of  $R$  does there exist a *sub-ring*  $T$  of  $R$  such that  $S \subseteq T$ ?
- (5) Given a field  $K$  and a subset  $F$  of  $K$  does there exist a *sub-field*  $K'$  of  $K$  such that  $F \subseteq K'$ .

Let us understand what exactly we want to find out. We discuss it in the context of groups, and the discussion is applicable verbatim to other objects. So, we are given a group  $G$  and a subset  $F$  of  $G$ . We have to find  $H$  with the following properties

- $H$  is a sub-group of  $G$
- $F \subseteq H$ , and
- if  $F'$  is any other subgroup of  $G$  such that  $H \subseteq F'$  then  $F \subseteq F'$ .

First, we discuss *Theme I* and its various examples, and then in section ??, we will discuss under what conditions we can find a positive solution to the problem written under *Theme II*. To begin with, we recall the definition of a sub-group.

**Definition 1.1 (Subgroup).** Let  $(G, \star)$  be a group. Let  $H$  be a subset of the group  $G$ . We say  $H$  is a subgroup of  $G$  and write this as  $H \leq G$  if  $(H, \star)$  is a subgroup in its own right. That is,  $H$  is closed under the binary operation  $\star$  and inverses.

- Let  $H \subset G$  and let  $H \leq G$  then inclusion map  $i : H \hookrightarrow G$  is indeed a group homomorphism.
- Note that, if  $H \leq G$ , then every equation between elements of  $H$  can be viewed as an equation between elements of  $G$ . Therefore,  $H$  *inherits the cancellation law from  $G$* .
- The inverse of any element in  $H$  is the same as that in  $G$ . This is an easy consequence of the 'cancellation law.'

**Lemma 1.1.** Prove that if  $(H, \star)$  is a subgroup of  $(G, \star)$ , then  $e_G = e_H$ .

PROOF. Let  $e_G$  and  $e_H$  be the identity element of  $G$  and  $H$ , respectively. As  $e_H$  is an identity element of  $H$ ,  $e_H \times e_H = e_H$ . Also,  $e_H \in G$  and therefore,  $e_G \times e_H = e_H$ . This gives  $e_H \times e_H = e_G \times e_H$ . Seeing this as an equation in  $G$ , and applying cancellation law yields  $e_H = e_G$ .  $\square$

**Proposition 1.1 (Characterization of Subgroup).** Let  $H$  be a non-empty subset of a group  $(G, \star)$ . Then  $H$  is a subgroup of  $G$  if and only if

$$(\forall a, b \in H) : a \star b^{-1} \in H.$$

PROOF. Assume that the stated condition is true. Given that  $H \neq \emptyset$ , therefore, there is at least an  $a \in H$ . Now, take  $b = a$  and by the stated condition  $e_G = a \star a^{-1} \in H$ . So,  $H$  has the same identity as  $G$ . Now, we know that for any  $a \in H$ , the stated condition gives  $e_G \star a^{-1} = a^{-1} \in H$ . So, for each  $b \in H$ , its inverse  $b^{-1} \in H$ . Using this we get  $a \star b = a \star (b^{-1})^{-1} \in H$ . So,  $H$  is closed under the binary operation  $\star$ . Moreover,  $H \subset G$  is closed under  $\star$ . Therefore, the associativity of  $\star$  in  $H$  is inherited from that in  $G$ .  $\square$

**Proposition 1.2 (Arbitrary intersection of subgroups is also a subgroup).** Let  $(H_\alpha)_{\alpha \in \mathcal{I}}$  be a family of subgroups of  $G$ . Then the intersection

$$H = \bigcap_{\alpha \in \mathcal{I}} H_\alpha$$

is a subgroup of the group  $G$ .

PROOF. We use Proposition 1.1 to prove this. As for each  $\alpha \in \mathcal{I}$ ,  $H_\alpha \leq G$ , therefore,  $e_G \in H_\alpha$  for all  $\alpha \in \mathcal{I}$  and hence,  $e_G \in H$ . Let  $a, b \in H$ . Then  $a, b \in H_\alpha$  for all  $\alpha \in \mathcal{I}$ . Since  $H_\alpha \leq G$ , we have  $ab^{-1} \in H_\alpha$  for all  $\alpha \in \mathcal{I}$ . This proves that  $ab^{-1} \in H$  for all  $a, b \in H$ .  $\square$

**Lemma 1.2.** Let  $\varphi : G \rightarrow H$  be a group homomorphism. Let  $H' \leq H$ . Then pullback image  $\varphi^{-1}(H') \leq G$ .

PROOF. Since  $H' \leq H$ , therefore,  $e_H \in H'$ . Hence  $e_G \in \varphi^{-1}(H')$ . Let  $a, b \in \varphi^{-1}(H')$ . There exist  $x, y \in H'$  such that  $\varphi(a) = x$  and  $\varphi(b) = y$ . Since  $H'$  is a group, we have  $xy = \varphi(a)\varphi(b) = \varphi(ab) \in H' \implies ab \in \varphi^{-1}(H')$ . Thus,  $a \in \varphi^{-1}(H') \implies \varphi(a) \in H' \implies \varphi(a)^{-1} \in H' \implies \varphi(a^{-1}) \in H' \implies a^{-1} \in \varphi^{-1}(H')$ . This proves our claim that the pullback image of a subgroup is also a subgroup.  $\square$

## 2. Derived or Commutator subgroup

**Definition 2.1 (Commutator subgroup).** Let  $G$  be a group. Then the collection  $\{ghg^{-1}h^{-1} \mid g \in G, h \in G\}$  has a group structure. This group is known as the commutator subgroup of  $G$ .

**Lemma 2.1.** (1) Prove that the commutant is a normal subgroup.

(2) Prove that the commutant coincides with the unit element  $\{e\}$  if and only if the group is commutative.

PROOF. (1) Let  $x$  be an element of the commutant. Then  $x = aba^{-1}b^{-1}$  for some  $a, b \in G$  and  $gxg^{-1} = gaba^{-1}b^{-1}g^{-1} = (gag^{-1})(gbg^{-1})(ga^{-1}g^{-1})(gb^{-1}g^{-1})$ . Note that,  $ga^{-1}g^{-1} = (gag^{-1})^{-1}$  and  $gb^{-1}g^{-1} = (gbg^{-1})^{-1}$ . Thus,  $(gag^{-1})(gbg^{-1})(ga^{-1}g^{-1})(gb^{-1}g^{-1})$  is in the form  $xyx^{-1}y^{-1}$ . Hence, it is a commutator.

(2)  $aba^{-1}b^{-1} = e$  iff  $aba^{-1}b^{-1}b = be = b$ . iff  $aba^{-1} = b$  iff  $aba^{-1}a = ba$  iff  $ab = ba$ .  $\square$

**Example 2.1.** Find the commutant for the following groups:

- (1) The group of symmetries of the triangle
- (2) The group of symmetries of the square
- (3) The group of quaternions.

**Solution 2.1.** (1)  $S_3$  is a non-abelian group. Therefore, following lemma 2.1, we know that the commutator subgroup of  $S_3$  is not just  $\{e\}$ . Let us consider an arbitrary commutator  $g_1g_2g_1^{-1}g_2^{-1}$  and we claim that it is a pure rotation. Note that, if  $a$  is a reflection, then  $a^{-1}$  is also a reflection. So there are 0, 2, or 4 reflections. For 0 reflections, clearly the commutant is a rotation. For 2 reflections, let us take  $g_1 = sr^i$  to be a reflection. Then  $g_1^{-1}$  is also a reflection. Assume  $g_2 = r^j$  is a rotation and hence,  $g_2^{-1}$  is a rotation. Now,  $g_1g_2g_1^{-1}g_2^{-1} = (sr^i)r^j(sr^i)^{-1}r^{-j} = sr^i r^j r^{-i} s^{-1} r^{-j} = sr^j s^{-1} r^j = r^{-j} s s^{-1} r^{-j} = r^{-2j}$ . Thus, in this case, the commutant is a rotation. For 4 reflections, we use the fact that the product of two reflections is a rotation. Therefore, if  $g_1, g_2$  and their inverses are reflection then the product  $g_1g_2g_1^{-1}g_2^{-1}$  is a rotation

because  $g_1^{-1}g_2^{-1}$  is a rotation and  $g_1g_2$  is also a rotation. We conclude that the commutant of  $S_3$  is the subgroup of rotations.

- (2) In fact, the previous argument holds true for any regular  $n$  gon. **Thus, the commutant of Symmetries of regular  $n$  gon, that is commutant of the dihedral group  $D_{2n}$ , is a subgroup of rotations and therefore, it is isomorphic to  $\mathbb{Z}_n$ .**
- (3) Note that,  $Q_8 = \{1, -1, i, j, k, -i, -j, -k\}$ . Commutator is  $g_1g_2g_1^{-1}g_2^{-1}$ . If any one of the  $g_1$  or  $g_2$  is 1 or  $-1$ . The commutator will be 1. If none of  $g_1$  and  $g_2$  is 1 or  $-1$ , then  $g_1^{-1} = -g_1$  and  $g_2^{-1} = -g_2$ . Thus,  $g_1g_2g_1^{-1}g_2^{-1} = g_1g_2g_1g_2 = (g_1g_2)^2$ . Moreover, the square of any element of  $Q_8$  is 1 or  $-1$ . Thus, the commutator is  $\{1, -1\}$ .

### 3. Centralizers and Normalizers, Stabilizers and Kernels

Let  $G$  be a group and  $A$  be a non-empty subset of  $G$ .

**Definition 3.1 (Centralizer).** The set  $C_G(A) := \{g \in G \mid g \star a \star g^{-1} = a \iff g \star a = a \star g \text{ for all } a \in A\}$  is known as centralizer of  $A$  in  $G$ . In essence,  $C_G(A)$  is the set of those elements in  $G$  which commute with each element of  $A$ .

**Definition 3.2 (Center).** In Definition 3.1, if  $A = G$ , then  $C_G(G)$  is known as center of the group  $G$  and is denoted by  $Z(G)$ .

**Definition 3.3 (Normalizer).** Define  $gAg^{-1} = \{gag^{-1} \mid g \in G\}$ . The set  $N_G(A) = \{g \in G \mid gAg^{-1} = A\}$  is known as the normalizer of set  $A$  in  $G$ .

**Definition 3.4 (Stabilizer).** Let  $A$  be a set, and  $G$  act on  $A$ . The stabilizer of  $x$  in  $G$  is  $G_x = \{g \in G \mid g \cdot x = x\}$ . By definition of group action, this set  $G_x$  is non-empty because  $e_G \in G_x$ .

**Theorem 3.1.** Prove that  $C_G(A), Z(G), N_G(A)$  are subgroups of group  $G$ .

PROOF. These claims follow immediately from Proposition 1.1 of this chapter.  $\square$

**Lemma 3.1.** Let  $G = S_n$ . Fix an  $i \in \{1, 2, \dots, n\}$  and  $G_i = \{\sigma \in G \mid \sigma(i) = i\}$ . This will be known as the Stabilizer of  $i$  in  $G$ . Use the definition of group action to prove that  $G_i$  is a subgroup of  $G$ . Find  $|G_i|$ .

PROOF. Note that,  $G_i$  is non empty set as  $e_G \in G_i$ . Let  $\sigma, \tau$  be some arbitrary elements of  $G_i$ . We have to prove that  $\sigma \circ \tau \in G_i$  and  $\sigma^{-1} \in G_i$ .

First, we prove that  $\sigma^{-1} \in G_i$ . Since  $\sigma \circ \sigma^{-1}(i) = i$  and  $\sigma(i) = i$ , we have  $\sigma \circ \sigma^{-1}(i) = \sigma(i)$ . As  $\sigma$  is a permutation, it is an injective map. Hence, the previous equation gives us  $\sigma^{-1}(i) = i$ .

Now, we prove that  $\sigma \circ \tau \in G_i$ . Note that,

$$\sigma \circ \tau(i) = \sigma(i) = i$$

because  $\tau, \sigma \in G_i$ . Hence, the set  $G_i$  is closed under composition, inverses, and has identity  $e_G$  in it. Therefore,  $G_i \leq G$ . Let  $i \in \{1, 2, \dots, n\}$ . Therefore, any permutation of letters  $\{1, 2, \dots, n\} \setminus \{i\}$  will fix  $i$ . Hence, the cardinality of  $G_i$  is  $(n-1)!$ . Note that, a permutation  $\sigma \in G(i)$  fixes element the  $i$ , so the rest of the  $n-1$  elements can be permuted in  $(n-1)!$  ways. Thus,  $|G_i| = (n-1)!$ .  $\square$

**Lemma 3.2.** For any subgroup  $H$  of  $G$  and any non-empty subset  $A$  of  $G$ , define  $N_H(A)$  to be the set  $\{h \in H \mid hAh^{-1} = A\}$ . This set is known as the normalizer of  $A$  in  $H$ . Show that  $N_H(A) = N_G(A) \cap H$  and deduce that  $N_H(A)$  is a subgroup of  $H$  (note that  $A$  need not be a subset of  $H$ ).

PROOF. • First, we prove that  $N_H(A) \subseteq N_G(A) \cap H$ . Let  $h \in N_H(A)$ . Then  $hAh^{-1} = A$  and  $h \in H \subseteq G$ ,  $h \in G$  and  $hAh^{-1} = A$ . Thus,  $h \in N_G(A) \cap H$ . Now, we prove that  $N_G(A) \cap H \subseteq N_H(A)$ . Let  $h \in N_G(A) \cap H$ . As  $h \in N_G(A)$ , we have  $hAh^{-1} = A$ . We also have  $h \in H$  and, hence,  $h \in N_H(A)$ . So we get,  $N_G(A) \cap H \subseteq N_H(A)$ . Therefore,  $N_H(A) = N_G(A) \cap H$ .

- Now, we prove that  $N_H(A) \leq H$ . To do so, note that  $N_H(A)$  is non-empty as  $e_G A e_G^{-1} = A$ . That is,  $e_G \in N_H(A)$ . Now, assume  $x, y$  be arbitrary elements of  $N_H(A)$  and we prove that  $xy \in N_H(A)$  and  $x^{-1} \in N_H(A)$ . Notice that

$$xyA(xy)^{-1} = x(yAy^{-1})x^{-1} = xAx^{-1} = A$$

as  $x, y \in N_H(A)$ . If  $x \in N_H(A)$ , then  $xAx^{-1} = A$ . By multiplying the right side by  $x$  and the left side by  $x^{-1}$ , we get

$$x^{-1}xAx^{-1}x = x^{-1}Ax \text{ which is the same as } x^{-1}Ax = A.$$

Thus,  $x^{-1} \in N_H(A)$ . This completes the proof that  $N_H(A)$  is a subgroup of  $H$ .  $\square$

**Remark 3.1.**  $H$  is a normal subgroup of group  $G$ . Normalizer measures how close a subgroup is to being a normal subgroup.

**Problem 3.1.** Let  $H$  be a subgroup of order 2 in  $G$ . Show that  $N_G(H) = C_G(H)$ . Deduce that if  $N_G(H) = G$  then  $H \leq Z(G)$ .

**Solution 3.1.** Since  $H$  has order 2, assume that  $H = \{1, h\}$ . Clearly,  $C_G(H) \subseteq N_G(H)$ . Now, we prove that  $N_G(H) \subseteq C_G(H)$ . Let  $g$  be an arbitrary element of  $N_G(H)$ . So,  $ghg^{-1} \in H$ . Case 1,  $ghg^{-1} = 1 \implies gh = g \implies h = 1$  a contradiction. Thus, we have  $ghg^{-1} = h \implies gh = hg$ . This means  $g \in C_G(H)$  and we are done. By Lagrange theorem, we have two left cosets and two right cosets of  $H$  in  $G$ . The right cosets are  $He_G = H$  and  $Hg$  for some  $g \in G$ . The left cosets are  $e_G H$  and  $gH$ . As the cosets partition  $G$ . Therefore, both  $Hg$  and  $gH$  contain all the elements of  $G - H$ . Thus,  $Hg = gH = G - H$ . This means that  $N_G(H) = G - H$ , the complement of  $H$  in  $G$ .

Next, we prove that  $C_G(A) = G$ . If  $N_G(H) = G$ , then  $H \trianglelefteq G \iff gHg^{-1} = H$  for all  $g \in G$ . So, assume that  $h \in H$  and  $H \trianglelefteq G$ . We need to show that for all  $g \in G, h \in H$ , we have  $hg = gh \iff h = ghg^{-1}$ . Since  $C_G(H) = \{g \in G | gh = hg \text{ for all } h \in H\}$ , we have  $C_G(H) = G$  as  $H$  is normal subgroup.

**Corollary 3.1.** Prove that  $Z(G) \leq N_G(A)$  for any subset  $A$  of  $G$ .

PROOF. First, we prove that  $Z(G)$  is a subset of  $N_G(A)$ . Let  $z$  be an arbitrary element of  $Z(G)$ . Now,

$$zAz^{-1} = \{zaz^{-1} | a \in A\} = \{a(zz^{-1}) | a \in A\} = \{a | a \in A\} = A.$$

So  $zAz^{-1} = A$  and hence  $z \in N_G(A)$ . That is,  $Z(G)$  is a subset of  $N_G(A)$ . Now, we prove that it is a subgroup. Let  $x \in Z(G), y \in Z(G)$  and  $g$  be an arbitrary element of  $G$ . Now, we prove that  $xy \in Z(G)$  and  $x^{-1} \in Z(G)$ . By the definition of center ( $Z(G)$ ), we have  $xg = gx$  and  $yg = gy$  for all  $g \in G$ . Using associativity and hypothesis, we get  $(xy)g = x(yg) = x(gy) = (xg)y = (gx)y = g(xy)$ . So  $xy \in Z(G)$ . As  $x \in Z(G)$ , we have  $xg = gx$ . By Multiplying both sides by  $x^{-1}$ , we get  $x^{-1}(xg)x^{-1} = x^{-1}(gx)x^{-1}$ . Simplifying the above expression, we get  $gx^{-1} = x^{-1}g$ . So  $x^{-1} \in Z(G)$  and consequently,  $Z(G) \leq G$ . Also, we proved that  $Z \subseteq N_G(A)$ . This implies that  $Z(G) \leq N_G(A)$ .  $\square$

**Corollary 3.2.** Prove that  $C_G(Z(G)) = G$  and deduce that  $N_G(Z(G)) = G$ .

PROOF. Since  $C_G(Z(G)) = \{g \in G | ag = ga \text{ for all } a \in Z(G)\}$  for all  $a \in Z(G)$ , it commutes with every element in  $G$ . Thus,  $C_G(Z(G)) = G$ . Also,  $N_G(Z(G)) = \{g \in G | gZ(G)g^{-1} = Z(G)\}$ . Consider an arbitrary element  $g \in G$ ,  $gZ(G)g^{-1} = \{gzg^{-1} | z \in Z(G)\}$ . As  $z \in Z(G)$ , we have  $gzg^{-1} = gg^{-1}z = z$ . So  $gZ(G)g^{-1} = \{z | z \in Z(G)\} = Z(G)$  all the  $g \in G$ . So the property that  $gZ(G)g^{-1} = Z(G)$  is true for all elements in the group  $G$ . Thus, by the definition of  $N_G(Z(G))$ , we get that  $N_G(Z(G)) = G$ .  $\square$

**Corollary 3.3.** Prove that if  $A$  and  $B$  are subsets of  $G$  with  $A \subseteq B$ , then  $C_G(B)$  is a subgroup of  $C_G(A)$ .

PROOF. Let  $g_1$  and  $g_2$  be the two elements of  $C_G(B)$ . We have to prove that  $g_1g_2 \in C_G(B)$  and  $g_1^{-1} \in C_G(B)$ . Note that, We have  $bg_1 = g_1b$  and  $bg_2 = g_2b$  for all  $b \in B$ . Now,  $b(g_1g_2) = (bg_1)g_2 = g_1(bg_2) = g_1(g_2b) = (g_1g_2)b$ . Thus,  $g_1g_2 \in C_G(B)$ . As  $bg_1 = g_1b$ , multiply the left and right sides by  $g_1^{-1}$  we get  $g_1^{-1}bg_1g_1^{-1} = g_1^{-1}g_1bg_1^{-1}$ . Thus,  $g_1^{-1}b = bg_1^{-1}$ . So,  $g_1^{-1} \in C_G(B)$ .  $\square$

**Proposition 3.1.** Let  $H$  be a subgroup of the group  $G$ .

- Show that  $H \leq N_G(H)$ . Give an example to show that this is not necessarily true if  $H$  is not a subgroup.
- Show that  $H \leq C_G(H)$  if and only if  $H$  is abelian.

PROOF. • Note that, we only need to prove that  $H \subseteq N_G(H)$ . Let  $h$  be an arbitrary element of  $H$ . We have to prove that  $h \in N_G(H)$ . That is, To Prove that  $hHh^{-1} = H$ . Let  $x \in hHh^{-1}$ , so  $x = hh_1h^{-1}$  for some  $h_1 \in H$ . As  $H$  is a subgroup of  $G$ ,  $x = hh_1h^{-1} \in H$ .  
 $(\supseteq)$  Let  $x \in H$ . Now, consider  $h_1 = h^{-1}xh$  which is in  $H$ . Now,  $h_1 \in H$  and  $x = hh_1h^{-1}$ . So  $x \in hHh^{-1}$ . Thus,  $hHh^{-1} = H$ .

- Let  $H$  be an abelian group. To prove that  $H$  is a subgroup of  $C_G(H)$ . It is enough to prove that  $H \subseteq C_G(H)$ . Let  $h \in H$ . As  $H$  is abelian it commutes with every element of  $H$ . So clearly  $h \in C_G(H)$ . Now, let  $H \leq C_G(H)$ . Let  $h_1 \in H$  and  $h_2 \in H$ . Now,  $h_1 \in C_G(H)$  and  $h_2 \in C_G(H)$ , by the definition of  $C_G(H)$ , we get  $h_1h_2 = h_2h_1$ .

□

#### 4. Cyclic Groups, and Cyclic Subgroups

My first introduction to the cyclic group was through the study of the group of roots of unity<sup>1</sup>. Basically, the idea is given an element  $g \in G$ , consider the set  $\langle g \rangle = \{g^m | m \in \mathbb{Z}\}$ . Then, this collection satisfies all axioms of the group and becomes a subgroup of  $G$ . In essence, each element of a cyclic group can be obtained by repeatedly applying the operation of  $G$  on either  $g$  or on its converse. The element  $g$  is called a generator of  $\langle g \rangle$ . Next, we prove major results in the theory of cyclic group, and then, towards the end of this section, we study some basic examples and applications.

**Theorem 4.1.** The order of a cyclic group is the same as the order of its generator. That is, if  $H = \langle x \rangle$  then  $|H| = |x|$ . So, if one of these two sides is infinite then so would be the other. In particular,

- (1) If  $|H| = n < \infty$ , then  $x^n = 1$  and  $1, x, x^2, \dots, x^{n-1}$  are all distinct powers of  $x$  and hence distinct elements of  $H = \langle x \rangle$ .
- (2) If  $|H| = \infty$ , then  $x^n \neq 1$  for all non zero integer  $n$ . Also, if integers  $n \neq m$ , then  $x^n \neq x^m$ .

PROOF. (1) First, we assume that the order  $|x|$  of generator  $x$  of  $H$  is  $n$  and prove that the order of  $H$  is  $n$ . But this is equivalent to saying that elements  $1, x, x^2, \dots, x^{n-1}$  are all distinct. To prove this, let us assume that  $x^i = x^j$  where  $0 \leq i < j \leq n-1$ . This gives  $x^{j-i} = 1$ . But this contradicts the minimality of positive integer  $n$  such that  $x^n = 1$ . Hence, elements  $1, x^1, x^2, \dots, x^{n-1}$  are all distinct and are in the group  $H$ . That is  $n \leq |H|$ . Now, we need to prove that these are the only elements of  $H$ . Since  $H = \langle x \rangle$ . Let  $a \in H$  and that  $a = x^m$ . By Euclid division algorithm,  $m = nq + r$ . This gives  $a = x^m = x^{nq+r} = x^{nq}x^r = x^r$ , where  $0 \leq r < n$ . That is, each element of  $H$  is written as  $x^m$  where  $0 \leq m < n$ . This proves that  $|H| = |x|$ .

- (2) If  $x^n = 1$  for some non-zero integer, then the order of the group can not be infinite, a contradiction. Further, if  $x^m = x^n$  for some integers  $m, n$ , then either  $x^{m-n} = 1$  or  $x^{n-m}$  and this again implies that the order of  $H = \langle x \rangle$  can not be an infinite, a contradiction.

□

**Proposition 4.1.** Let  $x$  be an element of an arbitrary group  $G$ . Let  $m, n \in \mathbb{Z}$  such that  $x^m = 1 = x^n$ . Then  $x^{(m,n)} = 1$ , where  $(m, n)$  denotes the greatest common divisor of  $m, n$ . In particular, if  $x^m = 1$  for some integer  $m \in \mathbb{Z}$ , then  $|x|$  divides  $m$ .

PROOF. Recall that, the greatest common divisor is the integer linear combination of  $m, n$ . That is, there exist integers  $a, b$  such that  $d = (m, n) = ma + nb$ . Using this  $x^d = x^{ma+nb} = x^{ma}x^{nb} = 1$ . For the second assertion, let us assume that  $|x| = n$ . If  $m = 0$ , then there is nothing to prove. Hence, we assume that  $m \neq 0$ . Let us consider  $d := (m, n)$ . Then we know that  $x^d = 1$ . However,  $d$  is the greatest common divisor of  $n$ ; therefore,  $1 \leq d \leq n$  and  $n$  are the smallest positive integer for which  $x^n = 1$ . Therefore,  $d = n$ . By definition of  $d$ , we have  $n = d|m$ .

□

<sup>1</sup>It happened in the seminar of my mentor—Yaashaa Golovanov at his premier program - AMMOC, a math circle

**Proposition 4.2.** Prove that

- (1) Any two cyclic groups of the same order are isomorphic.
- (2) Every cyclic group of order  $n$  is isomorphic to the group of the remainders of the division by  $n$  under addition modulo  $n$ .
- (3) Every infinite cyclic group is isomorphic to the group of integers under addition.
- (4) An infinite cyclic group has an infinite number of subgroups.

PROOF. (1) Let  $G$  and  $H$  be two cyclic groups of the same order. Let  $x$  and  $y$  be their generator. Let us assume that  $|x| = |y| = n$ . Define a map  $\varphi : G \rightarrow H$  such that  $\varphi(x^k) = y^k$ , where  $0 \leq k < n$ . We claim that it is a group isomorphism. First, we prove that this map is indeed *well defined*. Let us assume that  $x^i = x^j$  and  $i > j$ . This means  $x^{i-j} = 1$ . Therefore,  $\varphi(x^{i-j}) = e_H$ . Since  $|x| = n$ , therefore,  $n | i - j$ . This gives  $i - j = nq \implies i = nq + j$ . So,  $\varphi(x^i) = \varphi(x^{nq+j}) = y^{nq+j} = y^{nq}y^j = y^j = \varphi(x^j)$ . This proves that the map  $\varphi$  is well-defined. Now, we prove that this map is *homomorphism*. To do so,  $\varphi(x^i x^j) = \varphi(x^{i+j}) = y^{i+j} = y^i y^j = \varphi(x^i) \varphi(x^j)$ . This map is a surjection by its definition. Further, since the cardinality of  $G$  and  $H$  is the same, hence, the map also becomes injective. This proves that  $\varphi$  is a group isomorphism.

(2) Let  $\mathbb{Z}_n$  be additive group of integers modulo  $n$ . Let  $Z_n$  be a cyclic group of order  $n$ . Let  $Z_n = \langle x \rangle$ . Define the map  $\varphi : Z_n \rightarrow \mathbb{Z}_n$  as  $\varphi(x^i) = [i]_n$ , the congruence class of integer  $i$  modulo  $n$ . Following up on arguments of part (1), we know that  $\varphi$  is indeed a group isomorphism.

(3) Let  $G = \langle x \rangle$  be a group of infinite order. Define the map  $\mathbb{Z} \rightarrow G$  as  $k \mapsto x^k$ . By part (2) of Theorem 4.1, we know that this map is indeed an injection. By its definition, the map is a surjective homomorphism.

(4) By part (3), we know that every infinite cyclic group is isomorphic to the group of integers under addition. For each prime  $p$ , the set  $p\mathbb{Z} = \{pn \mid n \in \mathbb{Z}\}$  is a subgroup of  $\mathbb{Z}$  under addition. We know from the infamous result of Euler that there are infinitely many primes. Therefore,  $\mathbb{Z}$  has infinite many subgroups of the form  $p\mathbb{Z}$ . Since isomorphism preserves the group structure, therefore, every infinite cyclic group has infinitely many subgroups. □

**Proposition 4.3.** Let  $G$  be a group. Let  $x \in G$  and let  $a \in \mathbb{Z} - \{0\}$ .

- (1) If  $|x| = \infty$  then  $|x^a| = \infty$ .
- (2) If  $|x| = n < \infty$ , then  $|x^a| = \frac{n}{(n,a)}$ . Also, if  $a$  is a positive integer divisor of  $n$ , then  $|x^a| = \frac{n}{a}$ .

PROOF. (1) Let us assume that  $|x^a| = m < \infty$ . This means  $x^{am} = 1$  and also  $x^{-am} = 1$ . Now, either  $\infty > am > 0$  or  $\infty > -am > 0$ . Hence, one of these would imply that  $x$  has a finite order. However, this would contradict the given hypothesis.

(2) Let us take  $(n, a) = d$ . This means there exist integers  $b, c$  such that  $n = db$  and  $a = dc$ , where  $(b, c) = 1$ . First, note that  $x^{ab} = 1$  because  $b = \frac{n}{d}$ . This gives  $x^{ab} = x^{dcb} = x^{dbc} = x^{nc} = 1$ . By Proposition 4.1, we know that  $|x^a|$  must divide  $b$ . Let us assume that  $|x^a| = k$ .  $x^{ak} = 1$ . Now, the order of  $x$  is  $n$ . So,  $n | ak$ . This is equivalent to  $db | dck \iff b | ck \iff b | k$  because  $(b, c) = 1$ . This gives us  $b | k$  and  $k | b$  where  $k, b$  are positive integers. This means  $k = b$ . That is, order of  $x^a$  is  $b = \frac{n}{d} = \frac{n}{(a,n)}$ . □

**Proposition 4.4.** Let  $H = \langle x \rangle$ .

- Assume that  $|x| = \infty$ . Then  $H = \langle x^a \rangle$  if and only if  $a \pm 1$ . This basically means that an infinite cyclic group  $H = \langle x \rangle$  has only two possible generators,  $x$  and  $x^{-1}$ .
- Assume that  $|x| = n < \infty$ . Then  $H = \langle x^a \rangle$  if and only if  $(a, n) = 1$ . In particular, the number of generators of  $H$  is  $\varphi(n)$ , where  $\varphi$  is the Euler's  $\varphi$ -function.

PROOF. • Backward direction ( $\leftarrow$ ) Let  $a \pm 1$ . It is enough to show that  $H = \langle x^{-1} \rangle$ . To achieve this, we use the set inclusion. First, we prove  $\langle x^{-1} \rangle \subseteq H$ . Let  $(x^{-1})^a \in \langle x^{-1} \rangle$ . Then  $(x^{-1})^a = x^{-a}$  and  $(x^{-1})^a = x^{-a} \in H$  as  $H = \langle x \rangle$ . Now, we prove  $H \subseteq \langle x^{-1} \rangle$ . Let  $h$  be

an arbitrary element of  $H$ . Then  $h = x^k$  for some  $k$ . Thus,  $h = x^k = (x^{-1})^{-k} \in \langle x^{-1} \rangle$ . So  $H = \langle x^{-1} \rangle$ .

Forward direction ( $\rightarrow$ ). Assume  $H = \langle x^a \rangle$ . Consider arbitrary element  $x^k \in H$ . As  $H = \langle x^a \rangle$ ,  $x^k = x^{am}$  for some  $m$ . Since,  $|x| = \infty$  so  $k = am$  otherwise we would have  $x^{k-am} = 1$  or  $x^{am-k} = 1$ . Hence,  $k$  is a multiple of  $a$ . But  $k$  was arbitrarily chosen. So, any arbitrary power of  $x$  is a multiple of  $a$ . So  $a$  must be  $\pm 1$ .

- Assume that the order of  $x$  is  $n$ . To prove in the forward direction, assume that  $H = \langle x^a \rangle$ . Since we proved that the order of a finite cyclic group is the same as the order of its generator, therefore,  $|x^a| = |x|$  and by Proposition 4.3

$$\frac{n}{(n, a)} = n$$

$$(n, a) = 1.$$

To prove in the backwards direction, assume that  $(a, n) = 1$ . By Proposition 4.3,  $|x^a| = \frac{n}{(a, n)} = n = |x|$ . Hence,  $H = \langle x^a \rangle$ .

□

**Theorem 4.2 (Fundamental theorem of cyclic groups).** Let  $H = \langle x \rangle$  be a cyclic group.

- (1) Every subgroup of  $H$  is a cyclic. More precisely, if  $K \leq H$ , then either  $K = \{1\}$  or  $K = \langle x^d \rangle$ , where  $d$  is the smallest positive integer such that  $x^d \in K$ .
- (2) If  $|H| = \infty$ , then for any distinct non-negative integers  $a$  and  $b$ ,  $\langle x^a \rangle \neq \langle x^b \rangle$ . Furthermore, for every integer  $m$ ,  $\langle x^m \rangle = \langle x^{|m|} \rangle$ , where  $|m|$  denotes the absolute value of  $m$  so that the nontrivial subgroups of  $H$  correspond bijectively with the integers  $1, 2, 3, \dots$ .
- (3) If  $|H| = n < \infty$ , then for each positive integer  $a$  dividing  $n$ , there is a unique subgroup of  $H$  of order  $a$ . This subgroup is the cyclic group  $\langle x^d \rangle$ , where  $d = \frac{n}{a}$ . Furthermore, for every integer  $m$ ,  $\langle x^m \rangle = \langle x^{(n, m)} \rangle$ , so that the subgroups of  $H$  correspond bijectively with the positive divisors of  $n$ .

PROOF. (1) Obviously  $\{1\}$  is the subgroup of  $H$ . Let  $K \leq H$  and  $K \neq \{1\}$ . We have to prove that  $K = \langle x^d \rangle$ , where  $d$  is the smallest positive integer such that  $x^d \in K$ . As  $K \neq \{1\}$ , there is a positive integer  $p > 0$  such that  $x^p \in K$ . Now, let us consider the set

$$A = \{b \mid b \in \mathbb{Z}^+ \text{ and } x^b \in K\}.$$

$A$  has a minimum element, say  $d$ , and this is guaranteed by the Well-ordering principle. Since  $K$  is a subgroup and  $x^d \in K$ , therefore,  $\langle x^d \rangle \subseteq K$ . Now, we prove that  $K \subseteq \langle x^d \rangle$ . Towards this, let  $x^k$  be an arbitrary element of  $K$ . By Euclid's division lemma, we have

$$k = qd + r$$

where  $0 \leq r < d$ . This gives,  $x^r = x^k x^{-qd}$ . Now, if  $r \neq 0$ , then as both  $x^k \in K$  and  $x^d \in K$  we get  $x^r \in K$ . This contradicts the minimality of  $d$  such that  $x^d \in K$ . Therefore,  $r$  must be 0. Hence, each element  $x^k \in K$  is equal to some integer power of  $x^d$ . That is,  $K = \langle x^d \rangle$ .

- (2) Let  $a, b$  be non-negative integer. In fact, if one of these integers is zero and the other is positive, then the claim holds true. Also, if both of them are zero, then the claim is false. So, let us assume that  $a, b$  are positive integers such that  $\langle x^a \rangle = \langle x^b \rangle$ . Then there exists a non negative integer  $m$  such that  $x^{am} = b \implies x^{am-b} = 1$  or  $x^{b-am} = 1$ . This implies that  $x$  has finite order, a contradiction as  $|H| = \infty$ . The second part of the claim follows from the definition of a subgroup generated by  $x^m$  as a subgroup contains elements and their inverses. Hence  $\langle x^m \rangle = \langle x^{|m|} \rangle$ . The last part of the claim is obvious from the first part.
- (3) Consider  $d = \frac{n}{a}$ . This implies  $d$  divides  $n$ . So  $(n, d) = d$ . By Proposition 4.3, part (2), we know that  $|x^d| = \frac{n}{d} = a$ . This makes  $\langle x^d \rangle$  a subgroup of order  $a$ . To prove the uniqueness, let  $K$  be

another subgroup of order  $a$ . By part (1), order of  $K = \langle x^c \rangle$ , where  $c$  is the smallest integer such that  $x^c \in K$ . Again, by proposition 4.3, we have

$$\frac{n}{d} = a = |K| = |x^c| = \frac{n}{(n, c)}.$$

This gives  $(n, c) = d$  and  $d|c$ . Thus,  $x^c \in \langle x^d \rangle$ . So,  $\langle x^c \rangle \leq \langle x^d \rangle$  and both of these groups have the same order  $a$ . Hence, they must be equal.  $\square$

**Example 4.1 (Some classical results).**

- (1) Let  $G$  be a group with  $x, y \in G$ . Assume  $|x| = n$  and  $|y| = m$ . Suppose that  $x$  and  $y$  commute. Prove that  $|xy|$  divides the least common multiple of  $m$  and  $n$ . Need this be true if  $x$  and  $y$  do not commute? Give an example of commuting elements  $x$  and  $y$  such that the order of  $xy$  is not equal to the least common multiple of  $|x|$  and  $|y|$ .
- (2) We write  $Z_n = \langle x \rangle$ . Show that if  $H$  is any group and  $h \in H$  such that  $h^n = 1$ , then there is a unique homomorphism  $\varphi : Z_n \rightarrow H$  with  $x \mapsto h$ .
- (3) Show that if  $H$  is a group and  $h \in H$ , there exists a unique homomorphism  $\varphi : \mathbb{Z} \rightarrow H$  such that  $\varphi(1) = h$ .
- (4) Let  $p$  be an odd prime and  $n$  a positive integer. Use the Binomial theorem to show that  $(1+p)^{p^{n-1}} = 1 \pmod{p^n}$  but that  $(1+p)^{p^{n-2}} \not\equiv 1 \pmod{p^n}$ . Deduce that  $1+p$  is an element of order  $p^{n-1}$  in  $(\mathbb{Z}/(p^n))^\times$ .

**Solution 4.1.**

(1) Suppose that  $x$  and  $y$  be two commuting elements, i.e  $xy = yx$ . Then  $(xy)^k = x^k y^k$  for all integers  $k$ . Let  $\text{lcm}(|x|, |y|) = \text{lcm}(n, m) = d$ . By definition of LCM, there exist integers  $a, b$  such that  $d = na = mb$ . Now, consider,

$$(xy)^{\text{lcm}(|x|, |y|)} = (xy)^d = x^d y^d = x^{na} y^{mb} = (x^n)^a (y^m)^b = e^a e^b = e$$

Thus, by the defining property of order  $|xy|$  divides  $\text{lcm}(|x|, |y|)$ .

- (2) Define a map  $\varphi : Z_n \rightarrow H$  by  $\varphi(x^k) = h^k$ . Suppose  $x^p = x^q$  for some integers  $p$  and  $q$ . Then we get  $p \equiv q \pmod{n}$ . By Euclid division algorithm, we obtain  $p = yn + q$ . Therefore,

$$\varphi(x^p) = h^p = h^{yn+q} = h^q = \varphi(x^q).$$

In the above equation  $h^{yn+q} = h^q$  as  $n|(yn+q) - q$ . So,  $\varphi$  is well defined. Now, we have

$$\varphi(x^p x^q) = \varphi(x^{p+q}) = h^{p+q} = h^p h^q = \varphi(x^p) \varphi(x^q),$$

so  $\varphi$  is a homomorphism.

**Uniqueness:** Suppose we have another homomorphism  $\psi : Z_n \rightarrow H$  with  $\psi(x) = h$ . Then for all  $x^p \in Z_n$ ,

$$\varphi(x^p) = \varphi(x)^p = \psi(x)^p = \psi(x^p),$$

so that  $\varphi = \psi$  as the function match in values for all values in their domain. Thus,  $\varphi$  is unique.

- (3) Existence of such homomorphism: We define  $\varphi(n) = h^n$ . It is clearly a bijection. For all  $m, n \in \mathbb{Z}$ , we have

$$\varphi(m+n) = h^{m+n} = h^m h^n = \varphi(m) \varphi(n),$$

so that  $\varphi$  is a homomorphism.

For Uniqueness: Suppose we have another homomorphism  $\psi$  such that  $\psi(1) = h$ . Then

$$\psi(n) = \psi(n \cdot 1) = \psi(1)^n = \varphi(1)^n = \varphi(n),$$

So  $\varphi$  and  $\psi$  coincide on all the values.  $\psi = \varphi$  and so  $\varphi$  is unique.

- (4) We prove the first statement by induction on  $n$ . Base case :

For  $n = 1$ ,  $(1+p)^{p^0} = 1 \pmod{p}$ , which is obvious. Induction Hypothesis: Assume that

$$(1+p)^{p^{n-1}} = 1 \pmod{p^n}.$$

That is,  $(1+p)^{p^{n-1}} = 1 + kp^n$ . To prove that  $(1+p)^{p^n} = 1 \pmod{p^{n+1}}$ . Now,

$$\begin{aligned} (1+p)^{p^n} &= ((1+p)^{p^{n-1}})^p = (1+kp^n)^p \\ &= 1 + p(kp^n) + \binom{p}{2}(kp^n)^2 + \dots + \binom{p}{p}(kp^n)^p \\ &= 1 + p^{n+1}(k + \binom{p}{2}kp^{n-1} + \dots). \end{aligned}$$

Thus,  $(1+p)^{p^n} = 1 \pmod{p^{n+1}}$ .

## 5. Subgroup generated by subset of a group

This section discussed the solution (in the context of groups) to the problem written in *Theme II* (at the start of this chapter.) Nevertheless, there is a very nice document that sheds light upon the notion of mathematical structures – [Nicolas Bourbaki: Theory of Structures](#) – basically, the genesis of modern mathematical structure. So, we want to address the following.

Given a group  $G$  and a subset  $F$  does there exist a smallest subgroup  $H$  of  $G$  that contains  $F$ . The solution to this problem would be known as the subgroup generated by  $F$ . The existence of such *smallest object* is dependent on one observation, namely, if intersection of any collection of subgroup is also a group. But we have already proven this, hence, there is a solution to the problem of 'existence of smallest object.' In general, for other mathematical structures, say  $\mathcal{O}$ , such the smallest sub-object would exist if the intersection of any collection of sub-objects of  $\mathcal{O}$  is also a sub-object.

**Definition 5.1.** If  $A$  is any subset of the group  $G$ , define

$$\langle A \rangle = \bigcap_{\substack{A \subseteq H \\ H \leq G}} H.$$

This is called the *subgroup of  $G$  generated by  $A$* .

In essence,  $\langle A \rangle$  is the intersection of all subgroups of  $G$  which contain  $A$ . The question of whether it is a subgroup of  $G$  or not can be answered by the fact that the intersection of an arbitrary collection of subgroups of the group  $G$  is also a subgroup of  $G$ . This is very easy to verify. Hence,  $\langle A \rangle$  is indeed a subgroup of  $G$ .

**Lemma 5.1.** Let  $\mathcal{A} = \{H \leq G \mid A \subseteq H\}$ . Then  $\langle A \rangle$  is the minimal element of  $\mathcal{A}$ .

PROOF. First, note that  $\mathcal{A}$  is non-empty because  $G \in \mathcal{A}$ . Moreover,  $\langle A \rangle$  itself is a subgroup of  $G$  containing  $A$ , so  $\langle A \rangle \in \mathcal{A}$ . Now, any element (essentially group in their own right) of  $\mathcal{A}$  contains all elements in  $A$  and, therefore, will contain their product and inverses, i.e., contains  $\langle A \rangle$ . Thus,  $\langle A \rangle$  is indeed a minimal element in this construction.  $\square$

When  $A$  is the finite set  $\{a_1, a_2, \dots, a_n\}$  we write  $\langle a_1, a_2, \dots, a_n \rangle$  for the group generated by  $a_1, a_2, \dots, a_n$  instead of  $\langle \{a_1, a_2, \dots, a_n\} \rangle$ . If  $A$  and  $B$  are two subsets of  $G$  we shall write  $\langle A, B \rangle$  in place of  $\langle A \cup B \rangle$ .

This “top-down” approach to defining  $\langle A \rangle$  proves the existence and uniqueness of the smallest subgroup of  $G$  containing  $A$  but is not too enlightening as to how to construct the elements in it. As the word “generates” suggests, we now define the set, which is the closure of  $A$  under the group operation (and the process of taking inverses) and prove this set equals  $\langle A \rangle$ . Let

$$\overline{A} = \{a_1^{\epsilon_1} a_2^{\epsilon_2} \cdots a_n^{\epsilon_n} \mid n \geq 0, a_i \in A, \epsilon_i = \pm 1 \text{ for each } i\},$$

where  $\overline{A} = \{1\}$  if  $A = \emptyset$ , so that  $\overline{A}$  is the set of all finite products (called *words*) of elements of  $A$  and inverses of elements of  $A$ . Note that, the  $a_i$ 's need not be distinct, so  $a^2$  is written as  $aa$  in the notation defining  $\overline{A}$ . Note also that  $A$  is not assumed to be a finite (or even countable) set.

**Lemma 5.2.** Prove that if  $H$  is a subgroup of  $G$  then  $\langle H \rangle = H$ .

PROOF. First, we prove that  $H \subseteq \langle H \rangle$ . Note that,

$$\langle H \rangle = \bigcap_{\substack{H \subseteq K \\ K \leq G}} K$$

Let  $h \in H$  and  $H \subseteq K$ . Therefore,  $h \in K$  and  $h \in \langle H \rangle$ . Thus,  $H \subseteq \langle H \rangle$ . Next, we prove that  $\langle H \rangle \subseteq H$ . For this, first note that  $H$  is a subgroup and contains itself. Thus, one of the possible  $K$  is  $H$  itself. Now, let  $h \in \langle H \rangle$ . This means  $h \in \langle H \rangle = \bigcap_{\substack{H \subseteq K \\ K \leq G}} K$ , and by the previous statement since one of  $K$  is  $H$  itself.

Thus,  $h \in H$  and we get  $\langle H \rangle \subseteq H$ .  $\square$

Now, we prove an important result that repeats itself in the context of other algebraic structures. I have written an extensive document on Zorn's lemma, its various equivalences, and its application across the spectrum of mathematics.

**Definition 5.2 (Maximal subgroup).** A subgroup  $M$  of a group  $G$  is called a maximal subgroup if  $M \neq G$  and the only subgroups of  $G$  which contain  $M$  are  $M$  and  $G$ .

**Theorem 5.1 (Maximal subgroup of a finitely generated group).** Prove that every nontrivial finitely generated group possesses maximal subgroups. Let  $G$  be a finitely generated group, say  $G = \langle g_1, g_2, \dots, g_n \rangle$ , and let  $\mathcal{S}$  be the set of all proper subgroups of  $G$ . Then  $\mathcal{S}$  is partially ordered by inclusion ( $\subseteq$ ). Let  $\mathcal{C}$  be a chain in  $\mathcal{S}$ .

- Prove that the union,  $H$ , of all the subgroups in  $\mathcal{C}$  is a subgroup of  $G$ .
- Prove that  $H$  is a proper subgroup.
- Use Zorn's Lemma to show that  $\mathcal{S}$  has a maximal element (which is by the definition of a maximal subgroup).

PROOF. • Let  $\mathcal{C}$  be non-empty set, Define

$$H = \bigcup_{A \in \mathcal{C}} A.$$

Here  $A$  is a subgroup of  $G$ . Note that, each subgroup  $A$  in  $\mathcal{C}$  contains  $e$ . So,  $e \in H$ . Thus,  $H$  is a non-empty set. Now, let  $a, b \in H$ . Then there exists  $A_1, A_2 \in \mathcal{C}$  such that  $a \in A_1$  and  $b \in A_2$ . Now, as  $\mathcal{C}$  is a chain, so we have either  $A_1 \subseteq A_2$  or  $A_2 \subseteq A_1$ . Without loss of generality, let us assume that the latter is true. Thus,  $a, b$  are in  $A_1$ . As  $A_1$  is subgroup,  $ab^{-1}$  is in  $A_1$ . Thus, we have  $ab^{-1} \in H$ . So,  $H$  is indeed a subgroup.

Let  $g_1, g_2, \dots, g_n$  be generators of  $G$ . Assume for the sake of contradiction that  $H$  is not a proper subgroup. So,  $H = G$ . This means that  $H$  contains each generator  $g_i$  for  $1 \leq i \leq n$ . Consider the chain

$$\mathcal{C} = A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots \subseteq A_k.$$

This implies that,  $H$  contains all of the generators of  $G$  and  $A_k$  contains all the generators of  $G$ . Thus,  $A_k$  can generate every element of  $G$ . By lemma 5.2,  $A_k = \langle A_k \rangle$  and  $\langle A_k \rangle \subseteq G$ . This contradicts that  $A_k$  is a proper subgroup.

- The subgroup  $H$ , constructed as described, is also a proper subgroup. Thus,  $H$  belongs to  $\mathcal{S}$ . Each  $A_i$  in  $\mathcal{C}$  is a subset of  $H$ , as  $H$  is the union of all subgroups in  $\mathcal{C}$ , which makes  $H$  an upper bound for  $\mathcal{C}$ . We have demonstrated that every chain in the nonempty partially ordered set  $\mathcal{S}$  has an upper bound. According to Zorn's Lemma,  $\mathcal{S}$  must Therefore, contain a maximal element.  $\square$

**Corollary 5.1.** If  $G$  is a finite group and  $H$  is a proper subgroup of  $G$ , then there exists a maximal subgroup of  $G$  containing  $H$ .

**Remark 5.1.** This result does not remain true if we drop the assumption  $d = d(G)$ .

For example, let  $G = \mathbb{F}_2^d$  be the additive group of a vector space of dimension  $d \geq 2$  over the field  $\mathbb{F}_2$  with 2 elements and let

$$g_1 = (1, 0, \dots, 0), \dots, g_d = (0, \dots, 0, 1), g_{d+1} = (1, 0, 0, \dots, 0).$$

Here in general  $g_n$  has 1 in the  $n$ th coordinate and 0 elsewhere, and we define  $g_{d+1} = (1, 0, 0, \dots, 0)$ . Let  $M = \{(x_1, \dots, x_d) \in \mathbb{F}_2^d \mid a_1x_1 + \dots + a_dx_d = 0\}$  be a maximal subgroup of  $G$ . If  $i \leq d$ , then  $g_i \in M$  only when  $a_i = 0$ . Therefore,

$$\overline{M} = \{(x_1, \dots, x_d) \in \mathbb{F}_2^d \mid x_1 + \dots + x_d = 0\}$$

is the unique maximal subgroup of  $G$  with  $g_i \notin \overline{M}$  for every  $i \leq d$ . However,  $g_{d+1} \in \overline{M}$ . Hence, Every maximal subgroup of  $G$  contains at least one of the  $d+1$  elements  $g_1, \dots, g_d, g_{d+1}$ .

**Problem 5.1.** Prove that if  $A$  is a subset of  $B$ , then  $\langle A \rangle \leq \langle B \rangle$ . Give an example where  $A \subseteq B$  with  $A \neq B$  but  $\langle A \rangle = \langle B \rangle$ .

**Solution 5.1.**  $\langle A \rangle$  is the intersection of all subgroups of  $G$  containing  $A$ .  $\langle B \rangle$  is intersection of all subgroups of  $G$  containing  $B$ . Since  $A \subset B$ , any subgroup of  $G$  that contains  $B$  must contain  $A$ . Hence,  $\langle A \rangle \leq \langle B \rangle$ . For the second part of the problem, consider the additive modular group  $\mathbb{Z}/12\mathbb{Z}$ . Consider the set  $\{4\}$  and  $\{4, 8\}$ . Then subgroup generated by them is same  $\{0, 4, 8\}$  even though  $A \neq B$ .



## CHAPTER 5

# Group homomorphism and quotient structure

### 1. Normal Subgroups and Quotient Groups

Recall the definition 7.3 of the *kernel* associated with a group homomorphism and normal subgroups 1.9. The following lemma is a crucial result that will often resurface in many places in subsequent chapters.

**Lemma 1.1 (Kernel is a normal subgroup).** Let  $\varphi : G \rightarrow H$  be a group homomorphism and  $E \leq H$  be a subgroup. Prove that  $\varphi^{-1}[E]$  (the preimage of  $E$  under  $\varphi$ ) is a subgroup of  $G$ . If  $E \trianglelefteq H$ , then prove that  $\varphi^{-1}[E] \trianglelefteq G$ . Deduce that  $\ker \varphi \trianglelefteq G$ .

PROOF. Let  $x, y \in \varphi^{-1}[E]$ . Then  $\varphi(x), \varphi(y) \in E$ , and since  $E$  is a subgroup,

$$\varphi(x)\varphi(y)^{-1} = \varphi(xy^{-1}) \in E$$

by the characterization 1.1 of subgroups. So  $xy^{-1} \in \varphi^{-1}[E]$ . Again by the same characterization of subgroups, we have,  $\varphi^{-1}[E] \leq G$ .

To prove the second assertion, suppose  $E$  is normal in  $H$ . Let  $g \in G$  and  $x \in \varphi^{-1}[E]$ . Then

$$\varphi(gxg^{-1}) = \varphi(g)\varphi(x)\varphi(g)^{-1} \in \varphi(g)E\varphi(g)^{-1} = E,$$

where the last equality follows from  $E$  being normal in  $H$ . Thus,  $gxg^{-1} \in \varphi^{-1}[E]$ . Hence,  $\varphi^{-1}[E]$  is normal in  $G$ .

For the last assertion, note that trivial subgroup  $\{1\} \leq G$  is trivially normal for all  $G$ . By definition,  $\ker \varphi = \varphi^{-1}(1)$ . So  $\ker \varphi$  is normal in  $G$ .  $\square$

**Motivation for Cosets and Partition.** Let  $G$  be a group, and let  $H$  be a subgroup of  $G$ . For each  $g \in G$ , we define the left coset of  $H$  as

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

We are motivated to define the concept of a coset by the fact that, given a group homomorphism  $\varphi : G \rightarrow H$ , the pre-image  $\varphi^{-1}(y)$  of an element  $y \in \varphi(G)$  is a left coset of the kernel  $K = \ker(\varphi)$ . Specifically, we have  $\varphi^{-1}(y) = xK = \{xk \mid k \in K\}$  for any element  $x \in G$  such that  $\varphi(x) = y$ . This follows from the fact that for all  $k \in K$ ,  $\varphi(xk) = \varphi(x)\varphi(k) = \varphi(x) = y$ .

For any function  $f : G \rightarrow H$ , where  $G$  and  $H$  are sets, the preimages of elements in  $H$  form a partition of  $G$ . This means that the collection  $\{f^{-1}(h) \mid h \in H\}$  of these preimages is mutually, exclusive, and collectively cover  $G$ . In particular,

$$G = \bigcup_{h \in H} f^{-1}(h).$$

Thus, we conclude the following important observations.

- (1) When  $K$  is the kernel of a group homomorphism with domain  $G$ , the left cosets of  $K$  in  $G$  also partition  $G$ . This follows from our previous discussion that preimage of an element in group  $H$  is essentially the left coset of kernel  $K$  of the group homomorphism  $\varphi$ .
- (2) Since kernels are equivalent to normal subgroups, the left cosets of a normal subgroup partition the group.

Define  $G/H := \{gH \mid g \in G\}$ , the left cosets of  $H$  in  $G$ .

Similarly, we can define right cosets, however, we will only work with left cosets as all results can easily be translated into right cosets. It is just a matter of choice. Now, we prove that we can remove the normality condition from the previous statement:

**Theorem 1.1 (Cosets partition group).** Let  $H \leq G$ . Then:

- For all  $a, b \in G$ , either  $aH = bH$  or  $aH \cap bH = \emptyset$ .
- The left cosets of  $H$  in  $G$  partition  $G$ .

PROOF. • Let  $a, b \in G$ . Suppose that  $aH \cap bH \neq \emptyset$ . Then there exists a  $x \in aH \cap bH$ . So we can write  $x = ah_1 = bh_2$  for some  $h_1, h_2 \in H$ . From this, we get  $a = bh_2h_1^{-1}$  and  $b = ah_1h_2^{-1}$ .

Let  $y \in aH$ . Then  $y = ah$  for some  $h \in H$ . Now  $a = bh_2h_1^{-1}$ , so  $y = bh_2h_1^{-1}h$ . We have  $h_2h_1^{-1}h \in H$ . Thus,  $y \in bH$ . Thus,  $aH \subseteq bH$ .

Similarly, we can establish that  $bH \subseteq aH$ . So  $aH = bH$ . This proves that two left cosets are either mutually disjoint or they completely overlap. This guarantees that the *left cosets of  $H$  in  $G$  will partition  $G$* . More formally we must show that  $G$  can be written as union of left cosets. This is what we prove in the next part.

- Formally, we must show that  $G = \bigcup_{C \in L} C$ , where  $L$  is the set of all left cosets of  $H$  in  $G$ .

Let  $x \in G$ . Then  $x = x \cdot e \in xH$  (since  $e \in H$ ). This shows that  $G \subseteq \bigcup_{C \in L} C$ .

To show that  $\bigcup_{C \in L} C \subseteq G$ , we simply observe that each left coset  $C$  of  $H$  in  $G$  is a subset of  $G$ . By part (i), the union is disjoint. This completes the proof of (ii).  $\square$

Let  $G$  be a group, and  $H$  be a subgroup of  $G$ . Let  $C_1, \dots, C_k$  be the distinct cosets of  $H$ . This tells us that  $G$  is equal to the disjoint union  $G = C_1 \cup C_2 \cup \dots \cup C_k$ .

**Theorem 1.2 (Quotient group).** Let  $G$  be a group and  $H$  be a normal subgroup of  $G$ . Let

$$Q = \{xH : x \in G\}$$

the set of all left cosets of  $N$  in  $G$ . Define a binary operation  $\star$  on  $Q$  by

$$xH \star yH = (xy)H$$

where  $xy$  is the product taken in  $G$ . Then  $(Q, \star)$  becomes a group and will be known as the quotient group of  $G$  by  $H$ , and written  $G/H$ .

PROOF. We have  $H$  is a normal subgroup, i.e.,  $ghg^{-1} \in H$  for all  $g \in G$  and all  $h \in H$ . To prove the operation stated above is well-defined, let  $g_1, g'_1 \in g_1H$  and  $g_2, g'_2 \in g_2H$ . We may write

$$g'_1 = g_1h_1 \quad \text{and} \quad g'_2 = g_2h_2, \quad \text{for some } h_1, h_2 \in H.$$

To prove that  $g'_1g'_2 \in g_1g_2H$ :

$$g'_1g'_2 = (g_1h_1)(g_2h_2) = g_1(g_2g_2^{-1})h_1g_2h_2 = (g_1g_2)(g_2^{-1}h_1g_2)h_2$$

Now, as  $H$  is a normal subgroup  $(g_2^{-1}h_1g_2) \in H$  and as  $H$  is closed under products, so  $(g_2^{-1}h_1g_2)h_2 \in H$ . Thus  $g'_1g'_2 \in g_1g_2H$ . This proves that the operation is well defined. The identity of  $G/H$  is  $eH$ , and we simply write this as  $H$ . The inverse of an element  $gH$  is  $g^{-1}H$ . With this data in place, it is obvious that if  $H \trianglelefteq G$ , then  $G/H$  has a group structure.  $\square$

**Theorem 1.3.** Now, we prove the converse of theorem 1.2. i.e., if the operation  $(g_1H)(g_2H) = (g_1g_2)H$  is a well defined operation, then  $H$  is a normal subgroup.

PROOF. Now, we prove that the operation  $(g_1H)(g_2H) = (g_1g_2)H$  is a well defined operation. First we assume that the operation  $(g_1H)(g_2H) = (g_1g_2)H$  is well defined. To prove that  $H$  is normal subgroup. Operation is well defined, so it must follow the axiom of substitution. That is, for all  $g_1, g_2, g'_1, g'_2$

$$g_1, g'_1 \in g_1H \quad \text{and} \quad g_2, g'_2 \in g_2H \quad \text{then} \quad g_1g_2H = g'_1g'_2H.$$

Let  $g$  be an arbitrary element of  $G$  and  $h$  be an arbitrary element of  $H$ . Let

$$g_1 = 1, \quad g'_1 = h \quad \text{and} \quad g_2 = g^{-1}, \quad \text{and applying the assumption above, we get that}$$

$$1g^{-1} = H = hg^{-1}H \quad \text{i.e.,} \quad g^{-1}H = hg^{-1}H.$$

Since  $1 \in H$ ,  $hg^{-1} \cdot 1 \in hg^{-1}H$ . Thus,  $hg^{-1} \in g^{-1}H$ , hence  $hg^{-1} = g^{-1}h_1$ , for some  $h_1 \in H$ . Multiplying both sides on the left by  $g$  gives  $ghg^{-1} = h_1 \in H$ . Thus,  $H$  is the normal subgroup of  $G$ .  $\square$

**Theorem 1.4.** Let  $G$  be a group, and let  $H$  be a normal subgroup of  $G$ .

(a) The collection of cosets  $G/H$  is a group via the well-defined group operation:

$$g_1H \cdot g_2H = g_1g_2H.$$

(b) The map

$$\varphi : G \rightarrow G/H, \quad \varphi(g) = gH,$$

is a homomorphism whose kernel is

$$\ker(\varphi) = H.$$

(c) Let

$$\psi : G \rightarrow G_0$$

be a homomorphism with the property that  $H \subseteq \ker(\psi)$ . Then there is a unique homomorphism

$$f : G/H \rightarrow G_0$$

satisfying

$$f(gH) = \psi(g).$$

(d) If we take  $H = \ker(\psi)$  in (c), then the homomorphism

$$f : G/\ker(\psi) \rightarrow G_0$$

is injective. In particular, we get an isomorphism:

$$f : G/\ker(\psi) \xrightarrow{\sim} \text{Image}(\lambda) \subseteq G_0.$$

**Proof.**

(a) This we already proved.

(b) To verify that  $\varphi$  is a homomorphism, calculate:

$$\varphi(g_1)\varphi(g_2) = g_1H \cdot g_2H = g_1g_2H = \varphi(g_1g_2).$$

The kernel of  $\varphi$  is given by:

$$\ker(\varphi) = \{g \in G : \varphi(g) = eH\} = \{g \in G : gH = H\} = H.$$

(c) Define  $f : G/H \rightarrow G_0$  via the following procedure:

- (1) Let  $C \in G/H$  represent a coset.
- (2) Select any  $g \in G$  such that  $C = gH$ .
- (3) Set  $f(C) = \psi(g)$ .

A potential issue arises since step (2) permits multiple choices for  $g$ . To address this, we must prove the following statement:

(1) If  $g_0H = gH$ , then  $\psi(g_0) = \psi(g)$ .

The equality  $g_0H = gH$  implies  $g_0 = gh$  for some  $h \in H$ . Using this, compute:

$$\begin{aligned} \psi(g_0) &= \psi(gh) && (\text{since } g_0 = gh) \\ &= \psi(g) \cdot \psi(h) && (\text{as } \psi \text{ is a group homomorphism}) \\ &= \psi(g) \cdot e_0 && (\text{since } h \in H \text{ and } H \subseteq \ker(\psi)) \\ &= \psi(g). \end{aligned}$$

This confirms the validity of statement (1). Consequently, the 3-step procedure defines a well-defined mapping:

$$f : G/H \rightarrow G_0.$$

To verify that  $f$  is a homomorphism:

$$f(g_1g_2H) = \psi(g_1g_2) = \psi(g_1) \cdot \psi(g_2) = f(g_1H) \cdot f(g_2H).$$

Finally, for any given homomorphism  $\psi$ , there exists exactly one map  $f$  such that  $\psi(g) = f(gH)$ , as this equality uniquely determines  $f$  in terms of  $\psi$ .

(d) Taking  $H = \ker(\psi)$ , part (c) gives a well-defined homomorphism:

$$f : G/H \rightarrow G_0,$$

defined by  $f(gH) = \psi(g)$ . To establish that  $f$  is injective:

Assume  $gH \in \ker(f)$ . This implies:

$$f(gH) = \psi(g) = e_0,$$

so  $g \in \ker(\psi) = H$ . Thus,  $gH = H$ , which represents the identity element (identity coset) in  $G/H$ . Therefore,  $\ker(f)$  contains only the identity element of  $G/H$ . Thus,  $f$  is injective.

Lastly, as  $f : G/H \rightarrow G_0$  is injective and maps onto its image by definition, we conclude that:

$$f : G/H \rightarrow \text{Im}(f)$$

is a homomorphism that is bijective. Therefore  $f$  is an isomorphism.

## 2. Size of the Cosets, Lagrange's Theorem, and Consequences

**Lemma 2.1.** Let  $G$  be a group and  $H \subseteq G$ . Let  $a \in G$ . Then the function  $f : H \rightarrow aH$  given by  $h \mapsto ah$  is a bijection. Thus, all left cosets of  $H$  in  $G$  have the same size as  $H$ , and hence the same size as each other.

PROOF. Let  $x, y \in H$ , and suppose that  $f(x) = f(y)$ . Then  $ax = ay$ , so  $x = y$  by cancellation law. So,  $f$  is injective.

Let  $z \in aH$ . Then  $z = ah$  for some  $h \in H$ , and  $f(h) = z$ . So,  $f$  is surjective.  $\square$

**Definition 2.1.** Let  $G$  be a group,  $H \leq G$  and  $C$  be a coset of  $H$  in  $G$ . An element of  $C$  is called a representative of  $C$ , or a coset representative for  $C$ .

Elements in a coset do not depend upon the choice of the representative. To be more precise, suppose that  $H \leq G$  and  $a \in G$ . Let  $C = aH$ . Since  $e \in H$ , we always have  $a = ae \in aH$ , so  $a$  is a coset representative for  $C$ . Also, if  $b \in aH$ , then likewise  $b = be \in bH$ , so we must have  $bH = aH$ . Thus, any coset representative  $b$  for  $C$  has the property that  $C = bH$ .

**Definition 2.2 (Index).** Let  $H \leq G$ . The index of  $H$  in  $G$  is the number of left cosets of  $H$  in  $G$ . The index of  $H$  in  $G$  is denoted  $[G : H]$ .

The index of a subgroup in a group is either a positive integer or else it is infinite. In particular, a finite group can only have subgroups of finite index. However, an infinite group can have subgroups of finite or infinite index. Intuitively it is clear that the bigger a subgroup is, the smaller the index of that subgroup will be. A very special case of the index is 1 and the index of  $H$  in  $G$  is one if and only if  $H = G$ ;

We can form a complete set of coset representatives for  $H$  in  $G$  simply by choosing one element from every left coset of  $H$  in  $G$ ; this must work by Theorem 1.1. Moreover, for finite cases, we do not need to worry much but in case of an infinite set, the fact that it is possible in general to simultaneously choose an element from each left coset of  $H$  in  $G$  is actually equivalent to the Axiom of Choice.

**Lemma 2.2.** If  $G$  is a group and  $H \leq G$ , then the number of left cosets of  $H$  in  $G$  equals the number of right cosets of  $H$  in  $G$ .

PROOF. The left cosets of  $H$  in  $G$  are the sets of the form

$$gH = \{gh \mid h \in H\}, \quad g \in G.$$

The right cosets of  $H$  in  $G$  are the sets of the form

$$Hg = \{hg \mid h \in H\}, \quad g \in G.$$

**Construction of a bijection.** Now we construct an explicit bijection between the set of left cosets of  $H$  and the set of right cosets of  $H$ . Define the map  $\varphi : G/H \rightarrow H \backslash G$  by

$$\varphi(gH) = Hg^{-1},$$

where  $G/H$  denotes the set of left cosets and  $H \backslash G$  denotes the set of right cosets.

**Injectivity of  $\varphi$ .** Suppose  $\varphi(g_1H) = \varphi(g_2H)$ . Then

$$Hg_1^{-1} = Hg_2^{-1}.$$

This means that there exists  $h \in H$  such that  $g_1^{-1} = hg_2^{-1}$ , or equivalently,

$$g_1 = g_2h^{-1}.$$

Since  $h^{-1} \in H$ , it follows that  $g_1H = g_2H$ . Thus,  $\varphi$  is injective.

**Surjectivity of  $\varphi$ .** Let  $Hg$  be any right coset of  $H$ . We claim there exists a left coset  $g'H$  such that  $\varphi(g'H) = Hg$ . Set  $g' = g^{-1}$ . Then

$$\varphi(g'H) = \varphi(g^{-1}H) = Hg.$$

Thus,  $\varphi$  is surjective. □

Our previous work has rich implications for the structure of finite groups. Many of these implications follow immediately from the next result.

**Theorem 2.1 (Lagrange's Theorem).** Let  $G$  be a finite group. If  $H < G$ , then  $|H|$  divides  $|G|$ .

PROOF. Suppose that  $G$  is a finite group and  $H \leq G$ . By the second part of Theorem 1.1, we have  $G = \bigcup C_i$ , where  $C_i$  is the set of all left cosets of  $H$  in  $G$ . Therefore,

$$|G| = \sum_{i=1}^{[G:H]} |C_i| = |H|[G:H]$$

by Lemma 2.1, we have  $|C_i| = |H|$  for all  $i \in C$ . It follows that  $|G| = |H| \cdot [G:H]$ . The result now follows from the Definition 2.2. □

**Corollary 2.1.** If  $G$  is a finite group and  $H \trianglelefteq G$ , then  $|G| = |H| \cdot |G/H|$ .

PROOF. By the definitions of  $G/H$  and by Theorem 1.2 and  $|G:H|$  (definition 2.2), we have

$$|G/H| = |G:H|$$

By Lagrange's Theorem,

$$\frac{|G|}{|H|} = |G:H| = |G/H|.$$

□

Lagrange's Theorem is the first combinatorial result in the group theory and it provides one of the most basic and important tools in the study of finite groups. As an example of its power, we next show that any finite group of prime order has a very special form.

**Theorem 2.2.** Every group of prime order is cyclic.

PROOF. Suppose that  $G$  is a group of order  $p$ , where  $p$  is prime. Then  $|G| > 1$ , so there exists a  $x \neq e$ . Let  $H = \langle x \rangle$  and  $n = |H|$ . By Lagrange's Theorem, we have  $n \mid p$ . Also, we have  $n > 1$ , since  $x \neq e$  and  $e$  belongs to  $H$ . Therefore,  $n = p = |G|$ . Since  $H \leq G$ , it follows that  $H = G$ . Thus,  $\langle x \rangle = G$ , so  $G$  is cyclic.  $\square$

Our next application of Lagrange's theorem involves cyclic subgroups of a finite group. Before presenting this result, we make a standard definition.

**Definition 2.3.** Let  $G$  be a group and  $x \in G$ . The order of  $x$  is the order of the cyclic subgroup generated by  $x$ . We denote the order of  $x$  by  $|x|$ .

By definition of the order of  $x$ , we have  $|x| = |\langle x \rangle|$ . Just as with the index of a subgroup, the order of a group element may be a positive integer or it may be infinite.

Here finally is the promised result:

**Lemma 2.3.** Let  $G$  be a finite group, and let  $x \in G$ . Then,

$$|x| = \min\{n \in \mathbb{Z} \mid n > 0 \text{ and } x^n = e\},$$

and  $|x|$  divides the order  $|G|$  of the group  $G$ .

PROOF. Proof of this lemma becomes obvious if we just recall the following lemma and its proof from our discussion on order and its properties in the first chapter.

**Lemma 2.4.** Let  $G$  be a group, and  $a \in G$ , and suppose that  $a^k = e$  for some positive integer  $k$ . Let

$$m = \min\{n \in \mathbb{Z} : n \geq 1 \text{ and } a^n = e\}$$

be the smallest positive integer with this property.

- (a) Prove that  $|\langle a \rangle| = m$ .
- (b) Prove that for any integer  $j$ , we have  $a^j = e$  if and only if  $m \mid j$ .
- (c) Deduce that for two integers  $i$  and  $j$ , we have  $a^i = a^j$  if and only if  $i \equiv j \pmod{m}$ .

PROOF. (a)  $|a| = m$ . We prove that  $\langle a \rangle = \{e, a, \dots, a^{m-1}\}$ . i.e.,  $|\langle a \rangle| = m$ . Certainly, the elements  $e, a, \dots, a^{m-1}$  are in  $\langle a \rangle$ . Next, suppose that  $a^p$  is an arbitrary member of  $\langle a \rangle$ . By the division algorithm, there exist integers  $q$  and  $r$  such that

$$p = mq + r \text{ with } 0 \leq r < m.$$

Then  $a^p = a^{mq+r} = a^{mq}a^r = (a^m)^qa^r = e^qa^r = a^r$ , so that  $a^p \in \{e, a, a^2, \dots, a^{m-1}\}$ . This proves that  $|\langle a \rangle| = m$ .

- (b) ( $\implies$ ) Suppose  $a^j = e$ . By the division algorithm, we can write  $j = qm + r$  for some integers  $q$  and  $0 \leq r < m$ . Then  $e = a^j = a^{qm+r} = (a^m)^qa^r = e^qa^r = a^r$ . Since  $0 \leq r < m$  and  $m$  is the smallest positive integer such that  $a^m = e$ , we must have  $r = 0$ . Thus,  $j = qm$ , which means  $m \mid j$ .
- ( $\impliedby$ ) Suppose  $m \mid j$ . Then  $j = km$  for some integer  $k$ . Therefore,  $a^j = a^{km} = (a^m)^k = e^k = e$ .
- (c) ( $\implies$ ) Suppose  $a^i = a^j$ . Then  $a^i(a^j)^{-1} = e$ , so  $a^i a^{-j} = a^{i-j} = e$ . From part (b), this implies that  $m \mid (i - j)$ , which means  $i \equiv j \pmod{m}$ .
- ( $\impliedby$ ) Suppose  $i \equiv j \pmod{m}$ . Then  $i - j = km$  for some integer  $k$ . Therefore,  $a^i = a^{j+km} = a^j a^{km} = a^j (a^m)^k = a^j e^k = a^j$ .

$\square$

Since  $G$  is finite, there is a positive integer  $k$  such that  $x^k = e$ . Now, the formula  $|x|$  follows from the above lemma 2.4, and the definition of  $|x|$  (Definition 2.3). That  $|x|$  divides  $|G|$  follows from the definition of  $|x|$  (definition 2.3) and Lagrange's Theorem 2.1. This proves the Lemma.  $\square$

**Corollary 2.2.** . If  $G$  is a finite group and  $x \in G$ , then  $x^{|G|} = e$ .

PROOF. By Lemma 2.3, we can write  $|G| = |x| \cdot k$  for some  $k \in \mathbb{Z}$ . Thus,  $x^{|G|} = x^{|x| \cdot k} = (x^{|x|})^k = e^k = e$ .  $\square$

**Theorem 2.3** (Cauchy's Theorem due to James McKay). Let  $G$  be a finite group and  $p$  be a prime dividing  $|G|$ . Let  $\mathcal{S}$  denote the set of  $p$ -tuples of elements of  $G$  the product of whose coordinates is 1 :

$$\mathcal{S} = \{(x_1, x_2 \dots x_p) | x_i \in G \text{ and } x_1 x_2 \dots x_p = 1\}$$

- (1) Show that  $\mathcal{S}$  has  $|G|^{p-1}$  elements, hence have order divisible by  $p$ .  
Define a relation  $\sim$  on  $\mathcal{S}$  by letting  $\alpha \sim \beta$  if  $\beta$  is the cyclic permutation of  $\alpha$ .
- (2) Show that a cyclic permutation of an element of  $\mathcal{S}$  is again an element of  $\mathcal{S}$
- (3) Prove that  $\sim$  is an equivalence relation on  $\mathcal{S}$ .
- (4) Prove that an equivalence class contains a single element if and only if it is of the form  $(x, x \dots x)$  with  $x^p = 1$ .
- (5) Prove that every equivalence class has order 1 or  $p$  (this uses the fact that  $p$  is prime). Deduce that  $|\mathcal{S}| = |G|^{p-1} = k + pd$ , where  $k$  is the number of classes of size 1 and  $p$  is the number of classes of size  $p$ .

PROOF. (1) Let  $\mathcal{S}'$  be the set of all  $(p-1)$ -tuples of elements of  $G$ . Then  $\mathcal{S}'$  has  $|G|^{p-1}$  elements because each element has  $|G|$  choices. We prove that  $\mathcal{S}$  and  $\mathcal{S}'$  are in bijection. To do so, define  $\varphi : \mathcal{S}' \rightarrow \mathcal{S}$  by

$$\varphi(x_1, x_2, \dots, x_{p-1}) = (x_1, x_2, \dots, x_{p-1}, (x_1 x_2 \dots x_{p-1})^{-1}).$$

Let  $\alpha, \beta \in \mathcal{S}'$ . If  $\varphi(\alpha) = \varphi(\beta)$ . Then each of  $p$  co-ordinate must be same. In particular, their first  $p-1$  coordinates must be equal. Using this and the definition of  $\varphi$ , we see that  $\varphi$  is injective. Now, suppose  $\alpha = (u_1, u_2, \dots, u_{p-1}, u_p) \in \mathcal{S}$ . Then the image of the first  $p-1$  coordinates,  $(u_1, u_2, \dots, u_{p-1})$  of  $\alpha$  under  $\varphi$  will be  $\alpha$ , because the last coordinate is completely determined by the others (as it is written as the inverse of product of them and group inverses are unique). Therefore,  $\varphi$  is a bijection, so  $|\mathcal{S}| = |\mathcal{T}| = |G|^{p-1}$ .  $\square$

- (2) Let  $u = (x_1, x_2, \dots, x_p) \in \mathcal{S}$ . Then  $x_1 x_2 \dots x_p = 1$ . Now, left multiply by  $x_p$  to get  $x_p x_1 x_2 \dots x_{p-1} x_p = x_p 1$ . Now, right multiply by  $x_p^{-1}$  to get

$$x_p x_1 x_2 \dots x_{p-1} = x_p x_p^{-1} = 1.$$

Thus,  $(x_p, x_1, x_2, \dots, x_{p-1}) \in \mathcal{S}$ . We can continue this process  $p-1$  times and see that all cyclic permutations of  $u$  are in  $\mathcal{S}$ .

- (3) Let  $u, v, w$  be arbitrary.  $u$  is a cyclic permutation of itself. So  $u \sim u$ . Thus,  $\sim$  is *reflexive* relation on  $\mathcal{S}$ .

If  $u \sim v$ , then  $v$  is a cyclic permutation of  $u$ , so by applying the inverse of this permutation we see that  $u$  is a cyclic permutation of  $v$ . Therefore  $v \sim u$ , and  $\sim$  is *symmetric* relation on  $\mathcal{S}$ .

To prove this given relation is transitive, let  $u \sim v$  and  $v \sim w$ . This means,  $w$  is a cyclic permutation of  $v$  and  $v$  is a cyclic permutation of  $u$ . By applying the composition of these two permutations, we conclude that  $w$  is a cyclic permutation of  $u$ . Thus,  $\sim$  is *transitive*. Hence,  $\sim$  is an equivalence relation on  $\mathcal{S}$ .

- (4) Let  $[u]_{\sim}$  be the equivalence class of  $u$  under the relation  $\sim$ . By definition of the given equivalence relation, the equivalence class of  $u$  is essentially all possible cyclic permutations of  $u$ . Now, suppose that  $[u]_{\sim}$  contains only one element, namely  $u$  itself. This means every cyclic permutation of  $u$  is just  $u$  itself. Suppose that  $u = (u_1, u_2, \dots, u_p)$ . One of its cyclic permutation is  $(u_p, u_1, u_2, \dots, u_{p-1})$ . Since every cyclic permutation of  $u$  is equal to  $u$  itself (under the hypothesis), we have  $u_1 = u_p, u_2 = u_1, u_3 = u_2, \dots, u_{p-1} = u_p$ . This means  $u = (u_1, u_1, \dots, u_1)$ .

Claim in backward direction is trivial.

- (5) Suppose that  $[u]_{\sim}$  does not have a single element. Then, a  $p$ -tuple, can have  $p$  distinct cyclic permutation. In particular, the equivalence class of  $u = (u_1, u_2 \dots u_p)$  consists of following elements

$$\{u, (u_p, u_1, u_2, \dots, u_{p-1}), (u_{p-1}, u_p, u_1, \dots, u_{p-2}), \dots, (u_2, u_3, \dots, u_p, u_1)\}.$$

In order to obtain the number of elements in  $\mathcal{S}$ , it is enough to count how many distinct equivalence classes of these two types are. Two types here refer to the equivalence class with just one element the and equivalence class with  $p$  elements. To obtain the number of elements in  $\mathcal{S}$ ,

we need to add the number of elements in each equivalence class. This idea works perfectly here or in any setting as long as we are dealing with a mutually disjoint collection, and this is exactly the case here. So, if there are  $k$  equivalence classes of size 1 then this corresponds to  $k$  elements in  $\mathcal{S}$ . If there are  $d$  equivalence classes of size  $p$ , then they correspond to  $pd$  elements of  $\mathcal{S}$ . Thus,  $|\mathcal{S}| = |G|^{p-1} = k + pd$ .

- (6) **Culmination of Cauchy's theorem for finite group.** Since  $p$  divides  $|G|$ , it will divide any power of  $|G|$  and thus,  $p$  divides  $|G|^{p-1} \implies p|(k + pd)$ . This means,  $p|k$ . As  $p$  is prime and  $p|k$  so,  $k > 1$ . This means other than  $(1, 1, 1, \dots, 1)$  (a  $p$ -tuple), there is at least one more equivalence class which consists of exactly one element  $(x, x, x, \dots, x)$ . This equivalence class is an element of  $\mathcal{S}$  such that  $x \neq 1$ . By definition, elements in  $\mathcal{S}$  we must have  $x \cdot x \cdot x \cdot \dots \cdot x$  ( $p$ -times)  $= x^p = 1$ . That is,  $G$  contains a non-identity element of order  $p$ . This closes the proof. That is, given any finite group, if  $p$  is a prime divisor of  $|G|$ , then  $G$  contains an element and thus, a subgroup of order  $p$ . □

**Definition 2.4 (Product of subgroups).** Let  $H$  and  $K$  be subgroups of a group and define

$$HK = \{hk \mid h \in H, k \in K\}.$$

**Theorem 2.4.** If  $H$  and  $K$  are finite subgroups of a group, then

$$|HK| = \frac{|H||K|}{|H \cap K|}.$$

PROOF. First note that by definition 2.4,  $HK$  is just the union of left cosets of  $K$ . That is,

$$HK = \bigcup_{h \in H} hK.$$

Recall that the number of elements in a coset is the same as the size of the subgroup. So, each coset of  $K$  has  $|K|$  elements. Thus, in order to prove this claim it suffices to find the number of distinct left cosets of the form  $hK$ , where  $h \in H$ .

However, from the equality of two cosets, we know that  $h_1K = h_2K$  for  $h_1, h_2 \in H$  if and only if  $h_2^{-1}h_1 \in K$ . Also,  $h_2^{-1}h_1 \in H$ .

Therefore,

$$h_1K = h_2K \iff h_2^{-1}h_1 \in H \cap K \iff h_1(H \cap K) = h_2(H \cap K).$$

Hence, the number of distinct cosets of the form  $hK$ , for  $h \in H$ , is the number of distinct cosets  $h(H \cap K)$ , for  $h \in H$ . However, the latter number, by Lagrange's Theorem, equals

$$\frac{|H|}{|H \cap K|}.$$

This proves that  $HK$  consists of  $\frac{|H|}{|H \cap K|}$  distinct cosets of  $K$ . This gives the stated relation:

$$|HK| = \frac{|H||K|}{|H \cap K|}.$$

□

**Corollary 2.3 (Some important consequences).** (1) Let  $H \leq G$ . Prove that the map  $x \mapsto x^{-1}$  sends each left coset of  $H$  in  $G$  onto a right coset of  $H$  and gives a bijection between the set of left cosets and the set of right cosets of  $H$  in  $G$  (hence the number of left cosets of  $H$  in  $G$  equals the number of right cosets).

(2) Let  $G = S_n$  and for fixed  $i \in \{1, 2, \dots, n\}$  and  $G_i$  be the stabilizer of  $i$ . Prove that  $G_i \cong S_{n-1}$ .

(3) Use Lagrange's theorem in the multiplicative group  $(\mathbb{Z}/p\mathbb{Z})^*$  to prove Fermat's little theorem: If  $p$  is prime then  $a^p \equiv a \pmod{p}$  for all  $a \in \mathbb{Z}$ .

PROOF. (1) Consider an element  $kH \in K/H$ . Let  $\phi : \phi(gH) = (gH)^{-1} = Hg^{-1}$ . First, we prove that this map is indeed well defined. Let  $gH = g_1H$  for some  $g, g_1 \in G$ . To prove that  $Hg_1^{-1} = Hg^{-1}$ . Let  $x \in Hg^{-1}$ . So by definition  $x = hg^{-1}$  for some  $h \in H$ .  $gh^{-1} = x^{-1}$ , as  $x^{-1} = gh^{-1} \in gH$ .  $x^{-1} \in g_1H$ . So  $x^{-1} = g_1h_1$  for some  $h_1 \in H$ . Now,  $x^{-1} = g_1h_1$ , so  $x = h_1^{-1}g_1^{-1}$ . This means  $x \in Hg_1^{-1}$ . So,  $Hg^{-1} \subset Hg_1^{-1}$ . The proof of reverse inclusion is analogous.

Injective. Assume that  $Hg^{-1} = Hg_1^{-1}$  for some  $g, g_1$ . To prove that  $gH = g_1H$ .

Let  $x \in gH$ . Now, we have  $x = gh$  for some  $h$ .  $x^{-1} = h^{-1}g^{-1}$ . So  $x^{-1} \in Hg^{-1} = Hg_1^{-1}$ .

As  $x^{-1} \in Hg_1^{-1}$ ,  $x^{-1} = h_1g_1^{-1}$ . So  $x$  is  $g_1h_1^{-1}$ . So we have that  $x \in g_1H$ . So we get  $gH \subseteq g_1H$ .

Now, we prove that  $g_1H \subseteq gH$ . Let  $y$  be an arbitrary element of  $g_1H$ .  $y = g_1h_1$  for some  $h_1 \in H$ .  $y^{-1} = h_1^{-1}g_1^{-1}$ . So  $y^{-1} \in Hg_1^{-1} = Hg^{-1}$ .

As  $y^{-1} \in Hg^{-1}$ ,  $y^{-1} = hg^{-1}$  for some  $h \in H$ .  $y = gh^{-1}$ . So  $y \in gH$  and consequently  $g_1H \subset gH$ . Thus,  $gH = g_1H$ .

Surjective : Let  $Hg$  be a arbitrary right coset of  $H$ , it clearly has a preimage  $g^{-1}H$ , which is a left coset. As

$$\phi(g^{-1}H) = H((g^{-1})^{-1}) = Hg$$

- (2) Let  $G$  act on  $\{1, 2, \dots, n\}$ . Let  $\sigma$  be a element of  $G_i$ .  $\sigma$  is the product of disjoint cycles. Now by definition, if  $\sigma$  fixes  $i$  then  $i$  should not appear in any of the cycles of  $\sigma$ . That is the only condition  $\sigma$  needs to satisfy.  $G_i$  contains all of the permutations of  $\{1, 2, 3, \dots, n\} - \{i\}$ . Thus,  $G \cong S_{n-1}$ .
- (3) The multiplicative group  $G = (\mathbb{Z}/p\mathbb{Z})^*$  consists of equivalence classes  $\bar{0}, \bar{1}, \bar{2} \dots \overline{p-1}$ . We know that  $x^{|G|} = 1$ . for any element  $x \in G$ . i.e we get that

$$\bar{a}^{p-1} = \bar{1}$$

Thus, we get  $a^{p-1} \equiv 1 \pmod{p}$ . By multiplying both sides by  $a$  we arrive at the claim.  $\square$

**Theorem 2.5 (Subgroup property of HK).** Let  $H, K$  be subgroup of a group  $G$ .

- (1)  $HK$  is a subgroup if and only if  $HK = KH$ .
- (2) If  $H$  and  $K$  are subgroups of  $G$  and  $H \leq N_G(K)$ , then  $HK$  is a subgroup of  $G$ . In particular, if  $K \trianglelefteq G$ , then  $HK \leq G$  for any  $H \leq G$ .

PROOF. (1) Assume that  $HK$  is a subgroup of  $G$ . Since  $K \subseteq HK$  and  $H \subseteq HK$ , by the closure property of group,  $KH \subseteq HK$ . To show the containment in reverse direction, let  $hk \in HK$ . We assumed that  $HK$  to be a subgroup, so we can write  $hk = x^{-1}$ , for some  $x \in HK$ . If  $x = h_1k_1$ , then

$$hk = (h_1k_1)^{-1} = k_1^{-1}h_1^{-1} \in KH,$$

This proves the implication in one direction. To prove it conversely, assume that  $HK = KH$  and let  $a, b \in HK$ . We prove  $ab^{-1} \in HK$  so  $HK$  is a subgroup by the subgroup criterion proved in theorem 1.1. Let

$$a = h_1k_1 \quad \text{and} \quad b = h_2k_2,$$

for some  $h_1, h_2 \in H$  and  $k_1, k_2 \in K$ . This give,

$$b^{-1} = k_2^{-1}h_2^{-1}, \quad \text{and} \quad ab^{-1} = h_1k_1k_2^{-1}h_2^{-1}.$$

Assume  $k_3 = k_1k_2^{-1}$  and  $h_3 = h_2^{-1}$ . So,

$$ab^{-1} = h_1k_3h_3.$$

Since,  $HK = KH$ ,

$$k_3h_3 = h_4k_4, \quad \text{for some } h_4 \in H, \quad k_4 \in K.$$

Therefore,

$$ab^{-1} = h_1h_4k_4,$$

and since  $h_1h_4 \in H$  and  $k_4 \in K$ , we achieve  $ab^{-1} \in HK$ .

- (2) We use the part (1) to prove that  $HK = KH$ . Let  $h \in H$ ,  $k \in K$ . Since,  $K \trianglelefteq G$ ,  $hkh^{-1} \in K$ , hence

$$hk = (hkh^{-1})h \in KH.$$

This proves  $HK \subseteq KH$ . Similarly,  $kh = h(h^{-1}kh) \in HK$ . In same way we can prove the inclusion in other direction and the claim follows from previous part (1).  $\square$

**Problem 2.1.** If  $A$  is an abelian group with  $A \trianglelefteq G$  and  $B$  is any subgroup of  $G$ , prove that  $A \cap B \trianglelefteq AB$ .

**Solution 2.1.**  $AB$  is a subgroup of  $G$  by Corollary 15. Let  $x = a_1b_1$  be an element of  $AB$ , and  $c$  be an element of  $A \cap B$ . To prove that  $xcx^{-1} \in A \cap B$ . Now

$$xcx^{-1} = (a_1b_1)c(a_1b_1)^{-1} = a_1b_1cb_1^{-1}a_1^{-1}.$$

Now,  $c$  is an element of  $A$ ,  $b_1 \in G$  and as  $A \trianglelefteq G$ ,  $b_1cb_1^{-1} \in A$ . So  $b_1cb_1^{-1}$  commutes with  $a^{-1}$  (As  $A$  is an abelian group.) So we get

$$xcx^{-1} = a_1(b_1cb_1^{-1})a_1^{-1} = a_1a_1^{-1}(b_1cb_1^{-1}) = b_1cb_1^{-1}$$

Also, in  $b_1cb_1^{-1}$ ,  $b_1 \in B$  and  $c \in B$  and as  $B$  is a subgroup,  $b_1cb_1^{-1} \in B$ . So  $xcx^{-1} = b_1cb_1^{-1} \in A \cap B$ . So  $xcx^{-1} \in A \cap B$ , and we are done.

**Problem 2.2.** Use Lagrange's Theorem in the multiplicative group  $(\mathbb{Z}/n\mathbb{Z})^\times$  to prove Euler's Theorem:  $a^{\phi(n)} \equiv 1 \pmod{n}$  for every integer  $a$  relatively prime to  $n$ , where  $\phi$  denotes Euler's  $\phi$ -function.

**Solution 2.2.**  $(\mathbb{Z}/n\mathbb{Z})^\times$  has order equal to  $\phi(n)$ .  $|a|$  divides  $|G| = \phi(n)$ . Thus

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$

### 3. The isomorphism theorems

**Theorem 3.1 (First Isomorphism Theorem).** Let  $\phi : G_1 \rightarrow G_2$  be a homomorphism between groups, and suppose that  $\phi$  is onto  $G_2$  and has kernel  $K$ . Then the map  $H_1 \mapsto \phi(H_1)$  gives a one-one correspondence between (a) the subgroups  $H_1$  of  $G_1$  containing  $K$  and (b) the subgroups of  $G_2$ . Under this correspondence, normal subgroups correspond to normal subgroups. If  $H_1$  is normal in  $G_1$ , then  $gH_1 \mapsto \phi(g)\phi(H_1)$  is an isomorphism of  $G_1/H_1$  onto  $G_2/\phi(H_1)$ .

**PROOF.** First we shall show that, for every subgroup  $H_2$  of  $G_2$ , we shall show that  $\phi^{-1}(H_2)$  is a subgroup of  $G_1$ . For that let,  $g_1, g'_1 \in \phi^{-1}(H_2)$ , then  $\phi(g_1) \in H_2$  and  $\phi(g'_1) \in H_2$ . Let  $\phi(g_1) = h_2 \in H_2$  and  $\phi(g'_1) = h'_2 \in H_2$ . Then,  $\phi(g_1g'_1) = \phi(g_1)\phi(g'_1) = h_2h'_2 \in H_2$  by the closure property of  $H_2$  and  $\phi(g_1^{-1}) = \phi(g_1)^{-1} = h_2^{-1} \in H_2$  by the closure property of the inverse which shows that  $g_1g'_1 \in \phi^{-1}(H_2)$  and  $g_1^{-1} \in \phi^{-1}(H_2)$ .

Also, the subgroup  $\phi^{-1}(H_2)$  contains  $\phi^{-1}(1) = K$ . Therefore, the inverse image under  $\phi$  of a subgroup as in (b) is a subgroup as in (a). Since  $\phi$  is a function, we have  $\phi(\phi^{-1}(H_2)) = H_2$ . Thus, passing from (b) to (a) and back recovers the subgroup of  $G_2$ .

If  $H_1$  is a subgroup of  $G_1$  containing  $K$ , we have to prove that  $H_1 = \phi^{-1}(\phi(H_1))$ . Certainly,  $H_1 \subset \phi^{-1}(\phi(H_1))$ . For proving  $\phi^{-1}(\phi(H_1)) \subset H_1$ , consider  $g_1 \in \phi^{-1}(\phi(H_1))$ . Then  $\phi(g_1) \in \phi(H_1)$  i.e.,  $\phi(g_1) = \phi(h_1)$  for some  $h_1 \in H_1$ . Since  $\phi$  is a homomorphism,  $\phi(g_1h_1^{-1}) = \phi(g_1)\phi(h_1)^{-1} = \phi(g_1)\phi(g_1)^{-1} = 1$ . Thus,  $g_1h_1^{-1} \in \ker \phi = K$ , which is contained in  $H_1$  by assumption. Thus,  $h_1 \in H_1$  and  $g_1h_1^{-1} \in H_1$  and hence their product  $(g_1h_1^{-1})h_1 = g_1 \in H_1$  and so  $\phi^{-1}(\phi(H_1)) \subset H_1$ , and thus passing from (a) to (b) and then back recovers the subgroup of  $G_1$  containing  $K$ .

Now we show that normal subgroups correspond to the normal subgroups. If  $H_2$  is normal in  $G_2$ , let  $H_1$  be the subgroup of  $\phi^{-1}(H_2)$  of  $G_1$ . For  $h_1 \in H_1$  and  $g_1 \in G_1$ , we can write  $\phi(h_1) = h_2 \in H_2$ , and then  $\phi(g_1h_1g_1^{-1}) = \phi(g_1)h_2\phi(g_1)^{-1} \in \phi(g_1)H_2\phi(g_1)^{-1} = H_2$ . Hence  $g_1h_1g_1^{-1} \in \phi^{-1}(H_2) = H_1$ . In the reverse direction, let  $H_1$  be normal in  $G_1$ , and let  $g_2 \in G_2$ . Since  $\phi$  is an onto map  $G_2$ , we can write  $g_2 = \phi(g_1)$  for some  $g_1 \in G_1$ . Then  $g_2\phi(H_1)g_2^{-1} = \phi(g_1)\phi(H_1)\phi(g_1)^{-1} = \phi(g_1H_1g_1^{-1}) = \phi(H_1)$ . Thus,  $\phi(H_1)$  is normal.

Let  $H_2 = \phi(H_1)$ . Since  $\phi(H_1)$  is normal in  $G_2$ , hence  $G_2/H_2$  is a group. The mapping  $\Phi : G_1 \rightarrow G_2/H_2$  given by  $\Phi(g_1) = \phi(g_1)H_2$  is the composition of two homomorphisms and hence is a homomorphism. Its kernel is given by

$$\{g_1 \in G_1 \mid \phi(g_1) \in H_2\} = \{g_1 \in G_1 \mid \phi(g_1) \in \phi(H_1)\} = \phi^{-1}(\phi(H_1)).$$

It is equal to  $H_1$  by the first conclusion of the theorem. Then using the result mentioned below, we obtain the required isomorphism  $\bar{\Phi} : G_1/H_1 \rightarrow G_2/\phi(H_1)$ .  $\square$

**Example 3.1.** Let  $G = \text{Fr}(\{x\})$ , the free group generated by the element  $x$ . Then  $F = \{x^j : j \in \mathbb{Z}\}$ . Consider the map  $\sigma : F \rightarrow \mathbb{Z}$  by  $x^j \mapsto j$  for every  $j \in \mathbb{Z}$ . We now verify that  $\sigma$  is a group homomorphism. Let  $m, n \in \mathbb{Z}$ , then  $\sigma(x^m \cdot x^n) = \sigma(x^{m+n}) = m + n = \sigma(x^m) + \sigma(x^n)$ .

Now we will prove that  $\sigma$  is an injective map. For that consider  $\sigma(x^m) = \sigma(x^n)$ . Then  $m = n$  according to the definition of the map and so  $\sigma$  is an injective map. Also, given  $m \in \mathbb{Z}$ , there exists  $x^m \in F$  such that  $\sigma(x^m) = m$  and so the map  $\sigma$  is a surjective map.

We proved that the map  $\sigma$  is an isomorphism and so  $\text{Fr}(\{x\}) \cong (\mathbb{Z}, +)$ .

**Theorem 3.2 (Second Isomorphism Theorem).** Let  $H_1$  and  $H_2$  be subgroups of a group  $G$  with  $H_2$  normal in  $G$ . Then  $H_1 \cap H_2$  is a normal subgroup of  $H_1$ , the set  $H_1H_2$  of products is a subgroup of  $G$  with  $H_2$  as a normal subgroup, and the map  $h_1(H_1 \cap H_2) \rightarrow h_1H_2$  is a well-defined canonical isomorphism of groups  $H_1/(H_1 \cap H_2) \cong (H_1H_2)/H_2$ .

PROOF. Since  $H_1 \cap H_2$  is the intersection of two subgroups of  $G$ , hence  $H_1 \cap H_2$  is a subgroup of  $G$ . Now we shall show that  $H_1 \cap H_2$  is a normal subgroup of  $H_1$ . For that assume  $h_1 \in H_1$  and  $h \in H_1 \cap H_2$ . Then  $h \in H_1$  and  $h \in H_2$ . Since  $H$  is group, so by the closure property of the group,  $h_1hh_1^{-1} \in H_1$ . Also, since  $H_2$  is a normal subgroup of  $G$ , hence  $h_1hh_1^{-1} \in H_2$  and so  $h_1hh_1^{-1} \in H_1 \cap H_2$ . Hence  $H_1 \cap H_2$  is a normal subgroup of  $H_1$ .

Now, we shall show that  $H_1H_2$  is a subgroup of  $G$ . For that, assume  $h_1h_2, h'_1h'_2 \in H_1H_2$ . Then,  $h_1h_2h'_1h'_2 = h_1h'_1(h_1^{-1}h_2h'_1)h'_2 \in H_1H_2$  as  $h_1h'_1 \in H_1$  and  $(h_1^{-1}h_2h'_1)h'_2 \in H_2$ . Also,  $(h_1h_2)^{-1} \in H_2$  as  $(h_1h_2)^{-1} = h_1^{-1}(h_1h_2^{-1}h_1^{-1}) \in H_1H_2$ .  $(h_1h_2^{-1}h_1^{-1}) \in H_2$  as  $H_2$  is normal in  $H_1H_2$  and  $H_2$  is normal in  $G$ . Hence  $H_1H_2$  is a subgroup of  $G$ .

The function  $\phi(h_1(H_1 \cap H_2)) = h_1H_2$  is well defined since  $H_1 \cap H_2 \subseteq H_2$ , and  $\phi$  respects products. The domain of  $\phi$  is  $\{h_1(H_1 \cap H_2) \mid h_1 \in H_1\}$ , and the kernel is the subset of this such that  $h_1$  lies in  $H_2$  as well as  $H_1$ . For this to happen,  $h_1$  must be in  $H_1 \cap H_2$ , and thus the kernel is the identity coset of  $H_1/(H_1 \cap H_2)$ . Hence,  $\phi$  is one-to-one. To see that  $\phi$  is onto  $(H_1H_2)/H_2$ , let  $h_1h_2H_2$  be given. Then  $h_1(H_1 \cap H_2)$  maps to  $h_1H_2$ , which equals  $h_1h_2H_2$ . Hence,  $\phi$  is onto.  $\square$

**Theorem 3.3 (Third Isomorphism Theorem).** Let  $G$  be a group and let  $H$  and  $K$  be normal subgroups of  $G$  with  $H \leq K$ . Then  $K/H \trianglelefteq G/H$  and

$$(G/H)/(K/H) \cong G/K.$$

If we denote the quotient by  $H$  with a bar, this can be written

$$\bar{G}/\bar{K} \cong G/K.$$

PROOF. First we prove that  $K/H \trianglelefteq G/H$ . Let  $gH \in G/H$  and  $kH \in K/H$ . Then  $(gH)(kH)(gH)^{-1} = (gH)(kH)(g^{-1}H) = (gkg^{-1})H \in K/H$  because  $gkg^{-1} \in K$  as  $K \trianglelefteq G$ .

Define  $\phi : G/H \rightarrow G/K$ , by  $\phi(gH) = gK$ . We have to first show that  $\phi$  is well defined. For that, assume that  $g_1H = g_2H$  for some  $g_1 \in G$  and  $g_2 \in G$ . Since  $g_1 \in g_1H$ , so  $g_1 \in g_2H$  and so  $g_1 = g_2h_2$  for some  $h_2 \in H$ . Similarly,  $g_2 \in g_2H$ , so  $g_2 \in g_1H$  and so  $g_2 = g_1h_1$  for some  $h_1 \in H$ . We have to prove that  $g_1K = g_2K$ . For that assume that  $g \in g_1K$ , then  $g = g_1k_1$  for some  $k_1 \in K$  and so,  $g = g_1k_1 = (g_2h_2)k_1 = g_2(h_2k_1) \in g_2K$  as  $h_2k_1 \in K$  and so,  $g_1K \subseteq g_2K$ . Now, assume that  $g \in g_2K$ , then  $g = g_2k_2$  for some  $k_2 \in K$  and so,  $g = g_2k_2 = (g_1h_1)k_2 = g_1(h_1k_2) \in g_1K$  as  $h_1k_2 \in K$  and so,

$g_2K \subseteq g_1K$ . Hence  $g_1K = g_2K$  i.e.,  $\phi(g_1H) = \phi(g_2H)$  and so  $\phi$  is well-defined.

Also,  $\phi$  is a surjection because for  $gK \in G/K$ , there exists  $gH \in G/H$  such that  $\phi(gH) = gK$ .

$$\begin{aligned}\ker \phi &= \{gH \in G/H \mid \phi(gH) = 1K\} \\ &= \{gH \in G/H \mid gK = 1K\} \\ &= \{gH \in G/H \mid g \in K\}\end{aligned}$$

By the First Isomorphism Theorem,  $(G/H)/(K/H) \cong G/K$ .

□

**Problem 3.1.** Prove that if  $H$  is a normal subgroup of  $G$  of prime index  $p$ , then for all  $K \leq G$ , either

- $K \leq H$  or
- $G = HK$  and  $|K : K \cap H| = p$ .

**Solution 3.1.** Assume that  $p$  is the index of  $H$  ( $p$  is prime). We use the Second Isomorphism Theorem to get  $KH \leq G$  and  $H \triangleleft KH$ . Let us now consider  $HK$ 's index in  $G$ . By Problem 11 in section 3.2, we get that

$$|G : H| = |G : HK| \cdot |HK : H|.$$

However, since  $|G : H|$  is prime, there are just two options for  $|G : HK|$ . Either  $|G : HK| = p$  or  $|G : HK| = 1$ . In the first case,  $|HK : H| = 1$ , i.e.,  $H = HK$ . So  $K \leq H$ .

In the second case,  $|G : HK| = 1$ , i.e.,  $G = HK$  and  $|HK : H| = p$ . Now,

$$|HK : H| = \frac{|HK|}{|H|} = \frac{|K|}{|H \cap K|} = |K : K \cap H|$$

The above statement is true by 2.4. So, by above statement  $|K : K \cap H| = p$ . So we are done.

**Problem 3.2.** Let  $C$  be a normal subgroup of the group  $A$  and  $D$  be a normal subgroup of the group  $B$ . Prove that  $(C \times D) \trianglelefteq (A \times B)$  and  $(A \times B)/(C \times D) \cong (A/C) \times (B/D)$ .

**Solution 3.2.** We will construct a map  $\varphi : A \times B \rightarrow (A/C) \times (B/D)$  as  $(a, b) \mapsto (aC, bD)$ .

i.e., it is defined as  $\varphi(a, b) = (aC, bD)$ . First we will prove that  $\varphi$  has well defined. First, let  $(a_1, b_1) = (a_2, b_2)$  for some  $a_1, a_2 \in A$  and  $b_1, b_2 \in B$ . So,  $a_1 = a_2$  and  $b_1 = b_2$ ;  $a_1C = a_2C$  and  $b_1D = b_2D$ .  $(a_1C, b_1D) = (a_2C, b_2D)$ . Thus,  $\varphi(a_1, b_1) = \varphi(a_2, b_2)$ . So the map is well defined.

We will now show that this map is indeed a surjection. Take a random element  $(aC, bD) \in (A/C) \times (B/D)$ , its preimage is clearly  $(a, b)$ . Therefore,  $\varphi$  is surjective. Now for homomorphism consider,

$$\varphi((a, b)(a_1, b_1)) = \varphi(aa_1, bb_1) = (aa_1C, bb_1D) = ((aC)(a_1C), (bD)(b_1D))$$

$$\varphi((a, b)(a_1, b_1)) = ((aC)(a_1C), (bD)(b_1D)) = (aC, bD)(a_1C, b_1D) = \varphi(a, b)\varphi(a_1, b_1).$$

So it is a homomorphism. Now, we will find kernel of  $\varphi$ .

$$\ker \varphi = \{(a, b) \in A \times B \mid \varphi(a, b) = (1C, 1D)\}$$

$$\ker \varphi = \{(a, b) \in A \times B \mid (aC, bD) = (1C, 1D)\}$$

$$\ker \varphi = \{(a, b) \in A \times B \mid aC = 1C, bD = 1D\}$$

$$\ker \varphi = \{(a, b) \in A \times B \mid a \in C, b \in D\}$$

$$\ker \varphi = C \times D.$$

By the First isomorphism theorem 3.1, we conclude  $\ker \varphi = C \times D$  is normal in  $A \times B$ , and

$$(A \times B)/(C \times D) \cong (A/C) \times (B/D).$$

**Problem 3.3.** Let  $M$  and  $N$  be normal subgroups of  $G$  such that  $G = MN$ . Prove that  $G/(M \cap N) \cong (G/M) \times (G/N)$ .

**Solution 3.3.** Define a map  $\varphi : G \rightarrow (G/M) \times (G/N)$  as  $g \mapsto (gM, gN)$ , or  $\varphi(g) = (gM, gN)$ .

Well defined: Let  $g_1 = g_2$ . Clearly,  $(g_1M, g_1N) = (g_2M, g_2N)$ . So,  $\varphi(g_1) = \varphi(g_2)$ .

We will now prove that the map is surjective. Take a random element  $(gM, gN) \in (G/M) \times (G/N)$ , its preimage is clearly  $g$ . Thus,  $\varphi$  is surjective. For homomorphism,

$$\varphi(g_1g_2) = (g_1g_2M, g_1g_2N) = ((g_1M)(g_2M), (g_1N)(g_2N)) = (g_1M, g_1N)(g_2M, g_2N) = \varphi(g_1)\varphi(g_2).$$

So it is a homomorphism. Now we find the kernel of  $\varphi$ .

$$\ker \varphi = \{g \in G \mid \varphi(g) = (1M, 1N)\}$$

$$\ker \varphi = \{g \in G \mid (gM, gN) = (1M, 1N)\}$$

$$\ker \varphi = \{g \in G \mid gM = 1M, gN = 1N\}$$

$$\ker \varphi = \{g \in G \mid g \in M, g \in N\}$$

$$\ker \varphi = M \cap N.$$

By the first isomorphism theorem, we have

$$G/(M \cap N) \cong (G/M) \times (G/N).$$



## CHAPTER 6

### Group action

Recall that in the last section 10 of the first chapter, we already had a good start into the theory of group action. In this chapter, we will freely presume material covered therein. However, some of the material will be recalled so that discussion here remains complete as a chapter on the group action.

#### 1. Groups acting on sets

The elements of  $S_n$  are permutations of the set  $\{1, \dots, n\}$ , so we may view an element of  $S_n$  as giving a rule that takes each number in  $\{1, \dots, n\}$  and assigns it a number in  $\{1, \dots, n\}$ . Similarly, the elements of  $D_n$  are rigid rearrangements of the vertices of an  $n$ -gon, so we may view an element of  $D_n$  as giving a rule that takes each vertex of an  $n$ -gon and assigns it to a vertex of the  $n$ -gon.

We can axiomatize these examples by starting with an arbitrary group  $G$  and an arbitrary set  $X$  and having each element of  $G$  rearranges the elements of  $X$  (subject to suitable restrictions).

**Definition 1.1.** A group action of  $G$  on  $X$  is a rule assigning  $g \cdot x \in X$  to each  $g \in G$  and  $x \in X$  such that:

- Identity:  $e \cdot x = x$  for all  $x \in X$ .
- Associativity:  $(g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x)$  for all  $x \in X$  and  $g_1, g_2 \in G$ .

**Remark 1.1.** A group action corresponds to a homomorphism  $\alpha : G \rightarrow S_X$ , where  $\alpha(g)$  is a permutation of  $X$ , and  $g \cdot x = \alpha(g)(x)$ . Thus,  $\alpha$  sends each  $g \in G$  to a permutation  $\alpha(g) : X \rightarrow X$  of the set  $X$ , and the group action is  $g \cdot x = \alpha(g)(x)$ .

**Definition 1.2.** Given a group  $G$  acting on a set  $X$ , each element  $x \in X$  determines two natural objects that are worth studying:

- The orbit of  $x$ , denoted  $Gx$ , is the set of elements of  $X$  to which  $x$  is sent by the action of  $G$

$$Gx = \{g \cdot x : g \in G\}.$$

- The stabiliser of  $x$ , denoted  $G_x$ , is the set of elements of  $G$  that leave  $x$  unchanged

$$G_x = \{g \in G : g \cdot x = x\}.$$

The stabilizer  $G_x$  is a subgroup of  $G$ .

**Example 1.1.**

- Example 1: Consider the action of  $G = S_n$  on  $X = \{1, \dots, n\}$ . For  $G = S_n$  acting on  $X = \{1, \dots, n\}$ ,  $Gx = X$  as given any two elements  $x, y \in X$ , there is a permutation  $\pi$  that sends  $x$  to  $y$  and  $G_x$  consists of permutations fixing  $x$ , which form a subgroup isomorphic to  $S_{n-1}$ .
- Example 2: For  $G = D_n$  acting on the vertices of an  $n$ -gon,  $Gx = X$ . The stabilizer  $G_x$  consists of two elements: the identity and the flip about the axis through  $x$ . Thus,  $G_x$  is a cyclic group of order 2.
- Example 3: For  $G \leq S_5$  generated by  $\pi$ , with

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

the action of  $G$  on  $\{1, 2, 3, 4, 5\}$  has orbits  $\{1, 3, 4\}$  and  $\{2, 5\}$ .

**Proposition 1.1.** Let  $G$  be a group that acts on a set  $X$ .

- (a) The stabilizer  $G_x$  is a subgroup of  $G$ .
- (b) The relation  $x \sim y \iff y = g \cdot x$  is an equivalence relation with equivalence class  $Gx$ .

(c) There is a bijection  $\alpha : G/G_x \rightarrow Gx$ , defined by  $\alpha(C) = g \cdot x$  for any  $g \in C$ . If  $G$  is finite, then

$$|Gx| = \frac{|G|}{|G_x|}.$$

PROOF. (a) To show  $G_x$  is a subgroup of  $G$ :

The identity element  $e \in G_x$  because group actions satisfy  $e \cdot x = x$ . For  $g, g' \in G_x$ , associativity ensures  $(gg') \cdot x = g \cdot (g' \cdot x) = g \cdot x = x$ , so  $gg' \in G_x$ . For  $g \in G_x$ , applying  $g^{-1}$  gives  $g^{-1} \cdot x = g^{-1}(g \cdot x) = (g^{-1}g) \cdot x = e \cdot x = x$ , hence  $g^{-1} \in G_x$ .

Thus,  $G_x$  is a subgroup of  $G$ .

(b) To show  $\sim$  is an equivalence relation:

Reflexivity:  $x = e \cdot x$ , so  $x \sim x$ . Symmetry: If  $x \sim y$ , then  $y = g \cdot x$  for some  $g \in G$ . Applying  $g^{-1}$ ,  $g^{-1} \cdot y = g^{-1}(g \cdot x) = x$ , so  $y \sim x$ . Transitivity: If  $x \sim y$  and  $y \sim z$ , then  $y = g \cdot x$  and  $z = g' \cdot y$  for  $g, g' \in G$ . Substituting gives  $z = g' \cdot (g \cdot x) = (g'g) \cdot x$ , so  $x \sim z$ . This completes the proof that  $\sim$  is an equivalence relation. The equivalence class of an element  $x \in X$  is

$$\{y \in X : x \sim y\} = \{y \in X : y = gx \text{ for some } g \in G\} = \{gx : g \in G\} = Gx.$$

(c) To show  $\alpha$  is a well-defined bijection:

1. Well-defined: For any  $g, g' \in C$ ,  $g' = g \cdot h$  for some  $h \in G_x$ . Then:

$$g' \cdot x = (g \cdot h) \cdot x = g \cdot (h \cdot x) = g \cdot x \quad (\text{as } h \in G_x \text{ fixes } x).$$

Thus,  $\alpha(C) = g \cdot x$  is independent of the choice of  $g$ .

2. Surjective: Any element of  $Gx$  is  $g \cdot x$  for some  $g \in G$ , so  $\alpha(gG_x) = g \cdot x$ .

3. Injective: If  $\alpha(C_1) = \alpha(C_2)$ , then  $g_1 \cdot x = g_2 \cdot x$  for  $g_1, g_2 \in G$ . This implies

$$g_1^{-1} \cdot g_2 \cdot x = x \implies g_1^{-1}g_2 \in G_x \implies g_1G_x = g_2G_x \implies C_1 = C_2.$$

Thus,  $\alpha$  is a bijection.

For the last part of (c), we compute

$$\begin{aligned} |Gx| &= |(G/G_x)| \text{ from (a),} \\ &= \frac{|G|}{|G_x|} \text{ (by applying Lagrange's Theorem .)} \end{aligned}$$

This completes the proof of the Proposition. □

**Definition 1.3.** We say that  $G$  acts *transitively* on  $X$  if  $Gx = X$  for all  $x \in X$ .

**Theorem 1.1.** Let  $G$  be a finite group that acts on a finite set  $X$ . Choose  $x_1, \dots, x_k \in X$  so that  $Gx_1, \dots, Gx_k$  are the distinct orbits of elements of  $X$ . Then

$$|X| = \sum_{i=1}^k |Gx_i| = \sum_{i=1}^k \frac{|G|}{|G_{x_i}|}.$$

PROOF. There is an equivalence relation on  $X$  such that the equivalence class of  $x$  is exactly its orbit  $Gx$ . It follows from general properties of equivalence relations that  $X$  is the disjoint union of the distinct equivalence classes,

$$X = Gx_1 \cup Gx_2 \cup \dots \cup Gx_k,$$

so in particular

$$|X| = |Gx_1| + |Gx_2| + \dots + |Gx_k|.$$

this proves the first part of Theorem 1.1. For the second part, we use the third part of Proposition 1.1, which says that we can replace  $|Gx_i|$  with  $\frac{|G|}{|G_{x_i}|}$ . □

**Definition 1.4.** Let  $G$  be a group. The *center* of  $G$ , denoted  $Z(G)$ , is the set of elements of  $G$  that commute with every element of  $G$ :

$$Z(G) = \{g \in G : gg' = g'g \text{ for every } g' \in G\}.$$

Recall that earlier in the first chapter, we proved that center of the group is a subgroup.

**Remark 1.2.** We leave it to you to prove that the center of  $G$  is a subgroup of  $G$ , and indeed, it is a normal subgroup

**Theorem 1.2.** Let  $p$  be a prime, and let  $G$  be a finite group with  $p^n$  elements for some  $n \geq 1$ . Then  $Z(G) \neq \{e\}$ ; i.e.,  $G$  contains a non-identity element that commutes with every element of  $G$ .

PROOF. We take the set  $X$  to be a copy of the group  $G$ , and we let  $G$  act on  $X$  via the formula

$$g \in G \text{ sends } x \in X \text{ to } gxg^{-1} \in X.$$

This is called the *conjugation action* of  $G$  on itself. We leave it to you to check that this rule satisfies the group action axioms. Let  $x \in X$ . What is its stabilizer  $G_x$ ? It is the set

$$(1) \quad G_x = \{g \in G : gxg^{-1} = x\} = \{g \in G : gx = xg\}.$$

Thus the stabilizer  $G_x$  is the set of elements of  $G$  that commute with  $x$ . In particular, we see that

$$G_x = G \iff x \text{ commutes with every element of } G \iff x \in Z(G).$$

We choose elements  $x_1, \dots, x_k \in X$  that give the distinct orbits. Then Theorem 1.1 tells us that

$$(2) \quad |X| = \sum_{i=1}^k \frac{|G|}{|G_{x_i}|}.$$

We have  $|X| = |G| = p^n$ , and each  $G_{x_i}$  is a subgroup of  $G$ , so Lagrange tells us that each  $|G_{x_i}|$  divides  $|G|$ . Therefore for each  $i$  we have

$$|G_{x_i}| = p^{r_i} \text{ for some } 0 \leq r_i \leq n,$$

So the statement (2) becomes

$$(3) \quad p^n = \sum_{i=1}^k p^{n-r_i}.$$

The terms in the statement (3) with  $r_i = n$  are

$$r_i = n \iff G_{x_i} = G \iff x_i \in Z(G).$$

So if we separate out the terms in statement (3) that have  $r_i = n$ , we obtain the useful formula

$$p^n = \sum_{i=0, r_i=n}^k 1 + \sum_{i=0, r_i < n}^k p^{n-r_i} = |Z(G)| + \sum_{i=0, r_i < n}^k p^{n-r_i}.$$

It follows that  $p$  divides  $|Z(G)|$ . But we know that  $|Z(G)| \geq 1$ , since the identity element of  $G$  is in  $Z(G)$ . Hence  $|Z(G)| \geq p$ .  $\square$

**Corollary 1.1.** Let  $p$  be prime, and let  $G$  be a group with  $p^2$  elements. Then  $G$  is an abelian group.

PROOF. Let  $Z = Z(G)$  denote the center of  $G$ . By Lagrange's theorem, the order of  $Z$  must divide  $|G| = p^2$ , so we have three possibilities:

$$|Z| = 1, \quad |Z| = p, \quad \text{or} \quad |Z| = p^2.$$

Theorem 1.2 tells us that  $Z \neq \{e\}$ , so  $|Z| \neq 1$ .

Next, we assume  $|Z| = p$  and aim to derive a contradiction. Since  $Z$  is a normal subgroup of  $G$ , we can form the quotient group  $G/Z$ . By Lagrange's theorem,

$$|(G/Z)| = \frac{|G|}{|Z|} = \frac{p^2}{p} = p.$$

a group of prime order is cyclic, so  $G/Z$  is cyclic. Let  $hZ$  be a coset that generates  $G/Z$ , meaning

$$G/Z = \{Z, hZ, h^2Z, \dots, h^{p-1}Z\}.$$

This implies that every element of  $G$  belongs to a coset of  $Z$ , and we can write

$$G = Z \cup hZ \cup h^2Z \cup \dots \cup h^{p-1}Z.$$

Now, let  $g_1, g_2 \in G$  be arbitrary elements. Since  $g_1$  and  $g_2$  are in  $G$ , they belong to some cosets of  $Z$ , so we can express them as

$$g_1 = h^{i_1}z_1 \quad \text{and} \quad g_2 = h^{i_2}z_2,$$

where  $z_1, z_2 \in Z$  and  $0 \leq i_1, i_2 \leq p-1$ . Using the fact that elements of  $Z$  commute with all elements of  $G$ , we compute:

$$g_1g_2 = (h^{i_1}z_1)(h^{i_2}z_2) = h^{i_1}h^{i_2}z_1z_2 = h^{i_1+i_2}z_1z_2.$$

Since  $z_1$  and  $z_2$  commute and  $h^{i_1}$  and  $h^{i_2}$  are in the cyclic subgroup generated by  $h$ , the above simplifies to:

$$g_1g_2 = h^{i_1+i_2}z_2z_1 = h^{i_2+i_1}z_2z_1 = (h^{i_2}z_2)(h^{i_1}z_1) = g_2g_1.$$

Thus,  $g_1$  and  $g_2$  commute. Since  $g_1, g_2$  were arbitrary, this implies that every element of  $G$  commutes with every other element. Therefore,  $G$  is abelian, meaning  $Z = G$ , which contradicts our assumption that  $|Z| = p$ .

We know that  $|Z|$  is 1,  $p$ , or  $p^2$ . We have ruled out  $|Z| = 1$  and  $|Z| = p$ , so the only possibility is  $|Z| = p^2$ . Hence,  $Z = G$ , and  $G$  is abelian.  $\square$

**Problem 1.1.** Let  $G$  act on the set  $A$ . Prove that if  $a, b \in A$  and  $b = g \cdot a$  for some  $g \in G$ , then  $G_b = gG_ag^{-1}$  ( $G_a$  is the stabilizer of  $a$ ). Deduce that if  $G$  acts transitively on  $A$  then the kernel of the action is

$$\bigcap_{g \in G} gG_ag^{-1}.$$

**Solution 1.1.** First we will prove that  $G_b = gG_ag^{-1}$ . Let  $x$  be an arbitrary element of  $G_b$ . Define  $y := g^{-1}xg$ . We will prove now that  $y \in G_a$ . Note that we have

$$y \cdot a = (g^{-1}xg) \cdot a$$

$$y \cdot a = (g^{-1}x) \cdot (g \cdot a) = (g^{-1}x) \cdot b$$

$$y \cdot a = g^{-1} \cdot (x \cdot b) = g^{-1} \cdot b = g^{-1} \cdot (g \cdot a) = (g^{-1}g) \cdot a = a.$$

Thus,  $y \in G_a$ .  $g^{-1}xg \in G_a$ , this means  $x \in gG_ag^{-1}$ . So  $G_b \subseteq gG_ag^{-1}$ . Now will prove that  $gG_ag^{-1} \subseteq G_b$ . let  $x \in gG_ag^{-1}$ . So  $x = gyg^{-1}$  where  $y \in G_a$ . To prove that  $x \in G_b$

$$x = gyg^{-1}$$

$$x \cdot b = (gyg^{-1}) \cdot (g \cdot a) = (gyg^{-1}g) \cdot a = (gy) \cdot a$$

$$x \cdot b = (gy) \cdot a = g \cdot (y \cdot a) = g \cdot a = b$$

$$x \cdot b = b.$$

Thus we get,  $x \in G_b$  and  $gG_ag^{-1} \subseteq G_b$ . So,  $G_b = gG_ag^{-1}$

Claim: The kernel of the group is the intersection of the stabilizer of each element in  $A$ .

PROOF.

$$\ker \varphi := \{g \in G : \forall a \in A, g \cdot a = a\}$$

$$= \{g \in G : g \in G_a, \forall a \in A\}$$

$$= \bigcap_{a \in A} G_a$$

$\square$

As  $G$  is transitive, clearly

$$\ker \varphi = \bigcap_{g \in G} gG_ag^{-1}.$$

**Problem 1.2.** Let  $G$  be a permutation group on the set  $A$  (i.e.,  $G \leq S_A$ ) Let  $\sigma \in G$  and  $a \in A$ . Prove that  $\sigma G_a \sigma^{-1} = G_{\sigma(a)}$ . Deduce that if  $G$  acts transitively on  $A$  then

$$\bigcap_{\sigma \in G} \sigma G_a \sigma^{-1} = 1.$$

**Solution 1.2.** This is just Problem 1 with  $b = \sigma(a) = \sigma \cdot a$  and  $\sigma = g$ .

Thus,  $G_{\sigma(a)} = \sigma G_a \sigma^{-1}$ .

In this case,  $\ker G = 1$  as the permutation is always injective.

Thus,

$$\bigcap_{\sigma \in G} \sigma G_a \sigma^{-1} = 1.$$

**Problem 1.3.** Assume that  $G$  is an abelian, transitive subgroup of  $S_A$ . Show that  $\sigma(a) \neq a$  for all  $\sigma \in G - \{1\}$  and all  $a \in A$ . Deduce that  $|G| = |A|$ .

**Solution 1.3.** Assume that  $\sigma(a) = a$  for some  $\sigma \in G - \{1\}$ . As  $G$  is transitive,  $\forall b \in A$ ,  $b = \rho(a)$  for some  $\rho \in G$ .  $b = \rho(a) = \rho(\sigma(a)) = \sigma(\rho(a)) = \sigma(b)$ . Thus,  $\forall b \in A$ ,  $\sigma(b) = b \Rightarrow \sigma$  is an identity permutation, a contradiction.

Now,  $|G| = |G_a| \cdot |O_a|$ .

As  $G$  acts transitively and  $\sigma(a) \neq a \forall \sigma \in G - \{1\}$  means that  $|G_a| = 1$ .

Also, we have  $|O_a| = |A|$ .

$$|G| = \frac{|O_a|}{|G_a|} = |O_a| = |A|.$$

So  $|G| = |A|$ .

**Problem 1.4.** Let  $G$  be a transitive permutation group on the finite set  $A$ . A *block* is a nonempty subset  $B$  of  $A$  such that for all  $\sigma \in G$  either  $\sigma(B) = B$  or  $\sigma(B) \cap B = \emptyset$  (here  $\sigma(B)$  is the set  $\{\sigma(b) \mid b \in B\}$ ).

- (1) Prove that if  $B$  is a block containing the element  $a$  of  $A$ , then the set  $G_B$  defined by  $G_B = \{\sigma \in G \mid \sigma(B) = B\}$  is a subgroup of  $G$  containing  $G_a$ .
- (2) Show that if  $B$  is a block and  $\sigma_1(B), \sigma_2(B), \dots, \sigma_n(B)$  are all the distinct images of  $B$  under the elements of  $G$ , then these form a partition of  $A$ .

**Solution 1.4.** (1) Let  $\sigma$  be an arbitrary element of  $G_B$ . We prove that  $\sigma^{-1} \in G_B$ . To prove that  $\sigma^{-1}(B) = B$ . Now, we know that  $\sigma(B) = B$  i.e.,  $\{\sigma(b) \mid b \in B\} = B$ . So  $\sigma(b) \in B$  for all  $b \in B$ . As  $\sigma$  is a permutation,  $\sigma^{-1}(b) \in B$  for all  $b \in B$ . That is,  $\sigma^{-1}(B) = B$ .  $\sigma$  and  $\tau$  be arbitrary both in  $G_B$

$$(\sigma\tau)(B) = \sigma(\tau(B)) = \sigma(B) = B,$$

so  $\sigma\tau \in G_B$ . Thus  $G_B \leq G$

- (2) To show that  $\sigma_i(B)$  are pairwise disjoint. Claim that  $\sigma_i(B)$  and  $\sigma_j(B)$  are disjoint. Assume, for the sake of contradiction, they are not disjoint.

$$\exists x \in \sigma_i(B) \cap \sigma_j(B).$$

So, there exists  $b, b_1 \in B$  such that

$$x = \sigma_i(b) = \sigma_j(b_1)$$

$b = \sigma_i^{-1}(\sigma_j(b_1))$ . So  $\sigma_i^{-1}\sigma_j(B) \cap B$  is not empty (As  $b = \sigma_i^{-1}\sigma_j(b_1) \in B$ ). Thus,  $\sigma_i^{-1}\sigma_j(B) = B$ . So  $\sigma_i(B) = \sigma_j(B)$ .

## 2. Important Consequences of group action

**Theorem 2.1 (Cayley's Theorem).** If  $|G| = n < \infty$ , then  $G$  is isomorphic to a subgroup of  $S_n$ .

PROOF. Consider the map  $\varphi : G \rightarrow \text{Sym}(G)$  given by  $\varphi(g) = \varphi_g$  the permutation associated with group action as discussed in section 10. To show  $\varphi$  is a homomorphism we need to show for all  $g, h \in G$  that  $\varphi(gh) = \varphi(g)\varphi(h)$  or equivalently  $\varphi_{gh} = \varphi_g\varphi_h$ . But this is true, since for all  $a \in G$ ,

$$\varphi_{gh}(a) = (gh)a = g(ha) = \varphi_g(ha) = \varphi_g(\varphi_h(a)).$$

Finally,  $\varphi$  is an injection (monomorphism), since its kernel is trivial. Indeed, if  $g \in \ker \varphi$ , then  $\varphi(g) = 1_G$ , the identity map, so for all  $a \in G$  we have  $\varphi_g(a) = a$ . In particular,  $\varphi_g(1) = 1$  or equivalently  $g = 1$ . Thus, by the First Isomorphism Theorem 3.1

$$G \cong G/1 = G/\ker \varphi \cong \varphi(G) \leq \text{Sym}(G) \cong S_n.$$

Thus,  $G$  is isomorphic to a subgroup of  $S_n$ . In other words, every finite group can be embedded (homomorphism) inside symmetric group  $S_n$ .  $\square$

**Theorem 2.2 (Class Equation).** Let  $G$  be any finite group. Then

$$|G| = |Z(G)| + \sum_{i=1}^m [G : C_G(g_i)],$$

for some natural number  $m$  and some  $g_i \in G$  with  $[G : C_G(g_i)] > 1$ . Note that  $m = 0$  exactly when  $G$  is abelian.

PROOF. We proved this using the action of  $G$  on itself by left conjugation. Since orbits are equivalence classes on  $G$  and  $G$  is finite, we have

$$G = \bigcup_{i=1}^r Gg_i,$$

for some  $g_i \in G$  and some positive integer  $r$ . Some of these orbits may contain only a single element, and we have seen that  $|Gg_i| = 1$  iff  $g_i \in Z(G)$ . Therefore, all the singleton orbits of this action may be combined to obtain  $Z(G)$ . If  $G$  is abelian, then all the orbits are singletons and  $G = Z(G)$ . Otherwise, write the union as

$$G = Z(G) \cup \bigcup_{i=1}^m Gg_i,$$

for some  $g_i \in G$  with  $|Gg_i| > 1$  and some positive integer  $m$ . This yields the equation

$$|G| = |Z(G)| + \sum_{i=1}^m |Gg_i| = |Z(G)| + \sum_{i=1}^m [G : C_G(g_i)],$$

for some  $g_i \in G$  with  $[G : C_G(g_i)] > 1$  and some positive integer  $m$ . If the sum were empty, then  $G = Z(G)$  and the result follows trivially. So assume now that the sum has terms. Since for each  $i$ ,  $|G| = [G : C_G(g_i)]|C_G(g_i)|$ , this implies that  $[G : C_G(g_i)]$  divides  $|G|$ .

Now  $|G| = p^n$  and  $[G : C_G(g_i)] > 1$ , so it must be that  $[G : C_G(g_i)] = p^k$  for some  $1 \leq k \leq n$ . Regardless of what  $k$  is, we can conclude that each  $[G : C_G(g_i)]$  is a multiple of  $p$  and therefore  $\sum_{i=1}^m [G : C_G(g_i)]$  is a multiple of  $p$ .

So we can rewrite the class equation as

$$p^n = |Z(G)| + lp,$$

for some integer  $l$ . Therefore,

$$|Z(G)| = p(p^{n-1} - l),$$

and we can conclude that  $p$  divides  $|Z(G)|$ .  $\square$

**Corollary 2.1.** If  $G$  is a group with  $|G| = p^2$  for some prime  $p$ , then  $G$  is abelian.

PROOF. We know that  $p$  divides  $|Z(G)|$  and since  $|Z(G)|$  divides  $|G|$ , it must be that either  $|Z(G)| = p$  or  $p^2$ . But  $|Z(G)| \neq p$ , for if it were, then we would have

$$|G/Z(G)| = \frac{|G|}{|Z(G)|} = \frac{p^2}{p} = p.$$

This would imply that  $G/Z(G) \cong \mathbb{Z}_p$ , a cyclic group which implies that  $G$  is abelian, contradicting the fact that  $Z(G) \neq G$ . Hence,  $|Z(G)| = p^2$ , and so  $G$  must be abelian.  $\square$

**Theorem 2.3 (Cauchy's theorem using class equation).** If  $G$  is any finite group whose order is divisible by a prime  $p$ , then  $G$  has an element of order  $p$ .

PROOF. First let us assume that by induction that all groups of smaller size than  $G$  have the desired property of this theorem. Note that this theorem holds vacuously for the trivial group. If  $G$  is abelian, then this follows from Lagrange's theorem and the fact that if  $|x| = pn$ , then  $|x^n| = p$ . So we only focus on the case when  $G$  is nonabelian.

By the class equation, we know

$$|G| = |Z(G)| + \sum_{i=1}^m [G : C_G(g_i)],$$

for some positive integer  $m$  and some  $g_i \in G$  with  $[G : C_G(g_i)] > 1$ . Note that the sum is not empty, since  $G$  is assumed not to be abelian. Note also that for each  $g_i$ , we have

$$|G| = [G : C_G(g_i)] |C_G(g_i)|,$$

and since  $[G : C_G(g_i)] > 1$ , this makes  $C_G(g_i)$  a proper subgroup of  $G$ .

If  $p$  divides the order of one of the  $C_G(g_i)$ , then by induction such a  $C_G(g_i)$ , and therefore  $G$ , will have an element of order  $p$ , thus proving the our theorem. So we may assume that for all  $g_i$ , it isn't the case that  $p$  divides the order of  $C_G(g_i)$ .

Since  $p$  is prime and  $|G| = [G : C_G(g_i)] |C_G(g_i)|$ , it must then be true that for all  $g_i$ ,  $p$  divides  $[G : C_G(g_i)]$ . Therefore,  $p$  divides the sum

$$\sum_{i=1}^m [G : C_G(g_i)].$$

Since  $p$  divides this sum and also divides  $|G|$ , by the class equation it must be that  $p$  divides  $|Z(G)|$ . Since  $G$  is assumed to be nonabelian, this makes  $Z(G)$  a proper subgroup of  $G$ , and so by induction,  $Z(G)$  (and therefore  $G$ ) has an element of order  $p$ . Now we state and prove a lemma and using this we prove Ore's Theorem.

**Lemma 2.1.** Let  $G$  act on a set  $X$  and  $\varphi$  be the homomorphism associated with that group action. Let  $Y \subseteq X$  be a set of representatives for the distinct orbits of the group action and  $H = \prod_{x \in Y} G_x$ , then

$$\ker(\varphi) = \prod_{x \in X} G_x = \prod_{g \in G} H_g \leq H.$$

PROOF. The results are all contained in the following equivalent statements:  $k \in \ker(\varphi)$  if and only if  $\varphi(k) = 1_X$  if and only if  $\varphi(k)(x) = x$  for all  $x \in X$  if and only if  $k \cdot x = x$  for all  $x \in X$  if and only if  $k \in G_x$  for all  $x \in X$ .

This is equivalent to  $k \in G_{g \cdot y}$  for all  $g \in G$  and all  $y \in Y$ , which in turn is equivalent to  $k \in gG_y g^{-1}$  for all  $g \in G$  and all  $y \in Y$ . Thus,  $k \in \prod_{g \in G} H_g$ .

Certainly,  $\prod_{g \in G} H_g \leq H$ , since  $H = H_1$ .  $\square$

**Corollary 2.2.** If  $H \leq G$  is a group with  $[G : H] = n$ , then there exists  $K \trianglelefteq G$  with  $K \leq H$  such that  $[G : K]$  divides  $n!$ .

PROOF. Let  $G$  act on  $G/H$  by left multiplication and  $\varphi$  be the associated homomorphism for the group action. Set  $K = \ker(\varphi)$ , and we shall show that  $K$  is the desired normal subgroup we seek. First, by 2.1

$$K = \prod_{gH \in G/H} G_{gH} \subseteq G_H = H.$$

Second, due to Cayley's Theorem 2.1,  $\varphi(G)$  is isomorphic to a subgroup of  $S_n$ , which implies  $|\varphi(G)|$  divides  $|S_n| = n!$ . By the First isomorphism theorem 3.1,  $G/K \cong \varphi(G)$ , so that

$$[G : K] = |G/K| = |\varphi(G)|,$$

which divides  $n!$ . □

**Theorem 2.4 (Ore's Theorem).** If  $G$  is a finite group and  $H \leq G$  with  $[G : H] = p$ , where  $p$  is the smallest prime dividing  $|G|$ , then  $H \trianglelefteq G$ .

PROOF. By Corollary 2.2 of this section, there exists  $K \leq H$  with  $K \trianglelefteq G$  and  $[G : K]$  divides  $p!$ . Set  $m = [G : K]$ , and note that  $m$  divides  $|G|$ . Notice that for any prime  $q$  dividing  $m$ , we have  $q$  divides  $|G|$  and  $q$  divides  $p!$ .

But then, since  $p$  is the smallest prime dividing  $|G|$ , it must be that  $q = p$ . Therefore,  $m$  must be a power of  $p$ , say  $m = p^k$ . Now, since  $p^k$  divides  $p!$ , it must be that  $k = 1$ , and so  $m = p$ .

Thus,  $[G : K] = p = [G : H]$  and

$$|K| = \frac{|G|}{[G : K]} = \frac{|G|}{[G : H]} = |H|.$$

Since  $K \leq H$  and  $|K| = |H|$ , we have  $H = K \trianglelefteq G$ . □

**Remark 2.1.** In general, if  $H \leq G$  and  $[G : H] = p$  (a prime), then  $H$  need not be normal in  $G$ . For example, consider the group  $G = S_3$  and  $H = \langle \mu_1 \rangle$  where  $\mu = (2\ 3)$ . Notice that  $[G : H] = |G|/|H| = 6/2 = 3$  (a prime), yet  $H$  is not normal in  $G$ . Indeed, take  $\rho = (1\ 2\ 3) \in G$  and compute:

$$\rho^{-1}\mu\rho = (3\ 2\ 1)(2\ 3)(1\ 2\ 3) = (1\ 2) \notin H.$$

### 3. Sylow's theorems

**Theorem 3.1 (Sylow's first theorem).** Let  $|G| = p^l m$ , where  $p$  is a prime,  $l \geq 1$ , and  $(p, m) = 1$ . Then there exists a subgroup of  $G$  of order  $p^k$ , for any positive integer  $k \leq l$ .

PROOF.  $p^n$  is the largest power of  $p$  that divides  $|G|$ , hence  $|G| = p^n m$  where  $p \nmid m$ . If  $m = 1$ , then  $|G| = p^n$ , so  $G$  is the desired subgroup of  $G$ . To look for a subgroup of  $G$  of order  $p^n$  if  $m \geq 2$ , we will look for a subset  $A \subseteq G$  having  $p^n$  elements and will determine if  $A$  is a subgroup. For this we shall use a result that,

$A$  is a subgroup  $\iff aA = A$  for every  $a \in A$ .

Now consider, the collection of all  $p^n$ -element subsets of  $G$ , and then we will let  $G$  act on these subsets by left multiplication.

Consider the set

$$S = \{A \subseteq G : A \text{ is a subset of } G \text{ with } |A| = p^n\}.$$

The elements of  $S$  are themselves sets. Each of these element has size  $p^n$ . Therefore the number of elements of  $S$ ,  $|S|$  is equal to the number of ways to choose  $p^n$  elements from the  $|G| = p^n m$  elements of the group  $G$ . Therefore,

$$|S| = \text{Number of } p^n\text{-element subsets of } G = \binom{p^n m}{p^n}.$$

Now let  $G$  act on  $S$  by left multiplication. For  $g \in G$  and  $A \in S$ , we form

$$gA = \{ga : a \in A\} \in S.$$

Observe that  $gA$  is a subset of  $G$  and that  $gA$  contains  $p^n$  elements, and so  $gA$  is an element of  $S$ . Let  $A_1, A_2, \dots, A_r$  be elements of  $S$  that give the distinct orbits for the action of  $G$  on  $S$ . Then the Orbit-Stabilizer counting Theorem, applied to the group  $G$  acting on the set  $S$ , says that

$$|S| = \sum_{i=1}^r \frac{|G|}{|G_{A_i}|}$$

where,

$$G_{A_i} = \text{the stabilizer of } A_i = \{g \in G : gA_i = A_i\}.$$

Substitute the value of  $|S|$  and  $|G|$ , we get

$$\begin{aligned} & \binom{p^n m}{p^n} \\ &= \sum_{i=1}^r \frac{p^n m}{|G_{A_i}|} \end{aligned}$$

It can be proven that  $\binom{p^n m}{p^n}$  is not divisible by  $p$ . It follows that the sum in the last equation is not divisible by  $p$ , but each fraction  $p^n m / |G_{A_i}|$  is an integer, hence atleast one of these integers is not divisible by  $p$ . Hence there exists atleast one  $A_j \in S$  with the property that  $|G_{A_j}|$  is a multiple of  $p^n$ , since  $|G_{A_j}|$  needs to cancel the entire  $p^n$  in the numerator. Till now, we have shown that there is a subset  $A \subseteq G$  with  $|A| = p^n$  such that the order of the stabilizer  $G_A$  of  $A$  is divisible by  $p^n$ . We next select an element  $a \in A$ . By definition of  $G_A$ , for every  $g \in G$ , we have  $gA = A$ , so in particular

$$ga \in A \text{ for every } g \in G_A.$$

Applying  $a^{-1}$  on the right, we find that

$$g \in Aa^{-1} \text{ for every } g \in G_A.$$

Thus,  $G_A \subseteq Aa^{-1}$ , which implies that

$$|G_A| \leq |Aa^{-1}| = |A| = p^n.$$

We have shown that  $|G_A|$  is divisible by  $p^n$  and that  $|G_A|$  is at most  $p^n$ . It follows that

$$|G_A| = p^n,$$

so  $G_A$  is the desired subgroup of  $G$ .

□

**Example 3.1.** Let  $G = S_4$ . Since  $|S_4| = 4! = 24 = 2^3 \cdot 3$ , therefore by the Sylow's first theorem, there exists a Sylow 2-subgroup of order 8. This subgroup of order 8 is

$$H = \langle (12), (34), (13)(24) \rangle$$

**Definition 3.1 (Double Coset).** Let  $H_1 \subset G$  and  $H_2 \subset G$  be subgroups, and let  $g \in G$ . The *double coset* associated to  $g$  is the set

$$H_1 g H_2 = \{h_1 g h_2 : h_1 \in H_1 \text{ and } h_2 \in H_2\}.$$

Alternatively, we can define a *double coset equivalence relation* on  $G$  by saying that  $g \in G$  is equivalent to  $g' \in G$  if there is an  $h_1 \in H_1$  and an  $h_2 \in H_2$  satisfying  $g' = h_1 g h_2$ . The equivalence classes of  $g$  is precisely the double coset  $H_1 g H_2$ .

For the size of the double coset  $H_1 g H_2$ , we have the result

$$|H_1 g H_2| = \frac{|H_1| |H_2|}{|g^{-1} H_1 g \cap H_2|}.$$

**Theorem 3.2 (Sylow's second Theorem).** All  $p$ -Sylow subgroups of  $G$  are conjugate to one another; i.e., if  $H_1$  and  $H_2$  are  $p$ -Sylow subgroups of  $G$ , then there is some  $g \in G$  so that  $H_2 = g^{-1} H_1 g$ .

PROOF. We will use the double coset equivalence relation on  $G$  for the subgroups  $H_1$  and  $H_2$ . The equivalence classes are the double cosets  $H_1gH_2$ . Choose elements  $g_1, \dots, g_k \in G$  so that  $H_1g_1H_2, \dots, H_1g_kH_2$  are the distinct double cosets. Then since  $G$  is the disjoint union of these cosets. Then, we have

$$|G| = \sum_{i=1}^k |H_1g_iH_2| = \sum_{i=1}^k \frac{|H_1||H_2|}{|g_i^{-1}H_1g_i \cap H_2|}.$$

We now let  $|G| = p^n m$  with  $p \nmid m$ , so the  $p$ -Sylow subgroups satisfy  $|H_1| = |H_2| = p^n$ . This yields

$$p^n m = \sum_{i=1}^k \frac{p^{2n}}{|g_i^{-1}H_1g_i \cap H_2|}.$$

Each of the groups  $g_i^{-1}H_1g_i \cap H_2$  is a subgroups of  $H_2$ , and since order of the subgroup divide order of the group. So,  $|g_i^{-1}H_1g_i \cap H_2|$  divides  $|H_2| = p^n$ , which implies

$$|g_i^{-1}H_1g_i \cap H_2| = p^{n-t_i} \text{ with } 0 \leq t_i \leq n.$$

Substituting into the previous equation gives,

$$p^n m = \sum_{i=1}^k \frac{p^{2n}}{p^{n-t_i}} = \sum_{i=1}^k p^{n+t_i}, \text{ and hence } m = \sum_{i=1}^k p^{t_i}$$

Since  $p \nmid m$ ,  $t_i = 0$  for atleast one  $1 \leq j \leq k$  i.e.,  $t_j = 0$ . Then, we have,

$$|g_j^{-1}H_1g_j \cap H_2| = p^n.$$

Also, we know that each of  $g_j^{-1}H_1g_j$  and  $H_2$  have  $p^n$  elements, so we must have  $g_j^{-1}H_1g_j$  and  $H_2$ . This proves that  $H_1$  and  $H_2$  are conjugate subgroups.  $\square$

**Theorem 3.3 (Sylow's Third Theorem).** Let  $H_1, H_2, \dots, H_k$  be the distinct  $p$ -Sylow subgroups of  $G$ . Then:

- (1)  $k = \frac{|G|}{|N_G(H_1)|}$ , so in particular,  $k$  divides  $|G|$ .
- (2)  $k \equiv 1 \pmod{p}$ .

PROOF. (1) We know from second Sylow's theorem 3.2 that the  $p$ -Sylow subgroups are all conjugate to each other. So,  $H_1, \dots, H_k$  is the collection of conjugates of  $H_1$ . Since, we know that any subgroup  $H$  of  $G$  has  $\frac{|G|}{|N_G(H)|}$  distinct conjugates in  $G$ , hence  $k = \frac{|G|}{|N_G(H_1)|}$ .

- (2) Now, for the second part, let us assume  $H = H_1$  to be one of the  $p$ -Sylow subgroups. Consider the double coset decomposition of  $G$  using  $H$  on both the left and the right. We choose  $g_1, \dots, g_r \in G$  such that  $Hg_1H, \dots, Hg_rH$  are distinct double cosets of  $H$ . This facilitates the following computation.

$$|G| = \sum_{i=1}^r |Hg_iH| \quad \text{double coset decomposition,}$$

$$= \sum_{i=1}^r \frac{|H| \cdot |H|}{|g_i^{-1}Hg_i \cap H|}$$

$|G| = p^n m$  with  $p \nmid m$ , we know that  $|H| = p^n$ . Now, we aim to find the size of the intersections  $g_i^{-1}Hg_i \cap H$ . There are two possibilities.

- if  $g_i \in N_G(H)$ , then  $g_i^{-1}Hg_i = H$ , so

$$g_i \in N_G(H) \implies \frac{|H| \cdot |H|}{|g_i^{-1}Hg_i \cap H|} = \frac{p^{2n}}{p^n} = p^n.$$

- On the other hand, if  $g_i \notin N_G(H)$ , then  $g_i^{-1}Hg_i \cap H \subsetneq H$ . Thus,  $|g_i^{-1}Hg_i \cap H| < p^n$ . Hence,

$$g_i \notin N_G(H) \implies \frac{|H| \cdot |H|}{|g_i^{-1}Hg_i \cap H|} = \frac{p^{2n}}{p^k} = p^{2n-k} \quad \text{with } k < n, \text{ so } 2n - k \geq n + 1.$$

We use the last two equations to split the sum in the equation of  $|G|$  into two parts:

$$|G| = \sum_{g_i \in N_G(H)} \underbrace{\frac{|H| \cdot |H|}{|g_i^{-1}Hg_i \cap H|}}_{\text{equal to } p^n} + \sum_{g_i \notin N_G(H)} \underbrace{\frac{|H| \cdot |H|}{|g_i^{-1}Hg_i \cap H|}}_{\text{divisible by } p^{n+1}}.$$

We still need to count how many of the  $g_i$  are in  $N_G(H)$ . First, recall that  $g_1, \dots, g_r$  were chosen so that  $Hg_1H, \dots, Hg_rH$  are the distinct double cosets of  $H$ . By definition, every  $g \in N_G(H)$  satisfies  $Hg = gH$ . Thus, the double coset of  $g$  equals  $HgH = gHH = gH$ , which is simply the left coset  $gH$ . Hence  $g_i \in N_G(H)$  and  $g_j \in N_G(H)$  give the same double coset  $Hg_iH = Hg_jH$  if and only if they give the same left coset  $g_iH = g_jH$ . So, the number of distinct double cosets  $HgH$  as  $g$  is taken in  $N_G(H)$  is exactly the number of left  $H$  cosets in  $N_G(H)$ ; i.e.,

$$\{g_i : g_i \in N_G(H)\} = [N_G(H) : H] = \frac{|N_G(H)|}{|H|}.$$

Now we substitute (??) into (2) and then by using the fact that  $|H| = p^n$ , we get

$$|G| = |N_G(H)| + (\text{multiple of } p^{n+1}) :$$

The exact power of  $p$  dividing  $|N_G(H)|$  is  $p^n$ , so dividing by  $|N_G(H)|$  yields

$$\frac{|G|}{|N_G(H)|} = 1 + (\text{multiple of } p) :$$

(Note that the ratio  $\#G/\#N_G(H)$  is an integer from Lagrange's theorem, since  $N_G(H)$  is a subgroup of  $G$ .) Using  $k = \#G/\#N_G(H)$  from (c)(i), this gives (c)(ii).  $\square$

**Example 3.2.** We will show that if  $G$  is a group of order  $pq$  where  $p$  and  $q$  are prime with  $p < q$  and  $p$  does not divide  $q - 1$ , then  $G$  must be cyclic. Let  $H$  be a Sylow  $p$ -subgroup and  $K$  be a Sylow  $q$ -subgroup. Let  $n_p$  and  $n_q$  be the number of Sylow  $p$ -subgroups and Sylow  $q$ -subgroups, respectively. Using the Sylow third theorem  $n_p \equiv 1 \pmod{p}$  and  $n_p$  divides  $q$ . Thus,  $n_p = 1 + kp$ . Therefore,  $n_p = 1$  or  $q$ . But if  $n_p = q$ , then  $p|q - 1$ , which is a contradiction. Hence  $n_p = 1$ . In the same way  $n_q = 1$ .

Since there is only one of each Sylow subgroup, they are both normal in  $G$ . Take non-identity elements  $x \in H$  and  $y \in K$  and consider  $x^{-1}y^{-1}xy$ . By the normality of  $H, K$  we have that,  $x^{-1}y^{-1}xy = x^{-1}(y^{-1}xy) \in H$  and  $x^{-1}y^{-1}xy = (x^{-1}y^{-1}x)y \in K$ . By Lagrange's Theorem, we have  $H \cap K = \{1\}$ . Therefore  $x^{-1}y^{-1}xy = e$  and so  $xy = yx$ . Thus the order of  $xy$  is  $pq$  and so  $G$  is cyclic.

From this example, we can say that the groups of order 15, 33, 35, 65 are cyclic.

**Example 3.3.** Consider the Sylow 2-subgroups  $\{(1), (12)\}, \{(1), (23)\}, \{(1), (13)\}$  of  $S_3$ . Let us verify Sylow's third theorem for these Sylow 2-subgroups of  $S_3$ .

i.e., we will verify if we can obtain the last two subgroups by conjugation from the first subgroup.

$$\begin{aligned} (13)\{(1), (12)\}(13)^{-1} &= \{(1), (23)\}, \\ (23)\{(1), (12)\}(23)^{-1} &= \{(1), (13)\} \end{aligned}$$

Hence the Sylow's third theorem is valid for the above three subgroups of  $S_3$ .

**Example 3.4.** Consider the four Sylow 3-subgroups of  $A_4$ . These are:  
 $\{(1), (123), (132)\}, \{(1), (243), (234)\}, \{(1), (142), (124)\}, \{(1), (134), (143)\}$ . Then,

$$((12)(34))\{(1), (123), (132)\}((12)(34))^{-1} = \{(1), (142), (124)\}$$

$$((13)(24))\{(1), (123), (132)\}((13)(24))^{-1} = \{(1), (134), (143)\}$$

$$((14)(23))\{(1), (123), (132)\}((14)(23))^{-1} = \{(1), (243), (234)\}$$

We can verify that the number of Sylow 3-subgroups is 4 which is congruent to 1(mod3) and the Sylow 3-subgroups are conjugate.

**Example 3.5.** Let  $G$  be a group of order 12. Then we shall show that atleast one of it's Sylow subgroups is normal.

By the Sylow's third theorem,  $n_3 \equiv 1(\text{mod}3)$  and so  $n_3 = 1$  or 4. If  $n_3 = 1$ , then it's Sylow 3-subgroup is normal as for a given prime  $p$ , if there is a unique Sylow  $p$ -subgroup, then it must be normal.

If  $n_3 = 4$ , then the group  $G$  has eight elements of order 3 because the intersection of any two Sylow 3-subgroups contains only the identity. Therefore, the remaining four elements, except for the eight elements of order 3, form a Sylow 2-subgroup and so the Sylow 2-subgroup is unique and so it is normal.

**Example 3.6.** Determine all groups of order 21.

Let  $G$  be a group of order 21. Then by Cauchy's theorem, there exists  $a \in G, b \in G$  with  $|a| = 7$  and  $|b| = 3$ .  $\langle a \rangle \cap \langle b \rangle$  is a subgroup of both  $\langle a \rangle$  and  $\langle b \rangle$  and order of the subgroup divides order of the group, hence order of  $\langle a \rangle \cap \langle b \rangle$  divides both 7 and 3, i.e.,  $\langle a \rangle \cap \langle b \rangle = \{e\}$  and thus  $|\langle a \rangle \cap \langle b \rangle| = |a| \cdot |b| = 21$ , which means  $G = \langle a \rangle \cap \langle b \rangle$ .

Now,  $\langle a \rangle$  is a Sylow 7-subgroup of  $G$  and  $n_7 \equiv 1(\text{mod}7)$  and  $n_7 | 3$ . Hence  $n_7 = 1$  i.e., the Sylow 7-subgroup is unique and hence normal in  $G$ . So,  $\langle a \rangle \trianglelefteq G$  and  $bab^{-1} = a^i$  where  $1 \leq i < 7$ . If  $i = 1$ , then  $bab^{-1} = a$  and so  $G$  is abelian and  $|ab| = 21$ . For this,  $G = \langle ab \rangle \cong \mathbb{Z}_{21}$ .

Now assume  $i \neq 1$ , then  $b^3 = 1$  and  $bab^{-1} = a^i$ , with  $2 \leq i < 7$ , imply that  $a = b^3ab^{-3} = a^{i^3}$ . Therefore,  $7 | i^3 - 1$  and so  $i = 2$  or  $i = 4$ . Also,  $[2]_7 = [4]_7^{-1}$  because  $8 \equiv 1(\text{mod}7)$ . We can assume  $i = 2$  since, if required we can replace  $b$  by  $b^{-1}$ .

So far we can say that  $G = \{a^u b^v | 0 \leq u < 7, 0 \leq v < 3\}$  and that the relations  $a^7 = 1 = b^3, bab^{-1} = a^2$  hold. Infact there exists such a group. Take  $\alpha = (1234567)$  and  $\beta = (235)(476)$ : thus  $\langle \alpha, \beta \rangle$  is a subgroup of  $S_7$ . Observe that  $\alpha^7 = 1 = \beta^3$  and  $\beta\alpha\beta^{-1} = \alpha^2$ .

## Introduction to linear transformation

### 1. Linear Transformation, Null space, and Range

**Proposition 1.1.** Let  $V$  be a vector space over a field  $F$ .

- (1) If  $z \in V$  satisfies  $z + v = v$  for all  $v \in V$ , then  $z = 0_V$ .
- (2) If  $v, w \in V$  then there exists a unique element  $y \in V$  satisfying  $v + y = w$ .

PROOF. (1) Let  $z \in V$ . Since  $z + v = v$  is true for all  $v \in V$ , hence it is true in particular for  $v = 0_V$ , hence  $z + 0_V = 0_V$  or  $z = 0_V$ .

(2) Given  $v \in V$  and  $w \in V$ , take  $y = w - v$ , then  $v + y = v + (w - v) = v + (w + (-v)) = v + (-v + w) = (v + (-v)) + w = 0_V + w = w$ . If  $x \in V$  is another element such that  $v + x = w$ , then  $x = 0_V + x = (-v + v) + x = (-v) + (v + x) = (-v) + w = w + (-v) = w - v = y$ , hence  $y$  is unique.  $\square$

**Proposition 1.2.** Let  $V$  be a vector space over a field  $F$ . If  $v, w \in V$  and if  $a$  in  $F$ , then:

- (1)  $a0_V = 0_V$ ;
- (2)  $0v = 0_V$ ;
- (3)  $(-1)v = -v$ ;
- (4)  $(-a)v = -(av) = a(-v)$ ;
- (5)  $-(-v) = v$ ;
- (6)  $av = (-a)(-v)$ ;
- (7)  $-(v + w) = -v - w$ ;
- (8)  $a(v - w) = av - aw$ ;
- (9) If  $av = 0_V$  then either  $v = 0_V$  or  $a = 0$ .

PROOF. (1)  $a0_V + 0_V = a0_V = a(0_V + 0_V) = a0_V + a0_V$ . Add  $-(a0_V)$  both sides, we get  $(a0_V + 0_V) + (-a0_V) = (a0_V + a0_V) + (-a0_V)$  which implies  $a0_V + (0_V + (-a0_V)) = a0_V$  or  $a0_V + (-a0_V) = a0_V$  or  $a0_V = 0_V$ .

(2)  $0v + 0v = (0 + 0)v = 0v$ . Adding  $-0v$  from both sides, we get  $(0v + 0v) + (-0v) = 0v + (-0v)$  or  $0v + (0v + (-0v)) = 0_V$  or  $0v + 0_V = 0_V$  or  $0v = 0_V$ .

(3)  $0_V = 0v = (1 + (-1))v = 1v + (-1)v = v + (-1)v$ , which implies that the additive inverse of  $v$  is  $(-1)v$  i.e.,  $-v = (-1)v$ .

(4)  $0_V = 0v = (a + (-a))v = av + (-a)v$  which means the additive inverse of  $av$  is  $(-a)v$  i.e.,  $-(av) = (-a)v$ .  
 $0_V = a0_V = a(v + (-v)) = av + a(-v)$  which means the inverse of  $av$  is  $a(-v)$  i.e.,  $-(av) = a(-v)$ .

(5)  $v + (-v) = 0_V$  which means inverse of  $-v$  is  $v$  i.e.,  $-(-v) = v$ .

(6)  $(-a)(-v) = -a(-v) = -(-(av))$  using (4) and it is equal to  $av$  by (5).

(7)  $0_V = 0v + 0v = (v + (-v)) + (w + (-w)) = (-v + v) + (-w + w) = -v + (v + (-w)) + w = -v + ((-w) + v) + w = (-v + (-w)) + (v + w)$  which implies that the additive inverse of  $v + w$  is  $-v + (-w)$  or  $-(v + w) = -v - w$ .

(8)  $v - w = v + (-w)$ . So,  $a(v - w) = a(v + (-w)) = av + a(-w) = av - aw$  using (4).

(9) If  $a = 0$ , we are done, otherwise  $a \neq 0$ . Then  $a^{-1}$  exists and so  $a^{-1}(av) = a^{-1}(0_V)$  or  $(a^{-1}a)v = 0_V$  or  $1v = 0_V$  or  $v = 0_V$ .

□

**Proposition 1.3.** If  $V$  is a vector space over a field  $F$ , then a nonempty subset  $W$  of  $V$  is a subspace of  $V$  if and only if it is closed under addition and scalar multiplication.

PROOF. First, assume that  $W$  is a subspace of  $V$ . Then  $W$  is a vector space in its own right and so it is closed under addition and scalar multiplication.

Now, suppose that  $W$  is closed under addition and scalar multiplication. For any  $w \in W$ ,  $0_V \in W$  because  $0_V = 0w$ . Since  $0 \in F$ ,  $w \in W$  and  $W$  is closed under scalar multiplication, hence  $0_V = 0w \in W$ . Also,  $-w \in W$  because  $-w = (-1)w$ . Since  $-1 \in F$ ,  $w \in W$  and  $W$  is closed under scalar multiplication, hence  $-w = (-1)w \in W$ . □

**Problem 1.1.** Determine whether the following are the subspaces of  $\mathbb{R}^2$  or not:

- (1)  $\{(x, y) | x^2 + y^2 \leq 1, x, y \in \mathbb{R}\}$
- (2)  $\{(x, y) | xy \geq 0, x, y \in \mathbb{R}\}$
- (3)  $\{(x, y) | x, y \text{ are integers}\}$

**Solution 1.1.** (1) Let  $S = \{(x, y) | x^2 + y^2 \leq 1, x, y \in \mathbb{R}\}$ . Observe that  $(\frac{1}{2}, \frac{1}{2}) \in S$  and  $(0, 1) \in S$ , but  $(\frac{1}{2} + 0, \frac{1}{2} + 1) = (\frac{1}{2}, \frac{3}{2}) \notin S$  and so  $S$  is not a subspace of  $\mathbb{R}^2$ .  
 (2) Let  $S = \{(x, y) | xy \geq 0, x, y \in \mathbb{R}\}$ . Observe that  $(-\frac{1}{2}, -\frac{3}{2}) \in S$  and  $(1, 1) \in S$  but  $(-\frac{1}{2} + 1, -\frac{3}{2} + 1) = (\frac{1}{2}, -\frac{1}{2}) \notin S$  and so  $S$  is not a subspace of  $\mathbb{R}^2$ .  
 (3) Let  $S = \{(x, y) | x, y \text{ are integers}\}$ . Observe that  $(2, 3) \in S$  but for  $a = 0.3$ ,  $a(2, 3) = (0.6, 0.9) \notin S$  and so  $S$  is not closed under scalar multiplication and so  $S$  is not a subspace of  $\mathbb{R}^2$ .

**Problem 1.2.** Let  $V$  be the subspace of  $\mathbb{R}^4$  spanned by the 4-tuples

$$\alpha_1 = (1, 2, 3, 4), \alpha_2 = (2, 3, 4, 5), \alpha_3 = (3, 4, 5, 6), \alpha_4 = (4, 5, 6, 7).$$

Find a basis of  $V$  and  $\dim V$ .

**Solution 1.2.** Observe that  $\alpha_4 - \alpha_3 = (1, 1, 1, 1)$ ,  $\alpha_3 - \alpha_2 = (1, 1, 1, 1)$ ,  $\alpha_2 - \alpha_1 = (1, 1, 1, 1)$  and so  $\alpha_4 - \alpha_3 = \alpha_3 - \alpha_2 = \alpha_2 - \alpha_1$ . Therefore,  $\alpha_3 = 2\alpha_2 - \alpha_1$  and  $\alpha_4 = 3\alpha_2 - 2\alpha_1$ . We claim that  $\alpha_1$  and  $\alpha_2$  are linearly independent. Let  $a\alpha_1 + b\alpha_2 = 0$ . Then  $a(1, 2, 3, 4) + b(2, 3, 4, 5) = 0$  or  $(a + 2b, 2a + 3b, 3a + 4b, 4a + 5b) = 0$ , the solution of which is  $a = 0$  and  $b = 0$  and so  $\alpha_1$  and  $\alpha_2$  are linearly independent. Hence the basis of  $V = B = \{\alpha_1, \alpha_2\}$  and  $\dim V = 2$ .

**Proposition 1.4.** If  $V$  is a vector space over a field  $F$  and  $\{W_i | i \in \Omega\}$  is a collection of subspaces of  $V$ , then  $\cap_{i \in \Omega} W_i$  is a subspace of  $V$ .

PROOF. Let  $W = \cap_{i \in \Omega} W_i$ . Let  $w_1, w_2 \in W$ , then for each  $i \in \Omega$ ,  $w_1, w_2 \in W_i$  and since each  $W_i$  is a subspace,  $w_1 + w_2 \in W_i$  for each  $i$  and so  $w_1 + w_2 \in W$ . Also, if  $a \in F$  and  $w \in W$ , then  $w \in W_i$  for each  $i$  and since each  $W_i$  is a subspace, so  $aw \in \cap_{i \in \Omega} W_i$ . So,  $W = \cap_{i \in \Omega} W_i$  is a subspace of  $V$ . □

**Proposition 1.5.** Let  $D$  be a subset of a vector space  $V$  over a field  $F$ . Then

- (1)  $FD$  is a subspace of  $V$ .
- (2) Every subspace of  $V$  containing  $D$  also contains  $FD$ .

PROOF. (1) The trivial case is when  $D = \emptyset$ , then  $FD = \{0_V\}$  which is a subspace of  $V$ . Let  $D \neq \emptyset$ . Assume  $D = \{v_i | i \in I\}$  for some indexing set  $I$ . Let  $x = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n}$  and  $y = \beta_1 v_{j_1} + \beta_2 v_{j_2} + \cdots + \beta_m v_{j_m}$  be the two elements of  $FD$ . Then  $x + y = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n} + \beta_1 v_{j_1} + \beta_2 v_{j_2} + \cdots + \beta_m v_{j_m}$  is also a linear combination of the elements of  $D$  and so  $x + y$  is an element of  $FD$ . Also, for  $\alpha \in F$ ,  $\alpha x = \alpha(\alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n}) = \alpha\alpha_1 v_{i_1} + \alpha\alpha_2 v_{i_2} + \cdots + \alpha\alpha_n v_{i_n}$  which is also a linear combination of the elements of  $D$  and hence  $\alpha x \in FD$ . Hence  $FD$  is subspace of  $V$ .

- (2) Let  $U$  be a subspace of  $V$  such that  $D \subseteq U \subseteq V$ . We have to prove that  $FD \subseteq U$ . Let  $x = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n} \in FD$ . Since  $v_{i_1}, v_{i_2}, \dots, v_{i_n} \in D$ , hence  $v_{i_1}, v_{i_2}, \dots, v_{i_n} \in U$  and since  $U$  is a subspace,  $x = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n} \in U$  and so  $FD \subseteq U$ . □

**Proposition 1.6.** Let  $V$  be a vector space over a field  $F$ , and  $D_1$  and  $D_2$  be subsets of  $V$  satisfying  $D_1 \subseteq D_2 \subseteq FD_1$ . Then  $FD_1 = FD_2$ .

PROOF.  $FD_1$  is a subspace of  $V$  containing  $D_2$ . By Proposition 1.5(2),  $FD_2 \subseteq FD_1$ . Let  $x = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n} \in FD_1$ . Since  $D_1 \subseteq D_2$ , so  $x = \alpha_1 v_{i_1} + \alpha_2 v_{i_2} + \cdots + \alpha_n v_{i_n} \in FD_2$  and so  $FD_1 \subseteq FD_2$ . Hence  $FD_1 = FD_2$ .  $\square$

**Proposition 1.7 (Exchange Property).** Let  $V$  be a vector space over a field  $F$  and let  $v, w \in V$ . Let  $D$  be a subset of  $V$  satisfying  $v \in F(D \cup \{w\}) - FD$ . Then  $w \in F(D \cup \{v\})$ .

PROOF. Since  $v \in F(D \cup \{w\})$ , there exists scalars  $\alpha_1, \alpha_2, \dots, \alpha_n, \beta \in F$  and elements  $v_1, v_2, \dots, v_n \in D$  such that  $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n + \beta w$ . Also, since  $v \notin FD$ , so  $\beta \neq 0$  and so  $\beta w = v - (\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n)$  or  $w = b^{-1}v - (b^{-1}\alpha_1 v_1 + b^{-1}\alpha_2 v_2 + \cdots + b^{-1}\alpha_n v_n)$  which lies in  $F(D \cup \{v\})$ .  $\square$

**Proposition 1.8.** If  $V$  is a vector space over a field  $F$  and if  $\{W_i | i \in \Omega\}$  is a collection of subspaces of  $V$ , then

- (1)  $W_h$  is a subspace of  $\sum_{i \in \Omega} W_i$  for all  $h \in \Omega$ ;
- (2) If  $Y$  is a subspace of  $V$  satisfying the condition that  $W_h$  is a subspace of  $Y$  for all  $h \in \Omega$ , then  $\sum_{i \in \Omega} W_i$  is a subspace of  $Y$ .

PROOF. (1) Since  $W_h$  is a subspace of  $V$ , we have  $W_h \neq 0$ . Also, for each  $w \in W_h, w \in \sum_{i \in \Omega} W_i$ , so  $W_h$  is a subspace of  $\sum_{i \in \Omega} W_i$ .  
 (2)  $W_h \subseteq \sum_{i \in \Omega} W_i$  because for each element  $w \in W_h$ , it can be written as  $w + \sum_{i \neq h} w_i$  where  $w_i = 0$  for  $i \neq h$ . Since  $W_h$  is a subspace of  $V$ , we have  $W_h \neq \emptyset$ . Let  $w_1, w_2 \in W_h$ . Since  $W_h$  is a subspace, we have  $w_1 + w_2 \in W_h$ . Let  $a \in F$ . Since  $W_h$  is a subspace, we have  $aw_1 \in W_h$ . Hence,  $W_h$  is a subspace of  $V$ .  $\square$

**Proposition 1.9.** If  $V$  is a vector space over a field  $F$  and if  $W_1, W_2$ , and  $W_3$  are subspaces of  $V$ , then:

- (1)  $(W_1 + W_2) + W_3 = W_1 + (W_2 + W_3)$ ;
- (2)  $W_1 + W_2 = W_2 + W_1$ ;
- (3)  $W_3 \cap [W_2 + (W_1 \cap W_3)] = (W_1 \cap W_3) + (W_2 \cap W_3)$ ;

PROOF. (1) For  $w_1 \in W_1, w_2 \in W_2$  and  $w_3 \in W_3$ , since  $w_1, w_2, w_3 \in V$  also and because of the associativity property of addition of the vector space  $V$ ,  $(w_1 + w_2) + w_3 = w_1 + (w_2 + w_3)$  and so set inclusion is true on both sides.  
 (2) For  $w_1 \in W_1$  and  $w_2 \in W_2$ , since  $w_1, w_2 \in V$  and because of the commutativity property of addition of the vector space  $V$ ,  $w_1 + w_2 = w_2 + w_1$  and so the set inclusion is true on both sides i.e.,  $W_1 + W_2 \subseteq W_2 + W_1$  and  $W_2 + W_1 \subseteq W_1 + W_2$ .  
 (3) Let  $x \in W_3 \cap [W_2 + (W_1 \cap W_3)]$ . Then  $x \in W_3$  and  $x \in W_2 + (W_1 \cap W_3)$ . Since  $x \in W_2 + (W_1 \cap W_3)$ , so  $x = y + z$  for some  $y \in W_2$  and  $z \in W_1 \cap W_3$ . Since  $x \in W_3$  and  $z \in W_3$ , so  $y = x - z \in W_3$  and so  $y \in W_2 \cap W_3$  and  $z \in W_1 \cap W_3$ . Therefore,  $x = y + z \in (W_2 \cap W_3) + (W_1 \cap W_3) = (W_1 \cap W_3) + (W_2 \cap W_3)$ , by (2) and so  $W_3 \cap [W_2 + (W_1 \cap W_3)] \subseteq (W_1 \cap W_3) + (W_2 \cap W_3)$ .

Let  $x \in (W_1 \cap W_3) + (W_2 \cap W_3)$ . Then  $x = y + z$  for some  $y \in W_1 \cap W_3$  and  $z \in W_2 \cap W_3$ . Since  $z \in W_2$  and  $y \in W_1 \cap W_3$ , so  $x = y + z = z + y \in W_2 + (W_1 \cap W_3)$ . Also, since  $y \in W_3$  and  $z \in W_3$ , so  $x = y + z \in W_3$  and so  $x \in W_3 \cap [W_2 + (W_1 \cap W_3)]$ . Therefore,  $(W_1 \cap W_3) + (W_2 \cap W_3) \subseteq W_3 \cap [W_2 + (W_1 \cap W_3)]$ .  $\square$

## 2. Algebras over a Field

**Proposition 2.1.** If  $v$  and  $w$  are nonzero elements of  $\mathbb{R}^3$ , then  $v \times w = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  if and only if  $\mathbf{R}v = \mathbf{R}w$ .

PROOF. Let  $v = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$  and  $w = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ . It is given that the vectors  $v$  and  $w$  are non-zero vectors and so at least one of  $b_1, b_2$  or  $b_3$  is non-zero. Let us assume that  $b_3 \neq 0$ . Then,  $a_2b_3 - a_3b_2 = a_3b_1 - a_1b_3 = a_1b_2 - a_2b_1 = 0$  and so if we define  $c = a_1b_1^{-1}$ , then by the above relations,  $c = a_1b_1^{-1} = a_2b_2^{-1} = a_3b_3^{-1}$  and so  $a_1 = cb_1, a_2 = cb_2, a_3 = cb_3$  and so  $v = cw$ . Hence  $v \in \mathbf{R}w$ . Also,  $c \neq 0$  implies that  $w = c^{-1}v \in \mathbf{R}v$ , proving  $\mathbf{R}v = \mathbf{R}w$ .

Conversely, assume that  $\mathbf{R}v = \mathbf{R}w$ , then since  $w \in \mathbf{R}w$ , so  $w \in \mathbf{R}v$  and so  $w = dv$  for some  $d \neq 0$ . Then  $v \times w = v \times (dv) = d(v \times v) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$  □

**Proposition 2.2.** Let  $F$  be a field and  $p(X)$  be a polynomial in  $F[X]$ . Then an element  $c$  of  $F$  satisfies the condition that  $p(c) = 0$  if and only if there exists a polynomial  $u(X) \in F[X]$  satisfying  $p(X) = (X - c)u(X)$ .

PROOF. From the previous proposition, we know that there exists polynomials  $a(X)$  and  $b(X) \in F[X]$  such that  $p(X) = (X - c)a(X) + b(X)$ , where  $\deg(b(X)) < \deg(X - c) = 1$  and so  $\deg(b(X)) = 0$ . So,  $b(X) = d$  for some  $d \in F$  and  $p(X) = (X - c)a(X) + d$ .

If there exists an element  $c \in F$  such that  $p(c) = 0$ , then  $(c - c)a(c) + d = 0$  and so  $d = 0$ . Therefore,  $p(X) = (X - c)a(X)$  and so there exists  $u(X) = a(X) \in F[X]$  such that  $p(X) = (X - c)u(X)$ .

Now suppose that there exists a polynomial  $u(X) \in F[X]$  such that  $p(X) = (X - c)u(X)$ . Then  $p(c) = (c - c)u(c) = 0u(c) = 0$ . □

**Proposition 2.3.** A monic polynomial  $p(X) \in \mathbb{R}[X]$  is irreducible if and only if it is of the form  $X - a$  or  $(X - a)^2 + b^2$ , where  $a \in \mathbb{R}$  and  $0 \neq b \in \mathbb{R}$ .

PROOF. It is evident that every polynomial of the form  $X - a$  is irreducible. Now, let us consider the polynomial  $f(X) = (X - a)^2 + b^2$ , which expands to  $X^2 - 2aX + a^2 + b^2$ . If this polynomial were reducible, we would be able to identify real numbers  $c$  and  $d$  such that  $f(X) = (X - c)(X - d) = X^2 - (c + d)X + cd$ . This leads to the equations  $c + d = 2a$  and  $cd = a^2 + b^2$ . Consequently, we derive  $c^2 - 2ac + a^2 + b^2 = 0$ , resulting in  $c = \frac{1}{2}[2a \pm \sqrt{4a^2 - 4(a^2 + b^2)}] = a \pm \sqrt{-b^2}$ . This outcome contradicts the assumption that  $c$  is a real number, given that  $b$  is presumed to be nonzero. Therefore, we conclude that polynomials of both specified forms are indeed irreducible.

Conversely, let  $p(X) = \sum_{i=0}^n c_i X^i$  be a monic irreducible polynomial in  $\mathbb{R}[X]$  that does not conform to the form  $X - a$ . According to the Fundamental Theorem of Algebra, there exists a complex number  $z = a + bi$  such that  $p(z) = 0$ . Since the coefficients of  $p(X)$  are real, it follows that  $\overline{p(z)} = p(\overline{z}) = 0$  as well, leading to the equation  $0 = p(z) = \sum_{i=0}^n c_i z^i = p(\overline{z})$ . Thus, we can express  $p(X)$  as  $(X - z)(X - \overline{z})u(X)$ , where  $(X - z)(X - \overline{z}) = X^2 - (z + \overline{z})X + z\overline{z} = X^2 - 2aX + a^2 + b^2$ . Given that  $p(X)$  was assumed to be irreducible, we deduce that  $z$  is not a real number (i.e.,  $b \neq 0$ ) and that  $p(X)$  must equal  $X^2 - 2aX + a^2 + b^2$ , as required. □

### 3. Linear Independence and Dimension

**Proposition 3.1.** Let  $V$  be a vector space over a field  $F$ . A nonempty subset  $D$  of  $V$  is linearly dependent if and only if some element of  $D$  is a linear combination of the others over  $F$ .

PROOF. Suppose  $D$  is linearly dependent. Then there exists distinct vectors  $v_1, v_2, \dots, v_n \in D$  and scalars  $a_1, a_2, \dots, a_n \in F$ , not all equal to 0 such that  $\sum_{i=1}^n a_i v_i = 0_V$ . Suppose that  $a_m \neq 0$ . Then  $a_m v_m = -\sum_{i=1, i \neq m}^n a_i v_i$  i.e.,  $v_m = -a_m^{-1} \sum_{i=1, i \neq m}^n a_i v_i$  and so  $v_m$  is a linear combination of the elements of other elements.

Suppose there exists some element  $v_m \in V$ , which is a linear combination of the other elements of  $D$ . Let  $v_m = \sum_{i=1, i \neq m}^n a_i v_i$ . Then  $\sum_{i=1}^n a_i v_i = 0$ , where  $a_m = -1$  and so  $D$  is linearly dependent set.  $\square$

**Problem 3.1.** Let  $V$  and  $W$  be vector spaces over  $\mathbb{F}$ . Denote by  $V \times W$  the collection of all ordered pairs  $(v, w)$ , where  $v \in V$  and  $w \in W$ , and define

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$$

and

$$k(v, w) = (kv, kw), k \in \mathbb{F}.$$

- (1) Show that  $V \times W$  is a vector space over  $F$ .
- (2) Show that if  $V$  and  $W$  are finite dimensional, so is  $V \times W$ .
- (3) Find  $\dim(V \times W)$ , given that  $\dim V = m$  and  $\dim W = n$ .
- (4) Explain why  $\mathbb{R} \times \mathbb{R}^2$  can be identified with  $\mathbb{R}^3$ .
- (5) Find a basis for  $\mathbb{R}^2 \times M_2(\mathbb{R})$ .
- (6) What is the dimension of  $M_2(\mathbb{R}) \times M_2(\mathbb{R})$ .

**Solution 3.1.** (1) Let  $(v_1, w_1), (v_2, w_2), (v_3, w_3) \in V \times W$  and  $k, m \in \mathbb{F}$ . Then  $v_1 + v_2 \in V$  because  $V$  is a vector space and  $w_1 + w_2 \in W$  because  $W$  is a vector space and so  $(v_1 + w_1, v_2 + w_2) \in V \times W$ . So,  $V \times W$  is closed under vector addition.

Also,  $k(v_1, w_1) = (kv_1, kw_1) \in V \times W$  because  $kv_1 \in V, kw_1 \in W$  as  $V$  is a vector space and so  $V \times W$  is closed under scalar multiplication.

Now,  $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2) = (v_2 + v_1, w_2 + w_1) = (v_2, w_2) + (v_1, w_1)$  and so addition is commutative in  $V \times W$ .

$((v_1, w_1) + (v_2, w_2)) + (v_3, w_3) = (v_1 + v_2, w_1 + w_2) + (v_3, w_3) = ((v_1 + v_2) + v_3, (w_1 + w_2) + w_3) = (v_1 + (v_2 + v_3), w_1 + (w_2 + w_3))$  and so addition is associative in  $V \times W$ .

$(0_V, 0_W)$  is the identity element of  $V \times W$  because for any element  $(v, w) \in V \times W$ , we have  $(v, w) + (0_V, 0_W) = (v + 0_V, w + 0_W) = (v, w) = (0_V + v, 0_W + w) = (0_V, 0_W) + (v, w)$ .

Now,  $1$  is the multiplicative identity element of  $V \times W$  because for any  $(v, w) \in V \times W, 1(v, w) = (1v, 1w) = (v, w)$ .

$(km)(v, w) = ((km)v, (km)w) = (k(mv), k(mw)) = k(mv, mw)$  where we used this property in  $V$  and  $W$  individually to prove that this property is valid in  $V \times W$ .

$$k((v_1, w_1) + (v_2, w_2)) = k(v_1 + v_2, w_1 + w_2) = (k(v_1 + v_2), k(w_1 + w_2)) = (kv_1 + kv_2, kw_1 + kw_2) = (kv_1, kw_1) + (kv_2, kw_2)$$

$$(k + m)(v, w) = ((k + m)v, (k + m)w) = (kv + mv, kw + mw) = (kv, kw) + (mv, mw) = k(v, w) + m(v, w)$$

Since all the axioms of the vector space are valid, hence  $V \times W$  is a vector space.

- (2) Since  $V$  and  $W$  are finite dimensional, each of  $V$  and  $W$  has a finite basis. If  $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$  is a basis for  $V$  and  $\{\beta_1, \beta_2, \dots, \beta_n\}$  is a basis for  $W$ , then  $(\alpha_i, \beta_j), i = 1, 2, \dots, m; j = 1, 2, \dots, n$ , form a basis for  $V \times W$ .
- (3) Since  $V \times W$  has a basis which has  $mn$  elements, so  $\dim(V \times W) = mn$ .
- (4) There is a one-to-one correspondence between the elements  $(x, (y, z)) \in \mathbb{R} \times \mathbb{R}^2$  and  $(x, y, z) \in \mathbb{R}^3$  and so  $\mathbb{R} \times \mathbb{R}^2$  can be identified with  $\mathbb{R}^3$ .

- (5) The elements  $e_1 = (1, 0), e_2 = (0, 1) \in \mathbb{R}^2$  are the basis of  $\mathbb{R}^2$ . Let  $E_{ij}$  be the  $2 \times 2$  matrix with  $(i, j)$ th entry 1 and all remaining entries 0;  $i = 1, 2; j = 1, 2$ . Then  $E_{11}, E_{12}, E_{21}, E_{22}$  form a basis for  $M_2(\mathbb{R})$ .
- (6) There are 4 elements in the basis of  $M_2(\mathbb{R})$ , and so there will be  $4 \times 4 = 16$  elements in the basis of  $M_2(\mathbb{R}) \times M_2(\mathbb{R})$ . These 16 elements will be in the form of ordered pairs  $(E_{11}, E_{11}), (E_{11}, E_{12})$ , and so on.

**Proposition 3.2.** Let  $V$  be a vector space over a field  $F$ . Then the following conditions on a subset  $D$  of  $V$  are equivalent:

- (1)  $D$  is a minimal set of generators of  $V$ .
- (2)  $D$  is a maximal linearly-independent subset of  $V$ .
- (3)  $D$  is a linearly-independent set of generators of  $V$ .

PROOF. (1) Let  $D$  be a minimal set of generators of  $V$ . Suppose that  $D$  is linearly dependent. By the previous proposition, there exists an element  $v_0 \in D$ , which is a linear combination of elements of  $E = D - \{v_0\}$  over  $F$ . Let  $v_0 = \sum_{i=1}^n a_i u_i$ , where  $u_i \in E$  and the scalars  $a_i \in F$ . Let  $v \in V$ . Since  $D$  is a set of generators of  $V$ , there exists elements  $v_1, v_2, \dots, v_n \in E$  and scalars  $b_0, b_1, \dots, b_n \in F$  such that  $v = \sum_{j=0}^n b_j v_j$ . This implies that  $v = b_0 v_0 + \sum_{j=1}^n b_j v_j = \sum_{i=1}^n b_0 a_i u_i + \sum_{j=1}^n b_j v_j$  and so  $E$  is a set of generators of  $V$ , which contradicts the minimality of  $D$ . Hence  $D$  is linearly independent. If  $v \in V - D$ , then the set  $D \cup v$  is linearly dependent since  $v$  is a linear combination of the elements of  $D$ . Thus,  $D$  is maximal linearly independent set.

(2) Suppose that  $D$  is a maximal linearly independent subset of  $V$ . Let  $v_0 \in V - D$ . By (2),  $D \cup \{v_0\}$  is linearly dependent, and so  $0_V \in F(D \cup \{v_0\}) - FD$  which implies  $v_0 \in F(D \cup \{0_V\}) = FD$ . This proves that  $D$  is a set of generators of  $V$ .

(3) Let  $D$  be a linearly independent set of generators of  $V$  and  $E$  be a proper subset of  $D$  such that  $E$  generates  $D$ . Let  $v_0 \in D - E$ . Then there exists elements  $v_1, v_2, \dots, v_n \in E$  and scalars  $a_1, a_2, \dots, a_n \in F$  such that  $v_0 = \sum_{i=1}^n a_i v_i$ . By the previous proposition, this means that  $D$  is linearly dependent contradicting the statement of (3). Hence, no such  $E$  exists and so  $D$  is a minimal set of generators of  $V$ . □

**Proposition 3.3.** Let  $V$  be a vector space over a field  $F$  and  $D$  be a linearly-independent subset of  $V$ . If  $v_0 \in V \setminus D$ , then the set  $D \cup \{v_0\}$  is linearly independent.

PROOF. Consider that this collection is linearly dependent. In this case, there are elements  $v_1, \dots, v_n$  of  $D$  along with non-zero scalars  $a_0, a_1, \dots, a_n$ , not all of which are zero, such that  $\sum_{i=0}^n a_i v_i = 0_V$ . The non-zero scalar  $a_0$  is necessary not equal to 0, as if it were zero,  $D$  would be linearly dependent, which would lead to a contradiction. Hence,  $v_0 = \sum_{i=1}^n -a_0^{-1} a_i v_i \in FD$ , which contradicts the choice of  $v_0$ . Therefore, the set  $D \cup \{v_0\}$  must be linearly independent. □

**Proposition 3.4.** Let  $V$  be a vector space over a field  $F$  and let  $D$  be a nonempty subset of  $V$ . Then  $D$  is a basis of  $V$  if and only if every vector in  $V$  can be written as a linear combination of elements of  $D$  over  $F$  precisely in one way.

PROOF. Let us start by assuming that  $D$  forms a basis for  $V$  and that there is an element  $v$  in  $V$  that can be expressed as a sum of  $D$ 's elements over  $F$  in two distinct ways. In other words, there is a finite collection  $\{v_1, \dots, v_n\}$  of  $D$  and a set of numbers  $\{a_1, \dots, a_n, b_1, \dots, b_n\}$  in  $F$  such that

$$v = \sum_{i=1}^n a_i v_i = \sum_{i=1}^n b_i v_i,$$

where  $a_h \neq b_h$  for at least one index  $h$ . Then, the equation  $0_V = v - v$  simplifies to

$$v - v = \left( \sum_{i=1}^n a_i v_i \right) - \left( \sum_{i=1}^n b_i v_i \right) = \sum_{i=1}^n (a_i - b_i) v_i,$$

where at least one of the numbers  $a_i - b_i$  is not zero. This contradicts the idea that  $D$  is a basis, meaning it's linearly independent. Therefore, each vector in  $V$  can be uniquely expressed as a linear combination of  $D$ 's elements over  $F$ .

Now, suppose that every vector in  $V$  can be uniquely expressed as a linear combination of  $D$ 's elements over  $F$ , it is clear that  $D$  is a generating set for  $V$  over  $F$ . If  $\{v_1, \dots, v_n\}$  is a subset of  $D$  and  $\{a_1, \dots, a_n, b_1, \dots, b_n\}$  are numbers satisfying

$$\sum_{i=1}^n a_i v_i = 0_V$$

then we have

$$\sum_{i=1}^n a_i v_i = \sum_{i=1}^n 0 v_i$$

and thus, by the uniqueness of representation, we can conclude that  $a_i = 0$  for each  $1 \leq i \leq n$ . This demonstrates that  $D$  is linearly independent and thus serves as a basis.  $\square$

**Proposition 3.5.** Let  $D$  be a basis for a vector space  $V$  over a field  $F$ . Then any function  $f : D \times D \rightarrow V$  can be extended in a unique manner to a function  $V \times V \rightarrow V$ , which defines on  $V$  the structure of an  $F$ -algebra. Moreover, all  $F$ -algebra structures on  $V$  arise in this manner.

PROOF. Let  $D = \{y_i | i \in \Omega\}$ . Assume we are provided with a function  $f : D \times D \rightarrow V$ .

We establish an operation  $\cdot$  on  $V$  as follows: if  $v, w$  are in  $V$ , then, according to previous proposition, we can express  $v = \sum_{i \in \Omega} a_i y_i$  and  $w = \sum_{j \in \Omega} b_j y_j$  uniquely, where  $a_i$  and  $b_j$  are scalars, and only a finite number of which are nonzero; then,  $v \cdot w = \sum_{i \in \Omega} \sum_{j \in \Omega} a_i b_j f(y_i, y_j)$ . It is easy to demonstrate that this establishes the structure of an  $F$ -algebra on  $V$ . On the other hand, if  $(V, \cdot)$  is an  $F$ -algebra, we can define the function  $f : D \times D \rightarrow V$  by  $f(y_i, y_j) = i \cdot j$ .  $\square$

**Problem 3.2.** Show that  $\alpha_1 = (1, 1, 0)$ ,  $\alpha_2 = (1, 0, 1)$ , and  $\alpha_3 = (0, 1, 1)$  form a basis for  $\mathbb{R}^3$ . Find the coordinates of the vectors  $u = (2, 0, 0)$ ,  $v = (1, 0, 0)$ , and  $w = (1, 1, 1)$  under the basis  $\{\alpha_1, \alpha_2, \alpha_3\}$ .

**Solution 3.2.** Let  $a_1 \alpha_1 + a_2 \alpha_2 + a_3 \alpha_3 = 0$ . Then  $a_1(1, 1, 0) + a_2(1, 0, 1) + a_3(0, 1, 1) = 0$  which simplifies to  $a_1 + a_2 = 0$ ,  $a_1 + a_3 = 0$  and  $a_2 + a_3 = 0$ , for which the solution is  $a_1 = a_2 = a_3 = 0$ . Hence  $\alpha_1, \alpha_2$  and  $\alpha_3$  are linearly independent. Also, if  $(x, y, z) \in \mathbb{R}^3$  is an arbitrary element of  $\mathbb{R}^3$ , then  $(x, y, z) = (\frac{x+y-z}{2})(1, 1, 0) + (\frac{x-y+z}{2})(1, 0, 1) + (\frac{-x+y+z}{2})(0, 1, 1)$ . Hence  $\{\alpha_1, \alpha_2, \alpha_3\}$  span  $\mathbb{R}^3$ .

To determine the coordinates of  $u = (2, 0, 0)$ , put  $x = 2, y = 0, z = 0$  and we get  $u = (2, 0, 0) = 1(1, 1, 0) + 1(1, 0, 1) + (-1)(0, 1, 1) = 1\alpha_1 + 1\alpha_2 + (-1)\alpha_3$  and so the coordinates of  $u = (2, 0, 0)$  are  $(1, 1, -1)$ . Similarly, the coordinates of  $v = (1, 0, 0)$  is  $\frac{1}{2}(1, 1, -1)$  and the coordinates of  $w = (1, 1, 1)$  is  $\frac{1}{2}(1, 1, 1)$ .

**Proposition 3.6.** If  $V$  is a vector space finitely generated over a field  $F$  then every finite generating set of  $V$  over  $F$  contains a basis of  $V$ .

PROOF. Suppose  $V$  be a vector space finitely generated over a field  $F$ , and  $D$  be a finite generating set for  $V$  over  $F$ . If  $D$  is a minimal generating set for  $V$ , then  $D$  is a basis for  $V$ . If  $D$  is not a minimal generating set, then it properly contains other generating sets for  $V$  over  $F$ . One of those sets  $E$  will have the fewest elements and so  $E$  cannot contain any other generating set for  $V$  properly and so it must be the basis for  $V$ .  $\square$

**Proposition 3.7.** If  $V$  is a vector space finitely generated over a field  $F$ , then every linearly-independent subset  $B$  of  $V$  is contained in a basis of  $V$  over  $F$ .

PROOF. Since  $V$  is a finitely generated vector space, there exists a finite set  $\{v_1, \dots, v_n\}$  which generates  $V$  over  $F$ . Let  $B$  be a linearly independent subset of  $V$ . If  $\{v_1, \dots, v_n\} \subseteq FB$ , then  $FB = V$ , and so  $B$  is a basis for  $V$ . Otherwise, there exists at least one  $v_i$  such that  $v_i \notin FB$ . Let  $m = \min\{i | v_i \notin FB\}$ . Then the set  $D = B \cup \{v_m\}$  is linearly independent. If  $D$  is a generating set for the vector space  $V$ , then  $D$  is a

basis. If not, let  $k = \min\{i | v_i \in FD\}$ , and replace  $D$  by  $B \cup \{v_h, v_k\}$ . Proceeding similarly, after finitely many steps, we will obtain a basis for  $V$ .  $\square$

**Proposition 3.8.** Let  $V$  be a vector space over a field  $F$ , which is generated by a finite set  $B = \{v_1, \dots, v_n\}$  and let  $D$  be a linearly independent set of vectors in  $V$ . Then the number of elements in  $D$  is at most  $n$ .

PROOF. Consider a subset  $E = \{w_1, \dots, w_{n+1}\}$  of  $D$  having more than  $n$  elements. So,  $w_i \in D$  for all  $1 \leq i \leq n$ . Since  $D$  is linearly independent,  $w_i \neq 0$  for all  $1 \leq i \leq n$ . For each  $1 \leq k \leq n$ , set  $D_k = \{w_1, \dots, w_k, v_{k+1}, \dots, v_n\}$ .

The vector space  $V$  is generated by the set  $B$ , so we can find scalars  $a_1, \dots, a_n$ , not all equal to zero, such that  $w_1 = \sum_{i=1}^n a_i v_i$ . We will renumber the elements of  $B$  so that  $a_1 \neq 0$ . Then  $a_1 v_1 = w_1 - \sum_{i=2}^n a_i v_i$  and so  $v_1 = a_1^{-1} w_1 - \sum_{i=2}^n a_1^{-1} a_i v_i$  and so  $D \subseteq FD_1$ . Since  $D_1 \subseteq V = FD$ , so  $V = FD_1$ .

Suppose  $1 \leq k < n$  and that we have already shown that  $V = FD_k$ . We can find scalars  $b_1, \dots, b_n$ , not all equal to 0, such that  $w_{k+1} = \sum_{i=1}^k b_i w_i + \sum_{i=k+1}^n b_i v_i$ . At least one of  $b_{k+1}, \dots, b_n$  is non-zero because if all of them are zero would imply that  $D$  is linearly dependent which is not the case. by renumbering, we can assume that  $b_{k+1} \neq 0$ . Then,  $v_{k+1} = b_{k+1}^{-1} w_{k+1} - \sum_{i=1}^k b_{k+1}^{-1} b_i w_i - \sum_{i=k+2}^n b_{k+1}^{-1} b_i v_i$  and so  $V = FD_{k+1}$ . Following the same steps, we see after  $n$  steps, we obtain that  $V = FD_n = F\{w_1, \dots, w_n\}$ . But this implies that  $w_{n+1} \in F\{w_1, \dots, w_n\}$  and so  $E$  is linearly dependent, which contradicts our initial assumption. Hence, the set  $D$  can have at most  $n$  elements.  $\square$

**Proposition 3.9.** Let  $V$  be a vector space finitely generated over a field  $F$ . Then any two bases of  $V$  have the same number of elements.

PROOF. Since  $V$  is a finite generating vector space, hence it has a finite generating set having  $n$  elements (say). If  $B$  is a basis of  $V$ , then  $B$  has at most  $n$  elements and so  $B$  is finite. Now, suppose the  $B$  and  $B'$  are two bases of the vector space  $V$  having  $h$  and  $k$  elements, respectively. Then  $h \leq k$  because  $B$  is linearly independent and  $B'$  is the generating set. Similarly, since  $B'$  is linearly independent and  $B$  is the generating set, hence  $k \leq h$ , which implies  $h = k$ .  $\square$

**Proposition 3.10.** Let  $V$  be a vector space finitely generated over a field  $F$  and let  $W$  be a subspace of  $V$ . Then

- (1)  $W$  is finitely generated over  $F$ .
- (2) Every basis of  $W$  can be extended to a basis of  $V$ .
- (3)  $\dim(W) \leq \dim(V)$ , with equality when and only when  $W = V$ .

PROOF. Let  $n = \dim(V)$ .

- (1) Suppose to the contradiction that  $W$  is not finitely generated, then  $W$  has a linearly independent subset  $B$  having  $n+1$  elements. But  $B \subseteq V$  which contradicts  $\dim(V) = n$ .
- (2) If  $B$  is a basis of  $V$ , then  $B$  is linearly independent and can be extended to a basis of  $V$  by Proposition 3.7
- (3) By part (2), the number of elements of basis of  $W$  is less than or equal to the number of elements of basis of  $V$ . So,  $\dim(W) \leq \dim(V)$ . Also if  $\dim(W) = \dim(V)$ , then any basis of  $W$  is also a basis of  $V$ , so  $W = V$ .  $\square$

**Proposition 3.11.** Let  $V$  be a vector space over a field  $F$  and let  $W_1, \dots, W_n$  be distinct subspaces of  $V$ . Then the following conditions are equivalent

- (1)  $\{W_1, \dots, W_n\}$  is independent.
- (2) Every vector  $w \in \sum_{i=1}^n W_i$  can be written as  $w_1 + \dots + w_n$ , with  $w_i \in W_i$  for each  $1 \leq i \leq n$ , in exactly one way.
- (3)  $W_h$  and  $\sum_{i \neq h} W_i = \emptyset$ , for each  $1 \leq h \leq n$ .

PROOF. (1)  $\implies$  (2) Let  $w \in \sum_{i=1}^n W_i$  and suppose that  $w = w_1 + \dots + w_n = y_1 + \dots + y_n$ , where  $w_i, y_i \in W_i$  for each  $1 \leq i \leq n$  be the two representations of  $w$ . Then  $\sum_{i=1}^n (w_i - y_i) = 0_V$  and then

because of (1), we have  $w_i - y_i = 0_V$  for each  $1 \leq i \leq n$  and so (2) is true.

(2)  $\implies$  (3) Suppose to the contradiction that  $W_h$  and  $\sum_{i \neq h} W_i$  are not disjoint. Then there exists  $0_V \neq w_h \in W_h \cap \sum_{i \neq h} W_i$ . Then for each  $i \neq h$ , there exists an element  $w_i \in W_i$  satisfying  $w_h = \sum_{i \neq h} w_i$ , which contradicts (2).

(3)  $\implies$  (1) Let  $w_1 + \dots + w_n = 0_V$ , where  $w_i \in W_i$  for each  $1 \leq i \leq n$ , where  $w_h \neq 0_V$  for some  $h$ . Then  $w_h = -\sum_{i \neq h} w_i \in W_h \cap \sum_{i \neq h} W_i$ , which contradicts (3).  $\square$

**Proposition 3.12.** Let  $V$  be a vector space over a field  $F$ , let  $\{W_i \mid i \in \Omega\}$  be a pairwise disjoint collection of subspaces of  $V$  and for each  $i \in \Omega$ , let  $B_i$  be a basis of  $W_i$ . Then  $V = \bigoplus_{i \in \Omega} W_i$  if and only if  $B = \bigcup_{i \in \Omega} B_i$  is a basis of  $V$ .

PROOF. Let  $V = \bigoplus_{i \in \Omega} W_i$ . We shall prove that  $B$  is a basis for  $V$ . For that we shall show that  $B$  is linearly independent and  $B$  generates  $V$ . Let  $v \in V$ . Then there exists a finite subset  $A \subseteq \Omega$  such that  $v \in \bigoplus_{i \in A} W_i$  and therefore for each  $i \in A$  there exists an element  $w_i \in W_i$  satisfying  $v = \sum_{i \in A} w_i$ . Now, since  $B_i$  is a basis of  $W_i$  so each  $w_i$  is a linear combination of elements of  $B_i$ . Thus,  $v$  is a linear combination of the elements of  $B = \bigcup_{i \in \Omega} B_i$ . Therefore,  $B$  is a generating set for  $V$ .

Now, we shall show that  $B$  is linearly independent. Suppose to the contradiction that  $B$  is not a linearly independent set. Then there exists  $h \in \Omega$ , vectors  $v_1, \dots, v_t$  in  $B_h$  and scalars  $a_1, \dots, a_t$  in  $F$  not all equal to 0 such that  $\sum_{j=1}^t a_j v_j + u = 0_V$ , where  $u$  is a linear combination of elements of  $\bigcup_{i \neq h} B_i$ . Then  $\sum_{j=1}^t a_j v_j \in W_h \cap \sum_{i \neq h} W_i$ , which contradicts the fact that  $\{W_i \mid i \in \Omega\}$  is a pairwise disjoint collection of subspaces of  $V$ . Thus,  $B = \bigcup_{i \in \Omega} B_i$  is a basis of  $V$ .

Conversely, suppose  $B = \bigcup_{i \in \Omega} B_i$  is a basis of  $V$ , then every element of  $V$  can be written in a unique way as the element of  $B$  or  $\bigcup_{i \in \Omega} B_i$ . So, every element of  $V$  can be written in a unique way as  $\sum_{i \in A} w_i$ , where  $A$  is a finite subset of  $\Omega$ , which proves  $V = \bigoplus_{i \in \Omega} W_i$ .  $\square$

**Proposition 3.13.** Every subspace  $W$  of a vector space  $V$  over a field  $F$  has at least one complement in  $V$ .

PROOF. When  $W$  is an improper subspace of  $V$ , then  $\{0_V\}$  is a complement of  $W$  in  $V$ . In the same way,  $V$  is a complement of  $\{0_V\}$  in  $V$ . In other cases assume,  $B$  be a basis of  $W$ . Then there exists a linearly independent subset  $D$  of  $V$  such that  $B \cup D$  is a basis of  $V$ . Then  $FD$  is a complement of  $W$  in  $V$ .  $\square$

**Proposition 3.14.** Let  $F$  be a field which is not finite and let  $V$  be a vector space over  $F$  having dimension at least 2. Then every proper nontrivial subspace  $W$  of  $V$  has infinitely many complements in  $V$ .

PROOF. By the previous proposition,  $W$  has at least one complement  $U$  in  $V$ . Let  $B$  be a basis for  $U$ . Let  $0_V \neq w \in W$ . Since  $F$  is an infinite field,  $Fw$  will be an infinite subset of  $W$ . Thus,  $W$  is also infinite. For each  $w \in W$ , define  $Y_w = F\{u + w \mid u \in B\}$ . We claim that each of these subspaces is a complement of  $W$  in  $V$ . Assume  $v \in W \cap Y_w$ . Then there exists elements  $u_1, \dots, u_n \in B$  and scalars  $c_1, \dots, c_n \in F$  satisfying  $v = \sum_{i=1}^n c_i(u_i + w)$ . Then  $\sum_{i=1}^n c_i u_i = v - (\sum_{i=1}^n c_i)w \in W \cup U = 0_V$  and since  $\{u_1, \dots, u_n\}$  is a linearly independent set, so  $c_i = 0$  for all  $i$  and so  $v = 0_V$  and so  $W$  and  $Y_w$  are disjoint. Let  $v \in V$ , then  $v = x + (\sum_{i=1}^n c_i u_i)$ , where  $x \in W, u_1, \dots, u_n \in B$ , and the scalars  $c_1, \dots, c_n \in F$ . Then  $v = [x - (\sum_{i=1}^n c_i)w] + \sum_{i=1}^n c_i(u_i + w) \in W + Y_w$  and thus we have shown that  $V = W + Y_w$  and therefore  $Y_w$  is a complement of  $W$  in  $V$ .

Now, we shall show that all these complements are different from each other. For this let us take two elements  $w \neq x$  of  $W$  satisfying  $Y_w = Y_x$ . If  $u \in B$ , then there exists elements  $u_1, \dots, u_n \in B$  and scalars  $c_1, \dots, c_n$  such that  $u + w = \sum_{i=1}^n c_i(u_i + x)$ . So,  $u - \sum_{i=1}^n c_i u_i = (\sum_{i=1}^n c_i)x - w$  and this belongs to  $W \cap Y_w = \{0_V\}$ . Since  $B$  is linearly independent so  $u$  has to equal one of the  $u_h$  for some  $1 \leq h \leq n$  and therefore  $c_i = 0$  for  $i \neq h$  and  $c_h = 1$ . So,  $x - w = 0_V$ , namely  $x = w$ , which is a contradiction, hence all  $Y_w$  must be distinct.  $\square$

**Proposition 3.15.** Let  $V$  be a vector space over a field  $F$  and let  $W$  be a subspace of  $V$  having a complement  $Y$  in  $V$ . Let  $\{v_1, \dots, v_n\}$  be a subset of  $V$  and, for each  $1 \leq i \leq n$ , let  $v_i = w_i + y_i$ , where  $w_i \in W$  and  $y_i \in Y$ . If the vectors  $w_1, \dots, w_n$  are distinct and the set  $\{w_1, \dots, w_n\}$  is linearly independent, then so is the set  $\{v_1, \dots, v_n\}$ .

PROOF. Suppose there exists scalars  $a_1, \dots, a_n$  such that  $\sum_{i=1}^n a_i v_i = 0_V$ . Then  $\sum_{i=1}^n a_i w_i + \sum_{i=1}^n a_i y_i = 0_V$  and so  $\sum_{i=1}^n a_i w_i = \sum_{i=1}^n a_i y_i = 0_V$ . We know that the vectors  $w_1, \dots, w_n$  are all distinct and the set  $\{w_1, \dots, w_n\}$  is a linearly independent set, hence  $a_1 = \dots = a_n = 0$ , and so the set  $\{v_1, \dots, v_n\}$  is also linearly independent.  $\square$

#### 4. Linear Transformations

**Proposition 4.1.** Let  $V$  and  $W$  be vector spaces over a field  $F$  and let  $\alpha : V \rightarrow W$  be a function. Then  $\alpha$  is a linear transformation if and only if  $gr(\alpha)$  is a subspace of  $V \times W$ .

PROOF. Let  $\alpha$  be a linear transformation. Let  $v, v' \in V$  and  $c \in F$ . Then  $\alpha(v), \alpha(v') \in W$ , and so  $(v, \alpha(v)) + (v', \alpha(v')) = (v + v', \alpha(v) + \alpha(v')) = (v + v', \alpha(v + v')) \in gr(\alpha)$  and  $c(v, \alpha(v)) = (cv, c\alpha(v)) = (cv, \alpha(cv)) \in gr(\alpha)$ . This shows that  $gr(\alpha)$  is closed under taking sums and scalar multiples. Therefore,  $gr(\alpha)$  is a subspace of  $V \times W$ .

Now, suppose  $gr(\alpha)$  is a subspace of  $V \times W$ . Then for  $v, v' \in V$  and  $c \in F$ ,  $(v, \alpha(v)) + (v', \alpha(v')) = (v + v', \alpha(v) + \alpha(v')) \in gr(\alpha)$ . Therefore, we must have  $\alpha(v) + \alpha(v') = \alpha(v + v')$ . Also,  $c(v, \alpha(v)) = (cv, c\alpha(v)) \in gr(\alpha)$ , so  $c\alpha(v) = \alpha(cv)$  and therefore  $\alpha$  is a linear transformation.  $\square$

**Proposition 4.2.** Let  $V$  and  $W$  be vector spaces over a field  $F$ , and let  $B$  be a basis of  $V$ . If  $f \in W^B$  then there is a unique linear transformation  $\alpha \in \text{Hom}(V, W)$  satisfying the condition that  $\alpha(u) = f(u)$  for all  $u \in B$ .

PROOF. Since  $B$  is a basis of the vector space  $V$ , hence each vector  $v \in V$  can be written as a linear combination  $v = \sum_{i=1}^n a_i u_i$  of the elements of  $B$  in a unique way. Define a function  $\alpha : V \rightarrow W$  by  $\alpha(v) = \sum_{i=1}^n a_i f(u_i)$ . Because of the uniqueness of the representation of  $v$ , this function is well defined. It can be easily proved that  $\alpha$  is a linear transformation. If  $\beta : V \rightarrow W$  is a linear transformation satisfying the condition that  $\beta(u) = f(u)$  for all  $u \in B$ , then  $\beta(v) = \beta(\sum_{i=1}^n a_i u_i) = \sum_{i=1}^n a_i \beta(u_i) = \sum_{i=1}^n a_i f(u_i) = \alpha(v)$ , and therefore the two linear transformations  $\alpha$  and  $\beta$  are identical. Thus,  $\alpha$  is unique.  $\square$

**Proposition 4.3.** Let  $V$ ,  $W$ , and  $Y$  be vector spaces over a field  $F$  and  $\alpha : V \rightarrow W$  and  $\beta : W \rightarrow Y$  be linear transformations. Then  $\beta\alpha : V \rightarrow Y$  is a linear transformation.

PROOF. Let  $v_1, v_2 \in V$  and  $a \in F$ . Then  $(\beta\alpha)(v_1 + v_2) = \beta(\alpha(v_1 + v_2)) = \beta(\alpha(v_1) + \alpha(v_2)) = \beta(\alpha(v_1)) + \beta(\alpha(v_2)) = (\beta\alpha)(v_1) + (\beta\alpha)(v_2)$  and  $(\beta\alpha)(cv_1) = \beta(\alpha(cv_1)) = \beta(c\alpha(v_1)) = c\beta(\alpha(v_1)) = c(\beta\alpha)(v_1)$  and we are done.  $\square$

**Proposition 4.4.** Let  $V$  and  $W$  be vector spaces over a field  $F$  and let  $\alpha \in \text{Hom}(V, W)$ . Then  $\ker(\alpha)$  is a subspace of  $V$ , which is trivial if and only if  $\alpha$  is monic.

PROOF. Let  $v_1, v_2 \in \ker(\alpha)$  and  $a \in F$ . Then  $\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2) = 0_W + 0_W = 0_W$  and therefore  $v_1 + v_2 \in \ker(\alpha)$ . Similarly,  $\alpha(av_1) = a(\alpha v_1) = a0_W = 0_W$  and so  $av_1 \in \ker(\alpha)$ . Hence,  $\ker(\alpha)$  is a subspace of  $V$ .

Now suppose  $\alpha$  is monic, that is a one - one function, then  $\alpha^{-1}(w)$  is a singleton set for each  $w \in W$ . Therefore,  $\ker(\alpha) = \{0_V\}$ . Now, suppose that  $\ker(\alpha)$  is trivial. Consider two elements  $v_1 \neq v_2$  of  $V$  such that  $\alpha(v_1) = \alpha(v_2)$ . Then  $\alpha(v_1 - v_2) = \alpha(v_1) - \alpha(v_2) = 0_W$  and so  $v_1 - v_2 \in \ker(\alpha)$ . Then  $v_1 - v_2 = 0_V$  and so  $v_1 = v_2$ , which is a contradiction. Hence  $\alpha$  must be monic.  $\square$

**Proposition 4.5.** Let  $V$  and  $W$  be vector spaces over a field  $F$  and  $\alpha \in \text{Hom}(V, W)$ . Then  $\text{im}(\alpha)$  is a subspace of  $W$ , which is improper if and only if  $\alpha$  is onto.

PROOF. Let  $\alpha(v_1)$  and  $\alpha(v_2)$  be two elements of  $\text{im}(\alpha)$  and  $a \in F$ . Then  $\alpha(v_1) + \alpha(v_2) \in \text{im}(\alpha)$ . Also,  $a\alpha(v_1) = \alpha(av_1) \in \text{im}(\alpha)$ . Hence  $\text{im}(\alpha)$  is a subspace of  $W$ .

Now suppose  $\text{im}(\alpha)$  is a improper subspace of  $W$  i.e.,  $\text{im}(\alpha) = W$ , then  $\alpha$  is onto.  $\square$

**Proposition 4.6.** Let  $\alpha : V \rightarrow W$  be a linear transformation of vector spaces over a field  $F$  and let  $w \in \text{im}(\alpha)$ . For any  $v_0 \in \alpha^{-1}(w)$  we have  $\alpha^{-1}(w) = \{v + v_0 | v \in \ker(\alpha)\}$ .

PROOF. Let  $v \in \alpha^{-1}(w)$ , then  $v = (v - v_0) + v_0$ , where  $v - v_0 \in \ker(\alpha)$  because  $\alpha(v_1 - v_0) = \alpha(v_1) - \alpha(v_0) = w - w = 0_W$ .

Now suppose  $v \in \ker(\alpha)$ , then  $\alpha(v + v_0) = \alpha(v) + \alpha(v_0) = 0_W + w = w$  and so  $v + v_0 \in \alpha^{-1}(w)$ .  $\square$

**Proposition 4.7.** If  $\alpha : V \rightarrow W$  is an isomorphism of vector spaces over a field  $F$  then there exists an isomorphism  $\beta : W \rightarrow V$  satisfying  $\beta\alpha(v) = v$  and  $\alpha\beta(w) = w$  for all  $v \in V$  and all  $w \in W$ .

PROOF. Define a function  $\beta$  by  $\beta(w) = \alpha^{-1}(w)$ . Since  $\alpha$  is an isomorphism, hence  $\alpha^{-1}(w)$  is a unique element of  $V$  and so  $\beta$  is well defined. Then  $\beta\alpha(v) = \beta(\alpha(v)) = \alpha^{-1}\alpha(v) = v$ . Also,  $\alpha\beta(v) = \alpha(\beta(v)) = \alpha(\alpha^{-1}(v)) = v$

Now, we show that  $\beta$  is a linear transformation. Let  $w_1$  and  $w_2 \in W$  and let  $\alpha(v_1) = w_1$  and  $\alpha(v_2) = w_2$ , then  $\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2) = w_1 + w_2$ . So,  $\alpha^{-1}(w_1 + w_2) = v_1 + v_2 = \alpha^{-1}(w_1) + \alpha^{-1}(w_2)$ . Also,  $\alpha(cv_1) = c\alpha(v_1) = cw_1$  and therefore  $\alpha^{-1}(cw_1) = cv_1 = c\alpha^{-1}(w_1)$ . Therefore,  $\beta = \alpha^{-1}$  is a linear transformation.  $\square$

**Proposition 4.8.** Let  $V$  and  $W$  be vector spaces over a field  $F$  having bases  $B$  and  $D$ , respectively, and assume that there exists a bijective function  $f : B \rightarrow D$ . Then  $V \cong W$ .

PROOF. Since  $B$  is a basis for  $V$  and  $f : B \rightarrow D$ , hence there exists a linear transformation  $\alpha \in \text{Hom}(V, W)$  such that  $\alpha(v) = f(v)$  for all  $v \in B$ . As  $\alpha$  is onto because given an element of  $W$ , it can be written as the linear combination of the elements of  $D$  and each element of  $D$  is the image of some element of  $B$  under  $\alpha$  or  $f$ , hence the element of  $W$  can be written as the image of some element of  $V$  under  $f$ . Now, we shall show that  $\alpha$  is an injective function. Since  $\alpha$  is injective is equivalent to  $\ker(\alpha) = \{0_V\}$ , so we will prove that  $\ker(\alpha) = 0_V$ . Let  $v' = a_1v_1 + a_2v_2 + \cdots + a_nv_n \in \ker(\alpha)$ , then  $0_W = \alpha(v') = \alpha(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1\alpha(v_1) + a_2\alpha(v_2) + \cdots + a_n\alpha(v_n) = a_1f(v_1) + a_2f(v_2) + \cdots + a_nf(v_n)$ . Then  $a_i = 0$  for all  $1 \leq i \leq n$  because  $f(v_i) \in D$  for all  $1 \leq i \leq n$  and  $D$  is linearly independent.  $a_i = 0$  for all  $1 \leq i \leq n$  implies that  $v' = 0$  and so  $\ker(\alpha) = 0$  and so  $\alpha$  is an injective map and so  $\alpha$  is an isomorphism.  $\square$

**Proposition 4.9.** Let  $V$  and  $W$  be vector spaces over a field  $F$ , where  $V$  is finitely generated. Then  $\dim(V) = \dim(\text{im}(\alpha)) + \dim(\ker(\alpha))$  for any linear transformation  $\alpha \in \text{Hom}(V, W)$ .

PROOF. Since  $\ker(\alpha)$  is a subspace of a finite-dimensional vector space  $V$ , hence  $\ker(\alpha)$  is of finite dimension, say  $n$ . Let  $B = \{v_1, v_2, \dots, v_n\}$  be a basis of  $\ker(\alpha)$ . Then  $B$  can be extended to form a basis for  $V$ . Let  $B' = \{v_1, v_2, \dots, v_n, v_{n+1}, \dots, v_m\}$  be a basis for  $V$ . We claim that  $S = \{\alpha(v_{n+1}), \dots, \alpha(v_m)\}$  is a basis for  $\text{im}(\alpha)$ . Let  $a \in \text{im}(\alpha)$ , then  $a = \alpha(v)$  for some  $v \in V$ , then  $v = a_1v_1 + \cdots + a_mv_m$  and so  $a = \alpha(v) = a_1\alpha(v_1) + \cdots + a_m\alpha(v_m)$ . Now, we shall show that  $S$  is linearly independent. Assume that  $a_{n+1}\alpha(v_{n+1}) + \cdots + a_m\alpha(v_m) = 0_W$  which implies  $\alpha(a_{n+1}v_{n+1} + \cdots + a_mv_m) = 0_W$  and so  $a_{n+1}v_{n+1} + \cdots + a_mv_m \in \ker(\alpha)$  for which  $B$  is the basis and so  $a_{n+1}v_{n+1} + \cdots + a_mv_m = b_1v_1 + b_2v_2 + \cdots + b_nv_n$  or  $\sum_{i=1}^n b_i v_i + \sum_{i=n+1}^m a_i v_i = 0_V$ . But  $B'$  is a basis for  $V$ , so  $a_i = 0$  for all  $1 \leq i \leq n$  and so  $S$  is linearly independent.  $\square$

**Problem 4.1.** Let  $T : P(\mathbb{R}) \rightarrow P(\mathbb{R})$  be defined by  $T(f(x)) = f'(x)$ . Recall that  $T$  is linear. Prove that  $T$  is onto, but not one-to-one.

**Solution 4.1.** Consider  $f(x) \in P(\mathbb{R})$ . Take  $y = \int f(x)dx$ . Then  $T(\int f(x)dx) = \frac{d}{dx}(\int f(x)dx) = f(x)$  by Fundamental Theorem of Calculus

Now, we prove that  $T$  is not one-to-one. For that consider  $f(x) = x + 4$  and  $g(x) = x + 3$ . It is obvious that  $f(x) \neq g(x)$ . But  $T(f(x)) = T(g(x)) = 1$ .

**Problem 4.2.** Let  $V$  and  $W$  be finite-dimensional vector spaces and  $T : V \rightarrow W$  be linear.

- (1) Prove that if  $\dim(V) < \dim(W)$ , then  $T$  cannot be onto.
- (2) Prove that if  $\dim(V) > \dim(W)$ , then  $T$  cannot be one-to-one.

**Solution 4.2.** (1) Suppose to the contradiction that  $T$  is onto. Then  $R(T) = W$  which implies  $\dim(R(T)) = \dim(W)$ . Since  $\dim(N(T)) + \dim(W) = \dim(V)$ , hence  $\dim(V) \geq \dim(W)$ , which is a contradiction.

- (2) Suppose to the contradiction that  $T$  is one-to-one. This implies  $\dim(N(T)) = 0$ , which implies  $\dim(V) = \dim(R(T))$  but  $R(T)$  is a subspace of  $W$ . Thus,  $\dim(V) \leq \dim(W)$ , which is a contradiction.

**Problem 4.3.** Give an example of a linear transformation  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  such that  $N(T) = R(T)$ .

**Solution 4.3.** Define  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  by  $T(x, y) = (y, 0)$ . Then,

$$\begin{aligned} N(T) &= \{(x, y) \in \mathbb{R}^2 \mid T(x, y) = (0, 0)\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid (y, 0) = (0, 0)\} \\ &= \{(x, 0) \mid x \in \mathbb{R}\} = R(T) \end{aligned}$$

**Problem 4.4.** Let  $V$  and  $W$  be vector spaces with subspaces  $V_1$  and  $W_1$ , respectively. If  $T : V \rightarrow W$  is linear, prove that  $T(V_1)$  is a subspace of  $W$  and that  $\{x \in V : T(x) \in W_1\}$  is a subspace of  $V$ .

**Solution 4.4.** Since  $0_V \in V_1$ ,  $T(0_V) \in T(V_1)$  which implies  $0_W \in T(V_1)$ . Let  $y_1 \in T(V_1)$ ,  $y_2 \in T(V_1)$ . Then  $y_1 = T(x_1)$ ,  $y_2 = T(x_2)$  for  $x_1 \in V_1$ ,  $x_2 \in V_1$ . Now  $y_1 + y_2 = T(x_1) + T(x_2) = T(x_1 + x_2)$ . Since  $V_1$  is a subspace of  $V$ , thus  $y_1 + y_2 \in T(V_1)$ . Also for  $c \in \mathbb{R}$  and  $y_1 \in T(V_1)$ ,  $cy_1 = cT(x_1) = T(cx_1)$ , which implies  $cy_1 \in T(V_1)$ . Thus,  $T(V_1)$  is a subspace of  $W$ . Define  $A = \{x \in V : T(x) \in W_1\}$ .  $T(0_V) = 0_W \in W_1$ . Thus,  $0_V \in A$ . Let  $x_1 \in A$  and  $x_2 \in A$ . Then  $T(x_1) \in W_1$  and  $T(x_2) \in W_1$ . This implies  $T(x_1 + x_2) \in W_1$ , which implies  $x_1 + x_2 \in A$ . Let  $c \in \mathbb{R}$ ,  $x \in A$ . Then  $T(x) \in W_1$ . Since  $W_1$  is a subspace,  $cT(x) \in W_1$ , which implies  $T(cx) \in W_1$ . Therefore,  $cx \in A$ . Therefore,  $A$  is a subspace of  $V$ .

**Problem 4.5.** Let  $V$  be the vector space of an infinite tuple of sequences. Define the functions  $T : V \rightarrow V$  and  $U : V \rightarrow V$  by  $T(a_1, a_2, \dots) = (a_2, a_3, \dots)$  and  $U(a_1, a_2, \dots) = (0, a_1, a_2, \dots)$ .  $T$  and  $U$  are called the **left shift** and **right shift** operators on  $V$ , respectively.

- (1) Prove that  $T$  and  $U$  are linear.
- (2) Prove that  $T$  is onto, but not one-to-one.
- (3) Prove that  $U$  is one-to-one, but not onto.

**Solution 4.5.** (1) Let  $(a_1, a_2, \dots) \in V$ ,  $(b_1, b_2, \dots) \in V$ . We claim that  $T(\alpha(a_1, a_2, \dots) + (b_1, b_2, \dots)) = \alpha T((a_1, a_2, \dots)) + T(b_1, b_2, \dots)$ .  

$$\begin{aligned} &T(\alpha(a_1, a_2, \dots) + (b_1, b_2, \dots)) \\ &= T((\alpha a_1, \alpha a_2, \dots) + (b_1, b_2, \dots)) \\ &= T((\alpha a_1 + b_1, \alpha a_2 + b_2, \dots)) \\ &= (\alpha a_2 + b_2, \alpha a_3 + b_3, \dots) \\ &= (\alpha a_2, \alpha a_3, \dots) + (b_2, b_3, \dots) \\ &= \alpha(a_2, a_3, \dots) + (b_2, b_3, \dots) \\ &= \alpha T((a_1, a_2, \dots)) + T(b_1, b_2, \dots). \end{aligned}$$
 Hence  $T$  is linear.  
 Now, for  $a = (a_1, a_2, \dots)$  and  $b = (b_1, b_2, \dots)$   $U(\alpha a + b)$   

$$\begin{aligned} &= U((\alpha a_1, \alpha a_2, \dots) + (b_1, b_2, \dots)) \\ &= U(\alpha a_1 + b_1, \alpha a_2 + b_2, \dots) \\ &= (0, \alpha a_1 + b_1, \alpha a_2 + b_2, \dots) \\ &= (0, \alpha a_1, \alpha a_2, \dots) + (0, b_1, b_2, \dots) \\ &= \alpha(0, a_1, a_2, \dots) + (0, b_1, b_2, \dots) \\ &= \alpha U(a_1, a_2, \dots) + U(b_1, b_2, \dots) \end{aligned}$$
 Hence  $U$  is linear.

- (2) Now we prove that  $T$  is onto. For  $y = (a_1, a_2, \dots)$ , take  $x = (a_0, a_1, \dots)$  for any  $a_0 \in V$ . Then  $T(x) = T((a_0, a_1, a_2, \dots)) = (a_1, a_2, \dots) = y$ . For one-one, consider  $x = (a, a_1, a_2, \dots)$ ,  $y = (b, a_1, a_2, \dots)$  for two distinct elements  $a$  and  $b$  of  $V$ . Then  $x \neq y$  but  $T(x) = T(y) = (a_1, a_2, \dots)$ .
- (3) Now we prove that  $U$  is one-one. Let  $U(x) = U(y)$  for some  $x \in V$  and  $y \in V$ . Then  $x = (a_1, a_2, \dots)$  and  $y = (b_1, b_2, \dots)$  for some  $a_i \in V$  and  $b_i \in V$  for all  $i$ .  $U(x) = U(y)$  implies  $(0, a_1, a_2, \dots) = (0, b_1, b_2, \dots)$  which implies  $a_i = b_i$  for all  $i$  which means  $x = y$ .

Consider  $y = (1, a_1, a_2, \dots)$ . As the first coordinate of the image of any element is 0, this element does not have a preimage. So,  $U$  is not onto.

**Problem 4.6.** Let  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  be linear. Show that there exist scalars  $a, b$ , and  $c$  such that  $T(x, y, z) = ax + by + cz$  for all  $(x, y, z) \in \mathbb{R}^3$ . Can you generalize this result for  $T : \mathbb{F}^n \rightarrow \mathbb{F}^n$ ?

**Solution 4.6.**  $T(x, y, z) = T(x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1)) = xT(1, 0, 0) + yT(0, 1, 0) + zT(0, 0, 1)$   
 Choose  $a = T(1, 0, 0)$ ,  $b = T(0, 1, 0)$ ,  $c = T(0, 0, 1)$ , then the given equation will be satisfied.

We can generalize this result as: Let  $T : \mathbb{R}^n \rightarrow \mathbb{R}$  be linear. Then there exists scalar  $a_1 = T(1, 0, \dots)$ ,  $a_2 = T(0, 1, \dots)$ ,  $\dots$ ,  $a_n = T(0, 0, \dots, 1)$  such that  $T(x_1, x_2, \dots) = a_1x_1 + a_2x_2 + \dots + a_nx_n$ .

**Problem 4.7.** Let  $T : V \rightarrow W$  be linear,  $b \in W$ , and  $K = \{x \in V : T(x) = b\}$  be nonempty. Prove that if  $s \in K$ , then  $K = \{s\} + N(T)$ .

**Solution 4.7.** Since  $s \in K$ , thus  $T(s) = b$ .

$\{s\} + N(T) = \{x + y : x \in \{s\}, y \in N(T)\} = \{s + y : y \in N(T)\}$ . We have to prove that  $s + y \in K$  for every  $y \in N(T)$ .

$T(s + y) = T(s) + T(y) = b + 0 = b$ , which implies  $s + y \in K$ .

**Problem 4.8.** Let  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$

- (1) If  $T(a, b, c) = (a, b, 0)$ , show that  $T$  is the projection on the  $xy$ -plane along the  $z$ -axis.
- (2) Find a formula for  $T(a, b, c)$ , where  $T$  represents the projection on the  $z$ -axis along the  $xy$ -plane.
- (3) If  $T(a, b, c) = (a - c, b, 0)$ , show that  $T$  is the projection on the  $xy$ -plane along the line  $L = \{(a, 0, a) : a \in \mathbb{R}\}$ .

**Solution 4.8.** (1)  $T(a, b, c) = (a, b, 0)$  which lies in the  $xy$ -plane. Since the  $x$  and  $y$ -coordinates do not change during the transformation, hence the transformation is along the  $z$ -axis.

- (2)  $T(a, b, c) = (0, 0, c)$  which lies on the  $z$ -axis. Hence the projection is on the  $z$ -axis. Since the  $z$ -coordinate do not change during the transformation, hence the transformation is along the  $xy$ -plane.
- (3)  $T(a, b, c) = (a - c, b, 0) + (c, 0, c)$  where  $(a - c, b, 0)$  lies in the  $xy$ -plane and  $(c, 0, c) \in L$ . Therefore  $T$  is the projection on  $xy$ -plane along  $L$ .

**Problem 4.9.** Assume  $T : V \rightarrow V$  is the projection on  $W_1$  along  $W_2$ .

- (1) Prove that  $T$  is linear and  $W_1 = \{x \in V : T(x) = x\}$ .
- (2) Prove that  $W_1 = R(T)$  and  $W_2 = N(T)$ .
- (3) Describe  $T$  if  $W_1 = V$ .
- (4) Describe  $T$  if  $W_1$  is the zero subspace.

**Solution 4.9.**  $V = W_1 \oplus W_2$ . Let  $x \in V$ . Then  $x = x_1 + x_2$  where  $x_1 \in W_1$  and  $x_2 \in W_2$ .

- (1) Let  $x \in V, y \in V$  and  $c \in \mathbb{F}$ . Then  $x = x_1 + x_2$  and  $y = y_1 + y_2$  where  $x_1 \in W_1, x_2 \in W_2, y_1 \in W_1$  and  $y_2 \in W_2$ . Then  $T(x) = x_1$  and  $T(y) = y_1$  and  $T(cx + y) = T(c(x_1 + x_2) + (y_1 + y_2)) = T((cx_1 + y_1) + (cx_2 + y_2))$ . Since  $c \in \mathbb{F}$  and  $x_1 \in W_1, cx_1 \in W_1$  and so  $cx_1 + y_1 \in W_1$ . Similarly,  $cx_2 + y_2 \in W_2$ . Therefore,  $T(cx + y) = cx_1 + y_1 = cT(x) + T(y)$ , and so  $T$  is linear.

Now, we prove that  $W_1 = \{x \in V : T(x) = x\}$ . Let  $A = \{x \in V : T(x) = x\}$ . Let  $x \in W_1$ . Then  $T(x) = T(x + 0) = x$  as  $x \in W_1$  and  $0 \in W_2$ , which implies  $x \in A$ , and so  $W_1 \subset A$ . Now, suppose  $x \in A$  and  $x \notin W_1$ .  $x \in V$  implies  $x = x_1 + x_2$  where  $x_1 \in W_1$  and  $x_2 \in W_2$ . Since  $x \in W_1$  and  $x_1 \in W_1$  and  $W_1$  is a subspace, hence  $x_2 = x - x_1 \in W_1$ , which contradicts  $x_2 \in W_2$ . Hence  $x \in W_1$  and so  $A \subset W_1$ . Hence  $W_1 = A$ .

- (2) Let  $x \in W_1$ . Then by (a),  $T(x) = x$ . Since  $T(x) \in R(T)$ , hence  $x \in R(T)$  and so  $W_1 \subset R(T)$ . Now suppose  $x \in R(T)$ , then  $x = T(x_0)$  for some  $x_0 \in V$ .

Case I:  $x_0 \in W_1$ .

Then by (a),  $T(x_0) = x_0$  and so  $x = x_0$  which implies  $x \in W_1$ .

Case II:  $x_0 \notin W_1$

Since  $x_0 \in V$ , let  $x_0 = x_1 + x_2, x_1 \in W_1, x_2 \in W_2$ . Then  $T(x_0) = x_1$ , which implies  $x = x_1$  and so  $x \in W_1$ .

So,  $R(T) \subset W_1$ . Thus  $W_1 = R(T)$ .

Now we shall prove that  $W_2 = N(T)$ .

Let  $x \in N(T)$ . Then  $T(x) = 0$ . Since  $x \in V$ , we can write  $x = x_1 + x_2$  for some  $x_1 \in W_1$  and

$x_2 \in W_2$ . Then  $T(x) = T(x_1 + x_2) = x_1$ , which implies  $x_1 = 0$  or  $x = x_2$ . Since  $x_2 \in W_2$ , so  $x \in W_2$  and so  $N(T) \subset W_2$ .

Let  $x \in W_2$ . Then  $T(x) = T(0 + x) = 0$  by the definition of  $T$  as  $0 \in W_1$  and  $x \in W_2$ . So,  $x \in N(T)$  and  $W_2 \subset N(T)$ . Thus  $W_2 = N(T)$ .

- (3)  $W_1 = V$  implies  $\{x \in V : T(x) = x\} = V$ . This means that the map  $T$  is the identity map i.e.,  $T(x) = x$  for all  $x \in V$ .
- (4)  $W_1 = \{0\}$  implies  $\{x \in V : T(x) = x\} = 0$ . This means that the map  $T$  is the zero map i.e.,  $T(x) = 0$  for all  $x \in V$ .

**Problem 4.10.** Suppose that  $W$  is a subspace of a finite-dimensional vector space  $V$ . Prove that there exists a subspace  $W'$  and a function  $T : V \rightarrow V$  such that  $T$  is a projection on  $W$  along  $W'$ .

**Solution 4.10.** Let  $\beta = \{v_1, v_2, \dots, v_n\}$  be a basis of  $W$  where  $n = \dim(W)$ . Let  $m = \dim(V)$ . Extend the bases  $\beta$  to the basis  $\beta' = \{v_1, v_2, \dots, v_n, v_{n+1}, \dots, v_m\}$ . Let  $W' = \text{span}\{v_{n+1}, v_{n+2}, \dots, v_m\}$ . For  $v \in V$ , it can be written as  $v = \alpha_1 v_1 + \dots + \alpha_n v_n + \alpha_{n+1} v_{n+1} + \dots + \alpha_m v_m = w + w'$  where  $w = \alpha_1 v_1 + \dots + \alpha_n v_n$  and  $w' = \alpha_{n+1} v_{n+1} + \dots + \alpha_m v_m$ . Define  $T$  as: If  $v \in V$  and  $v = w + w'$  then  $T(v) = w = \alpha_1 v_1 + \dots + \alpha_n v_n \in W$ .

**Problem 4.11.** Prove that the subspaces,  $\{0\}$ ,  $V$ ,  $R(T)$ , and  $N(T)$  are all  $T$ -invariant.

**Solution 4.11.** Since  $T(0) = 0 \in \{0\}$ , so  $\{0\}$  is  $T$ -invariant. Also,  $T(v) \in V$  for all  $v \in V$ , hence  $V$  is also  $T$ -invariant. Since  $R(T) = \{T(v) : v \in V\}$ , so  $T(v) \in R(T)$  for every  $v \in R(T)$ . Also, if  $v \in N(T)$  then  $T(v) = 0$  which implies  $T(T(v)) = T(0) = 0$  and so  $T(v) \in N(T)$ . Hence,  $N(T)$  is  $T$ -invariant.

**Problem 4.12.** If  $W$  is  $T$ -invariant, prove that  $T_W$  is linear.

**Solution 4.12.** Since  $T_W(x) \in W$ ,  $T_W : W \rightarrow W$  is well-defined. Also, since  $T : V \rightarrow V$  is linear and  $T_W(x) = T(x)$  for every  $x \in W$ , so  $T_W(\alpha x + y) = \alpha T_W(x) + T_W(y)$  for every  $x \in W$  and  $y \in W$ . Hence  $T_W : W \rightarrow W$  is linear.

**Problem 4.13.** Suppose that  $T$  is the projection on  $W$  along some subspace  $W'$ . Prove that  $W$  is  $T$ -invariant and  $T_W = I_W$ .

**Solution 4.13.** For  $x \in W$ ,  $T(x) = T(x + 0) = x$  because  $T$  is a projection of  $W$  along  $W'$ , so  $T(x) \in W$  for all  $x \in W$  and so  $W$  is  $T$ -invariant. Also,  $T_W(x) = x = I_W(x)$  for all  $x \in x$  and so  $T_W = I_W$ .

**Problem 4.14.** Suppose that  $V = R(T) \oplus W$  and  $W$  is  $T$ -invariant. Prove that  $W \subseteq N(T)$ .

**Solution 4.14.** Let  $w \in W$ . Since  $W$  is  $T$ -invariant, so  $T(w) \in W$ . Now,  $T(w) = T(0 + w) = 0$ , where  $0 \in R(T)$  and  $w \in W$ . Since  $T(w) = 0$ ,  $w \in N(T)$  and so  $W \subseteq N(T)$ .

**Problem 4.15.** Suppose that  $W$  is  $T$ -invariant. Prove that  $N(T_W) = N(T) \cap W$  and  $R(T_W) = T(W)$ .

**Solution 4.15.** Let  $x \in N(T_W)$ , then  $T_W(x) = 0$  which implies  $T(x) = 0$  or  $x \in N(T)$ . Also since  $T_W : W \rightarrow W$ , hence  $x \in N(T_W) \subseteq W$ . Hence  $x \in N(T) \cap W$  i.e.,  $N(T_W) \subseteq N(T) \cap W$ .

Let  $x \in N(T) \cap W$ . Then  $x \in W$  and  $T_W(x) = T(x) = 0$  and so  $x \in N(T_W)$ . Now,  $R(T_W) = \{T_W(x) : x \in W\} = \{T(x) : x \in W\}$ , so  $R(T_W) = T(W)$ .

**Problem 4.16.** Prove the following theorem: Let  $V$  and  $W$  be vector spaces, and let  $T : V \rightarrow W$  be linear. If  $\beta = \{v_1, v_2, \dots\}$  is an infinite basis for  $V$  then  $R(T) = \text{span}(\{T(v) : v \in \beta\})$ .

**Solution 4.16.** It is obvious that  $\text{span}(\{T(v) : v \in \beta\}) \subseteq R(T)$ , so we have to prove that  $R(T) \subseteq \text{span}(\{T(v) : v \in \beta\})$ . Let  $y \in R(T)$ . Then there exists  $x \in V$  such that  $T(x) = y$ . Since  $x \in V$ ,  $x = \sum_{i=1}^n a_i v_i$  for some  $v_i$  in  $\beta$ . Then  $T(\sum_{i=1}^n a_i v_i) = y$  which implies  $\sum_{i=1}^n a_i T(v_i) = y$ . Clearly,  $y \in \text{span}(\{T(v) : v \in \beta\})$  and so  $R(T) \subseteq \text{span}(\{T(v) : v \in \beta\})$ .

**Problem 4.17.** Prove the following: Let  $V$  and  $W$  be vector spaces over a common field, and let  $\beta$  be a basis for  $V$ . Then for any function  $f : \beta \rightarrow W$ , there exists exactly one linear transformation  $T : V \rightarrow W$  such that  $T(x) = f(x)$  for all  $x \in \beta$ .

**Solution 4.17.** Let  $x \in V$ . Then  $x = \sum_{i=1}^n a_i v_i$  where  $v_i \in \beta$ . Define  $T : V \rightarrow W$  by  $T(x) = \sum_{i=1}^n a_i f(v_i)$ . Then, for  $x = v_i = 0v_1 + 0v_2 + \dots + 0v_{i-1} + 1v_i + 0v_{i+1} + \dots + 0v_n = \sum_{j=1}^n a_j v_j$ , where  $a_j = 0$  for  $j \neq i$  and  $a_i = 1$ . Then  $T(x) = T(v_i) = T(\sum_{j=1}^n a_j v_j) = \sum_{j=1}^n a_j f(v_j) = a_i f(v_i) = 1f(v_i) = f(x)$ . Hence

$T(x) = f(x)$  for all  $x \in \beta$ .

For uniqueness, observe that for given  $x \in V$ ,  $a_1, a_2, \dots, a_n$  are unique, hence  $T(x)$  is also unique, so the map  $T$  is unique.

### 5. The Endomorphism Algebra of a Vector Space

**Proposition 5.1.** If  $V$  is a nontrivial vector space over a field  $F$ , then  $\text{End}(V)$  is an associative unital  $F$ -algebra with  $\sigma_0$  being the identity element for addition and  $\sigma_1$  being the identity element for multiplication.

PROOF. (1) Left Distributive: Let  $T, S, K \in \text{End}(V)$ . Let  $v \in V$ . Then  $(T(S + K))(v) = (T(S + K))(v) = T((S + K)(v)) = T(S(v) + K(v)) = T(S(v)) + T(K(v)) = (TS)(v) + (TK)(v)$  and therefore  $T(S + K) = TS + TK$   
 (2) Right Distributive: Let  $T, S, K \in \text{End}(V)$ . Let  $v \in V$ . Then  $((S + K)(T))(v) = (S + K)(T(v)) = S(T(v)) + K(T(v)) = (ST)(v) + (KT)(v)$  and therefore  $(S + K)T = ST + KT$   
 (3) Let  $a \in F, T, S \in \text{End}(V)$ . Then  $a(TS)(v) = aT(S(v)) = (aT)(S(v))$  and therefore  $a(TS) = (aT)S$ . Also,  $a(TS)(v) = aT(S(v)) = T(aS(v)) = T(aS)(v)$  and therefore  $a(TS) = T(aS)$ .  
 (4) Additive Identity: Let  $T \in \text{End}(V)$  and  $v \in V$ . Then  $(\sigma_0 + T)(v) = \sigma_0(v) + T(v) = 0_V + T(v) = T(v) + 0_V = T(v) + \sigma_0(v) = (T + \sigma_0)(v)$  and therefore  $\sigma_0 + T = T + \sigma_0$  and so  $\sigma_0$  is the identity element for  $\text{End}(V)$ .  
 (5) Multiplicative identity: Let  $T \in \text{End}(V)$  and  $v \in V$ . Then  $(\sigma_1 T)(v) = (\sigma_1(T(v))) = T(v) = (T(\sigma_1(v))) = (T\sigma_1)(v)$  and therefore  $\sigma_1 T = T\sigma_1$ .  $\square$

**Proposition 5.2.** Let  $V$  be a vector space over a field  $F$ . Then for all  $\alpha \in \text{End}(V)$  and all  $c \in F$  we have  $\alpha\sigma_c = \sigma_c\alpha$ .

PROOF. Let  $v \in V$ . Then  $(\alpha\sigma_c)(v) = \alpha(\sigma_c(v)) = \alpha(cv) = c\alpha(v) = \sigma_c(\alpha(v)) = (\sigma_c\alpha)(v)$  and therefore  $\alpha\sigma_c = \sigma_c\alpha$   $\square$

**Proposition 5.3.** Let  $V$  be a vector space of finite dimension  $n$  over a field  $F$ . Then the following conditions on an endomorphism  $\alpha$  of  $V$  are equivalent:

- (1)  $\alpha$  is an automorphism of  $V$ .
- (2)  $\alpha$  is one-one.
- (3)  $\alpha$  is onto.

PROOF. First, suppose that  $\alpha$  is an automorphism of  $V$ . Then  $\alpha$  is one-one by the definition of an automorphism. Now suppose that  $\alpha$  is one-one. Then  $\text{Ker}(\alpha) = \{0_V\}$  and so  $\dim(\text{ker}(\alpha)) = 0$ . By Rank-Nullity Theorem,  $\text{rank}(\alpha) = n$  and so  $\text{Im}(\alpha) = V$  and so  $\alpha$  is onto. Now suppose that  $\alpha$  is onto. Then  $\text{Im}(\alpha) = V$  and so  $\text{rank}(\alpha) = n$ . By rank-nullity theorem,  $\dim(\text{ker}(\alpha)) = 0$  and so  $\text{ker}(\alpha) = \{0_V\}$ . Therefore  $\alpha$  is one-one. Since  $\alpha$  is an automorphism which is one-one and onto, hence  $\alpha$  is an automorphism.  $\square$

**Proposition 5.4.** Let  $V$  be a finite-dimensional vector space over a field  $F$  and  $\alpha \in \text{End}(V)$ . If there exists a  $\beta \in \text{End}(V)$  satisfying  $\alpha\beta = \sigma_1$  or  $\beta\alpha = \sigma_1$ , then  $\alpha \in \text{Aut}(V)$  and  $\beta = \alpha^{-1}$ .

PROOF. If  $\beta\alpha = \sigma_1$ , then  $\text{ker}(\alpha) \subseteq \text{ker}(\sigma_1) = \{0_V\}$  and so by the previous proposition  $\alpha \in \text{Aut}(V)$ . Now, suppose, if  $\alpha\beta = \sigma_1$ , then  $\text{im}(\alpha) \subseteq \text{im}(\sigma_1) = V$ , and so by the previous proposition,  $\alpha \in \text{Aut}(V)$ . Also, if  $\alpha\beta = \sigma_1$ , we see that  $\alpha^{-1} = \alpha^{-1}\sigma_1 = \alpha^{-1}(\alpha\beta) = \beta$  and similarly  $\alpha^{-1} = \beta$  when  $\beta\alpha = \sigma_1$   $\square$

**Proposition 5.5.** Let  $V$  be a vector space over a field  $F$  and  $\alpha \in \text{End}(V)$  be a projection. Then  $V = \text{im}(\alpha) \oplus \text{ker}(\alpha)$ .

PROOF. First, we show that  $\text{im}(\alpha)$  and  $\text{ker}(\alpha)$  are disjoint sets. Let  $v \in \text{im}(\alpha) \cap \text{ker}(\alpha)$ . Then since  $v \in \text{im}(\alpha)$ , there exists an element  $x \in V$  such that  $v = \alpha(x)$ . Since  $v \in \text{ker}(\alpha)$ , hence  $\alpha(v) = 0_V$  and so  $0_V = \alpha(v) = \alpha(\alpha(x)) = \alpha(x) = v$  and so  $v = 0_V$  and therefore  $\text{im}(\alpha) \cap \text{ker}(\alpha) = \{0_V\}$ . Thus  $\text{im}(\alpha)$  and  $\text{ker}(\alpha)$  are disjoint.

Now assume that  $v \in V$ , then  $v - \alpha(v) \in \ker(\alpha)$  as  $\alpha(v - \alpha(v)) = \alpha(v) - \alpha(\alpha(v)) = \alpha(v) - \alpha(v) = 0_V$  and so  $v = (v - \alpha(v)) + \alpha(v) \in \ker(\alpha) + \text{im}(\alpha)$ . Therefore,  $V = \text{im}(\alpha) \oplus \ker(\alpha)$   $\square$

**Proposition 5.6.** Let  $V$  be a vector space over a field  $F$  and  $\alpha \in \text{End}(V)$ . A subspace  $W$  of  $V$  is invariant under  $\alpha$  if and only if  $\beta\alpha\beta = \alpha\beta$  for each projection  $\beta$  of  $V$  the image of which is  $W$ .

PROOF. Let  $W$  be a subspace of  $V$  which is invariant under  $\alpha$ . Let  $\beta$  be a projection of  $V$ , the image of which is  $W$ . By the previous proposition,  $V = W \oplus \ker(\beta)$ . Let  $v \in V$ . Then  $v$  can be written as  $v = w + y$  for some  $w \in W$  and  $y \in \ker(\beta)$ . Hence  $\alpha\beta(v) = \alpha\beta(w + y) = \alpha\beta(w) + \alpha\beta(y) = \alpha(w) + 0_V = \alpha(w) = \beta\alpha(w) = \beta\alpha\beta(v)$  and hence  $\beta\alpha\beta = \alpha\beta$ .

Now, suppose that  $\beta\alpha\beta = \alpha\beta$  for each projection  $\beta$  of  $V$ , the image of which is  $W$ .  $\square$

**Proposition 5.7.** Let  $V$  be a vector space over a field  $F$  and  $\{W_1, \dots, W_n\}$  be a set of subspaces of  $V$ . Then the following are equivalent:

- (1)  $V = W_1 \oplus \dots \oplus W_n$ .
- (2) There exists projections  $\alpha_1, \dots, \alpha_n$  in  $\text{End}(V)$  with  $W_i = \text{im}(\alpha_i)$  for all  $1 \leq i \leq n$ , which satisfy the conditions  $\alpha_i\alpha_j = \sigma_0$  for  $i \neq j$  and  $\alpha_1 + \dots + \alpha_n = \sigma_1$ .

PROOF. (1)  $\implies$  (2) Since  $V = W_1 \oplus \dots \oplus W_n$ , hence for any  $v \in V$ ,  $v = \sum_{i=1}^n w_i$ , where  $w_i \in W_i$  for all  $1 \leq i \leq n$ . By definition this representation of  $v$  is unique. Define  $\alpha_i : V \rightarrow W_i$  defined by  $\alpha_i(v) = w_i$ . Since the representation of  $v$  is unique, these maps are well-defined. Since  $\alpha_i : V \rightarrow W_i$ , hence  $\text{im}(\alpha_i) \subseteq W_i$ . Also if  $w_i \in W_i$ , then  $\alpha_i((0_V, \dots, w_i, \dots, 0_V)) = w_i$  and so  $W_i \subseteq \text{im}(\alpha_i)$ . Hence,  $W_i = \text{im}(\alpha_i)$ .

Now, for  $i \neq j$ ,  $(\alpha_i\alpha_j)(v) = \alpha_i(\alpha_j(v)) = \alpha_i(\alpha_j(\sum_{i=1}^n w_i)) = \alpha_i(w_j) = 0_V$  and so  $\alpha_i\alpha_j = \sigma_0$ .

Also, for  $(\alpha_1 + \dots + \alpha_n)(v) = (\alpha_1 + \dots + \alpha_n)(\sum_{i=1}^n w_i) = \alpha_1(\sum_{i=1}^n w_i) + \dots + \alpha_n(\sum_{i=1}^n w_i) = w_1 + \dots + w_n = v$  and so  $\alpha_1 + \dots + \alpha_n = \sigma_1$ .

- (2)  $\implies$  (1) Let  $v \in V$ . Then  $v = \sigma_1(v) = (\alpha_1 + \dots + \alpha_n)(v) = \alpha_1(v) + \dots + \alpha_n(v) \in \text{im}(\alpha_1) + \dots + \text{im}(\alpha_n) \subseteq \sum_{i=1}^n \text{im}(\alpha_i) = \sum_{i=1}^n W_i$  and so  $V = \sum_{i=1}^n W_i$ .

Now we have to show that  $W_h \cap \sum_{j \neq h} W_j = \phi$ . Suppose to that contradiction that  $W_h \cap \sum_{j \neq h} W_j \neq \phi$ . Then, there exists  $0_V \neq v \in W_h \cap \sum_{j \neq h} W_j$  for some  $i \neq h$ . Since  $v \in \alpha_i(v)$ , hence  $\alpha_i(v) \neq 0_V$ . But  $\alpha_h(v) = v$  so  $\alpha_i\alpha_h \neq \sigma_0$  which is a contradiction. Therefore,  $W_h \cap \sum_{j \neq h} W_j = \{0_V\}$  for each  $1 \leq h \leq n$ .  $\square$

## 6. The Matrix Representation of a Linear Map

**Problem 6.1.** Let  $V$  and  $W$  be finite-dimensional vector spaces with ordered bases  $\beta$  and  $\gamma$ , respectively, and let  $T : V \rightarrow W$  be a linear transformation. Then  $[aT]_\beta^\gamma = a[T]_\beta^\gamma$  for all scalars  $a$ .

**Solution 6.1.** Let  $\beta = \{v_1, v_2, \dots, v_n\}$  and  $\gamma = \{w_1, w_2, \dots, w_m\}$ . There exists a unique scalars  $a_{ij}$  ( $1 \leq i \leq m, 1 \leq j \leq n$ ) such that

$$T(v_j) = \sum_{i=1}^m a_{ij} w_i \text{ for } 1 \leq j \leq n.$$

Hence  $aT(v_j) = a \sum_{i=1}^m a_{ij} w_i = \sum_{i=1}^m aa_{ij} w_i$ . Thus  $[aT]_\beta^\gamma = a[T]_\beta^\gamma$ .

**Problem 6.2.** Let  $V$  be a  $n$ -dimensional vector space with an ordered basis  $\beta$ . Define  $T : V \rightarrow F^n$  by  $T(x) = [x]_\beta$ . Prove that  $T$  is linear.

**Solution 6.2.** Let  $\beta = \{u_1, u_2, \dots, u_n\}$ . Let  $x \in V$  and  $y \in V$ . Then  $T(cx + y) = [cx + y]_\beta$  where  $x = \sum_{i=1}^n a_i u_i$  and  $y = \sum_{i=1}^n b_i u_i$ . Then  $cx + y = \sum_{i=1}^n (ca_i + b_i) u_i$ .

$$[cx + y]_\beta = \begin{pmatrix} ca_1 + b_1 \\ ca_2 + b_2 \\ \vdots \\ ca_m + b_m \end{pmatrix} = c \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = c[x]_\beta + [y]_\beta.$$

So,  $T(cx + y) = cT(x) + T(y)$ .

**Problem 6.3.** Let  $V$  be the vector space of complex numbers over the field  $\mathbb{R}$ . Define  $T : V \rightarrow V$  by  $T(z) = \bar{z}$ , where  $\bar{z}$  is the complex conjugate of  $z$ . Prove that  $T$  is linear, and compute  $[T]_\beta$ , where  $\beta = \{1, i\}$ .

**Solution 6.3.** We have to prove that  $T(cz_1 + z_2) = cT(z_1) + T(z_2)$  for all complex numbers  $z_1$  and  $z_2$ . Consider two arbitrary complex numbers  $z_1$  and  $z_2$ . Then  $z_1 = x_1 + iy_1$  and  $z_2 = x_2 + iy_2$  for some real numbers  $x_1, y_1, x_2$  and  $y_2$ .

$$T(cz_1 + z_2) = T(cx_1 + x_2 + i(cy_1 + y_2)) = cx_1 + x_2 - i(cy_1 + y_2) = c(x_1 - iy_1) + (x_2 - iy_2) = cT(z_1) + T(z_2).$$

Hence  $T$  is linear.

Now we shall find the matrix representation of  $T$  with respect to  $\beta$  i.e.,  $[T]_\beta$ .

$$T(1) = \bar{1} = 1 = 1 + 0i$$

$$T(i) = \bar{i} = -i = 0 + (-1)i$$

$$\Rightarrow [T]_\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

**Problem 6.4.** Let  $V$  be a vector space with ordered basis  $\beta = \{v_1, v_2, \dots, v_n\}$ . Define  $v_0 = 0$ . By Theorem 2.6, there exists a linear transformation  $T : V \rightarrow V$  such that  $T(v_j) = v_j + v_{j-1}$  for  $j = 1, 2, \dots, n$ . Compute  $[T]_\beta$ .

$$\textbf{Solution 6.4. } T(v_1) = v_1 + v_0 = 1v_1 + 0v_2 + \dots + 0v_n$$

$$T(v_2) = v_2 + v_1 = 1v_1 + 1v_2 + \dots + 0v_n$$

$$T(v_3) = v_3 + v_2 = 0v_1 + 1v_2 + 1v_3 + \dots + 0v_n$$

$\vdots$

$$T(v_n) = 0v_1 + \dots + 1v_{n-1} + 1v_n$$

$$\Rightarrow [T]_\beta = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 1 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 1 \end{pmatrix}.$$

**Problem 6.5.** let  $V$  be an  $n$ -dimensional vector space, and let  $T : V \rightarrow V$  be a linear transformation. Suppose the  $W$  is a  $T$ -invariant subspace of  $V$  having dimension  $k$ . Show that there is a basis  $\beta$  for  $V$  such that  $[T]_\beta$  has the form

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

where  $A$  is a  $k \times k$  matrix and  $O$  is the  $(n - k) \times k$  zero matrix.

**Solution 6.5.** Let  $W$  have the basis  $\{v_1, v_2, \dots, v_n\}$ . Extend this to a basis of  $V = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ . Since  $W$  is  $T$ -invariant subspace of  $V$ , we can define the restriction of  $T$  on  $W$  to be the function  $T_W : W \rightarrow W$  such that  $T_W(x) = T(x)$  for all  $x \in W$ .

For  $j = 1, 2, \dots, k, T(w_j) = \sum_{s=1}^n b_{sj} w_s = \sum_{s=1}^k b_{sj} w_s \in W$ . For  $j = k+1, \dots, n, T$  maps  $w_j$  to vector outside of  $W$  (because otherwise  $w_j \in W$ , contradicting the linear independence of  $w_1, \dots, w_k, w_j$ ). Thus,  $T(w_j) = \sum_{s=k+1}^n b_{sj} w_s$ .

$$[T]_\beta = \begin{pmatrix} b_{11} & b_{12} \dots b_{1k} & 0 & 0 & \dots & 0 \\ b_{21} & b_{22} \dots b_{2k} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} \dots b_{kk} & b_{k+1,k+1} & b_{k+1,k+2} & \dots & b_{k+1,n} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 \dots 0 & b_{n,k+1} & b_{n,k+2} & \dots & b_{n,n} \end{pmatrix}$$

**Problem 6.6.** Let  $\beta = \{v_1, v_2, \dots, v_n\}$  be a basis for a vector space  $V$  and  $T : V \rightarrow V$  be a linear transformation. Prove that  $[T]_\beta$  is upper triangular if and only if  $T(v_j) \in \text{span}(\{v_1, v_2, \dots, v_j\})$  for  $j = 1, 2, \dots, n$ .

**Solution 6.6.** Assume  $[T]_\beta$  is upper triangular. Then  $a_{ij} = 0$  for  $i > j$ . Then  $T(v_j) = \sum_{i=1}^n a_{ij}v_i$ . As  $a_{ij} = 0$  for  $i > j$ , we have  $T(v_j) = \sum_{i=1}^j a_{ij}v_i$ . Thus clearly  $T(v_j) \in \text{span}(\{v_1, v_2, \dots, v_j\})$  for  $j = 1, 2, \dots, n$ . Now assume  $T(v_j) \in \text{span}(v_1, v_2, \dots, v_j)$ . We have to prove that  $a_{ij} = 0$  for all  $i > j$ .

$$T(v_j) = \sum_{i=1}^j b_i v_i$$

As  $v_1, \dots, v_n$  is a basis of  $V$ ,

$$T(v_j) = \sum_{i=1}^n a_i v_i$$

Then,  $\sum_{i=1}^n b_i v_i = \sum_{i=1}^n a_i v_i \implies \sum_{i=1}^j (b_i - a_i) v_i + \sum_{i=j+1}^n (-a_i) v_i = 0$

$v_i$  are linearly independent  $\implies b_i = a_i$  for  $i = 1, 2, \dots, j$  and  $a_{j+1} = a_{j+2} = \dots = a_n = 0$ . Thus  $a_{ij} = 0$  for  $i > j$ .

**Problem 6.7.** Let  $T$  and  $W$  be vector spaces, and let  $S$  be a subset of  $V$ . Define  $S^0 = \{T \in L(V, W) : T(x) = 0 \text{ for all } x \in S\}$ . Prove the following statements.

- (1)  $S^0$  is a subspace of  $L(V, W)$ .
- (2) If  $S_1$  and  $S_2$  are subsets of  $V$  and  $S_1 \subseteq S_2$ , then  $S_2^0 \subseteq S_1^0$ .
- (3) If  $V_1$  and  $V_2$  are the subspaces of  $V$ , then  $(V_1 \cup V_2)^0 = (V_1 + V_2)^0$ .

**Solution 6.7.** (1) Let  $T_1 \in S^0$  and  $T_2 \in S^0$ . Then  $T_1(x) = 0$  and  $T_2(x) = 0$  for all  $x \in S$ . Then  $(T_1 + T_2)(x) = T_1(x) + T_2(x) = 0$  for all  $x \in S$ . Also,  $(cT_1)(x) = cT_1(x) = 0$  for all  $x \in S$ .

(2) Let  $T \in S_2^0$ . Then  $T(x) = 0$  for all  $x \in S_2$ . Since  $S_1 \subseteq S_2$ ,  $T(x) = 0$  for all  $x \in S_1$  and so  $T \in S_1^0$ .

(3) Let  $V_1$  and  $V_2$  be the subspaces of  $V$ . We claim that  $(V_1 \cup V_2)^0 = (V_1 + V_2)^0$ . Let  $T \in (V_1 + V_2)^0$ , then  $T(x) = 0$  for all  $x \in V_1 + V_2$ . For any  $x \in V_1 \cup V_2$ , either  $x \in V_1$  or  $x \in V_2$ . Then  $T(v_1) = 0$  and  $T(v_2) = 0$  as  $v_1 \in V_1 \cup V_2$  and  $v_2 \in V_1 \cup V_2$ . Then  $T(x) = T(v_1 + v_2) = T(v_1) + T(v_2) = 0 + 0 = 0$ , which implies  $T \in (V_1 \cup V_2)^0$ . Hence  $(V_1 + V_2)^0 \subseteq (V_1 \cup V_2)^0$ .

Let  $T \in (V_1 \cup V_2)^0$ , then  $T(x) = 0$  for all  $x \in V_1 \cup V_2$ . Let  $x \in V_1 + V_2$ , then  $x = v_1 + v_2$  for some  $v_1 \in V_1$  and  $v_2 \in V_2$ . Then  $T(v_1) = 0$  as  $v_1 \in V_1 \cup V_2$  and  $T(v_2) = 0$  as  $v_2 \in V_1 \cup V_2$  and so  $T(x) = T(v_1 + v_2) = T(v_1) + T(v_2) = 0 + 0 = 0$ , which means  $T \in (V_1 + V_2)^0$ . Hence  $(V_1 \cup V_2)^0 \subseteq (V_1 + V_2)^0$ .

We proved that  $(V_1 \cup V_2)^0 = (V_1 + V_2)^0$ .

**Problem 6.8.** Let  $A$  be an  $m \times n$  matrix,  $B$  and  $C$  be  $n \times p$  matrices, and  $D$  and  $E$  be  $q \times m$  matrices. Then,

- (1)  $A(B + C) = AB + AC$  and  $(D + E)A = DA + EA$ .
- (2)  $a(AB) = (aA)B = A(aB)$  for any scalar  $a$ .
- (3)  $I_m A = A = A I_n$ .

**Solution 6.8.** (1) We shall prove that  $(D + E)A = DA + EA$ . Observe that  $((D + E)A)_{ij} = \sum_{k=1}^m (D + E)_{ik} A_{kj} = \sum_{k=1}^m D_{ik} A_{kj} + \sum_{k=1}^m E_{ik} A_{kj} = \sum_{k=1}^m D_{ik} A_{kj} + \sum_{k=1}^m E_{ik} A_{kj} = (DA)_{ij} + (EA)_{ij} = (DA + EA)_{ij}$

(2) First we will prove that  $(aA)B = A(aB)$ .

$$((aA)B)_{ij} = \sum_{k=1}^n (aA)_{ik} (B)_{kj} = \sum_{k=1}^n (A)_{ik} a (B)_{kj} = \sum_{k=1}^n (A)_{ik} (aB)_{kj} = A(aB)_{ij}$$

Now, we shall prove that  $a(AB) = (aA)B$ .

$$(a(AB))_{ij} = \sum_{k=1}^m a (A_{ik} B_{kj}) = \sum_{k=1}^m (aA)_{ik} B_{kj}, \text{ which implies } a(AB) = (aA)B.$$

(3)  $(A I_n)_{ij} = \sum_{k=1}^n A_{ik} (I_n)_{kj} = \sum_{k=1}^n A_{ik} \delta_{kj} = A_{ij}$

**Problem 6.9.** Let  $A$  be an  $m \times n$  matrix,  $B_1, B_2, \dots, B_k$  be  $n \times p$  matrices,  $C_1, C_2, \dots, C_k$  be  $q \times m$  matrices, and  $a_1, a_2, \dots, a_k$  be scalars. Then

$$A(\sum_{i=1}^k a_i B_i) = \sum_{i=1}^k a_i AB_i$$

and

$$(\sum_{i=1}^k a_i C_i)A = \sum_{i=1}^k a_i C_i A$$

**Solution 6.9.**  $(A(\sum_{i=1}^k a_i B_i))_{ij}$

$$= \sum_{p=1}^n A_{pk} (\sum_{i=1}^k a_i B_i)_{kp}.$$

By (c) of Theorem 2.12, the above is

$$= \sum_{p=1}^n \sum_{i=1}^k a_i A_{pk} (B_i)_{kp}$$

$$= (\sum_{i=1}^k a_i AB_i)_{ij}$$

**Problem 6.10.** Let  $A$  be a  $m \times n$  matrix with entries from  $F$ . Then the left-multiplication transformation  $L_A : F^n \rightarrow F^m$  is linear. Furthermore, if  $B$  is any other  $m \times n$  matrix (with entries from  $F$ ) and  $\beta$  and  $\gamma$  are the ordered bases for  $F^n$  and  $F^m$ , respectively, then we have the following properties.

(1)  $L_{A+B} = L_A + L_B$  and  $L_{aA} = aL_A$  for all  $a \in F$ .

(2) If  $m = n$ , then  $L_{I_n} = I_{F_n}$ .

**Solution 6.10.** (1)  $L_{A+B}(v) = L_A(v) + L_B(v)$  for all  $v \in F^n$   $L_{A+B}(v) = (A+B)v = Av + Bv = L_A(v) + L_B(v)$   
 $L_{aA}(v) = (aA)v = a(Av) = aL_A(v)$  for all  $v \in F^n$   
 which implies  $L_{aA} = aL_A$

**Problem 6.11.** Let  $V$  be a vector space, and let  $T : V \rightarrow V$  be linear. Prove that  $T^2 = T_0$  if and only if  $R(T) \subseteq N(T)$ .

**Solution 6.11.** Let  $V$  be a vector space and let  $T : V \rightarrow V$  be a linear. Assume  $T^2 = T_0$ . Let  $x \in R(T)$ , then there exists  $v \in V$  such that  $T(v) = x$ . Then  $T(x) = T(T(v)) = T^2(v) = T_0(v) = 0$  which implies  $x \in N(T)$ . Hence  $R(T) \subseteq N(T)$ .

Now, assume that  $R(T) \subseteq N(T)$ . We have to prove that  $T^2 = T_0$ . Let  $v \in V$ . Then  $T(v) \in R(T)$  which implies  $T(v) \in N(T)$  and  $T(T(v)) = 0$  or  $T^2(v) = 0$ . Hence  $T^2 = T_0$ .

**Problem 6.12.** Let  $A$  and  $b$  be  $n \times n$  matrices. Recall that the trace of  $A$  is defined by

$$\text{tr}(A) = \sum_{i=1}^n A_{ii}.$$

Prove that  $\text{tr}(AB) = \text{tr}(BA)$  and  $\text{tr}(A) = \text{tr}(A^T)$ .

**Solution 6.12.**  $\text{tr}(A) = \sum_{i=1}^n a_{ii} = \text{tr}(A^T)$   
 $\text{tr}(AB) = \sum_{i=1}^n (AB)_{ii} = \sum_{i=1}^n \sum_{k=1}^n A_{ik} B_{ki} = \sum_{k=1}^n \sum_{i=1}^n B_{ki} A_{ik} = \sum_{i=1}^n (BA)_{kk} = \text{tr}(BA)$ .

**Problem 6.13.** Let  $A$  and  $B$  be matrices for which the product matrix  $AB$  is defined, and let  $u_j$  and  $v_j$  denote the  $j$ th columns of  $AB$  and  $B$ , respectively. If  $v_p = c_1 v_{j_1} + c_2 v_{j_2} + \dots + c_k v_{j_k}$  for some scalars  $c_1, c_2, \dots, c_k$ , prove that  $u_p = c_1 u_{j_1} + c_2 u_{j_2} + \dots + c_k u_{j_k}$ .

**Solution 6.13.** Let  $A$  be a  $m \times n$  and  $B$  be a  $n \times p$  matrix. By Theorem 2.13,  $u_j = Av_j$  for  $j = 1, 2, \dots, p$ .  
 $u_p = Av_p = c_1 Av_{j_1} + c_2 Av_{j_2} + \dots + c_k Av_{j_k}$ .

Again by Theorem 2.13(a),  $Av_{j_1} = u_{j_1}$ ,  $Av_{j_2} = u_{j_2}$

Hence  $u_p = c_1 u_{j_1} + c_2 u_{j_2} + \dots + c_k u_{j_k}$

**Problem 6.14.** Let  $V$  be a finite-dimensional vector space, and let  $T : V \rightarrow V$  be linear.

(1) If  $\text{rank}(T) = \text{rank}(T^2)$ , prove that  $R(T) \cap N(T) = \{0\}$ . Deduce that  $V = R(T) \oplus N(T)$ .

**Solution 6.14.** We have  $\dim(\text{Im}(T)) = \dim(\text{Im}(T^2))$ . Let  $V' = \text{Im}(T)$ ,  $W' = T(V') = \text{Im}(T^2)$ . Let  $T' : V' \rightarrow W'$  be such that for all  $v \in V'$ ,  $T'(v) = T(v)$ .

By Rank-Nullity Theorem,  $\dim W' + \dim \text{Ker } T' = \dim V'$ . As  $\dim W' = \dim V'$ ,  $\dim \text{Ker } T' = 0$ .

Let  $u \in \text{Im } T = V'$  and  $u \in \text{Ker } T$ ,  $T(u) = 0$ , so  $T'(u) = 0$ . As  $\dim \text{Ker } T' = 0$ , we must have  $u = 0$ .

To prove,  $V = R(T) \oplus N(T)$ , observe that since  $\dim \text{Im } T + \dim \text{Ker } T = \dim V$  and  $\text{Im } T \cap \text{Ker } T = \{0\}$ , we have  $\text{Im } T \oplus \text{Ker } T = V$ .

**Problem 6.15.** Find linear transformation  $U, T : F^2 \rightarrow F^2$  such that  $UT = T_0$  (the zero transformation) but  $TU \neq T_0$ . Use your answer to find matrices  $A$  and  $B$  such that  $AB = O$  but  $BA \neq O$ .

**Solution 6.15.** We will use the fact that every linear transformation is uniquely determined by its action on basis vectors. Define  $U : F^2 \rightarrow F^2$  by

$$U(x, y) = (x, 0)$$

and define  $T : F^2 \rightarrow F^2$  by

$$T(x, y) = (0, x).$$

$$\text{Then } (U \circ T)(x, y) = u(0, x) = (0, 0)$$

$$(T \circ U)(x, y) = T(x, 0) = (0, x)$$

Consider the matrices

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

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

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

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

**Problem 6.16.** Using only the definition of matrix multiplication, prove that, the multiplication of matrices is associative.

**Solution 6.16.** Let  $A$  be a  $m \times n$ ,  $B$  be a  $n \times r$  and  $C$  be a  $r \times p$  matrix.

$$((AB)C)_{ij} = \sum_{k=1}^r (AB)_{ik} C_{kj}$$

$$= \sum_{k=1}^r \left( \sum_{l=1}^n A_{il} B_{lk} \right) C_{kj}$$

$$= \sum_{k=1}^r \sum_{l=1}^n A_{il} B_{lk} C_{kj}$$

$$= \sum_{l=1}^n \sum_{k=1}^r A_{il} B_{lk} C_{kj}$$

$$= \sum_{l=1}^n A_{il} \left( \sum_{k=1}^r B_{lk} C_{kj} \right)$$

$$= \sum_{l=1}^n A_{il} (BC)_{lj}$$

$$= (A(BC))_{ij}$$

**Problem 6.17.** Let  $V$  and  $Z$  be vector spaces and  $T : V \rightarrow Z$  be a linear transformation that is onto. Define the mapping

$$\bar{T} : V/N(T) \rightarrow Z \text{ by } \bar{T}(v + N(T)) = T(v)$$

for any coset  $v + N(T)$  in  $V/N(T)$ .

- (1) Prove that  $\bar{T}$  is well-defined. That is, prove that if  $v + N(T) = v' + N(T)$ , then  $T(v) = T(v')$ .
- (2) Prove that  $\bar{T}$  is linear.
- (3) Prove that  $\bar{T}$  is an isomorphism.

**Solution 6.17.** (1) First let,  $v + N(T) = v' + N(T)$ .

Claim:  $v - v' \in N(T)$ . Let  $t \in v + N(T)$ , then we have  $t \in v' + N(T)$ . Then  $t = v + x_1 = v' + x_2$ , where  $x_1, x_2 \in N(T)$ . Thus,  $v - v' \in N(T)$ . Then  $T(v - v') = 0$  and so  $T(v) - T(v') = 0$  or  $T(v) = T(v')$ .

(2) We have to prove that  $\bar{T}(v_1 + N(T) + c(v_2 + N(T))) = \bar{T}(v_1 + N(T)) + c\bar{T}(v_2 + N(T))$ . Now,  $v_1 + N(T) + c(v_2 + N(T)) = v_1 + cv_2 + N(T)$  and so  $\bar{T}(v_1 + N(T) + c(v_2 + N(T))) = \bar{T}(v_1 + cv_2 + N(T)) = \bar{T}(v_1 + cv_2) = T(v_1) + cT(v_2) = \bar{T}(v_1 + N(T)) + c\bar{T}(v_2 + N(T))$ .

(3) We have to prove that  $\bar{T}$  is a bijection. Let  $\bar{T}(v_1 + N(T)) = \bar{T}(v_2 + N(T))$ . But  $\bar{T}(v_1 + N(T)) = T(v_1)$  and  $\bar{T}(v_2 + N(T)) = T(v_2)$  and so  $T(v_1) = T(v_2)$  which implies  $T(v_1 - v_2) = 0$  and so  $v_1 - v_2 \in N(T)$ . Thus,  $v_1 + N(T) = v_2 + N(T)$ .

Now, we shall prove that the map  $\bar{T}$  is surjective.  $\bar{T} : V/N(T) \rightarrow Z$ . Consider  $z \in Z$ . As  $T$  is onto, there exists  $v \in V$  such that  $T(v) = z$  and so  $\bar{T}(v + N(T)) = T(v) = z$ . Thus, the map  $\bar{T}$  is surjective.

## CHAPTER 8

### Rings and Fields

This is a joint project with Rama Lakshmi Chunduri, a junior at the University of California, San Diego. The genesis of this chapter lies in the fact that after my training camp for the second round of Indian national mathematical olympiad, I began to study Babrbeau's book on the polynomial. My mentor discussed lot of problems from the PUTNAM and IMO. He proved many advanced theorems. However, I felt that I needed to study polynomials over a field in much details. This is how I started with the study of *rings of polynomials, fields of extensions*. This chapter is more of elementary problems that I and Rama solved together and these problems were taken from [Golan]

**History.** Nineteenth century Heinrich Weber used earlier works of Richard Dedekind and Leopold Kronecker for the development of the abstract theory of fields. Importance of such properties as associativity, distributivity, and so forth were largely studied by Augustus De Morgan and François Joseph Servois. The final axioms of a field are due to the twentieth century German mathematician Ernst Steinitz.

#### 1. Main definitions and theorems

**Definition 1.1 (Rings).** Let  $R$  be a set with two binary operations

$$+ : R \times R \rightarrow R \text{ and } \cdot : R \times R \rightarrow R$$

We call  $R$  a *ring* (under  $+$  and  $\cdot$ ) if:

- (1)  $(R, +)$  is an additive group (with identity written 0 or  $0_R$ ).
- (2)  $(R, \cdot)$  is a monoid (with identity written 1 or  $1_R$ ).
- (3)  $R$  satisfies the distributive laws, i.e., for all  $a, b, c \in R$ , we have:
  - (a)  $a \cdot (b + c) = a \cdot b + a \cdot c$ .
  - (b)  $(b + c) \cdot a = b \cdot a + c \cdot a$

If (2) is replaced

(2')  $(R, \cdot)$  satisfies associativity (i.e., possibly no 1),  
then, we call  $R$  (under  $+$  and  $\cdot$ ) a ring.

A ring satisfying

- (4)  $(R - \{0\}, \cdot)$  is a group.  
is called a division ring.

If  $R$  is a ring with  $1 \neq 0$ , then

$R^\times := \{x \in R \mid \text{There exists } x^{-1} \in R \text{ such that } xx^{-1} = 1 = x^{-1}x\} = \{x \in R \mid \text{There exists } a, b \in R \text{ such that } ax = xb\}$

is called the *group of units* of  $R$ . In particular, a ring is division ring if and only if  $R^\times = R - \{0\}$ .

**Definition 1.2.** A ring  $R$  is called a commutative ring if for all  $a, b \in R$ , we have

$$ab = ba$$

A commutative ring  $R$  is called an (*integral*) *domain* if

- (1)  $0 \neq 1$ .
- (2) Let  $a, b \in R$ . If  $ab = 0$ , then either  $a = 0$  or  $b = 0$ .

A commutative division ring is called a *field*.

**Theorem 1.1.** Let  $R$  be a ring and  $a, b, c \in R$ . Then we have the following are true:

- (1)  $a \cdot 0 = 0 = 0 \cdot a$ .
- (2)  $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$ .
- (3) If  $R$  is a ring, then for all  $a \in R$ :
  - (a)  $(-1) \cdot a = -a$ .
  - (b)  $(-1) \cdot (-1) = 1$ .
  - (c)  $|R| \geq 1$  if and only if  $1 \neq 0$ .
- (d) If  $R$  is a commutative ring with  $1 \neq 0$ , then  $R$  is a domain if and only if  $R$  satisfies the *Cancellation Law*: If  $a, b, c$  lie in  $R$ , then
 
$$a \cdot b = a \cdot c \text{ with } a \neq 0 \text{ implies } b = c.$$

**Definition 1.3.** A map  $\phi : R \rightarrow S$  of rings is called a *ring homomorphism* if for all  $a, b \in R$ , the map satisfies:

- (1)  $\phi(0_R) = 0_S$ .
- (2)  $\phi(a + b) = \phi(a) + \phi(b)$ .
- (3)  $\phi(1_R) = 1_S$ .
- (4)  $\phi(a \cdot b) = \phi(a) \cdot \phi(b)$ .

A ring homomorphism  $\phi : R \rightarrow S$  is called

- (1) a ring monomorphism or monic or mono if  $\phi$  is injective.
- (2) a ring epimorphism or epic or epi if  $\phi$  is surjective.
- (3) a ring isomorphism if  $\phi$  is bijective with inverse a ring homomorphism.
- (4) a ring automorphism if  $R = S$  and  $\phi$  is a ring isomorphism.

Of course, if  $R$  and  $S$  are fields, etc., then  $\phi$  is called a field homomorphism, etc. If  $\phi : R \rightarrow S$  is a ring homomorphism, we let

$\ker \phi := \{x \in R \mid \phi(x) = 0\} \subset R$

called the kernel of  $\phi$ , and

$\text{im } \phi := \{\phi(x) \mid x \in R\} \subset S$  called the image of  $\phi$ .

**Definition 1.4.** Let  $R$  be a ring,  $S \subset R$  a subset with  $S$  a ring. We say  $S$  is a subring of  $R$  if the inclusion map  $\text{inc} : S \rightarrow R$  (often written  $\text{inc} : S \hookrightarrow R$ ) is a ring monomorphism, so the restriction of the identity map on  $R$  to  $S$  is the inclusion map of  $S$  into  $R$  with  $1_S = 1_R$  and  $S$  is closed under the ring operations  $+$  and  $\cdot$  on  $R$ , meaning that the restriction of the operations on  $R$  to  $S$  have image in  $S$ , viz.,  $+\big|_{S \times S} : S \times S \rightarrow S$  and  $\cdot\big|_{S \times S} : S \times S \rightarrow S$ .

**Definition 1.5.** Let  $R$  be a ring. The center of  $R$  is the set

$$Z(R) := \{a \in R \mid ax = xa \text{ for all } x \in R\}.$$

**Definition 1.6.** Let  $R$  be a ring and  $A$  a nonempty subset of  $R$ . We call  $A$  a (2-sided) ideal of  $R$  if  $(A; +) \subset (R; +)$  is a subgroup and for all  $r \in R$  and  $a \in A$ , we have both  $ar$  and  $ra$  lie in  $A$ , i.e.,  $\cdot : R \times A \rightarrow A$  and  $\cdot : A \times R \rightarrow A$ . So an ideal of  $R$  is a subring of  $R$  closed under multiplication by  $R$  on both the right and left. If  $A$  only satisfies being a subring and closed under multiplication by  $R$  on the left (respectively, on the right), we call  $A$  a left ideal (respectively, right ideal). Of course, if  $R$  is a commutative ring, the notions of (2-sided) ideal, right ideal, and left ideal coincide.

**Definition 1.7.** Let  $R$  be a nontrivial ring. 1.  $0 = \{0\}$  is an ideal of  $R$  called the trivial ideal and  $R$  is an ideal of  $R$  called the unit ideal. If these are the only ideals of  $R$ , then we call  $R$  a simple ring.

**Definition 1.8.** Let  $a \in R$ . Define

$$(a) = RaR := \left\{ \sum_{i=1}^n x_i a y_i \mid n \in \mathbb{Z}^+; x_i, y_i \in R; 1 \leq i \leq n \right\},$$

an ideal of  $R$  called the principal ideal generated by  $a$ . It is the smallest ideal in  $R$  containing  $a$ . If  $R$  is commutative, this is much simpler, as (using the distributive laws), it is just

$$(a) = Ra := \{ra \mid r \in R\}$$

**Definition 1.9.** Let  $a_1, \dots, a_n \in R$ . The smallest ideal generated by  $a_1, \dots, a_n$  is

$$(a_1, \dots, a_n) = Ra_1R + \dots + Ra_nR := \left\{ \sum_{i=1}^n \sum_{j=1}^{m_i} x_{ij} a_i y_{ij} \mid x_{ij}, y_{ij} \in R \text{ all } i, j, \text{ some } m_i \in \mathbb{Z}^+ \right\}.$$

If  $R$  is commutative, this is simply

$$(a_1, \dots, a_n) = Ra_1 + \dots + Ra_n := \left\{ \sum_{i=1}^n r_i a_i \mid r_i \in R, 1 \leq i \leq n \right\},$$

i.e., it is all  $R$ -linear combinations of  $a_1, \dots, a_n$ .

**Definition 1.10.** Let  $I$  be an (indexing) set,  $A = \{a_i\}_{i \in I}$ . Then the smallest ideal in  $R$  containing  $A$  is

$$\langle A \rangle = (a_i)_{i \in I} := \{a \in R \mid \text{There exist } a_{i_1}, \dots, a_{i_n} \in A, \text{ some } n \text{ and } i_j \in I; \text{ satisfying } a \in (a_{i_1}, \dots, a_{i_n})\}.$$

We say that  $A$  generates  $I$ .  $\langle A \rangle$  is the union of all the ideals generated by  $a_{i_1}, \dots, a_{i_n}$  in  $A$  with  $i_1, \dots, i_n \in I$ , for some  $n$ .

**Definition 1.11.** Let  $A$  and  $B$  be ideals in  $R$ . Then

$$\begin{aligned} A + B &:= (A + B) = \{a + b \mid a \in A \text{ and } b \in B\} \\ A \cap B &:= \{x \mid x \in A \text{ and } x \in B\} \\ AB &:= (\{ab \mid a \in A, b \in B\}) \\ &= \{\sum_{i=1}^n a_i b_i \mid a_i \in A, b_i \in B, 1 \leq i \leq n, \text{ some } n\} \end{aligned}$$

are ideals in  $R$ . Moreover, we have  $AB \subset A \cap B \subset A + B$

**Definition 1.12.** Let  $R$  be a commutative ring,  $a, b \in R$  with  $a \neq 0$ . We say that  $a$  divides  $b$  and write  $a|b$ , if there exists an  $x \in R$  satisfying  $b = ax$ .

**Theorem 1.2.** Let  $R$  be a commutative ring and  $a, b, c$  elements in  $R$ .

- (1) If  $a|b$  and  $a|c$ , then  $a|bx + cy$  for all  $x, y \in R$ .
- (2) If  $R$  is a domain, then  $a|b$  and  $b|a$  if and only if there exists a unit  $u$  in  $R$  satisfying  $a = ub$ .

**Theorem 1.3.** Let  $R$  be a commutative ring. Then the relationship between division and principal ideal is given by the following: Let  $a, b$ , and  $c$  be nonzero elements in  $R$ .

1.  $a|b$  if and only if  $b \in (a)$  if and only if  $(a) \subset (b)$ .
2.  $(a) = (b)$  if and only if  $a|b$  and  $b|a$ .
3.  $a|b$  and  $a|c$  if and only if  $(b, c) \subset (a)$ .
4.  $(b, c) = (a)$  if and only if  $a|b$  and  $a|c$  and there exist  $x$  and  $y$  in  $R$  satisfying  $a = bx + cy$ .
5. Suppose that  $R$  is a domain. Then (a)  $a \sim b$  if and only if  $(a) = (b)$ . (b) If  $a = rb$  in  $R$ , then  $a \sim b$  if and only if  $r \in R^\times$ . In particular, if  $a \sim b$  and  $a = rb$  in  $R$ , then  $(a) < (b)$ . Indeed, if  $ub = a = rb$  with  $u \in R^\times$ , then  $u = r$  as  $R$  is a domain.

**Definition 1.13.** Let  $R$  be a commutative ring and  $A < R$  an ideal. We call  $A$  a prime ideal in  $R$  if  $ab \in A$  implies that  $a \in A$  or  $b \in A$ . It is called a maximal ideal in  $R$  if for any ideal  $B$  such that  $A < B \subset R$ , then  $B = R$ .

**Theorem 1.4.** Let  $R$  be a nontrivial commutative ring.

- (1)  $R$  is a domain if and only if  $0 = (0)$  is a prime ideal in  $R$ .
- (2) Every maximal ideal  $\mathfrak{m} \subset R$  is a prime ideal.
- (3) Suppose that  $R$  is a PID. Then every nonzero prime ideal in  $R$  is maximal.
- (4) The prime ideals in  $\mathbb{Z}$  are:

$$\begin{aligned} 0 = (0), & \text{ prime, not maximal;} \\ p\mathbb{Z} = (p), & p \text{ a prime, maximal.} \end{aligned}$$

**Lemma 1.1.** Let  $R$  be a commutative ring and  $p \subset R$  an ideal. Then  $p$  is a prime ideal if and only if for all ideals  $A$  and  $B$  in  $R$  satisfying  $AB \subset p$ , either  $A \subset p$  or  $B \subset p$ . Moreover, if  $A \subset R$  is not a prime ideal, then there exist ideals  $A \subset B_i$  in  $R$  with  $i = 1, 2$  satisfying  $B_1 B_2 \subset A$ .

**Definition 1.14.** Let  $R$  be a ring,  $A$  an ideal in  $R$ . If  $a$  and  $b$  are elements in  $R$ , write  $a \equiv b \pmod{A}$  if  $a - b$  is an element of  $A$ .

**Proposition 1.1.** Let  $R$  be a ring and  $A$  an ideal in  $R$ . Then  $\equiv \pmod{A}$  is an equivalence relation. Suppose that the elements  $a, a_0, b, b_0$  in  $R$  satisfy:

$$a \equiv a_0 \pmod{A} \text{ and } b \equiv b_0 \pmod{A};$$

then

$$a + b \equiv a_0 + b_0 \pmod{A}; \quad a \cdot b \equiv a_0 \cdot b_0 \pmod{A}.$$

Let  $R = R/A := R/(\equiv \pmod{A})$ . If  $a \in R$ , set

$$\bar{a} = \{b \in R \mid b \equiv a \pmod{A}\} = a + A.$$

If

$$\phi : R \rightarrow R/A \text{ given by } a \mapsto \bar{a}$$

denotes the canonical surjection and  $a, b$  lie in  $R$ , define

$$\begin{aligned} (1) \quad \bar{a} + \bar{b} &:= \overline{a + b} \\ (2) \quad \bar{a} \cdot \bar{b} &:= \overline{a \cdot b}. \end{aligned}$$

Then under these operations,  $R/A$  is a ring with  $0_{R/A} = A$  and  $1_{R/A} = 1_R + A$  with  $R/A$  commutative if  $R$  is. The canonical surjection is an epimorphism with kernel  $A$ . We call  $R/A$  the factor or quotient ring of  $R$  by  $A$ .

**Theorem 1.5 (First isomorphism theorem).** . Let  $R$  and  $S$  be rings and let  $\varphi : R \rightarrow S$  be a homomorphism. Then:

- (1) The kernel of  $\varphi$  is an ideal of  $R$ ,
- (2) The image of  $\varphi$  is a subring of  $S$ ,
- (3) The map

$$\phi : R/\ker \varphi \rightarrow \text{im } \varphi \subset S, \quad r + \ker \varphi \mapsto \varphi(r)$$

is a well-defined isomorphism.

**Theorem 1.6 (Second isomorphism theorem).** . Let  $R$  be a ring, let  $S \subset R$  be a subring, and let  $I$  be an ideal of  $R$ . Then:

- (1)  $S + I := \{s + a : s \in S, a \in I\}$  is a subring of  $R$ ,
- (2)  $S \cap I$  is an ideal of  $S$ , and
- (3)  $(S + I)/I$  is isomorphic to  $S/(S \cap I)$ .

**Theorem 1.7 (Third isomorphism theorem).** . Let  $R$  be a ring and let  $J \subset I$  be ideals of  $R$ . Then  $I/J$  is an ideal of  $R/J$  and

$$\frac{R/J}{I/J} \cong R/I.$$

**Theorem 1.8 (Correspondence Principle).** Let  $\phi : R \rightarrow S$  be a ring epimorphism. Then

$$\{B \mid B \text{ an ideal in } R \text{ with } \ker \phi \subset B\} \rightarrow \{C \mid C \text{ an ideal in } S\}$$

given by  $B \mapsto \phi(B)$  is an order preserving bijection.

**Theorem 1.9.** Let  $R$  be a nontrivial commutative ring and  $A$  an ideal of  $R$ . Then

- (1)  $A$  is a prime ideal of  $R$  if and only if  $R/A$  is a domain.
- (2)  $A$  is a maximal ideal of  $R$  if and only if  $R/A$  is a field.

**Theorem 1.10.** Let  $R$  be a ring and  $A \subset B \subset R$  ideals. Then

$$B/A := \{b + A \mid b \in B\} \subset R/A$$

is an ideal in  $R/A$  and

$$R/B \cong (R/A)/(B/A).$$

## 2. Problems and solutions

Each problem of this section is taken from exercises given in the book, *The Linear Algebra a Beginning Graduate Student Ought to Know* by Jonathan S. Golan, an Israeli Mathematicians.

**Definition 2.1.** A field is a set  $F$  together with two binary operations  $+$  and  $\cdot$  on  $F$  such that  $(F, +)$  is an abelian group and  $(F - \{0\}, \cdot)$  is also an abelian group, and the following distributive law holds:

$$a.(b + c) = (a.b) + (a.c) \text{ for all } a, b, c \in F.$$

**Definition 2.2. [Dummit and Foote]** A commutative ring with identity  $1 \neq 0$  is called an integral domain if it has no zero divisors.

**Theorem 2.1.** Every field is an integral domain.

**Example 2.1.** Let  $F = \mathbb{R}$ , and consider the following elements in  $G$  :

$$(a, b) = (1, 0), (c, d) = (0, 1)$$

Both  $(a, b)$  and  $(c, d)$  are nonzero elements in  $G$ , because neither is equal to  $(0, 0)$ .

Now compute their product under the given multiplication operation:

$$(a, b) \cdot (c, d) = (ac, bd) = (1 \cdot 0, 0 \cdot 1) = (0, 0).$$

We have found two nonzero elements  $(1, 0)$  and  $(0, 1)$  in  $G$  such that their product is zero. Therefore,  $G$  is not an integral domain. Since  $G$  is not an integral domain, it cannot be a field.

**Problem 2.1.** Let  $r \in \mathbb{R}$  and let  $0 \neq s \in \mathbb{R}$ . Define operations  $\boxplus$  and  $\boxtimes$  on  $\mathbb{R} \times \mathbb{R}$  by  $(a, b) \boxplus (c, d) = (a + c, b + d)$  and  $(a, b) \boxtimes (c, d) = (ac - bd(r^2 + s^2), ad + bc + 2rbd)$ . Do these operations, considered as addition and multiplication, respectively, define the structure of a field on  $\mathbb{R} \times \mathbb{R}$ ?

**Solution 2.1.** Let  $(a, b), (c, d), (e, f) \in \mathbb{R} \times \mathbb{R}$  for  $a, b, c, d, e, f \in \mathbb{R}$ .

**Associativity of addition and multiplication:**

$$\begin{aligned} ((a, b) \boxplus (c, d)) \boxplus (e, f) &= (a + c, b + d) \boxplus (e, f) \\ &= (a + c + e, b + d + f) \\ (a, b) \boxplus ((c, d) \boxplus (e, f)) &= (a, b) \boxplus (c + e, d + f) \\ &= (a + c + e, b + d + f) \\ ((a, b) \boxplus (c, d)) \boxplus (e, f) &= (a, b) \boxplus ((c, d) \boxplus (e, f)) \end{aligned}$$

Addition is associative.

$$\begin{aligned} ((a, b) \boxtimes (c, d)) \boxtimes (e, f) &= (ac - bd(r^2 + s^2), ad + bc + 2rbd) \boxtimes (e, f) \\ &= (ace - bde(r^2 + s^2) - (adf + bcf + 2rbdf)(r^2 + s^2), acf \\ &\quad - bdf(r^2 + s^2) + ade + bce + 2rbde + 2r(adf + bcf + 2rbdf)) \\ &= (ace - (bde + adf + bcf + 2rbdf)(r^2 + s^2), acf + ade + bce \\ &\quad - bdf(r^2 + s^2) + 2r(bde + adf + bcf + 2rbdf)) \end{aligned}$$

$$\begin{aligned} (a, b) \boxtimes ((c, d) \boxtimes (e, f)) &= (a, b) \boxtimes (ce - df(r^2 + s^2), cf + de + 2rdf) \\ &= (ace - adf(r^2 + s^2) - (bcf + bde + 2rbdf)(r^2 + s^2), acf + ade \\ &\quad + 2radf + bce - bdf(r^2 + s^2) + 2r(bcf + bde + 2rbdf)) \\ &= (ace - (bde + adf + bcf + 2rbdf)(r^2 + s^2), acf + ade + bce \\ &\quad - bdf(r^2 + s^2) + 2r(bde + adf + bcf + 2rbdf)) \\ ((a, b) \boxtimes (c, d)) \boxtimes (e, f) &= (a, b) \boxtimes ((c, d) \boxtimes (e, f)) \end{aligned}$$

Multiplication is associative.

**Commutativity of addition and multiplication:**

$$(a, b) \boxplus (c, d) = (a + c, b + d)$$

$$(c, d) \boxplus (a, b) = (a + c, b + d).$$

Addition is commutative.

$$(a, b) \boxdot (c, d) = (ac - bd(r^2 + s^2), ad + bc + 2rbd)$$

$$(c, d) \boxdot (a, b) = (ac - bd(r^2 + s^2), ad + bc + 2rbd).$$

Multiplication is commutative.

**Identity of addition and multiplication:**

$$(a, b) \boxplus (0, 0) = (a + 0, b + 0) = (a, b)$$

$$(a, b) \boxdot (1, 0) = (a - 0, 0 + b + 0) = (a, b).$$

**Inverse of addition and multiplication:**

$$(a, b) \boxplus (-a, -b) = (0, 0)$$

$$(a, b) \boxdot (c, d) = (ac - bd(r^2 + s^2), ad + bc + 2rbd) = (1, 0)$$

$$ac - bd(r^2 + s^2) = 1$$

$$ad + bc + 2rbd = 0$$

Solving these equations for  $c$  and  $d$  (given  $a \neq 0$  or  $b \neq 0$ ) shows that  $(c, d)$  exists, and hence, every nonzero element has a multiplicative inverse.

**Distributivity of multiplication over addition:**

$$(a, b) \boxdot ((c, d) \boxplus (e, f)) = (a, b) \boxdot (c + e, d + f)$$

$$= (ac + ae - (bd + bf)(r^2 + s^2), ad + af + bc + be + 2r(bd + bf))$$

$$((a, b) \boxdot (c, d)) \boxplus ((a, b) \boxdot (e, f)) = (ac - bd(r^2 + s^2), ad + bc + 2rbd)$$

$$\boxplus (ae - bf(r^2 + s^2), af + be + 2rbf)$$

$$= (ac + ae - (bd + bf)(r^2 + s^2), ad + af + bc + be + 2r(bd + bf))$$

The structure  $(\mathbb{R} \times \mathbb{R}, \boxplus, \boxdot)$  satisfies all the axioms of a field. Thus, it defines the structure of a field.

**Problem 2.2.** Define a new operation  $\dagger$  on  $\mathbb{R}$  by setting  $a \dagger b = a^3 b$ . Show that  $\mathbb{R}$ , equipped with the usual addition and this new operation as multiplication, satisfies all the axioms of a field except one.

**Solution 2.2.** Let  $a, b, c \in \mathbb{R}$ . A field must satisfy the following axioms:

**Associativity of addition and multiplication:**

$$a + (b + c) = a + b + c$$

$$(a + b) + c = a + b + c$$

$$a + (b + c) = (a + b) + c$$

Addition is associative

$$a \dagger (b \dagger c) = a \dagger (b^3 c) = a^3 b^3 c$$

$$(a \dagger b) \dagger c = a^3 b \dagger c = a^9 b^3 c$$

$$a \dagger (b \dagger c) \neq (a \dagger b) \dagger c$$

Multiplication ( $\dagger$ ) is not associative

**Commutativity of addition and multiplication:**

Since  $a, b \in \mathbb{R}$

$$a + b = b + a$$

Addition is commutative

$$\begin{aligned} a \dagger b &= a^3 b \\ b \dagger a &= b^3 a \\ a \dagger b &\neq b \dagger a \end{aligned}$$

Multiplication ( $\dagger$ ) is not commutative

Distribution of multiplication over addition:

$$\begin{aligned} a \dagger (b + c) &= a^3(b + c) = a^3 b + a^3 c \\ (a \dagger b) + (a \dagger c) &= a^3 b + a^3 c \\ a \dagger (b + c) &= (a \dagger b) + (a \dagger c) \end{aligned}$$

So multiplication is distributive over addition.

Existence of identities of addition and multiplication:

$$a + 0 = 0 + a = a$$

So 0 is the identity for addition. Multiplication has to be commutative for the existence of multiplicative identity and multiplicative inverse.

**Problem 2.3.** Let  $1 < t \in \mathbb{R}$  and let  $F = \{a \in \mathbb{R} \mid a < 1\}$ . Define the operations  $\oplus$  and  $\odot$  in  $F$  as follows:

- (1)  $a \oplus b = a + b - ab$  for all  $a, b \in F$ ;
- (2)  $a \odot b = 1 - t^{\log_t(1-a) \log_t(1-b)}$  for all  $a, b \in F$ . For which values of  $t$  does  $F$ , together with these operations, form a field?

**Solution 2.3.** If  $F$  forms a field, then both operations must be closed, associative, and commutative, along with the existence of additive and multiplicative identities and inverses.

Thus, for  $a, b, c \in F$ , we proceed as follows:

**Commutativity of Operation  $\oplus$ :**

$$\begin{aligned} a \oplus b &= a + b - ab \\ b \oplus a &= b + a - ba = a + b - ab \\ a \oplus b &= b \oplus a \end{aligned}$$

Operation  $\oplus$  is commutative.

**Commutativity of Operation  $\odot$ :**

$$\begin{aligned} a \odot b &= 1 - t^{\log_t(1-a) \log_t(1-b)} \\ b \odot a &= 1 - t^{\log_t(1-b) \log_t(1-a)} = 1 - t^{\log_t(1-a) \log_t(1-b)} \\ a \odot b &= b \odot a \end{aligned}$$

Operation  $\odot$  is commutative.

**Associativity of Operation  $\oplus$ :**

$$\begin{aligned} a \oplus (b \oplus c) &= a \oplus (b + c - bc) \\ &= a + b + c - bc - a(b + c - bc) \\ &= a + b + c - bc - ab - ac + abc \\ (a \oplus b) \oplus c &= (a + b - ab) \oplus c \\ &= a + b + c - ab - c(a + b - ab) \\ &= a + b + c - bc - ab - ac + abc \\ (a \oplus b) \oplus c &= a \oplus (b \oplus c) \end{aligned}$$

Operation  $\oplus$  is associative.

Notice that we can simplify the  $\odot$  expression as

$$a \odot b = 1 - t^{\log_t(1-a) \log_t(1-b)} = 1 - (t^{\log_t(1-a)})^{\log_t(1-b)}$$

Notice that  $t^{\log_t(1-b)} = 1 - b$ . So we have,

$$a \odot b = 1 - (1 - a)^{\log_t(1-b)}$$

Associativity of Operation  $\odot$ :

$$\begin{aligned} a \odot (b \odot c) &= a \odot (1 - (1 - b)^{\log_t(1-c)}) \\ &= 1 - (1 - a)^{\log_t(1 - (1 - (1 - b)^{\log_t(1-c)}))} \\ &= 1 - (1 - a)^{\log_t((1-b)^{\log_t(1-c)})} \\ &= 1 - (1 - a)^{\log_t(1-b) \log_t(1-c)} \\ &= 1 - t^{\log_t(1-a) \log_t(1-b) \log_t(1-c)} \\ (a \odot b) \odot c &= (1 - t^{\log_t(1-a) \log_t(1-b)}) \odot c \\ &= 1 - (t^{\log_t(1-b)})^{\log_t(1-a)} \odot c \\ &= 1 - (1 - b)^{\log_t(1-a)} \odot c \\ &= 1 - t^{\log_t(1 - (1 - (1 - b)^{\log_t(1-a)})) \log_t(1-c)} \\ &= 1 - t^{\log_t(1-b) \log_t(1-a) \log_t(1-c)} \\ &= 1 - t^{\log_t(1-a) \log_t(1-b) \log_t(1-c)} \end{aligned}$$

Thus, we have:

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

This shows that the operation  $\odot$  is associative.

#### Additive Identity:

Let  $b$  be the identity for the operation  $\oplus$  in the field  $F$

Then  $a \oplus b = a$

$$a \oplus b = a + b - ab = a \implies b = 0$$

So  $b = 0$  is the additive identity.

Multiplicative Identity:

Let  $b$  be the identity for the operation  $\odot$  in the field  $F$

$$a \odot b = a \implies 1 - t^{\log_t(1-a) \log_t(1-b)} = a \implies t^{\log_t(1-a) \log_t(1-b)} = 1 - a$$

Applying  $\log_t$  on both sides

$$\begin{aligned} \log_t(t^{\log_t(1-a) \log_t(1-b)}) &= \log_t(1 - a) \\ (\log_t(1 - a) \log_t(1 - b)) \log_t t &= \log_t(1 - a) \\ \log_t(1 - a) \log_t(1 - b) &= \log_t(1 - a) \\ \log_t(1 - b) &= 1 \implies 1 - b = t \end{aligned}$$

$b = 1 - t$  is the multiplicative identity of the field  $F$ .

#### Additive inverse:

Let  $b$  be the additive inverse of  $a$

Then  $a \oplus b = 0 \implies a + b - ab = 0$

$$a + b - ab = 0 \implies b(1 - a) = -a \implies a = \frac{a}{a - 1}$$

The additive inverse of  $a$  would be  $\frac{a}{a-1}$ .

Multiplicative inverse:

Let  $b$  be the multiplicative inverse of  $a$

Then  $a \odot b = 1 - b$

$$1 - t^{\log_t(1-a) \log_t(1-b)} = 1 - (1 - b)^{\log_t(1-a)} = 1 - (1 - b)^{\log_{1-b}(1-a)}$$

$$1 - (1 - a) = 1 - b \implies b = 1 - a$$

So the multiplicative inverse of  $a$  is  $1 - a$

Since  $F$  has associativity of addition and multiplication, commutativity of addition and multiplication, distributivity of multiplication over addition, existence of additive and multiplicative identities and existence of additive and multiplicative inverses. Therefore,  $G$  is a field.

**Definition 2.3. [Dummit and Foote]** If  $K$  is a field containing the sub field  $F$ , then  $K$  is said to be an extension field (or simply an extension) of  $F$ , denoted  $K/F$ .

**Definition 2.4. [Dummit and Foote]** The degree (or relative degree or index) of a field extension  $K/F$ , denoted  $[K : F]$ , is the dimension of  $K$  as a vector space over  $F$  (i.e.,  $[K : F] = \dim_F K$ ). The extension is said to be finite if  $[K : F]$  is finite and is said to be infinite otherwise.

**Problem 2.4.** Show that the set of all real numbers of the form  $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ , where  $a, b, c, d \in \mathbb{Q}$ , forms a subfield of  $\mathbb{R}$ .

**Solution 2.4.** The set  $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\}$ . This is the field obtained by adjoining  $\sqrt{2}$  and  $\sqrt{3}$ . The degree of the field extension  $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$ , because  $\sqrt{2}$  and  $\sqrt{3}$  are algebraic of degree 2 over  $\mathbb{Q}$ . To prove that the set  $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\}$  forms a subfield of  $\mathbb{R}$ , we must show that the set satisfies the following field axioms.

Let  $a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6}, a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}, a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6} \in \mathbb{R}$  for  $a_i, b_i, c_i \in \mathbb{Q}$  for  $1 \leq i \leq 3$ .

**Associativity of addition and multiplication:**

$$\begin{aligned} & (a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6} + a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}) + a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6} \\ &= a_1 + a_2 + (b_1 + b_2)\sqrt{2} + (c_1 + c_2)\sqrt{3} + (d_1 + d_2)\sqrt{6} + a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6} \\ &= a_1 + a_2 + a_3 + (b_1 + b_2 + b_3)\sqrt{2} + (c_1 + c_2 + c_3)\sqrt{3} + (d_1 + d_2 + d_3)\sqrt{6} \\ & a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6} + (a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6} + a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6}) \\ &= a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6} + (a_2 + a_3) + (b_2 + b_3)\sqrt{2} + (c_2 + c_3)\sqrt{3} + (d_2 + d_3)\sqrt{6} \\ &= (a_1 + a_2 + a_3) + (b_1 + b_2 + b_3)\sqrt{2} + (c_1 + c_2 + c_3)\sqrt{3} + (d_1 + d_2 + d_3)\sqrt{6} \\ & (a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6} + a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}) + a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6} \\ &= a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6} + (a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6} + a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6}) \end{aligned}$$

Addition is associative

Multiplication is associative

**Commutativity of addition and multiplication:**

$$\begin{aligned} a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6} + a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6} &= (a_1 + a_2) + (b_1 + b_2)\sqrt{2} \\ &+ (c_1 + c_2)\sqrt{3} + (d_1 + d_2)\sqrt{6} \\ a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6} + a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6} &= (a_1 + a_2) + (b_1 + b_2)\sqrt{2} \\ &+ (c_1 + c_2)\sqrt{3} + (d_1 + d_2)\sqrt{6} \end{aligned}$$

Addition is commutative

$$\begin{aligned}
 (a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6}) \cdot (a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}) &= a_1a_2 + 2b_1b_2 + 3c_1c_2 + 6d_1d_2 \\
 &\quad + (a_1b_2 + b_1a_2)\sqrt{2} + (a_1c_2 + c_1a_2)\sqrt{3} \\
 &\quad + (a_1d_2 + d_1a_2)\sqrt{6} + (b_1c_2 + c_1b_2)\sqrt{2}\sqrt{3} \\
 &\quad + (b_1d_2 + d_1b_2)\sqrt{2}\sqrt{6} + (c_1d_2 + d_1c_2)\sqrt{3}\sqrt{6} \\
 (a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}) \cdot (a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6}) &= a_1a_2 + 2b_1b_2 + 3c_1c_2 + 6d_1d_2 \\
 &\quad + (a_1b_2 + b_1a_2)\sqrt{2} + (a_1c_2 + c_1a_2)\sqrt{3} \\
 &\quad + (a_1d_2 + d_1a_2)\sqrt{6} + (b_1c_2 + c_1b_2)\sqrt{2}\sqrt{3} \\
 &\quad + (b_1d_2 + d_1b_2)\sqrt{2}\sqrt{6} + (c_1d_2 + d_1c_2)\sqrt{3}\sqrt{6}
 \end{aligned}$$

Multiplication is commutative

**Distributivity of multiplication over addition:**

$$\begin{aligned}
 (a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6}) \cdot ((a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}) + (a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6})) \\
 = (a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6}) \cdot (a_2 + b_2\sqrt{2} + c_2\sqrt{3} + d_2\sqrt{6}) \\
 + (a_1 + b_1\sqrt{2} + c_1\sqrt{3} + d_1\sqrt{6}) \cdot (a_3 + b_3\sqrt{2} + c_3\sqrt{3} + d_3\sqrt{6})
 \end{aligned}$$

Multiplication is distributive over addition

**Existence of identity elements for addition and multiplication:**

$a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} + 0 + 0\sqrt{2} + 0\sqrt{3} + 0\sqrt{6} = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ . So 0 is identity element for addition.

Similarly, 1 is identity element for multiplication.

**Existence of additive inverses:**

If  $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$  then  $-a - b\sqrt{2} - c\sqrt{3} - d\sqrt{6} \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$   $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} + (-a - b\sqrt{2} - c\sqrt{3} - d\sqrt{6}) = 0$ .

So  $-a - b\sqrt{2} - c\sqrt{3} - d\sqrt{6}$  is the additive inverse of  $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ .

If  $a \neq 0 \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$  then  $\frac{1}{a} \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$

$$a \cdot \frac{1}{a} = 1$$

Since G has associativity of addition and multiplication, commutativity of addition and multiplication, distributivity of multiplication over addition, existence of additive and multiplicative identities and existence of additive and multiplicative inverses. Therefore G is a field.

**Problem 2.5.** Is  $\{a + b\sqrt{15} \mid a, b \in \mathbb{Q}\}$  a sub-field of  $\mathbb{R}$ ?

**Solution 2.5.** The set  $\mathbb{Q}(\sqrt{15}) = \{a + b\sqrt{15} \mid a, b \in \mathbb{Q}\}$ . This is the field of degree 2 extension of  $\mathbb{Q}$ . To prove that the set  $\mathbb{Q}(\sqrt{15})$  forms a subfield of  $\mathbb{R}$ , we must show that the set satisfies the following field axioms.

Let  $a_1 + b_1\sqrt{15}, a_2 + b_2\sqrt{15}, a_3 + b_3\sqrt{15} \in \mathbb{Q}(\sqrt{15})$  where  $a_1, a_2, a_3, b_1, b_2, b_3 \in \mathbb{R}$ .

**Associativity of Addition and Multiplication:**

$$\begin{aligned}
 (a_1 + b_1\sqrt{15} + a_2 + b_2\sqrt{15}) + a_3 + b_3\sqrt{15} &= (a_1 + a_2) + (b_1 + b_2)\sqrt{15} + a_3 + b_3\sqrt{15} \\
 &= (a_1 + a_2 + a_3) + (b_1 + b_2 + b_3)\sqrt{15} \\
 (a_1 + b_1\sqrt{15} + (a_2 + b_2\sqrt{15} + a_3 + b_3\sqrt{15})) &= a_1 + b_1\sqrt{15} + (a_2 + a_3) + (b_2 + b_3)\sqrt{15} \\
 &= (a_1 + a_2 + a_3) + (b_1 + b_2 + b_3)\sqrt{15}
 \end{aligned}$$

Thus, we have:

$$(a_1 + b_1\sqrt{15} + a_2 + b_2\sqrt{15}) + a_3 + b_3\sqrt{15} = a_1 + b_1\sqrt{15} + (a_2 + b_2\sqrt{15} + a_3 + b_3\sqrt{15}).$$

Addition is associative.

$$\begin{aligned}
 ((a_1 + b_1\sqrt{15}) \cdot (a_2 + b_2\sqrt{15})) \cdot (a_3 + b_3\sqrt{15}) &= (a_1 \cdot a_2 + (a_1b_2 + a_2b_1)\sqrt{15} + 15b_1b_2) \cdot (a_3 + b_3\sqrt{15}) \\
 &= a_1a_2a_3 + \sqrt{15}(a_1a_2b_3 + a_1b_2a_3 + a_2b_1a_3) \\
 &\quad + 15\sqrt{15}b_1b_2b_3 + 15(b_1b_2a_3 + a_1b_2b_3 + a_2b_1b_3) \\
 &\quad (a_1 + b_1\sqrt{15}) \cdot ((a_2 + b_2\sqrt{15}) \cdot (a_3 + b_3\sqrt{15})) \\
 &= (a_1 + b_1\sqrt{15}) \cdot (a_2a_3 + \sqrt{15}(a_2b_3 + a_3b_2) + 15b_2b_3) \\
 &= a_1a_2a_3 + \sqrt{15}(a_1a_2b_3 + a_1a_3b_2 + b_1a_2a_3) \\
 &\quad + 15(b_2b_3a_1 + b_1a_2b_3 + b_1a_3b_2) + 15\sqrt{15}b_1b_2b_3
 \end{aligned}$$

Thus, we have:

$$((a_1 + b_1\sqrt{15}) \cdot (a_2 + b_2\sqrt{15})) \cdot (a_3 + b_3\sqrt{15}) = (a_1 + b_1\sqrt{15}) \cdot ((a_2 + b_2\sqrt{15}) \cdot (a_3 + b_3\sqrt{15})).$$

Multiplication is associative.

**Commutativity of Addition and Multiplication:**

$$\begin{aligned}
 (a_1 + b_1\sqrt{15}) + (a_2 + b_2\sqrt{15}) &= (a_1 + a_2) + (b_1 + b_2)\sqrt{15}, \\
 (a_2 + b_2\sqrt{15}) + (a_1 + b_1\sqrt{15}) &= (a_2 + a_1) + (b_2 + b_1)\sqrt{15} = (a_1 + a_2) + (b_1 + b_2)\sqrt{15}.
 \end{aligned}$$

Thus,

$$(a_1 + b_1\sqrt{15}) + (a_2 + b_2\sqrt{15}) = (a_2 + b_2\sqrt{15}) + (a_1 + b_1\sqrt{15}).$$

Addition is commutative.

$$\begin{aligned}
 (a_1 + b_1\sqrt{15}) \cdot (a_2 + b_2\sqrt{15}) &= a_1a_2 + (a_1b_2 + a_2b_1)\sqrt{15} + 15b_1b_2, \\
 (a_2 + b_2\sqrt{15}) \cdot (a_1 + b_1\sqrt{15}) &= a_2a_1 + (a_2b_1 + a_1b_2)\sqrt{15} + 15b_2b_1 \\
 &= a_1a_2 + (a_1b_2 + a_2b_1)\sqrt{15} + 15b_1b_2.
 \end{aligned}$$

Thus,

$$(a_1 + b_1\sqrt{15}) \cdot (a_2 + b_2\sqrt{15}) = (a_2 + b_2\sqrt{15}) \cdot (a_1 + b_1\sqrt{15}).$$

Multiplication is commutative.

**Distributivity of Multiplication over Addition:**

$$\begin{aligned}
 (a_1 + b_1\sqrt{15}) \cdot ((a_2 + b_2\sqrt{15}) + (a_3 + b_3\sqrt{15})) &= (a_1 + b_1\sqrt{15}) \cdot ((a_2 + a_3) + (b_2 + b_3)\sqrt{15}) \\
 &= (a_1a_2 + a_1a_3) \\
 &\quad + (a_1b_2 + a_1b_3 + b_1a_2 + b_1a_3)\sqrt{15} \\
 &\quad + 15(b_1b_2 + b_1b_3) \\
 (a_1 + b_1\sqrt{15}) \cdot (a_2 + b_2\sqrt{15}) + (a_1 + b_1\sqrt{15}) \cdot (a_3 + b_3\sqrt{15}) &= (a_1a_2 + (a_1b_2 + a_2b_1)\sqrt{15} + 15b_1b_2) \\
 &\quad + (a_1a_3 + (a_1b_3 + a_3b_1)\sqrt{15} + 15b_1b_3) \\
 &= ((a_1a_2 + a_1a_3) \\
 &\quad + (a_1b_2 + a_1b_3 + b_1a_2 + b_1a_3)\sqrt{15} \\
 &\quad + 15(b_1b_2 + b_1b_3)
 \end{aligned}$$

Thus,

$$(a_1 + b_1\sqrt{15}) \cdot ((a_2 + b_2\sqrt{15}) + (a_3 + b_3\sqrt{15})) = (a_1 + b_1\sqrt{15}) \cdot (a_2 + b_2\sqrt{15}) + (a_1 + b_1\sqrt{15}) \cdot (a_3 + b_3\sqrt{15}).$$

Multiplication is distributive over addition

### Existence of Identity Elements for Addition and Multiplication

For addition:

Since  $0 \in \mathbb{Q}$ , we have when  $a_2 = 0$  and  $b_2 = 0$ :

$$a_2 + b_2\sqrt{15} = 0 + 0\sqrt{15} = 0.$$

Thus,

$$(a_1 + b_1\sqrt{15}) + 0 = (a_1 + b_1\sqrt{15}),$$

$$0 + (a_1 + b_1\sqrt{15}) = (a_1 + b_1\sqrt{15}).$$

So 0 is the identity element for addition.

For multiplication:

Since  $1 \in \mathbb{Q}$ , we have when  $a_2 = 1$  and  $b_2 = 0$ :

$$a_2 + b_2\sqrt{15} = 1 + 0\sqrt{15} = 1.$$

Thus,

$$(a_1 + b_1\sqrt{15}) \cdot 1 = (a_1 + b_1\sqrt{15}),$$

$$1 \cdot (a_1 + b_1\sqrt{15}) = (a_1 + b_1\sqrt{15}).$$

So 1 is the identity element for multiplication.

### Existence of Additive Inverses:

If  $a, b \in \mathbb{Q}$ , then  $-a, -b \in \mathbb{Q}$ :

$$(a + b\sqrt{15}) + (-a - b\sqrt{15}) = (a + (-a) + (b + (-b))\sqrt{15}) = 0 + 0\sqrt{15} = 0.$$

Thus,  $-a - b\sqrt{15}$  is the additive inverse of  $a + b\sqrt{15}$ .

### Existence of Multiplicative Inverses:

For  $a + b\sqrt{15}$ :

$$(a + b\sqrt{15}) \cdot \frac{1}{a + b\sqrt{15}} = 1.$$

The multiplicative inverse is given by:

$$\frac{1}{a + b\sqrt{15}} = \frac{a - b\sqrt{15}}{(a + b\sqrt{15})(a - b\sqrt{15})} = \frac{a - b\sqrt{15}}{a^2 - 15b^2}.$$

This simplifies to:

$$= \frac{a}{a^2 - 15b^2} + \sqrt{15} \left( -\frac{b}{a^2 - 15b^2} \right).$$

Since  $\frac{a}{a^2 - 15b^2}, -\frac{b}{a^2 - 15b^2} \in \mathbb{Q}$ , we conclude that  $\frac{1}{a + b\sqrt{15}}$  is the multiplicative inverse of  $a + b\sqrt{15}$ . Since  $\{a + b\sqrt{15} \mid a, b \in \mathbb{Q}\}$  has associativity of addition and multiplication, commutativity of addition and multiplication, distributivity of multiplication over addition, existence of additive and multiplicative identities and existence of additive and multiplicative inverses. Therefore it is a field.

**Problem 2.6.** Show that the field  $\mathbb{R}$  has infinitely-many distinct subfields.

**Solution 2.6.** Consider the collection  $\mathbb{Q}[p] = \{a + b\sqrt{p} \mid a, b \in \mathbb{Q}\}$ , where  $p$  is any prime. Then following sets of solution given to the previous problem we can prove that  $\mathbb{Q}[p]$  is a field. Since, due to Euler we know that there are infinitely many primes, hence there are infinitely many subfield of  $\mathbb{R}$ . Also, each of these subfield contains rational number  $\mathbb{Q}$  as a subfield.

**Definition 2.5. [Dummit and Foote]** The characteristic of a field  $F$ , denoted  $ch(F)$ , is defined to be the smallest positive integer  $p$  such that  $p \cdot 1_F = 0$  if such a  $p$  exists and is defined to be 0 otherwise.

**Problem 2.7.** Let  $F$  be a field and define a new operation  $*$  on  $F$  by setting  $a * b = a + b + ab$ . When is  $(F, +, *)$  a field?

**Solution 2.7.** For  $(F, +, *)$  there should be additive and multiplicative inverses and identities.

So there exists multiplicative identity  $x$  such that  $a * x = x * a = a$

$$a * x = a + x + ax$$

Note that  $x = 0$ , is the multiplicative identity as :

$$a * 0 = a + 0 + a(0) = a$$

$$0 * a = 0 + a + 0(a) = a$$

So, the multiplicative identity in this field is 0.

Distributivity of  $*$  over  $+$

We need to check if:

$$a * (b + c) = a * b + a * c.$$

Calculate  $a * (b + c)$ :

$$a * (b + c) = a + (b + c) + a(b + c) = a + b + c + ab + ac.$$

Now calculate  $a * b + a * c$ :

$$a * b = a + b + ab, \quad a * c = a + c + ac,$$

thus,

$$a * b + a * c = (a + b + ab) + (a + c + ac) = 2a + b + c + ab + ac.$$

This is equal if  $2a = 0$  for all  $a$ , which holds if the characteristic of  $F$  is 2.

Conclusion

The structure  $(F, +, *)$  is a field if and only if:

- The field  $F$  has characteristic 2.

**Problem 2.8.** Let  $F$  be a field and let  $G_n$  be the subset of  $F$  consisting of all elements which can be written as a sum of  $n$  squares of elements of  $F$ .

(1) Is the product of two elements of  $G_2$  again an element of  $G_2$ ?

(2) Is the product of two elements of  $G_4$  again an element of  $G_4$ ?

**Solution 2.8.** (1) Let  $a_1^2 + b_1^2, a_2^2 + b_2^2 \in G_2$  for  $a_1, a_2, b_1, b_2 \in F$

$$(a_1^2 + b_1^2)(a_2^2 + b_2^2) = a_1^2 b_1^2 + a_1^2 b_2^2 + a_2^2 b_1^2 + a_2^2 b_2^2 = (a_1 b_1)^2 + (a_1 b_2)^2 + (a_2 b_1)^2 + (a_2 b_2)^2$$

$$a_1 b_1, a_2 b_1, a_1 b_2, a_2 b_2 \in F$$

So the product of two elements in  $G_2$  is the sum of four elements and hence it is an element of  $G_4$

(2), the product of two elements in  $G_4$  would give us an element in  $G_{16}$

**Problem 2.9.** Let  $t = \sqrt[3]{2} \in \mathbb{R}$  and let  $S$  be the set of all real numbers of the form  $a + bt + ct^2$ , where  $a, b, c \in \mathbb{Q}$ . Is  $S$  a subfield of  $\mathbb{R}$ ?

**Solution 2.9.** To determine whether  $S = \{a + bt + ct^2 \mid a, b, c \in \mathbb{Q}, t = \sqrt[3]{2}\}$  is a subfield of  $\mathbb{R}$ , we need to check the following conditions:

Let  $a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3 \in \mathbb{Q}$ . Then  $a_1 + b_1t + c_1t^2, a_2 + b_2t + c_2t^2, a_3 + b_3t + c_3t^2 \in S$ .

**Commutativity of Addition:**

$$\begin{aligned} a_1 + b_1t + c_1t^2 + a_2 + b_2t + c_2t^2 &= (a_1 + a_2) + t(b_1 + b_2) + t^2(c_1 + c_2) \\ a_2 + b_2t + c_2t^2 + a_1 + b_1t + c_1t^2 &= (a_2 + a_1) + t(b_2 + b_1) + t^2(c_2 + c_1) \\ &= (a_1 + a_2) + t(b_1 + b_2) + t^2(c_1 + c_2) \end{aligned}$$

Thus,

$$a_1 + b_1t + c_1t^2 + a_2 + b_2t + c_2t^2 = a_2 + b_2t + c_2t^2 + a_1 + b_1t + c_1t^2,$$

we conclude that addition is commutative.

**Commutativity of Multiplication:**

$$\begin{aligned} (a_1 + b_1t + c_1t^2)(a_2 + b_2t + c_2t^2) &= a_1a_2 + t(a_1b_2 + a_2b_1) + t^2(c_1a_2 + c_2a_1 + b_1b_2) \\ &\quad + t^3(b_1c_2 + b_2c_1) + c_1c_2t^4 \\ (a_2 + b_2t + c_2t^2)(a_1 + b_1t + c_1t^2) &= a_2a_1 + t(a_2b_1 + a_1b_2) + t^2(c_2a_1 + c_1a_2 + b_2b_1) \\ &\quad + t^3(b_2c_1 + b_1c_2) + c_2c_1t^4 \\ &= a_1a_2 + t(a_1b_2 + a_2b_1) + t^2(c_1a_2 + c_2a_1 + b_1b_2) \\ &\quad + t^3(b_1c_2 + b_2c_1) + c_1c_2t^4. \end{aligned}$$

Multiplication is commutative.

**Associativity of Addition:**

$$\begin{aligned} ((a_1 + b_1t + c_1t^2) + (a_2 + b_2t + c_2t^2)) + (a_3 + b_3t + c_3t^2) &= ((a_1 + a_2) + t(b_1 + b_2) + t^2(c_1 + c_2)) \\ &\quad + (a_3 + b_3t + c_3t^2) \\ &= (a_1 + a_2 + a_3) + t(b_1 + b_2 + b_3) \\ &\quad + t^2(c_1 + c_2 + c_3) \\ (a_1 + b_1t + c_1t^2) + ((a_2 + b_2t + c_2t^2) + (a_3 + b_3t + c_3t^2)) &= (a_1 + b_1t + c_1t^2) + (a_2 + a_3) + t(b_2 + b_3) \\ &\quad + t^2(c_2 + c_3) \\ &= (a_1 + a_2 + a_3) + t(b_1 + b_2 + b_3) \\ &\quad + t^2(c_1 + c_2 + c_3). \end{aligned}$$

Thus, addition is associative.

**Associativity of Multiplication:**

$$\begin{aligned} ((a_1 + b_1t + c_1t^2)(a_2 + b_2t + c_2t^2))(a_3 + b_3t + c_3t^2) &= (a_1a_2 + t(a_1b_2 + a_2b_1) + t^2(a_1c_2 + a_2c_1 + b_1b_2) \\ &\quad + t^3(c_1b_2 + c_2b_1) + c_1c_2t^4)(a_3 + b_3t + c_3t^2). \end{aligned}$$

The expansion will yield terms involving products of all three variables, confirming associativity. **Additive Identity:** Let  $a_1 = 0, b_1 = 0, c_1 = 0$ . Then:

$$a_1 + b_1t + c_1t^2 = 0 \in S,$$

and

$$(a_1 + b_1t + c_1t^2) + 0 = 0 + (a_1 + b_1t + c_1t^2) = a_1 + b_1t + c_1t^2.$$

Thus, 0 is the additive identity.

**Multiplicative Identity:** Let  $a_1 = 1, b_1 = 0, c_1 = 0$ . Then:

$$(a_1 + b_1t + c_1t^2) \cdot 1 = 1 \cdot (a_1 + b_1t + c_1t^2) = a_1 + b_1t + c_1t^2.$$

Therefore, 1 is the multiplicative identity.

**Additive Inverse:** For  $a_1 + b_1t + c_1t^2 \in S$ , the additive inverse is:

$$-(a_1 + b_1t + c_1t^2).$$

Then,

$$(a_1 + b_1t + c_1t^2) + (-(a_1 + b_1t + c_1t^2)) = 0.$$

**Distributivity of Multiplication over Addition:**

$$\begin{aligned}
& (a_1 + b_1t + c_1t^2) \cdot (a_2 + b_2t + c_2t^2 + a_3 + b_3t + c_3t^2) \\
&= (a_1 + b_1t + c_1t^2)((a_2 + a_3) + (b_2 + b_3)t + (c_2 + c_3)t^2) \\
&= (a_1a_2 + a_1a_3) + t(b_1a_2 + b_1a_3 + a_1b_2 + a_1b_3) \\
&\quad + t^2(c_1a_2 + c_1a_3 + b_1b_2 + b_1b_3 + a_1c_2 + a_1c_3) \\
&\quad + t^3(c_1b_2 + c_1b_3 + b_1c_2 + b_1c_3) + t^4(c_1c_2 + c_1c_3) \\
& (a_1 + b_1t + c_1t^2)(a_2 + b_2t + c_2t^2) \\
&= a_1a_2 + t(a_1b_2 + a_2b_1) + t^2(c_1a_2 + c_2a_1 + b_1b_2) + t^3(b_1c_2 + b_2c_1) + c_1c_2t^4 \\
& (a_2 + b_2t + c_2t^2)(a_3 + b_3t + c_3t^2) \\
&= a_2a_3 + t(a_2b_3 + a_3b_2) + t^2(c_2a_3 + c_3a_2 + b_2b_3) + t^3(b_2c_3 + b_3c_2) + c_2c_3t^4 \\
& (a_1 + b_1t + c_1t^2)(a_2 + b_2t + c_2t^2) + (a_2 + b_2t + c_2t^2)(a_3 + b_3t + c_3t^2) \\
&= (a_1a_2 + t(a_1b_2 + a_2b_1) + t^2(c_1a_2 + c_2a_1 + b_1b_2) + t^3(b_1c_2 + b_2c_1) + c_1c_2t^4) \\
&\quad + (a_2a_3 + t(a_2b_3 + a_3b_2) + t^2(c_2a_3 + c_3a_2 + b_2b_3) + t^3(b_2c_3 + b_3c_2) + c_2c_3t^4) \\
&= (a_1a_2 + a_1a_3) + t(b_1a_2 + b_1a_3 + a_1b_2 + a_1b_3) + \\
&\quad t^2(c_1a_2 + c_1a_3 + b_1b_2 + b_1b_3 + a_1c_2 + a_1c_3) + t^3(c_1b_2 + c_1b_3 + b_1c_2 + b_1c_3) + t^4(c_1c_2 + c_1c_3) \\
& (a_1 + b_1t + c_1t^2)(a_2 + b_2t + c_2t^2 + a_3 + b_3t + c_3t^2) \\
&= (a_1 + b_1t + c_1t^2)(a_2 + b_2t + c_2t^2) + (a_2 + b_2t + c_2t^2)(a_3 + b_3t + c_3t^2)
\end{aligned}$$

Therefore, multiplication is distributive over addition in  $S$ . Since  $F$  satisfies the properties of associativity of addition, the existence of additive and multiplicative identities, and the existence of additive and multiplicative inverses, we conclude that  $F$  is a field.

**Definition 2.6.** A permutation of a set  $A$  is simply a bijection from  $A$  to itself.

**Problem 2.10.** Let  $F$  be a field. Show that the function  $a \rightarrow a^{-1}$  is a permutation of  $F \setminus \{0_F\}$ .

**Solution 2.10.** To prove that the function  $\varphi : F \setminus \{0_F\} \rightarrow F \setminus \{0_F\}$  defined by  $\varphi(a) = a^{-1}$ , where  $F$  is a field, is a permutation of  $F \setminus \{0_F\}$ , we must show that  $\varphi$  is a bijection. First, we will prove that the function is well defined. Assume that  $a_1 = a_2$  for some  $a_1, a_2 \in F \setminus \{0_F\}$ .

$$\begin{aligned}
a_1 &= a_2 \\
1 &= a_2a_1^{-1} \\
a_2^{-1} &= a_1^{-1} \\
\varphi(a_1) &= \varphi(a_2)
\end{aligned}$$

So the function is well defined.

Injective:

Assume that  $\varphi(a_1) = \varphi(a_2)$  for some  $a_1, a_2 \in F \setminus \{0_F\}$ . So,  $a_2^{-1} = a_1^{-1} \implies 1 = a_2a_1^{-1} \implies a_1 = a_2$ .

Surjective:

Take an element  $a \in F \setminus \{0_F\}$  its preimage is  $a^{-1}$  as

$$\varphi(a^{-1}) = (a^{-1})^{-1} = a$$

So the function is bijective, thus a permutation.

**Problem 2.11.** Show that every  $z \in C$  satisfies  $z^4 + 4 = (z - 1 - i)(z - 1 + i)(z + 1 + i)(z + 1 - i)$

**Solution 2.11.** We aim to show that the equation  $z^4 + 4 = (z - 1 - i)(z - 1 + i)(z + 1 + i)(z + 1 - i)$  holds for all  $z \in \mathbb{C}$

First, we consider the R.H.S

$$\begin{aligned}
 (z - 1 - i)(z - 1 + i)(z + 1 + i)(z + 1 - i) &= [(z - 1)^2 - i^2][(z + 1)^2 - i^2] \\
 &= [z^2 + 1 - 2z - (-1)][z^2 + 1 + 2z - (-1)] \\
 &= [z^2 + 1 - 2z + 1][z^2 + 1 + 2z + 1] \\
 &= [z^2 + 2 - 2z][z^2 + 2 + 2z] \\
 &= [(z^2 + 2)^2 - (2z)^2] \\
 &= [z^4 + 4 + 4z^2 - 4z^2] \\
 &= z^4 + 4
 \end{aligned}$$

**Problem 2.12.** In each of the following, find the set of all complex numbers  $z = a + bi$  satisfying the given relation. Note that this set may be empty or may be all of  $\mathbb{C}$ . Justify your result in each case.

- (a)  $z^2 = \frac{1}{2}(1 + i\sqrt{3})$
- (b)  $(\sqrt{2})|z| \geq |a| + |b|$
- (c)  $|z| + z = 2 + i$
- (d)  $z^4 = 2 - (\sqrt{12})i$
- (e)  $z^4 = -4$

**Solution 2.12.** (a)  $z^2 = \frac{1}{2}(1 + i\sqrt{3}) = \frac{1}{4}(2 + 2i\sqrt{3}) = \frac{(\sqrt{3}+i)^2}{4}$

$$z = \frac{\sqrt{3} + i}{2}, -\frac{\sqrt{3} + i}{2}$$

(b) Square both sides to get

$$\begin{aligned}
 2|z|^2 &\geq (|a| + |b|)^2 \\
 2(|a|^2 + |b|^2) &\geq |a|^2 + |b|^2 + 2|a||b| \\
 |a|^2 + |b|^2 &\geq 2|a||b| \\
 (|a| - |b|)^2 &\geq 0
 \end{aligned}$$

Thus all the complex number satisfy the relation. The set is  $\mathbb{C}$ .

(c) Let  $z = a + ib$  then  $|z| = \sqrt{a^2 + b^2}$

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

Comparing constant term and coefficient of  $i$ ,

$$b = 1$$

$$\sqrt{a^2 + b^2} + a = 2$$

$$\sqrt{a^2 + 1} + a = 2$$

$$\sqrt{a^2 + 1} = 2 - a$$

Squaring both sides of the equation,

$$a^2 + 1 = (2 - a)^2 = 4 + a^2 - 4a$$

$$4a = 3$$

So,  $a = 3/4, b = 1$

Therefore  $z = \frac{3}{4} + i$

$$(d) \ z^4 = 2 - (\sqrt{12})i = 2 - 2\sqrt{3}i = (\sqrt{3} - i)^2$$

$$z^2 = \sqrt{3} - i, -(\sqrt{3} - i) = i - \sqrt{3}$$

$$\text{Case 1) } z^2 = \sqrt{3} - i$$

Let  $z = a + ib$ , then we have

$$(a + ib)^2 = \sqrt{3} - i$$

$$a^2 - b^2 + 2iab = \sqrt{3} - i$$

$$\text{So, } a^2 - b^2 = \sqrt{3} \text{ and } 2ab = -1. \text{ So } a = -\frac{1}{2b}$$

$$a^2 - b^2 = \sqrt{3}$$

$$\left(-\frac{1}{2b}\right)^2 - b^2 = \sqrt{3}$$

$$\frac{1}{4b^2} - b^2 = \sqrt{3}$$

$$1 - 4b^4 = 4\sqrt{3}b^2$$

$$4b^4 + 4\sqrt{3}b^2 - 1 = 0$$

Let  $m = b^2$ , then the equation will be

$$4m^2 + 4\sqrt{3}m - 1 = 0$$

$$\begin{aligned} m &= \frac{-4\sqrt{3} \pm \sqrt{(4\sqrt{3})^2 - 4(4)(-1)}}{8} \\ &= \frac{-4\sqrt{3} \pm \sqrt{48 + 16}}{8} \\ &= \frac{-4\sqrt{3} \pm \sqrt{64}}{8} \\ &= \frac{-4\sqrt{3} \pm 8}{8} \\ &= \frac{-\sqrt{3} \pm 2}{2} \end{aligned}$$

Since  $m = b^2$ ,  $m$  is positive,

$$\text{So } m = \frac{-\sqrt{3} + 2}{2}$$

$$b^2 = \frac{-2\sqrt{3} + 4}{4}$$

$$b^2 = \frac{(\sqrt{3} - 1)^2}{4}$$

$$b = \frac{\sqrt{3} - 1}{2} \text{ or } b = \frac{1 - \sqrt{3}}{2}$$

$$\text{As, } a = -\frac{1}{2b}$$

$$a = \frac{-1 - \sqrt{3}}{2} \text{ or } a = \frac{1 + \sqrt{3}}{2}$$

Thus  $z$  is ,

$$z = \frac{1 - \sqrt{3}}{2} + i \left( \frac{1 - \sqrt{3}}{2} \right) \text{ or } z = \frac{-1 - \sqrt{3}}{2} + i \left( \frac{\sqrt{3} - 1}{2} \right)$$

- (e)  $z^4 = -4$ , So  $z^2 = 2i, -2i$   
 Let  $z = a + ib$  where  $a, b \in \mathbb{R}$   
 Case 1: When  $z^2 = 2i$

$$\begin{aligned} z^2 &= 2i \\ (a + ib)^2 &= 2i \\ a^2 - b^2 + 2abi &= 2i. \end{aligned}$$

So  $a^2 - b^2 = 0$  and  $2ab = 2$ , we have  $a^2 = b^2$  and  $ab = 1$ . This implies,  $a = b = 1$  or  $a = b = -1$ .  
 So  $z = 1 + i, -1 - i$   
 Case 2: When  $z^2 = -2i$

$$\begin{aligned} z^2 &= -2i \\ (a + ib)^2 &= -2i \\ a^2 - b^2 + 2abi &= -2i. \end{aligned}$$

So  $a^2 - b^2 = 0$  and  $2ab = -2$ , we have  $a^2 = b^2$  and  $ab = -1$ . This implies  $a = 1, b = -1$  or  $a = -1, b = 1$ . So  $z = 1 - i, -1 + i$   
 Therefore  $z = 1 + i, 1 - i, -1 - i, -1 + i$

**Problem 2.13.** Let  $z_1, z_2$  and  $z_3$  be complex numbers satisfying the condition that  $|z_i| = 1$  for  $i = 1, 2, 3$ . Show that  $|z_1 z_2 + z_1 z_3 + z_2 z_3| = |z_1 + z_2 + z_3|$ .

**Solution 2.13.** Let  $z_1, z_2, z_3 \in \mathbb{C}$  such that  $|z_1| = |z_2| = |z_3| = 1$ . We want to show that:

$$|z_1 z_2 + z_1 z_3 + z_2 z_3| = |z_1 + z_2 + z_3|.$$

Divide the R.H.S by  $|z_1 z_2 z_3| = 1$

$$|z_1 z_2 + z_1 z_3 + z_2 z_3| = \frac{|z_1 z_2 + z_1 z_3 + z_2 z_3|}{|z_1 z_2 z_3|} = \left| \frac{1}{z_1} + \frac{1}{z_2} + \frac{1}{z_3} \right|$$

So,

$$\begin{aligned} |z_1 z_2 + z_1 z_3 + z_2 z_3|^2 &= \left| \frac{1}{z_1} + \frac{1}{z_2} + \frac{1}{z_3} \right|^2 \\ &= \left( \frac{1}{z_1} + \frac{1}{z_2} + \frac{1}{z_3} \right) \left( \frac{1}{\bar{z}_1} + \frac{1}{\bar{z}_2} + \frac{1}{\bar{z}_3} \right) \end{aligned}$$

$$|z_1 + z_2 + z_3|^2 = (z_1 + z_2 + z_3)(\bar{z}_1 + \bar{z}_2 + \bar{z}_3)$$

As,  $|z_1| = 1$ ,  $z_1 = \frac{1}{\bar{z}_1}$  and  $\bar{z}_1 = \frac{1}{z_1}$ , So ,

$$|z_1 + z_2 + z_3|^2 = \left( \frac{1}{\bar{z}_1} + \frac{1}{\bar{z}_2} + \frac{1}{\bar{z}_3} \right) \left( \frac{1}{z_1} + \frac{1}{z_2} + \frac{1}{z_3} \right)$$

**Problem 2.14.** For any  $z_1, z_2 \in \mathbb{C}$ , show that  $|z_1|^2 + |z_2|^2 - z_1 \bar{z}_2 - \bar{z}_1 z_2 = |z_1 - z_2|^2$ .

**Solution 2.14.** Note that

$$\begin{aligned} |z_1 - z_2|^2 &= (z_1 - z_2)(\bar{z}_1 - \bar{z}_2) \\ &= z_1 \bar{z}_1 - z_1 \bar{z}_2 - \bar{z}_1 z_2 + z_2 \bar{z}_2 \\ &= |z_1|^2 + |z_2|^2 - z_1 \bar{z}_2 - \bar{z}_1 z_2 \end{aligned}$$

**Problem 2.15.** Show that  $|z + 1| \leq |z + 1|^2 + |z|$  for all  $z \in \mathbb{C}$ .

**Solution 2.15.** Now,

$$\begin{aligned} |z+1| &\leq |z+1|^2 + |z| \\ |a+ib+1| &\leq |a+ib+1|^2 + |a+ib| \\ \sqrt{(a+1)^2 + b^2} &\leq (a+1)^2 + b^2 + \sqrt{a^2 + b^2} \end{aligned}$$

square on both sides, we get

$$\begin{aligned} (a+1)^2 + b^2 &\leq ((a+1)^2 + b^2 + \sqrt{a^2 + b^2})^2 \\ &\leq ((a+1)^2 + b^2)^2 + a^2 + b^2 + 2((a+1)^2 + b^2 + \sqrt{a^2 + b^2}) \\ ((a+1)^2 + b^2)^2 + a^2 + 2b^2 + 2(a+1)^2 + 2\sqrt{a^2 + b^2} &\geq 0. \end{aligned}$$

This is true for all  $a, b \in \mathbb{R}$ .

**Problem 2.16.** If  $z \in \mathbb{C}$ , find  $w \in \mathbb{C}$  satisfying  $w^2 = z$ .

**Solution 2.16.** Let  $z = a + ib$ . We will find  $w = x + iy$  such that  $w^2 = z$  we have

$$\begin{aligned} a + ib &= (x + iy)^2 \\ \Rightarrow a + ib &= x^2 + (iy)^2 + i2xy \\ \Rightarrow a + ib &= x^2 - y^2 + 2ixy \text{ [Because } i^2 = -1] \end{aligned}$$

Comparing real and imaginary parts of the above equation, we have

$$a = x^2 - y^2, b = 2xy$$

We know that  $(x^2 + y^2)^2 = (x^2 - y^2)^2 + 4x^2y^2$

$$\Rightarrow (x^2 + y^2)^2 = a^2 + b^2$$

$$\Rightarrow x^2 + y^2 = \sqrt{a^2 + b^2}$$

Now, we have  $x^2 + y^2 = \sqrt{a^2 + b^2}$  and  $a = x^2 - y^2$ . Solving for these two values, we have

$$x = \pm \sqrt{\left[\sqrt{a^2 + b^2} + a\right] / 2} \text{ and } y = \pm \sqrt{\left[\sqrt{a^2 + b^2} - a\right] / 2}$$

**Problem 2.17.** Define new operations  $\circ$  and  $\diamond$  on  $\mathbb{C}$  by setting  $y \circ z = |y|z$  and  $y \diamond z = \frac{yz}{|y|}$  for all  $y, z \in \mathbb{C}$ . Is it true that  $w \diamond (y \circ z) = (w \diamond y) \circ (w \diamond z)$  and  $w \circ (y \diamond z) = (w \circ y) \diamond (w \circ z)$  for all  $w, y, z \in \mathbb{C}$ ?

**Solution 2.17.**  $y \circ z = |y|z$  and  $y \diamond z = \frac{yz}{|y|}$

$$w \diamond (y \circ z) = w \diamond (|y|z) = \frac{w|y|z}{|w|}$$

$$w \diamond y = \frac{wy}{|w|} \text{ and } w \diamond z = \frac{wz}{|w|}$$

$$\begin{aligned} (w \diamond y) \circ (w \diamond z) &= \frac{wy}{|w|} \circ \frac{wz}{|w|} \\ &= \frac{|wy|}{|w|} \times \frac{wz}{|w|} = \frac{|w||y|wz}{|w||w|} = \frac{|y|wz}{|w|} \end{aligned}$$

So,  $w \diamond (y \circ z) = (w \diamond y) \circ (w \diamond z)$

$$w \circ (y \diamond z) = w \circ \left(\frac{yz}{|y|}\right) = \frac{|w|yz}{|y|}$$

$$(w \circ y) \diamond (w \circ z) = (|w|y) \diamond (|w|z) = \frac{|w|y|w|z}{|w||z|} = \frac{y|w|z}{|y|}$$

So,  $w \circ (y \diamond z) = (w \circ y) \diamond (w \circ z)$

**Problem 2.18.** Let  $0 \neq z \in \mathbb{C}$ . Show that there are infinitely-many complex numbers  $y$  satisfying the condition  $y\bar{y} = z\bar{z}$ .

**Solution 2.18.** Now,  $y\bar{y} = |y|^2$  and  $z\bar{z} = |z|^2$ . So,  $|y|^2 = |z|^2$ .  $|y| = |z|$ . Clearly there are infinitely many complex numbers of a given modulus ( $|z|$ ).  $y = |z|e^{i\theta}$  where  $\theta$  is angle with positive  $x$  axis. And there are infinite values of  $\theta$

**Problem 2.19.** (Abel's inequality) Let  $z_1, \dots, z_n$  be a list of complex numbers and, for each  $1 \leq k \leq n$ , let  $s_k = \sum_{i=1}^k z_i$ . For real numbers  $a_1, \dots, a_n$  satisfying  $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$ , show that  $|\sum_{i=1}^n a_i z_i| \leq a_1 (\max_{1 \leq k \leq n} |s_k|)$ .

**Solution 2.19.** We have  $z_i = s_i - s_{i-1}$

$$\sum_{i=1}^n a_i z_i = \sum_{i=1}^n a_i (s_i - s_{i-1})$$

here, we take  $s_0 = 0$ , simplify this;

$$\begin{aligned} \sum_{i=1}^n a_i z_i &= \sum_{i=1}^n a_i s_i - \sum_{i=1}^n a_i s_{i-1} \\ &= \sum_{i=1}^n a_i s_i - \sum_{i=1}^{n-1} a_{i+1} s_i \end{aligned}$$

Here, we use  $a_{n+1} = 0$ . Thus,

$$\sum_{i=1}^n a_i z_i = \sum_{i=1}^n (a_i - a_{i+1}) s_i$$

Take the absolute value of both sides:

$$|\sum_{i=1}^n a_i z_i| = |\sum_{i=1}^n (a_i - a_{i+1}) s_i|$$

By the triangle inequality:

$$|\sum_{i=1}^n a_i z_i| \leq \sum_{i=1}^n |(a_i - a_{i+1})| |s_i|.$$

Since  $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$ , it follows that  $a_i - a_{i+1} \geq 0$  for all  $1 \leq i \leq n$ . Therefore:

$$|\sum_{i=1}^n a_i z_i| \leq \max_{1 \leq k \leq n} |s_k| \sum_{i=1}^n (a_i - a_{i+1})$$

Observe that:

$$\sum_{i=1}^n (a_i - a_{i+1}) = a_1 - a_{n+1}$$

Since  $a_{n+1} = 0$ , we have:

$$\sum_{i=1}^n (a_i - a_{i+1}) = a_1$$

Thus:

$$|\sum_{i=1}^n a_i z_i| \leq a_1 \max_{1 \leq k \leq n} |s_k|$$

**Problem 2.20.** If  $p$  is a prime positive integer, find all subfields of  $GF(p)$ .

**Solution 2.20.** Note that  $GF(p)$  is just a abelian group with  $+$ ,  $\cdot$ . Now, a group with a prime order will never have a non proper subgroup. Only subfield is identity.

**Problem 2.21.** Find  $10^{-1}$  in  $GF(33)$ .

**Solution 2.21.** Assume  $x = 10^{-1}$ , we now have

$$10x \equiv 1 \pmod{33}$$

$10x = 33k + 1$ . Now,  $10x \equiv 1 \pmod{3}$  so we have  $x \equiv 1 \pmod{3}$ . Also,  $10x \equiv 1 \pmod{11}$  so we have  $x \equiv 10 \pmod{11}$ . Solving  $x \equiv 1 \pmod{3}$  and  $x \equiv 10 \pmod{11}$  by Chinese remainder theorem, we get  $x \equiv 10 \pmod{33}$

**Problem 2.22.** Find elements  $c, d \neq \pm 1$  in the field  $\mathbb{Q}(\sqrt{5})$  satisfying  $cd = 19$ .

**Solution 2.22.** Let  $c = a + b\sqrt{5} \in \mathbb{Q}(\sqrt{5})$ , then also  $d = a - b\sqrt{5} \in \mathbb{Q}(\sqrt{5})$

$$cd = (a + b\sqrt{5})(a - b\sqrt{5}) = a^2 - (b\sqrt{5})^2 = a^2 - 5b^2 = 19$$

When  $a = 8$ ,

$$a^2 - 5b^2 = 8^2 - 5b^2 = 64 - 5b^2 = 19$$

$$5b^2 = 45$$

$$b^2 = 9$$

$$b = \pm 3$$

So,  $c = 8 + 3\sqrt{5}$  and  $d = 8 - 3\sqrt{5}$ .

**Problem 2.23.** Let  $F$  be the set of all real numbers of the form

$$a + b(\sqrt[3]{5}) + c(\sqrt[3]{5})^2,$$

where  $a, b, c \in \mathbb{Q}$ . Is  $F$  a subfield of  $\mathbb{R}$ ?

**Solution 2.23.** Clearly it is closed under addition. Now, we need to verify that it is closed under multiplication. Given two elements from  $F$ :

$$x_1 = a_1 + b_1\sqrt[3]{5} + c_1(\sqrt[3]{5})^2$$

$$x_2 = a_2 + b_2\sqrt[3]{5} + c_2(\sqrt[3]{5})^2,$$

Consider  $x_1x_2$ :

$$x_1x_2 = (a_1 + b_1\sqrt[3]{5} + c_1(\sqrt[3]{5})^2)(a_2 + b_2\sqrt[3]{5} + c_2(\sqrt[3]{5})^2).$$

Expanding this product:

$$x_1x_2 = a_1a_2 + a_1b_2\sqrt[3]{5} + a_1c_2(\sqrt[3]{5})^2 + b_1a_2\sqrt[3]{5} + b_1b_2(\sqrt[3]{5})^2 + b_1c_2(\sqrt[3]{5})^3 + c_1a_2(\sqrt[3]{5})^2 + c_1b_2(\sqrt[3]{5})^3 + c_1c_2(\sqrt[3]{5})^4.$$

We know  $\sqrt[3]{5}^3 = 5$  and  $\sqrt[3]{5}^4 = 5\sqrt[3]{5}$ . Therefore, we can substitute these back into the expression:

$$x_1x_2 = a_1a_2 + (a_1b_2 + b_1a_2 + 5(c_1b_2 + b_1c_2))\sqrt[3]{5} + (a_1c_2 + b_1b_2 + c_1a_2)(\sqrt[3]{5})^2 + (c_1c_2 \cdot 5) + (b_1c_2 + c_1b_2)5\sqrt[3]{5}.$$

$$x_1x_2 = (a_1a_2 + 5c_1c_2) + (a_1b_2 + b_1a_2 + b_1c_2 + c_1b_2)\sqrt[3]{5} + (a_1c_2 + b_1b_2 + c_1a_2)(\sqrt[3]{5})^2.$$

Now let:  $d = a_1a_2 + 5c_1c_2$ ,  $e = a_1b_2 + b_1a_2 + b_1c_2 + c_1b_2$ ,  $f = a_1c_2 + b_1b_2 + c_1a_2$ .

Thus, we can write:

$$x_1x_2 = d + e\sqrt[3]{5} + f(\sqrt[3]{5})^2,$$

where  $d, e, f \in \mathbb{Q}$ . We can conclude that  $F$  is closed under multiplication. To find the inverse of the element

$$x = a + b\sqrt[3]{5} + c(\sqrt[3]{5})^2,$$

we want to find  $y$  in the form

$$y = d + e\sqrt[3]{5} + f(\sqrt[3]{5})^2$$

such that

$$xy = 1.$$

Expanding the product  $xy$ :

$$xy = (a + b\sqrt[3]{5} + c(\sqrt[3]{5})^2)(d + e\sqrt[3]{5} + f(\sqrt[3]{5})^2).$$

Using distributive properties, we compute this:

$$\begin{aligned} xy &= ad + ae\sqrt[3]{5} + af(\sqrt[3]{5})^2 + bd\sqrt[3]{5} + be(\sqrt[3]{5})^2 + bf(\sqrt[3]{5})^3 + cd(\sqrt[3]{5})^2 + ce(\sqrt[3]{5})^3 + cf(\sqrt[3]{5})^4. \\ xy &= ad + (ae + bd + 5cf + 5ce)\sqrt[3]{5} + (af + be + cd)(\sqrt[3]{5})^2 + 5(bf + ce). \end{aligned}$$

comparing coefficients:

1. Constant term:

$$ad + 5bc = 1$$

2. Coefficient of  $\sqrt[3]{5}$ :

$$ae + bd + 5(cf) = 0$$

3. Coefficient of  $(\sqrt[3]{5})^2$ :

$$af + be + cd = 0.$$

This gives us a system of equations:

$$(13) \quad ad + 5bc = 1; ae + bd + 5cf = 0; af + be + cd = 0$$

Three equations and three variables so we can solve this system of equations to find the multiplicative inverse of  $x$  in  $F$ . Therefore,  $F$  is indeed a subfield of  $\mathbb{R}$ .

**Problem 2.24.** Let  $p$  be a prime positive integer and let  $a \in GF(p)$ . Does there necessarily exist an element  $b$  of  $GF(p)$  satisfying  $b^2 = a$ ?

**Solution 2.24.** Consider  $p = 5$  and  $a = 3$ . It is easy to verify that

$$b^2 \not\equiv 3 \pmod{5}$$

has no solutions. So there do not necessarily exists an element  $b$  satisfying  $b^2 = a$ .

**Problem 2.25.** Let  $F = GF(11)$  and let  $G = F \times F$ . Define operations of addition and multiplication on  $G$  by setting  $(a, b) + (c, d) = (a + c, b + d)$  and  $(a, b)(c, d) = (ac + 7bd, ad + bc)$ . Do these operations define the structure of a field on  $G$ ?

**Solution 2.25.** Obviously  $(G, +)$  is abelian group. Now Commutativity of Multiplication:

$$(a, b) \cdot (c, d) = (ac + 7bd, ad + bc)$$

$$(c, d) \cdot (a, b) = (ac + 7bd, bc + ad)$$

So multiplication is commutative. Now we will check associativity of Multiplication.

$$\begin{aligned} ((a, b) \cdot (c, d)) \cdot (e, f) &= (ac + 7bd, ad + bc) \cdot (e, f) = ((ac + 7bd)e + 7(ad + bc)f, (ac + 7bd)f + (ad + bc)e) \\ (ac + 7bd, ad + bc) \cdot (e, f) &= (ace + 7bde + 7adf + 7bcf, ade + bce + acf + 7bdf) \end{aligned}$$

Now

$$\begin{aligned} (a, b) \cdot ((c, d) \cdot (e, f)) &= (a, b) \cdot (ce + 7df, cf + de) = ((ce + 7df)a + 7b(cf + de), (cf + de)a + b(ce + 7df)) \\ (a, b) \cdot ((c, d) \cdot (e, f)) &= (ace + 7bde + 7adf + 7bcf, ade + bce + acf + 7bdf) \end{aligned}$$

Thus, multiplication is associative in  $G$ .

Multiplicative identity is  $(1, 0)$  as

$$(a, b) \cdot (1, 0) = (a \cdot 1 + 7b \cdot 0, a \cdot 0 + b \cdot 1) = (a, b)$$

Multiplicative inverse

$$\begin{aligned}(a, b)(c, d) &= (1, 0) \\ (ac + 7bd, ad + bc) &= (1, 0) \\ ac + 7bd &= 1 \\ ad + bc &= 1\end{aligned}$$

Solving these equations for  $c$  and  $d$  (given  $a \neq 0$  or  $b \neq 0$ ) shows that  $(c, d)$  exists, and hence, every nonzero element has a multiplicative inverse.

Distributivity of multiplication over addition:

$$\begin{aligned}(a, b) \cdot ((c, d) + (e, f)) &= (a, b) \cdot (c + e, d + f) \\ &= (ac + ae + 7bd + 7bf, ad + af + bc + be) \\ ((a, b) \cdot (c, d)) + ((a, b) \cdot (e, f)) &= (ac + 7bd, ad + bc) + (ae + 7bf, af + be) \\ &= (ac + ae + 7bd + 7bf, ad + af + bc + be)\end{aligned}$$

Therefore, multiplication is distributive over addition. Since  $G$  satisfies the properties of associativity of addition, the existence of additive and multiplicative identities, and the existence of additive and multiplicative inverses, we conclude that  $G$  is a field.

**Proposition 2.1.** [Dummit and Foote] A finite subgroup of the multiplicative group of a field is cyclic. In particular, if  $F$  is a finite field, then the multiplicative group  $F^\times$  of nonzero elements of  $F$  is a cyclic group.

PROOF. We give a proof of this result using the Fundamental Theorem of Finitely Generated Abelian Groups (Theorem 3 in Section 5.2) [Dummit and Foote]. A more number-theoretic proof is outlined in the exercises, or Proposition 5 in Section 6.1 [Dummit and Foote] may be used in place of the Fundamental Theorem. By the Fundamental Theorem, the finite subgroup can be written as the direct product of cyclic groups

$$\mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$$

where  $n_k | n_{k-1} | \cdots | n_2 | n_1$ . In general, if  $G$  is a cyclic group and  $d || |G|$  then  $G$  contains precisely  $d$  elements of order dividing  $d$ . Since  $n_k$  divides the order of each of the cyclic groups in the direct product, it follows that each direct factor contains  $n_k$  elements of order dividing  $n_k$ . If  $k$  were greater than 1, there would therefore be a total of more than  $n_k$  such elements. But then there would be more than  $n_k$  roots of the polynomial  $x^{n_k} - 1$  in the field  $F$ , contradicting Proposition 17 [Dummit and Foote]. Hence  $k = 1$  and the group is cyclic.  $\square$

**Problem 2.26.** Is  $\mathbb{Z}/(6)$  an integral domain?

**Solution 2.26.** Now consider the elements  $[2]_6$  and  $[3]_6$  (i.e equivalence class of 2 and 3). Note that  $[2]_6[3]_6 = [6]_6 = [0]_6$ , but  $[2]_6 \neq 0$  and  $[3]_6 \neq 0$ . Therefore,  $\mathbb{Z}/(6)$  is not an integral domain.

**Problem 2.27.** Let  $F = \{a + b\sqrt{5} \in \mathbb{Q}(\sqrt{5}) | a, b \in \mathbb{Z}\}$ . Is  $F$  an integral domain?

**Solution 2.27.** The inverse of  $1 + \sqrt{5}$  is  $\frac{1}{1+\sqrt{5}} = \frac{\sqrt{5}-1}{24}$ . Clearly  $\frac{1}{24}, \frac{-1}{24}$  are not integers. Therefore, the inverse is not in  $F$ . So, it is not an integral domain.

**Problem 2.28.** Let  $F$  be an integral domain and let  $a \in F$  satisfy  $a^2 = a$ . Show that  $a = 0$  or  $a = 1$ .

**Solution 2.28.** Now,  $a^2 = a$ .  $a^2 - a = 0$ ,  $a(a - 1) = 0$ . As  $F$  is a integral domain,  $a = 0$  or  $a - 1 = 0$ . i.e  $a = 0$  or  $a = 1$

**Problem 2.29.** Let  $a$  be a nonzero element in an integral domain  $F$ . If  $b \neq c$  are distinct elements of  $F$ , show that  $ab \neq ac$ .

**Solution 2.29.** Assume for contradiction.  $ab = ac$ . So  $a(b - c) = 0$  and as  $F$  is a integral domain,  $a = 0$  or  $b = c$ . But  $a \neq 0$  and  $b \neq c$ . This is a contradiction.

**Problem 2.30.** Show that every integral domain having only finitely-many elements is a field.

**Solution 2.30.** Let  $R$  be a finite integral domain and let  $a$  be a nonzero element of  $(R)$ . By the cancellation law, the map  $\varphi : (x \mapsto ax)$  is an injective. As  $(R)$  is finite, this map is also surjective. So for any  $1 \in R$  there is a  $b \in F$  such that  $\varphi(b) = 1$ , i.e.,  $ab = 1$ .  $a$  is a unit in  $(R)$ .  $(R)$  is a field.

**Problem 2.31.** A field  $F$  is orderable if and only if there exists a subset  $P$  closed under addition and multiplication such that for each  $a \in F$  precisely one of the following conditions holds: (i)  $a = 0$ ; (ii)  $a \in P$ ; (iii)  $-a \in P$ . Show that  $\text{GF}(5)$  is not orderable.

**Solution 2.31.** Assume for sake of contradiction  $\text{GF}(5)$  is orderable. If  $-1 \in P$ , observe that  $(-1)(-1) = 1$  therefore if  $-1 \in P$ , both  $x, -x \in P$ , a contradiction. So  $-1 = 4 \notin P$ .

Claim :  $2 \notin P$ . Assume  $2 \in P$ , observe that  $(2)(2) = 4 = -1 \in P$  contradiction.

Claim :  $-2 \notin P$ . Assume  $-2 \in P$ , observe that  $(-2)(-2) = 4 = -1 \in P$  contradiction. So,  $2 \neq 0, 2 \notin P$  and  $-2 \notin P$

**Problem 2.32.** If  $a$  is a nonzero element of a field  $F$ , show that  $(-a)^{-1} = -(a^{-1})$ .

**Solution 2.32.** Now,  $(-a)(-a^{-1}) = aa^{-1} = 1$  and  $(-a^{-1})(-a) = a^{-1}a = 1$ . So  $-(a^{-1})$  is the inverse of  $(-a)$

**Problem 2.33.** Exercise 43 Is every nonzero element of the field  $\text{GF}(5)$  in the form  $2^i$  for some positive integer  $i$ ? What happens in the case of the field  $\text{GF}(7)$ ?

**Solution 2.33.** No, 3 is not power of 2. In  $\text{GF}(7)$  3, 5, 6, 7 are not powers of 2



## Sets Relations, Functions, and Partitions

### 1. Sets, Functions, Equivalence Relation and Partition

This appendix is part of the 'Introduction to Proofs' at AMMOC.

**Definition 1.1 (Partition of Set).** Let  $X$  be a set. Let  $\mathcal{C} = \{X_i | i \in \mathcal{I}, X_i \subset X\}$  be some collection of non empty subsets of set  $X$ . We say  $\mathcal{C}$  defines a **partition** of  $X$  if followings are true

- for every  $i, j \in \mathcal{I}$  whenever  $i \neq j$  we must have  $X_i \cap X_j = \Phi$
- $\cup_{i \in \mathcal{I}} X_i = X$

**Definition 1.2 (Equivalence Classes).** Let  $X$  be a set. Let  $\sim$  be an equivalence relation on  $X$ . Define **equivalence class** of an element  $x_o \in X$  as collection :  $[x_o]_{\sim} = \{x \in X | x \sim x_o\}$ . Since  $\sim$  is an equivalence, and thus reflexive relation on  $X$ . Therefore, for each  $x \in X$ ,  $x \sim x$ , and hence  $x_o \in X$ ,  $x_o \in [x_o]_{\sim}$ . So, equivalence class of any element is never an empty set.

**Theorem 1.1 (Equivalence Classes are mutually disjoint).** Let  $x_1, x_2 \in X$ . Let  $[x_1]_{\sim}$  and  $[x_2]_{\sim}$  be respective equivalence classes. We claim, if  $x' \in [x_1]_{\sim} \cap [x_2]_{\sim}$  then  $[x_1]_{\sim} = [x_2]_{\sim}$

PROOF. Let  $x' \in [x_1]_{\sim} \cap [x_2]_{\sim}$ . So, by definition 1.2 of equivalence class,  $x' \sim x_1$  and  $x' \sim x_2$ . This means  $x_1 \sim x' \sim x_2 \implies x_1 \sim x_2 \implies [x_1]_{\sim} \subset [x_2]_{\sim}$ . Verbatim argument also implies  $[x_2]_{\sim} \subset [x_1]_{\sim}$ . So, if one element is common in two equivalence classes then they coincide. Which, contra-positively means, if two equivalence classes do not coincide then they have no element in common at all. That means any two equivalence classes are mutually disjoint.  $\square$

**Theorem 1.2 (Equivalence classes form a set theoretical cover).** Let  $X$  be a set. Let  $\sim$  be an equivalence relation on  $X$ . Then  $X = \cup_{i \in \mathcal{I}} [x_i]_{\sim}$ .

PROOF. Note that by definition 1.2, for each  $x \in X$  and for each  $i \in \mathcal{I}$ ,  $[x_i]_{\sim} \subset X$ . So,  $\cup_{i \in \mathcal{I}} [x_i]_{\sim} \subset X$ . For other side inclusion, note that for each  $x \in X$ ,  $x \in [x]_{\sim}$  therefore,  $X \subset \cup_{i \in \mathcal{I}} [x_i]_{\sim}$ . This closes the proof.  $\square$

So, theorem 1.2 establishes that corresponding to each equivalent relation we have a partition of the underlying set. In what follows we will prove that converse is true as well. That is "given a partition there exist a corresponding equivalent relation on the given set".

**Theorem 1.3 (Partition induces an equivalence relation).** Let  $\mathcal{P} = \{X_j | j \in \mathcal{J}, X_j \subset X\}$  be a partition of the set  $X$ . We claim that this partition induces an equivalence relation on set  $X$ .

PROOF. Define a relation  $\mathcal{R}$  on the set  $X$  as :  $x_1 \mathcal{R} x_2$  iff there exist a  $j \in \mathcal{J}$  such that  $x_1, x_2 \in X_j$ .

- **Reflexive.** For each  $x \in X$ , there exists exactly one  $X_j \in \mathcal{P}$  such that  $x \in X_j$ . So,  $x$  is  $\mathcal{R}$  related to itself and hence  $\mathcal{R}$  is a reflexive relation on  $X$ .
- **Symmetric.** Let  $x_1 \mathcal{R} x_2$ . This means, by definition of  $\mathcal{R}$ , that there exist a  $j \in \mathcal{J}$  and corresponding  $P_j \in \mathcal{P}$  such that  $x_1 \in P_j$  and  $x_2 \in P_j$ . However, this effectively means  $x_2 \mathcal{R} x_1$ . So,  $\mathcal{R}$  is symmetric relation on  $X$ .
- **Transitive.** Let  $x_1 \mathcal{R} x_2$ , and  $x_2 \mathcal{R} x_3$ . By definition of  $\mathcal{R}$ , there exist  $j, k \in \mathcal{J}$  and corresponding  $P_j, P_k \in \mathcal{P}$ , such that  $x_1, x_2 \in P_j$  and  $x_2, x_3 \in P_k$ . However,  $\mathcal{P}$  is a partition of  $X$  and  $x_2 \in P_j \cap P_k$ . So,  $j$  must be equal to  $k$ . That is,  $P_j = P_k$  and this gives  $x_1, x_3 \in P_j$ . Hence, the relation  $\mathcal{R}$  is transitive on  $X$ .

$\square$

**Example 1.1** (Every function defines an equivalence relation).  $f : X \rightarrow Y$  be any mapping. Define a relation  $\sim$  on  $X$  as  $x_1 \sim x_2 \iff f(x_1) = f(x_2)$ . This  $\sim$  obviously is an equivalence relation because equality implies equivalence. Though converse is not true.

What are equivalence classes of this equivalence relation? Well, all elements which are mapped to a common image form an equivalence class. That is  $[x_0]_{\sim} = \{x \in X \mid f(x) = f(x_0)\}$ .

**Example 1.2.** Every equality defines an equivalence relation. However, converse is not true.

Once can ask a basic question can we turn an equivalence into equality. The answer is yes. It is possible in certain suitable sense - equivalence becomes equality in quotient set, quotient group or say in quotient object. By this we mean that if  $\sim$  is an equivalence relation on  $X$  and  $\mathcal{C}$  is collection of equivalence classes corresponding to this relation, then  $[x]_{\sim} = [x']_{\sim}$ , (equality of equivalence classes) if and only if  $x \sim x'$ , Theorem 1.1. In this sense one can say that equivalence relation induced equality in quotient set or quotient object.

## 2. Mapping between sets

In this section, we will prove the necessary and sufficient condition for a mapping between two sets to be **injective, surjective and bijective**.

**Definition 2.1 (Injection).** Let  $\phi : X \rightarrow Y$  be a mapping. We say  $\phi$  is an injections of  $X$  into  $Y$  if any of the following are true

- (1) whenever  $x_1 \neq x_2$  then  $\phi(x_1) \neq \phi(x_2)$ ; that is, distinct element have distinct images. Contrapositively, we say that  $\phi$  is an injection if  $\phi(x_1) = \phi(x_2) \implies x_1 = x_2$

**Definition 2.2 (Left Inverse).** Let  $\varphi : X \rightarrow Y$  be a function. We say that  $f : Y \rightarrow X$  is a left inverse of  $\varphi$  if  $f \circ \varphi = I_X$ , the identity mapping on set  $X$ .

**Theorem 2.1 (Injection and Left Inverse).** Let  $\varphi : X \rightarrow Y$  be a mapping. Followings are equivalent

- (1).  $\varphi$  is an injection
- (2).  $\varphi$  has a left inverse

PROOF. We prove both implications using Definitions 2.1, 2.2.

- (2)  $\rightarrow$  (1) Let us assume that  $f : Y \rightarrow X$  is left inverse of  $\varphi : X \rightarrow Y$ . By definition,  $f \circ \varphi = I_X$ . Now, let  $\varphi(x_1) = \varphi(x_2)$ . Given that  $f$  is a function defined on  $Y$ , so equal inputs to  $f$  must give equal outputs. That is,  $f(\varphi(x_1)) = f(\varphi(x_2))$ . However, we know that  $f \circ \varphi : X \rightarrow X$  is an identity map of  $X$ . Therefore,  $f(\varphi(x_1)) = x_1 = f(\varphi(x_2)) = x_2$ . This proves that  $\varphi : X \rightarrow Y$  is indeed an injective map, Definition 2.1.
- (1)  $\rightarrow$  (2) Let us assume that  $\varphi$  is an injection. We need to construct a function  $f : Y \rightarrow X$  such that  $f \circ \varphi = I_X$ . So, by definition, we have to ensure that for each  $y \in Y$ ,  $f(y)$  is uniquely defined. Recall that inverse image of an element under any map, and thus for  $\varphi$ , is well defined as set  $\varphi^{-1}(y) := \{x \in X \mid \varphi(x) = y\}$ . Furthermore, if  $\varphi$  is an injective map then  $\varphi^{-1}(y) := \{x \in X \mid \varphi(x) = y\}$  is singleton. Hence we define map  $f$  as, for any arbitrary but fixed element  $x_o \in X$ ,

$$(14) \quad f(y) := \begin{cases} x & \text{if } \varphi(x) = y \\ x_o & \text{if } y \notin \text{Im}(\varphi) \end{cases}$$

Indeed, it is obvious that  $f \circ \varphi = I_X$ . □

**Definition 2.3 (Surjection).** Let  $\varphi : X \rightarrow Y$  be a mapping. We say  $\varphi$  is a surjection of  $X$  into  $Y$  if any of the followings are true

- (1)  $\forall y \in Y, \exists x \in X$  such that  $\varphi(x) = y$ . That is, if each element in  $Y$  is image of at least one element in  $X$ .
- (2) for any  $y \in Y$ , the collection  $\varphi^{-1}(y)$  is never an empty set,  $\Phi$ .

**Definition 2.4 (Right Inverse).** Let  $\varphi : X \rightarrow Y$  be a function. We say that  $f : Y \rightarrow X$  is a right inverse of  $\varphi$  if  $\varphi \circ f = I_Y$ , the identity mapping on set  $Y$ .

**Theorem 2.2 (Surjections and Right Inverse).** Let  $\varphi : X \rightarrow Y$  be a mapping. Followings are equivalent

- (1).  $\varphi$  is an Surjection
- (2).  $\varphi$  has a Right inverse

PROOF. We prove both implications using Definitions 2.3, 2.4.

- (2)  $\rightarrow$  (1). Let us assume that  $f : Y \rightarrow X$  be right inverse of  $\varphi : X \rightarrow Y$ . That is,  $\varphi \circ f(y) = I_Y$ . In sense of Definition 2.3, we need to show that for each  $y \in Y$ ,  $\varphi^{-1}(y) \neq \Phi$ . So, let  $y_0$  be an arbitrary element of  $Y$ . We can choose such element as set  $Y \neq \Phi$ . Since  $f : Y \rightarrow X$  is a function so  $f(y_0) = x_0$  is unique image of  $y_0$  under  $f$ . Also, given that  $\varphi \circ f = I_Y$ , so  $\varphi \circ f(y_0) = \varphi(f(y_0)) = \varphi(x_0) = y_0$ . That is,  $x_0 \in \varphi^{-1}(y_0)$ . This closes the proof.
- (1)  $\rightarrow$  (2). Let us assume that  $\varphi : X \rightarrow Y$  is a surjection. Then we have to construct a right inverse  $f : Y \rightarrow X$  of  $\varphi$ . That is,  $\varphi \circ f = I_Y$ . Let  $y_0 \in Y$  be an arbitrary element. Since,  $\varphi$  is surjection, therefore by Theorem 2.3, there exists at least one  $x_0 \in X$  such that  $\varphi(x_0) = y_0$ . Now, the question is why can't we just define  $f(y_0) = x_0$ ? Well, it is not necessary that  $\varphi^{-1}(y_0)$  is exactly one element. There may be more than one preimages of  $y_0$  under  $\varphi$ . Therefore, the preceding definition of  $f(y_0) = x_0$  would not be a well defined single valued function. To circumvent this problem, we decompose  $X$  into family of non empty mutually disjoint sets as  $X = \bigcup_{y \in Y} \varphi^{-1}(y)$ . This forms the partitions, in sense of Definition 1.1, of  $X$ . We can use axiom of choice to choose an element from each member of the collection  $\mathcal{C} = \bigcup_{y \in Y} \varphi^{-1}(y)$  and we define  $f(y) := x$  such that  $x$  is arbitrarily fixed element of  $\varphi^{-1}(y)$ . An important point of this construction is that it right inverse of a function is not unique because we have multiple elements available, in each sub-collection  $\varphi^{-1}(y)$ , if  $\varphi$  were not necessarily injection, however, surely a surjection. □

**Definition 2.5 (Bijection).** Let  $\varphi : X \rightarrow Y$  be a mapping. We say,  $f$  is a bijection if it is injection and surjection both.

**Theorem 2.3.** Let  $\phi : X \rightarrow Y$  be a mapping. We say  $\varphi$  is a bijection if and only if  $\varphi$  has left as well as right inverse.

PROOF. This immediately follows from the Definition 2.5 and Theorems 2.1, 2.2. □

**Theorem 2.4.** If a function  $f : X \rightarrow Y$  has left and right inverse both then these two inverses must coincide.

PROOF. Let  $f : X \rightarrow Y$  be a map and let  $g : Y \rightarrow X$  be its left inverse and  $h : Y \rightarrow X$  be its right inverse. That is,  $g \circ f = I_X$  and  $f \circ h = I_Y$ . Now, let  $g \neq h$  then there exist at least one  $y_0 \in Y$  such that  $g(y_0) = x_1 \neq h(y_0) = x_2$ . Now,  $f$  has right inverse so it is a surjection and hence there exist an element  $x_0 \in X$  such that  $f(x_0) = y_0$ . Now,  $g$  is a function and left inverse of  $f$ , so  $g(f(x_0)) = x_0 = g(y_0) = x_1$ . Also,  $f$  is a function and  $h$  is its right inverse, so  $f(h(y_0)) = y_0 = f(x_2)$ . So,  $f(x_0) = f(x_2)$ , but  $f$  is an injection so  $x_0 = x_2$ . We conclude that  $x_1 = x_0 = x_2$ . But this contradicts our initial assumption. Hence,  $g(y_0) = x_1 \neq h(y_0) = x_2$ . Hence, if map is bijective then it has unique left and right inverse and these left and right inverses coincide with each other. □

**Example 2.1 (Canonical projection).** Let  $\sim$  be an equivalence relation on set  $X$ . Then the mapping  $X \rightarrow X/\sim$  defined by  $x \mapsto [x]_\sim$  is known as canonical projection and is obviously a surjective map.

To see why it is obvious, we invoke the Definition 2.3 of surjection. Collection  $\mathcal{C} = \{[x]_\sim | x \in X\}$  forms partition of  $X$  and this partition is induced by equivalence relation of Example 1.1. In example 1.1, the equivalence class of an element  $x'$  consists of essentially those and only those elements in  $X$  which have the same image, under the mapping  $f$ , in set  $Y$ .

The object  $X/\sim$ , in example 2.1, is called as **quotient set** induced by equivalence relation  $\sim$ .

Injection and Surjection maps, defined in Definitions 2.1 2.3, form the basic building blocks out of which any map can be constructed. Theorem 2.5 establishes this assertion and is one of the significant result with deep manifestation in proof of first isomorphism theorem of the group.

**Theorem 2.5 (Function decomposition and Quotient set).** Let  $f : X \rightarrow Y$  be any function. Define equivalence relation on  $X$  as defined in example 1.1. That is  $x_1 \sim x_2$  if and only if  $f(x_1) = f(x_2)$ . Then we have following elegant decomposition of function

- $\pi : X \rightarrow X/\sim$  is a natural surjection in the sense of Definition 2.3
- $\varphi : X/\sim \rightarrow \text{Im}(f)$  defined by  $\varphi([x]_\sim) = f(x)$  is a bijection in the sense of Definition 2.5
- $\iota : \text{Im}(f) \rightarrow Y$  given by is a natural injection, also called as inclusion, in the sense of Definition 2.1

PROOF. I learnt this non trivial beautiful theorem and its proof from my mentor [Yaashaa Golovanov](#) I will explain the proof here.

Note that for any  $[x_0]_\sim \in X/\sim$ ,  $x_0 \in X$ , so  $\pi(x_0) = [x_0]_\sim$  is, by definition, a surjection.  $\iota$  also is, by definition, an injection. The difficult part is to prove that  $\varphi([x]_\sim) = f(x)$  is a bijection. To prove this, we need to show three things

- $\varphi([x]_\sim) = f(x)$  is **well defined**. To prove that it is well defined we have to basically establish that equal inputs gives equal output. So, let us assume that  $[x_1]_\sim = [x_2]_\sim$ . Then by Theorem 1.1,  $x_1 \sim x_2$  and by definition of equivalence relation, given in statement of this theorem,  $f(x_1) = f(x_2)$  which, by definition of  $\varphi$ , implies that  $\varphi([x_1]_\sim) = \varphi([x_2]_\sim)$
- $\varphi$  is an **injective map**. We have to prove that equal output  $\varphi([x_1]_\sim) = \varphi([x_2]_\sim)$  gives equal input  $[x_1]_\sim = [x_2]_\sim$ . To begin with, let  $\varphi([x_1]_\sim) = \varphi([x_2]_\sim)$ . By definition of  $\varphi$ ,  $\varphi([x_1]_\sim) = f(x_1)$  and  $\varphi([x_2]_\sim) = f(x_2)$ . So,  $f(x_1) = f(x_2)$ . By definition of equivalence relation under consideration,  $f(x_1) = f(x_2) \implies x_1 \sim x_2$  and this means, by Theorem 1.1,  $[x_1]_\sim = [x_2]_\sim$ .
- $\varphi$  is a **surjective map**. Let  $y_0 \in \text{Im}(f)$ . This means there exist an  $x_0 \in X$  such that  $f(x_0) = y_0$ . Since, equivalence classes form the partition of  $X$ , one candidate for an equivalence class which contains  $x_0$  is  $[x_0]_\sim$  itself and image of this class under  $\varphi$  is  $f(x_0) = y_0$ . This closes the proof.  $\square$

**Remark 2.1.** Notice that this theorem says that Every set theoretical map,  $f : X \rightarrow Y$ , is composition of surjective map, followed by bijective map, followed by injective map. Moreover, the bijection  $\varphi$  induced by  $f$  itself.

## Bibliography

- [1] Tom W. Korner, *A Companion to Analysis - A Second First and First Second Course in Analysis*, Graduate Studies in Mathematics, Vol 62, American Mathematical Society, 2011.
- [2] Georgi Shilov, *Elementary Real and Complex Analysis*, Dover Books on Mathematics, 1996.
- [3] Terence Tao, *Analysis I and Analysis II*, Springer, 2023.
- [4] V.B. Alekseev and V.I. Arnold, *Abel's theorem in problem and solution*, Springer, 1986.
- [5] Joseph Silverman, *Abstract Algebra - An Integrated Approach*, Pure and Applied Undergraduate Texts, 55, American Mathematical Society, 2022.
- [6] Ralph Abraham, Jerrold Eldon Marsden, Tudor Ratiu *Manifolds, Tensor Analysis, and Applications*, Applied Mathematical Sciences, 75, Springer. (I used the PDF version accessible at Caltech's repositories.)
- [7] David S. Dummit and Richard M. Foote (2004), *Abstract Algebra*, John Wiley and Sons edition.
- [8] Steven J. Rosenberg, *Abstract Algebra*, John Wiley and Sons edition.
- [9] Jonathan S. Golan *Linear Algebra Every Graduate Student Ought to Know*, Springer, 2012.
- [10] Leslie Lamport (1994) *L<sup>A</sup>T<sub>E</sub>X: a document preparation system*, Addison Wesley, Massachusetts, 2nd ed.
- [11] K. Conrad, *Expository Papers at <https://kconrad.math.uconn.edu/blurbs/>*.