Lu Sun, and many more.

A Notebook on Algebra

To my family, friends and communities members who have been dedicating to the presentation of this notebook, and to all students, researchers and faculty members who might find this notebook helpful.

Contents

Foreword

If software and e-books can be made completely open-source, why not a notebook?

This brings me back to the summer of 2009 when I started my third year as a high school student in Harbin No. 3 High School. In the end of August when the results of Gaokao (National College Entrance Examination of China, annually held in July) were released, people from photocopy shops would start selling notebooks photocopies that they claimed to be from the top scorers of the exam. Much curious as I was about what these notebooks look like, never have I expected myself to actually learn anything from them, mainly for the following three reasons.

First of all, some (in fact many) of these notebooks were more difficult to read than the textbooks. I guess we cannot blame the top scorers for being so smart that they sometimes made things extremely brief or overwhelmingly complicated.

Secondly, why would I want to adapt to notebooks of others when I had my own notebooks which in my opinion should be just as good as theirs.

And lastly, as a student in the top-tier high school myself, I knew that the top scorers were probably my schoolmates. Why would I pay money to a stranger in a photocopy shop for my friends' notebooks, rather than requesting a copy from them directly?

However, my mind changed after becoming an undergraduate student in 2010. There were so many modules and materials to learn for a college student, and as an unfortunate result, students were often distracted from digging deeply into a module (and for those who were still able to do so, you have my highest respect). The situation became worse when I started pursuing my Ph.D. in 2014. As I had to focus on specific research areas entirely, I could hardly split enough time on other irrelevant but still important and interesting contents.

To make a difference, I enforced myself reading articles beyond my comfort zone, which ended up motivating me to take notes to consolidate the knowledge. I used to work with hand-written notebooks. My very first notebook was on Numerical Analysis, an entrance-level module for engineering background graduate students. Till today I still have dozens of these notebooks on my bookshelf. Eventually, it came to me: why not digitizing them, making them accessible online and open-source and letting everyone read and edit it?

Similar with most open-source software, this notebook does not come with any "warranty" of any kind, meaning that there is no guarantee that every-

thing in this notebook is correct, and it is not peer reviewed. **Do NOT cite this notebook in your academic research paper or book!** If you find anything helpful here with your research, please trace back to the origin of the knowledge and confirm by yourself.

This notebook is suitable as:

- a quick reference guide;
- a brief introduction for beginners of an area;
- a "cheat sheet" for students to prepare for the exam or for lecturers to prepare the teaching materials.

This notebook is NOT suitable as:

- a direct research reference;
- a replacement of the textbook.

The notebook is NOT peer reviewed, thus is more of a notebook than a book. It is meant to be easy to read, not to be comprehensive and very rigorous.

Although this notebook is open-source, the reference materials of this notebook, including textbooks, journal papers, conference proceedings, etc., may not be open-source. Very likely many of these reference materials are licensed or copyrighted. Please legitimately access these materials and properly use them, should you decided to trace the origin of the knowledge.

Some of the figures in this notebook are plotted using Excalidraw, a very convenient tool to emulate hand drawings. The Excalidraw project can be found on GitHub, *excalidraw/excalidraw*. Other figures may come from MATLAB, R, Python, and other computation engines. The source code to reproduce the results are intended to be included in the same repository of the notebook, but there might be exceptions.

This work might have benefited from the assistance of large language models, which are used exclusively for editing purposes such as correcting grammar and rephrasing sentences, without introducing new content, generating novel information, or changing the original intent of the text.

Preface

This notebook is on algebra. The first part of the notebook is about linear algebra, one of the first few modules a science and engineering undergraduate student would take in his first semester in the collage. It is absolutely the fundamental of almost all the mathematical tools he would use in the future. The second part of the notebook is about abstract algebra, a far more advanced topic and yet still a lot of fun to learn. The third part of the notebook is on number theory, which is a natural expansion to abstract algebra.

The key reference of this notebook is listed below. During the development of the notebook, this list may become longer and longer.

Book *Basic Algebra Second Edition (I and II)* by Nathan Jacobson, published by DOVER [1]

List of Figures

List of Tables

Part I

Linear Algebra

1

Linear System

CONTENTS

This chapter studies the concepts and basic properties of linear equations and linear systems. They serve as the basis of linear algebra.

1.1 Linear System

This section introduced linear system and its notations.

1.1.1 Linear System and Solution Set

A **linear equation** refers to an equation of the following form.

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

which is a linear equation of variables x_i. The number of x_i is an integer denoted by n and $n \geq 1$. Variables a_i and b real or complex numbers known in advance. Variables a_i are known as the **coefficients** of the equation.

A system of linear equations on the same variable set is known as a **linear system**. For example,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

is a linear system of variables x_i.

A linear system has solution if there is at least one set of variables x_i that fulfills all the linear equations in the system. In practice, a linear system may:

- have no solution;
- have one unique solution;
- have infinite number of solutions.

The solution(s) to a linear system forms its **solution set**. The cardinality of the solution set of a linear system can be 0, 1 or infinity.

Can a linear system have 2 or more finite number of solutions?

A linear system cannot have 2 or more finite number of solutions. This can be proved easily by proof by contradiction.

Assume that a linear system has a finite number of $N \geq 2$ solutions. From those solutions, select two distinct set of solutions x_i^a and x_i^b. Since the solutions are distinct, at least one of them is non-zero. Without loosing generality, let us assume that x_i^a is an non-zero solution.

Substituting x_i^a into the linear system gives

$$\begin{aligned} a_{11}x_1^a + a_{12}x_2^a + \ldots + a_{1n}x_n^a &= b_1 \\ &\vdots \\ a_{m1}x_1^a + a_{m2}x_2^a + \ldots + a_{mn}x_n^a &= b_m \end{aligned}$$

Multiplying p on both side of the equations gives

$$\begin{aligned} a_{11}px_1^a + a_{12}px_2^a + \ldots + a_{1n}px_n^a &= pb_1 \\ &\vdots \\ a_{m1}px_1^a + a_{m2}px_2^a + \ldots + a_{mn}px_n^a &= pb_m \end{aligned}$$

Substituting x_i^b into the linear system and multiplying $(1-p)$ on both side gives

$$\begin{aligned} a_{11}(1-p)x_1^b + a_{12}(1-p)x_2^b + \ldots + a_{1n}(1-p)x_n^b &= (1-p)b_1 \\ &\vdots \\ a_{m1}(1-p)x_1^b + a_{m2}(1-p)x_2^b + \ldots + a_{mn}(1-p)x_n^b &= (1-p)b_m \end{aligned}$$

Adding the two set of equations gives

$$\begin{aligned} a_{11}\left[px_1^a + (1-p)x_1^b\right] + \ldots + a_{1n}\left[px_n^a + (1-p)x_n^b\right] &= b_1 \\ &\vdots \\ a_{m1}\left[px_1^a + (1-p)x_1^b\right] + \ldots + a_{mn}\left[px_1^a + (1-p)x_1^b\right] &= b_m \end{aligned}$$

Therefore, $px_i^a + (1-p)x_i^b$ must also be a solution to the original linear system.

Let $p \in P = \{0, \frac{1}{N}, \frac{2}{N} \ldots, \frac{N}{N}\}$. The cardinality of P is $N+1$. It can be easily proved that with non-zero x_i^a, $px_i^a + (1-p)x_i^b$ are distinct values for each different values p taken from P. All of these $N+1$ distinct $px_i^a + (1-p)x_i^b$ with different p are solutions to the original linear system. This contradicts with the assumption that the linear system has a finite number of N solutions.

A linear system with one for infinite solutions is said to be **consistent**, whereas a linear system with no solution is said to be **inconsistent**. Two linear systems are said to be **equivalent** if they have the same solution set.

1.1.2 Matrix Notation

A linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

can be denoted by

$$Ax = b \tag{1.1}$$

where x is a $n \times 1$ vector, A a $m \times n$ **coefficient matrix**, and b a $m \times 1$ vector given by

$$\begin{aligned} x &= \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T \\ A &= \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \\ b &= \begin{bmatrix} b_1 & \ldots & b_m \end{bmatrix}^T \end{aligned}$$

1.1.3 Solution to a Linear System

Matrix A and vector b contains all the information necessary to solve the linear system. **Elimination algorithm** can be used to find the solution set of a linear system. The details are too basic hence not giving in this notebook.

As mentioned earlier, a linear system may have zero, one or infinite solutions. It is interesting to see how the numbers of solutions are determined by (certain patterns) in A and b. Detailed discussions are given in later part of this notebook. A quick review is given below.

There are mainly two factors that decide the number of solutions of a linear system.

- The number of linearly independent equations (the definition of linearly independent will be introduced in later part of the notebook) among the m equations versus the number of independent variables n.
- The consistency of linearly dependent equations, if any.

For a linear system to be consistent, all the linear dependent equations must be consistent. For a consistent linear system to have a unique solution, the number of linearly independent equations must match the number of independent variables.

An extension to the question is given below. Consider the **homogeneous system** of (1.1) which is given by

$$Ax = 0$$

with $b = 0$ substituted into the original linear system. It is obvious that the system must be consistent, and at least the zero vector $x = 0$ is a solution. The question is whether $x = 0$ is the unique solution. This is determined by the number of linearly independent rows of A, and whether it matches with the number of the independent variables.

Let A be $m \times n$, where the number of equations is m and the number of independent variables, n.

For $m < n$, even if all the equations are linearly independent, there are still fewer linearly independent equations than the number of independent variables. Therefore, the linear system has infinite solutions. For $m = n$, the linear system has a unique solution if and only if all the equations are linearly independent. For $m > n$, there must be linearly dependent rows. There can be at most n linearly independent rows and at least $m - n$ rows are linearly dependent. The linear system has a unique solution if and only if the number of the linearly independent rows, after removing those linearly dependent rows, matches n.

The number of linearly independent rows, in this context, is denoted by the rank of the matrix. (The formal definition of the rank of a matrix will be introduced in later part of the notebook.) For matrix A which is $m \times n$, $Ax = 0$ has unique solution if and only if $\text{rank}(A) = n$.

1.2 Vector and Matrix Calculations

Basic rules for calculations with vectors and matrices are introduced in this section.

1.2.1 Vector Calculations

“nobreak

1.2.2 Matrix Calculations

Part II

Abstract algebra

2

Abstract Algebra Basics

CONTENTS

Abstract algebra studies the principals and algorithms used in different algebraic systems. The basics of abstract algebra including the scope and commonly seen definitions and concepts are introduced in this chapter.

2.1 Scope

Linear algebra has been introduced in the earlier part of this notebook. Linear algebra as well as other classic algebraic algorithms solve practical problems using algebra theory, whereas abstract algebra studies these theories. While classic algebra performs calculations on numbers, vector and matrices, **abstract algebra** studies the concepts, tools, derivations and logic we use in the calculation, and tries to explain why they work in the way they do. Ab-

stract algebra also develops new concepts, tools and algorithms that we can use to solve more complicated algebraic problems.

As an example, consider the following equation

$$Ax = y$$

where x, y are vectors and A a matrix. Solving x given particular A and y falls into the classic algebra domain. It is obvious that x does not necessarily exist or being unique for different y and A. Studying the general rules when x exists and when it is unique for a set of y and A becomes an abstract algebra problem.

One may have seen the following expressions in elementary schools

$$\begin{aligned} a + b &= b + a \\ ab &= ba \end{aligned}$$

which are often used to demonstrate the commutative property of calculations (summation and multiplication, in this example). In classic algebra, they are considered as ground truth. In abstract algebra, however, the focus shifts to a more formal and generalized understanding of the property. We need to dig deeper into how commutative property is defined, and why it holds true for summation and multiplication, but not for some other operations such as division.

Here is another great motivating example. One of the most famous applications of abstract algebra is to study the analytical solution to the following polynomial equation

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n = 0 \tag{2.1}$$

where $n \geq 1$ is the order of the polynomial and $a_1, \ldots, a_n$ are any arbitrary values. The analytical solutions to (2.1) for $n = 1$ and $n = 2$ are obvious. With some effort, the analytical solutions for $n = 3$ and $n = 4$ were found in the 16-th century. Since then, people have been struggling to find the analytical solution to the polynomial with fifth and higher orders.

In the 18th and 19th century, Euler, Lagrange and Gaussian attended this problem. Their conclusion was that there is no analytical solution to polynomial equations of fifth order or higher, but they could not give a very solid proof to the statement. Nevertheless, the methods they used inspired a lot of people that would work on this problem.

In the 19th century, Abel was able to prove that there is no analytical solution to general polynomial equations of fifth degree or higher with arbitrary coefficients (see Abel–Ruffini theorem). Furthermore, he discussed a set of special cases with specific coefficient patterns that can have analytical solutions, and he came up with a sufficient condition for a fifth order polynomial equation to have the analytical solution.

The necessary and sufficient condition for a fifth or higher order polynomial

to have an analytical solution is finally fully discovered by the genius Galois at a remarkably young age. Galois was able to create his theorem (known as the Galois theorem) and use it to find the ultimate answer to this problem that people have been studied for centuries. His theorem goes far beyond that. Galois theorem will find its usefulness in many areas to come, and eventually it becomes an important building block of abstract algebra.

In the remaining of this part of the notebook, abstract algebra including the contribution from Abel, Galois and other mathematicians are introduced.

2.2 Algebraic System

An **algebraic system** is a mathematical system consisting of a non-empty set and a series of operations defined on the set. Abstract algebra studies the properties of different algebraic systems. Depending on the properties of the algebraic systems, we can categorize them as "groups", "rings", "fields", "vector spaces", etc.

The commonly seen definitions and concepts used in algebraic system is introduced in the reminder of this section.

2.2.1 Domain and Mapping

Set is one of the most commonly used terms across different mathematical subjects. It is also one of the fundamental concepts in abstract algebra. A **set** usually refers to a collection of distinct objects. Given a set U and an object x, one and only one of the following two statements must be true:

- Object x is a member of set U, denoted by $x \in U$;
- Object x is not a member of set U, denoted by $x \notin U$.

However, notice that due to the Russell's paradox, it is challenging to give a rigorous mathematical definition to a set that fulfill the above statement. Therefore, we usually introduce set by its features, not its definition.

An algebraic system must contains a non-empty set. The set is known as the **domain** of the system.

Mapping, or **function**, is used to describe the association of elements in two sets. The two terms are used interchangeably in the this notebook.

An example of mapping is given as follows. Let A be a set, and $A_0 \subset A$ a subset of A. For any element $x \in A_0$, define mapping

$$i \quad : \quad A_0 \to A$$

where

$$i(x) \quad = \quad x$$

In this example, mapping i is called the **embedding mapping** from A_0 to A.

Let A, B be two sets, and $A_0 \subset A$ a subset of A. Let $f : A \to B$, and $g : A_0 \to B$. If $\forall x \in A_0$, $f(x) = g(x)$, function f is known as an **extension** of function g, and function g a **restriction** of function f (on A_0). This is denoted by $g = f|_{A_0}$.

Mappings can be chained together. For example, consider several mappings

$$\begin{aligned} f_1 &: A_1 \to A_2 \\ f_2 &: A_2 \to A_3 \\ f_3 &: A_3 \to A_4 \end{aligned}$$

With the above, we can denote $f = f_3 \circ f_2 \circ f_1$ a mapping from A_1 to A_4. Meantime, consider other mappings

$$\begin{aligned} g_1 &: A_1 \to B \\ g_2 &: B \to A_4 \end{aligned}$$

Clearly, $g = g_2 \circ g_1$ is also a mapping from A_1 to A_4. The mappings above can be illustrated intuitively using the **mapping diagram** in Fig. 2.1. Mapping

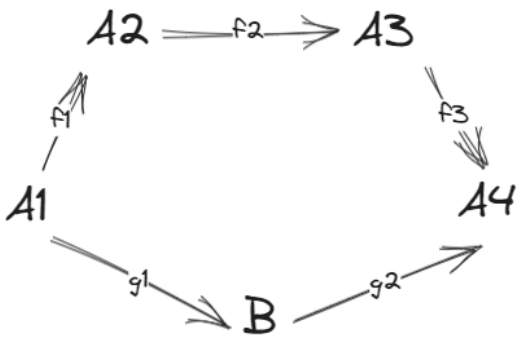

FIGURE 2.1
Mapping diagram that demonstrates f and g.

diagram can become handy with complicated mappings.

The embedding mapping, extension function and restriction function introduced earlier can also be represented by a mapping diagram as shown in Fig. 2.2, where $A_0 \subset A$ and $i : A_0 \to A$ a embedding mapping. Function g defined on the subset A_0 is a restriction of f defined on the superset, whereas f is an extension of g, i.e., $f(x) = g(x)$ for $x \in A_0$.

Multiple sets can be "combined" and "augmented" to form new sets. For example, the **Cartesian product** of two sets, A_1 and A_2, is defined as follows.

$$A_1 \times A_2 = \{(a, b) | a \in A_1, b \in A_2\}$$

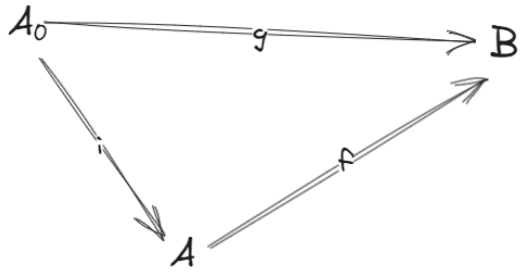

FIGURE 2.2
Mapping diagram of embedding mapping.

where the tuple (a, b) can be interpreted as an ordered combination of elements in A_1 and A_2. The similar idea can be applied to Cartesian product of 3 or more sets.

2.2.2 Operation

Summation, multiplication, etc., are operations. In abstract algebra, we are more interested in the broad definition of operations from set perspective, instead of listing down individual operations and study how they work.

In an algebraic system, in addition to the domain, an operation must also be defined. An **operation** describes the rule of deriving or mapping from one or multiple elements to a new element in the domain.

Take binary operation as an example which maps two elements to one element. Let A, B, D be three non-empty sets. Let f be a mapping

$$f \quad : \quad A \times B \to D \tag{2.2}$$

Then f is called an **algebraic operation** from A, B to D. The operation can be denoted by $f(a, b)$ where $a \in A$ and $b \in B$. For convenience, binary operation is often denoted in the form of $a \oplus b$, $a \otimes b$, etc. Notice that there are conventions on how to use the symbols. For example, $+$, $-$, $\times$, etc., already have clear meanings.

The following is an example of operations. Let the domain of interest be $\mathbb{R}$, which is the real number set. Let v be a vector space (a formal definition of vector space will be given later; for now, take vector space as a set of vectors with summation operation defined on it) defined under $\mathbb{R}^n$. We can then define summation $+$ as a binary operation

$$+ \quad : \quad v \times v \to v$$

which indicates that the summation takes in two elements in the vector space, and generates a new vector that also belongs to the same vector space.

In the case where the domain and range of the operation come from the same set, i.e., in (2.2) $A = B = D$, the operation is said to be **closed** under this operation. In this example, the summation operation $+$ defined on v is closed, and we can simply say "$+$ is a binary operation defined on v".

Operations may have the following properties. Commonly seen ones are introduced below. Let binary operation $\oplus$ be defined on A, i.e., $\oplus : A \times A \to A$. If

$$a \oplus b \quad = \quad b \oplus a, \forall a, b \in A$$

then the operation is said to have **commutative property**. If

$$a \oplus (b \oplus c) \quad = \quad (a \oplus b) \oplus c, \forall a, b, c \in A \tag{2.3}$$

then the operation is said to have **associative property**, and (2.3) can be simply denoted by $a \oplus b \oplus c$. Let two binary operations defined on A, denoted by $\oplus$ and $\otimes$ respectively. If

$$a \otimes (b \oplus c) \quad = \quad (a \otimes b) \oplus (a \otimes c)$$

then the operation $\otimes$ is said to have the left-hand **distributive property** on operation $\oplus$. Similarly, right-hand distributive property can also be defined.

Commutative property, associative property and distributive property are commonly seen and widely discussed properties of operations. When operations have some of the three properties, simplified notation may apply. For example, if $\oplus$ operation has associative property, then

$$a^n \quad \equiv \quad \overbrace{a \oplus \cdots \oplus a}^{n}$$

can be used when there is no ambiguity. With this notation, it can be easily proved that for an operation that has both commutative and associative properties,

$$a^n b^n \quad = \quad (ab)^n$$

There are several different ways to represents the result of an operation. When the cardinality of the domain is finite, a simple way is to list all the input-output associations in a table. An example is given in Fig. 2.3 which exhaustively shows "AND" logical operation results. For operations defined on a domain with infinite cardinality, other methods must be used.

2.2.3 Relation

Consider the relation of two elements a and b. **Relation** is essentially a property defined on tuple (a, b). A straight forward way of defining a relation is to construct a set that contains some tuples, and if the specified tuple (a, b) is

AND	0	1
0	0	0
1	0	1

FIGURE 2.3
Result of "and" logical operation.

in the set, we say that a, b have the corresponding relation, and equivalently (a, b) has that associated property.

For example, consider $a, b \in \mathbb{Z}$ where $\mathbb{Z}$ denotes the integer set. Define relation

$$R \quad = \quad \{(a, b) \mid a - b = kn,\ k, n \in \mathbb{Z}, n > 1\} \tag{2.4}$$

If a, b are such that $(a, b) \in R$ in (2.4), we say a, b are congruent modulo n. This defines a **congruence relation** and it can be denoted by

$$a \quad \equiv \quad b \pmod{n}$$

Congruence relation is a commonly used relation in abstract algebra and number theory.

To summarize, let $a \in A$, $b \in B$. Let $R \subset A \times B$ be a subset of $A \times B$. the following statements can be considered equivalent.

- Two elements a and b have relation R;
- Tuple (a, b) has the property associated with relation R;
- Tuple $(a, b) \in R$;

and in that case we can use aRb to signify the relation. From above, we can see that each relation is associated with a subset R defined on the Cartesian product of the sets that the two elements belong to, and vise versa.

Relations may have the following properties. Commonly seen ones are introduced below. Let relation R be defined on $A \times A$. If $\forall a \in A$,

$$aRa$$

then R is said to be **reflexive**. If $\forall a, b \in A$,

$$aRb \quad \Rightarrow \quad bRa$$

then R is said to be **symmetric**. If $\forall a, b, c \in A$,

$$aRb, bRc \quad \Rightarrow \quad aRc$$

then R is said to be **transitive**. Finally, if $\forall a, b \in A$,

$$aRb, bRa \quad \Rightarrow \quad a = b$$

then R is said to be **antisymmetric**.

From the above definitions, we can see that the commonly seen "=" relation defined on $\mathbb{R} \times \mathbb{R}$ is reflexive, symmetric and transitive. Similarly, "$\leq$" is reflexive, transitive and antisymmetric.

2.2.4 Equivalence Relation and Equivalence Class

If a relation is simultaneously reflexive, symmetric and transitive, it is called an **equivalence relation**. The equal "=" relation and congruence relation introduced earlier are examples of equivalence relations.

We can define a **partition of a (non-empty) set** by grouping its elements into non-empty subsets in such a way that every element is included in exactly one subset, i.e.

$$\begin{aligned} A &= \bigcup_{i \in I} A_i \\ \text{s.t.} \quad & A_i \neq \varnothing \\ & \forall i, j \in I, i \neq j, A_i \bigcap A_j = \varnothing \end{aligned}$$

A partition of a set can form a new set $\{A_i\}$. Furthermore, we can define a relation $R \subset A \times A$ on top of that partition as follows

$$R = \{(a, b) \mid \exists i, a, b \in A_i\}$$

which intuitively means "the two elements are from the same subset in the partition" and it is clear that such R is an equivalence relation. A partition of a set A can determine an equivalence relation $R \subset A \times A$ using the above method. Vise versa, an equivalence relation $R \subset A \times A$ can also determine a partition of set A as follows. For $\forall a \in A$, define

$$[a] \quad = \quad \{b \in A | aRb\}$$

where R is the equivalence relation, and $[a]$ the **equivalence class** of a.

With the definition of an equivalence relation and a set of equivalence class, a partition of a set A can then be derived from the equivalence relation R by

$$A/R \quad = \quad \{[a] | a \in A\} \text{ (remove duplication)} \tag{2.5}$$

where A/R is the partition of set A (this can be proved easily) derived from equivalence relation R. The partition of set obtained using the above method is known as the **quotient set**.

There is a one-to-one correspondence between an equivalent relation and its quotient set. Given an equivalence relation, the following mappings can be defined.

$$\pi \quad : \quad A \to A/R,\ \pi(a) = [a]$$

which is known as the **canonical projection** or **quotient map**.

2.2.5 Congruent Modulo

Congruent modulo has already been introduced in (2.4). In that example, when two integers a and b are congruent modulo n, their equivalence relation persists even if $k_1 n$, $k_2 n$ are added or subtracted from both of them respectively.

The concept of congruent modulo can be further generalized in the context of operation and equivalence relation as follows. Let A be an non-empty set. Let $\oplus$ be a closed binary operation defined on A, i.e., $\oplus : A \times A \to A$. Let R be an equivalence relation defined on A, i.e., $R \subset A \times A$, and A/R its corresponding quotient set. Let $a_1, a_2, b_1, b_2 \in A$. If operation $\oplus$ and equivalence relation R satisfy the following condition

$$a_1 R a_2,\ b_1 R b_2 \quad \Rightarrow \quad (a_1 \oplus b_1) R (a_2 \oplus b_2) \tag{2.6}$$

then we say that R is **congruent modulo** $\oplus$. This is demonstrated by Fig. 2.4 the left plot. From Fig. 2.4 the left plot, congruent modulo can be intuitively interpreted as follows. If elements a_1, a_2 are congruent modulo $\oplus$, their equivalence relation persists even if operation $\oplus b_1$, $\oplus b_2$ are applied to them respectively, so long as b_1, b_2 are also congruent modulo $\oplus$.

If R is congruent modulo $\oplus$ indeed, we can define an operation on the quotient set A/R as follows. Let $[\cdot]$ denote the equivalence class of an element in A. Define $\bar{\oplus}$ as an operation on A/R as follows.

$$[a] \bar{\oplus} [b] \quad = \quad [a \oplus b],\ a \in [a],\ b \in [b] \tag{2.7}$$

where a, b are arbitrarily selected elements in the equivalence class $[a]$ and $[b]$ respectively, as shown in Fig. 2.4 right plot. Notice that $[a \oplus b]$ should be well-defined regardless of the choice of a and b so long as they are selected in their associated equivalence class $[a]$ and $[b]$ respectively since R is congruent modulo $\oplus$. In this context, $\bar{\oplus}$ is known as the **included operation** or **quotient operation** of $\oplus$ on the quotient set A/R.

To conclude, in the algebraic system with non-empty set A and an operation $\oplus$ (denoted by $\{A; \oplus\}$), if there is a equivalent relation R which is congruent modulo $\oplus$, we can construct a new algebraic system $\{A/R; \bar{\oplus}\}$, where A/R is the quotient set of A on R from (2.5), and $\bar{\oplus}$ the included operation of $\oplus$ on A/R from (2.7).

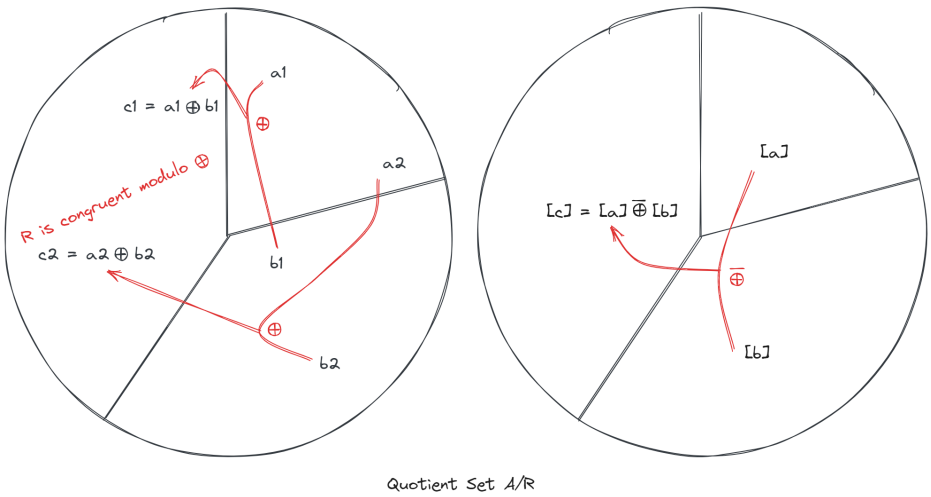

FIGURE 2.4
A relation is congruent modulo an operation.

2.3 Semi-Group and Group

As introduced earlier, a general algebraic system shall contain a set and one or more operations defined on the set. The operations can be unary (involving one operand), binary (involving two operands), ternary (involving three operands), etc. Relations are defined as sets of tuples, with each element of the tuple coming from the set of interest. Likewise, relation can also be binary, ternary, etc.

Special properties can apply to both operations and relations. Binary operations, for instance, can exhibit commutativity, associativity, and distributivity. Binary relations, on the other hand, can be reflexive, symmetric, and transitive. If a relation is simultaneously reflexive, symmetric and transitive, it is called an equivalent relation. Each equivalent relation is corresponded with a quotient set which is a partition of the original set by equivalence classes of the relation. When a relation is congruent modulo an operation (recall (2.6) and Fig. 2.4), we can establish a new algebraic system. This system comprises the quotient set and a well-defined induced operation on that set. The new system retains structural properties from the original set, reflecting the consistency and compatibility of the congruence relation with the operation.

Moving forward, our focus will shift to specialized algebraic systems characterized by distinctive properties. This exploration will pave the way to understanding the fundamental algebraic structures such as semi-groups, groups, and rings. We will see the critical impact the operation possess over the alge-

braic system. Even with the same set, different operations often leads to very different features of the algebraic system.

2.3.1 Semi-Group

Let S be a non-empty set, and $\oplus : S \times S \to S$ a closed binary operation defined on S. If $\oplus$ has associative property, i.e. $a \oplus (b \oplus c) = (a \oplus b) \oplus c$, then algebraic system $\{S; \oplus\}$ (for simplicity, S, when there is no ambiguity) is called a **semi-group**.

Examples of semi-group include $\{\mathbb{N}; +\}$, $\{\mathbb{N}; \times\}$ where $\mathbb{N}$ refers to the (positive) natural numbers $1, 2, 3, \ldots$

Another example is given as follows. Let A be an non-empty set, $\mathcal{M}(A)$ the set of all the closed mappings defined on A, and $\circ$ the compound operation. Then $\{\mathcal{M}(A); \circ\}$ is also a semi-group. Additionally, $\{\mathcal{P}(A); \bigcup\}$, $\{\mathcal{P}(A); \bigcap\}$ are also semi-groups, where $\mathcal{P}(A)$ represents the power set of A.

Is Zero Included in Natural Numbers?

The definition of natural numbers may or may not include zero "0" depending on the context. In the context of number theory and abstract algebra, natural numbers often exclude zero.

2.3.2 Monoid

Consider semi-group $\{S; \oplus\}$. If $\exists e \in S$ so that

$$e \oplus a \quad = \quad a, \ \forall a \in S$$

then e is known as the **left identity**. Similarly, we can define **right identity**. Notice that a semi-group may have many distinct left and right identities.

If an element e is both left identity and right identity, it is called the **identity element**. If a semi-group has an identity element, the identity element must be unique. This can be easily illustrated using proof by contradiction as follows. Assume that e_1, e_2 are two distinct identity elements, thus

$$e_1 \oplus e_2 = e_1 \quad \text{and} \quad e_1 \oplus e_2 = e_2$$

indicating $e_1 = e_2$, which contradicts with the assumption.

A semi-group with the identity element is knows as a **monoid**. From the definition, we know that $\{\mathbb{N}; \times\}$, $\{\mathcal{M}(A); \circ\}$, $\{\mathcal{P}(A); \bigcup\}$ and $\{\mathcal{P}(A); \bigcap\}$ are all monoids, with their identity elements being 1, $f : A \to A, f(a) = a$, $\varnothing$ and A, respectively.

2.3.3 Group

Let $\{S; \oplus\}$ be a monoid with identity element e. If for $a \in S$, there is

$$a' \oplus a \quad = \quad e$$

then a' is called the left inverse of a. Likewise, right inverse can be defined. Notice that an element may have many left and right inverses.

If a' is both the left and the right inverse of a, then a' is known as the inverse of a, ands a' must be unique. The proof can be obtained similarly using proof by contradiction. The inverse of a is often denoted by a^{-1}.

If in a monoid $\{S; \oplus\}$, there is the inverse for all its elements, the monoid is called a group, usually denoted by $\{G; \oplus\}$, or simply G when without ambiguity.

The relationship of general algebraic system, semi-group, monoid and group are demonstrated in Fig. 2.5.

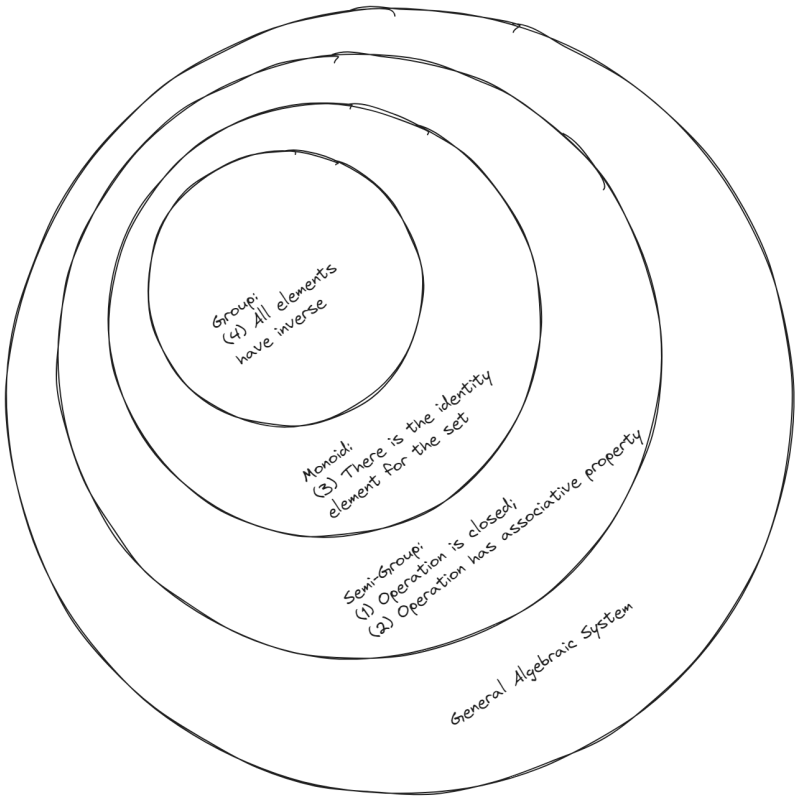

FIGURE 2.5
General Algebraic System, Semi-Group, Monoid and Group.

Notice that group does not pre-assume commutativity. If a group's operation is commutative, it is called an **abelian group** named after Niels Henrik Abel. Examples of abelian group include $\{\mathbb{Z}; +\}$, $\{\mathbb{R}; +\}$, $\{\mathbb{C}; +\}$ where $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ represent integer set, real number set and complex number set, respec-

tively. Additionally, $\{\mathbb{R}^*;\times\}$, $\{\mathbb{C}^*;\times\}$ are also abelian groups, where $\mathbb{R}^*$, $\mathbb{C}^*$ denotes the corresponding sets excluding zero.

We know that $\{\mathcal{M}(A);\circ\}$ introduced earlier is not a group. This is because not all mappings defined on a set A necessarily have an inverse. However, if define $\{\mathcal{S}(A);\circ\}$ where $\mathcal{S}(A)$ represents the set of all bijections (invertible mappings) on A, $\{\mathcal{S}(A);\circ\}$ becomes a group. Notice that $\mathcal{S}(A)$ is not necessarily an abelian group.

To conclude, a group must meet the following requirement:

- There is a non-empty set G;
- There is a closed binary operation $\oplus : G \times G \to G$;
- The operation $\oplus$ has associative property, i.e. $\forall a, b, c \in G, (a \oplus b) \oplus c = a \oplus (b \oplus c)$;
- There is the identity element, i.e. $\exists e \in G$, $\forall a \in G, e \oplus a = a \oplus e = a$;
- There is the inverse for all elements, i.e., $\forall a \in G$, $\exists a^{-1}$, $a \oplus a^{-1} = a^{-1} \oplus a = e$.

Furthermore, if

- Operation $\oplus$ has commutative property, i.e. $\forall a, b \in G, a \oplus b = b \oplus a$;

the group is called an abelian group.

It can be proved that if a semi-group has left identity and all its elements have left inverse, the left identity must also be its right identity and all its elements must also have right reverse, thus making the semi-group a group. The proof is given below. Let e be the left identity of semi-group $\{S;\oplus\}$, while b be the left inverse of any arbitrary element $a \in S$. Let c be the left inverse of b.

$$\begin{aligned} a \oplus b &= e \oplus a \oplus b \\ &= c \oplus (b \oplus a) \oplus b \text{ (associativity)} \\ &= c \oplus (e \oplus b) \text{ (associativity)} \\ &= c \oplus b \\ &= e \end{aligned}$$

Therefore, b is also the right inverse of a. Furthermore,

$$\begin{aligned} a \oplus e &= (a \oplus b) \oplus a \text{ (associativity)} \\ &= e \oplus a \\ &= a \end{aligned}$$

Therefore, e is also the right identity. All the above derivations also apply when a semi-group has right identity and all its elements have right inverse.

2.3.4 Properties of Semi-group and Group

Commonly used properties and notations of semi-group and group are introduced in this section.

If $\{S;\oplus\}$ is a semi-group and $a,b \in S$, then the following criteria can be used to further determine that $\{S;\oplus\}$ is a group.

- For $\forall a,b \in S$, if both $a \oplus x = b$ and $x \oplus a = b$ have solution, then S is a group.
- If $|S| < \infty$ (cardinal number of S is finite), and $\forall a,b,c \in S$, $a \oplus c = b \oplus c \Rightarrow a = b$, $c \oplus a = c \oplus b \Rightarrow a = b$, then S is a group.

If G is a group and $a,b,c \in G$, then the following properties apply.

- $a \oplus c = b \oplus c \Rightarrow a = b$;
- $c \oplus a = c \oplus b \Rightarrow a = b$;
- $a \oplus x = b$ has the unique solution $x = a^{-1} \oplus b$; $x \oplus a = b$ has the unique solution $x = b \oplus a^{-1}$;

Given group $\{G;\oplus\}$ and $a \in G$. The following notations apply.

$$\begin{aligned} a^n &\equiv \overbrace{a \oplus \cdots \oplus a}^{n} && (2.8)\\ a^0 &\equiv e && (2.9)\\ a^{-n} &\equiv \overbrace{a^{-1} \oplus \cdots \oplus a^{-1}}^{n} && (2.10) \end{aligned}$$

where n is a positive integer, e the identity element of G and a^{-1} the inverse of a. With this notation, instead of (2.8), (2.9) and (2.10), the following properties apply.

- $a^m a^n = a^{m+n}$
- $(a^m)^n = a^{mn}$

When studying abelian groups, "+" is most commonly used to represent the operation, and the following notations apply.

$$\begin{aligned} na &\equiv \overbrace{a + \cdots + a}^{n}\\ 0a &\equiv e && (2.11)\\ (-n)a &\equiv \overbrace{(-a) + \cdots + (-a)}^{n} \end{aligned}$$

where $-a$ is used to denote the inverse of a. In this case, the identity element e in (2.11) is often denoted by "0". Note that in this context, 0 represents the

identity element of the abelian group and does not have to be the numerical number zero. When the group G is $\mathbb{Z}$, $\mathbb{R}$ or $\mathbb{C}$, and the operation is indeed summation, the identity element is indeed numerical zero.

The idea of group goes beyond number theory and mathematics. Generally speaking, any system that can move from states to states with movements revertible can be described by a group or something similar. When a system shows symmetry, it can probably be described by a group.

2.3.5 Order of Elements in a Group

Let G be a group, e its identity element, and $a \in G$. If there exists a smallest positive integer n so that $a^n = e$, we say that the **order** of a is n. If no such n exists, the order of a is infinity. The identity element e is the only element that has the order of 1. For an element with order n, its inverse also has the order of n.

Let G be a group, e its identity element, and $a \in G$. Consider the sequence of a^n as follows.

$$\ldots \quad a^{-3} \quad a^{-2} \quad a^{-1} \quad a^0(e) \quad a^1 \quad a^2 \quad a^3 \quad \ldots$$

If a has the order of infinity, then $a^m \neq a^n$ for any integers $m \neq n$, and vise versa. This can be easily proved with proof by contradiction. If a has the order of d, then $a^m = a^n$ if and only if $m - n = kd, k \in \mathbb{Z}$.

If $a \in G$ has order d, then $a^k, k \neq 0$ has order $d/(d, |k|)$, where $(d, |k|)$ is the greatest common divisor of d and $|k|$. This implies that a^k has order no more than d, and it has order d if and only if $(d, |k|) = 1$.

If $a, b \in G$ has order m and n respectively, and $ab = ba$, then ab and ba have the order of $[m, n]$, where $[m, n]$ stands for the least common multiple of m and n.

2.4 Subgroup and Quotient Set

Subgroup and quotient set are both fundamental concepts in group theory. They not only form an important part of abstract algebra by themselves, but also help with better understanding group and its properties. Notice that quotient set has been mentioned earlier during the introduction of the equivalence relation and equivalence class. More details are given here.

2.4.1 Subgroup

Let $\{G; \oplus\}$ be a group, and $H \subseteq G, H \neq \varnothing$ is a non-empty subset of the group. If $\{H; \oplus\}$ forms a group under the same operation $\oplus$, then H is called

a **subgroup** of G, denoted by $H \leq G$. Notice that if $H \neq G$, H is called a **proper subgroup** of G and can be denoted by $H < G$. That implies that

- The operation $\oplus$ which is closed in G is also closed in H;
- The identity element of G, denoted by e, is in H;
- For every element $a \in H$, its inverse a^{-1} is also in H.

Here is an example of a subgroup. Consider group $\{\mathbb{R}^*; \cdot\}$ where $\mathbb{R}^*$ denotes real number set excluding zero, and $\cdot$ the multiplication operation. Then group $\{\mathbb{R}^+; \cdot\}$ where $\mathbb{R}^+$ denoting positive real number set is a subgroup of $\{\mathbb{R}^*; \cdot\}$.

Here is another example of a subgroup. The **general linear group** of degree n consists of $n \times n$ invertible matrices with the operation of ordinary matrix multiplication. It is denoted by $GL_n(F)$ where F is the field / ring of interest such as the complex numbers field, integer ring, etc. More about field and ring are introduced in later sections and chapters. Obviously, $GL_n(F)$ is a group. The **special linear group** of degree n, denoted by $SL_n(F)$ which consists of matrices with determinant of 1, therefore, is a subgroup of $GL_n(F)$.

In practice, an invertible matrix A is equivalent with a linear transformation $x \mapsto Ax$. Therefore, $GL_n(F)$ can be taken as a group of linear transformation as well (hence the name), and from that perspective it is a subgroup of the **general affine group** $x \mapsto Ax + b$.

The following criteria can be used to determine whether a subset is a subgroup under the operation. Let $\{G; \oplus\}$ be a group, and $H \subseteq G$ a nonempty subset. The following three statements are equivalent.

(i) $H \leq G$;

(ii) $\forall a, b \in H$, $a \oplus b \in H$, $a^{-1} \in H$;

(iii) $\forall a, b \in H$, $a \oplus b^{-1} \in H$.

Furthermore, if $|H| < \infty$, the following two statements are equivalent.

(i) $H \leq G$;

(ii) $\forall a, b \in H$, $a \oplus b \in H$.

The proof is neglected in this notebook.

Subgroup has the following properties.

- If $H_1 < G$, $H_2 < G$, $H_1 \bigcap H_2 \neq \varnothing$, then $H_1 \bigcap H_2 < G$;
- If $H_1 < G$, $H_2 < G$, then $H_1 \bigcup H_2$ is not necessarily a subgroup of G.

2.4.2 Coset of a Subgroup

Let $\{G;\oplus\}$ be a group, $H < G$ a subgroup, and $a \in G$. Define **left coset** of H about a, aH, as follows.

$$aH \quad = \quad \{a \oplus h | h \in H\}$$

Similarly, the **right coset** Ha can be defined.

Let $\{G;\oplus\}$ be a group and $H < G$. Define the following relation for any two elements in G as follows.

$$aRb \quad = \quad \left\{(a,b)\middle| a^{-1} \oplus b \in H\right\} \tag{2.12}$$

It can be proved that the above relation is an equivalence relation and the equivalence class of a, $[a]$, is nothing but $[a] = aH$. The proof is neglected here.

This theorem associate a subgroup H with an equivalence relation (2.12).

2.4.3 Quotient Set of a Subgroup

Recall that in an algebraic system an equivalence relation is corresponding with a quotient set. Given a subgroup $H < G$, an equivalence relation (2.12) can be defined, hence the quotient set G/R. This quotient set is known as the left quotient set of subgroup H.

Part III

Number Theory

Bibliography

[1] Nathan Jacobson. *Basic algebra I and II.* Courier Corporation, 2012.